

Werk

Titel: Articles

Ort: Wiesbaden

Jahr: 2005

PURL: https://resolver.sub.uni-goettingen.de/purl?666048797_0009|LOG_0032

Kontakt/Contact

[Digizeitschriften e.V.](#)
SUB Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen

✉ info@digizeitschriften.de

□
□
□
□
□
□
□

Consonant assimilations: A possible parameter for the classification of Turkish dialects

Bernt Brendemoen

Brendemoen, Bernt 2005. Consonant assimilation: A possible parameter for the classification of Turkish dialects. *Turkic Languages* 9, 173-187.

This article poses the question whether preservation of the consonant group *-rl-* or assimilation to *-ll-* can be used as a parameter for classifying Anatolian and Balkan Turkish dialects. The material used is made up of different dialect texts, and also dialect studies from different parts of Anatolia and the Balkans. Aorist (and in some cases, also present tense) 3rd person plural forms constitute the main bulk of the material. The survey is complicated by the weak pronunciation of *-r* in syllable final position over a large area close to the Aegean. In Anatolia, the consonant group *-rl-* is preserved in the Eastern Black Sea dialects (i.e. the dialects in Trabzon except for the westernmost part, and the western parts of Rize) and in the area of Erzurum adjacent to Rize. In addition, preservation seems to be the main rule in an area roughly following the Euphrates (Malatya, parts of Elazığ, Diyarbakır, and Adana), and also in parts of Sivas. Preservation of *-rl-*, which may be interpreted as an archaism, is also found in the most archaic Turkish Balkan dialects, i.e. the West Rumelian dialects. That assimilation of this consonant group is recent in Anatolia and the Balkans also becomes obvious from the fact that it is not found in older (Ottoman or transcription) texts.

Bernt Brendemoen, Department of Culture Studies and Oriental Languages, University of Oslo, P.O. Box 1010, Blindern, NO-0315 Oslo, Norway. E-mail: bernt.brendemoen@ikos.uio.no

Introduction

Most of the parameters which are generally used in descriptions of Anatolian dialects were suggested by Kowalski, the father of modern Turkish dialectology, (1929-30) and later (1934) used by him not actually for a classification, since he probably thought too little material was available to attempt one, but for a survey of variable features in the Balkan and Anatolian dialects that had been studied by then. These parameters were later developed and discussed by Kral in his unpublished thesis from 1981, which was made known to the world by Boeschoten (1991), who also discusses briefly the significance and usefulness of the parameters. It is basically these parameters which that are used by Karahan in her important attempt to classify the Anatolian dialects (1996). In the present article I am going to take up one possible parameter which in fact was suggested earlier, but which has not been followed

up by any scholars because of certain difficulties it implies, i.e. the parameter of consonant assimilations.

Various kinds of consonant assimilations are found in most languages in the Turkic family to a greater or lesser extent. The most common one is systematic voicedness assimilations at morpheme boundaries, the so-called consonant harmony, but assimilations (and also dissimilations) as to *manner* of articulation are also found, such as the systematic alternations shown by the plural suffix {-LAR} in Kazakh, where the plural of *at* ‘horse’ is *at-tar*, while the plural of *köl* ‘lake’ is *kölder*, cf. the plural of *taw* ‘mountain’ in Bashkir, which is *taw-đar* and of *nom* ‘book’ in Tuvan, which is *nomnar* (see Johanson 1998: 34). What I am especially going to concentrate on here, is the regressive assimilation of *-r+l-* to *-ll-* particularly at morpheme boundaries, which is very well illustrated by the different shapes of third person plural aorist forms, as the aorist stem invariably ends in *-r* and the plural suffix {-LAR} starts with an *-l*. For example, the aorist stem of the verbal root *yap-* ‘to make’ is *yapar*; 3rd person plural is *yaparlar* without assimilation, but with assimilation *yapallar*. In the same way, in the present tense most Turkish dialects have a suffix in *-yor*, so that the stem signifying ‘is doing’ is *yapıyor*; 3rd person plural *yapıyorlar* without assimilation, but with assimilation *yapıyollar*. However, since present tense formations without an *-r-* are very old and may partly represent an archaism, present tense forms are less suitable as evidence for assimilation or non-assimilation than aorist forms are. Similar assimilations occur frequently within polysyllabic stems, too, e.g. Standard Turkish and some dialects have *tarla*, while other dialects have *talla* ‘field’. That the assimilation *-rl-* to *-ll-* is old at least in some Turkic languages is obvious from the name of the Turkic tribe *Qarluq* cited as *Xallux* by the Persian historian Gardīzī (middle of 11th century, Golden 1992: 198), but as we shall show, it is probably quite recent in West Anatolia.

Assimilation to a following *-l-* is not the only thing that may happen to a syllable-final *-r* in Turkish; in some areas it disappears, giving forms such as *yapa:lar*, *gide:ler*, *gidiyo:lar*, sometimes with lengthening of the preceding vowel, sometimes without. The title of this paper could in fact have been “The destiny of syllable final *-r-* before *-l-* in Turkish dialects”.

Already in 1911 the Bulgarian philologist Gadžanov, who studied the different Turkish dialects of Bulgaria, suggested that the different treatments of *-r-* in syllable-final position should be used as a parameter in Turkish dialectology (passim, especially p. 42). Although the different treatments of *-r-* were listed as possible parameters for a classification by Kowalski in his 1930 study (p. 273, 274-275), Kowalski points out in his 1934 survey (p. 2003) that most consonant assimilations occur only sporadically, and that “eine Bestimmung der Gebiete, auf denen sie auftreten, lässt sich gegenwärtig noch nicht ausführen [it is currently not yet possible to determine the areas where they occur]”. However, his wording “gegenwärtig noch nicht [currently not yet]” implies that he does not exclude that they may have a certain geographic distribution.

If we look at studies on individual dialects, we see that in most of them, examples of assimilation *-r/- > -ll- >* are mentioned, and very often these examples are 3rd person plural forms, but quite rarely does the scholar in question provide us with any information on whether the feature occurs *as a rule* or only *sporadically*. If we go through the text material the different scholars present, we very often find unassimilated forms alongside assimilated ones, which may of course very well be the case even in one and the same speaker; the frustrating point is that this variation is not mentioned or evaluated by most of the scholars. Thus, descriptions of a dialect without texts may not be trusted. Another factor that makes the matter difficult is that because of the weak articulation *-r-* has in syllable-final position, it may be difficult to decide acoustically whether it actually is there or not. Kowalski remarks about consonant assimilations that: “In den Volksdialekten treten sie schon deswegen häufiger auf, weil da die normierende Wirkung des Schriftbildes fehlt. [They occur more frequently in the spoken dialects because of the absence of the normative effect of the written language.]” (1934: 2003). In fact we may turn this statement upside down and say that because of the normative effect of the written language, a dialectologist or would-be dialectologist will tend to transcribe the forms as unassimilated, unless he or she is especially focused on consonant assimilations, which nobody seems to have been up to now.

A third complicating factor is the lack of dialect material from all areas of Turkey; if our intention is to give a survey of a linguistic phenomenon in Anatolia as a whole, there will be quite a number of blank spots on the map, but this is of course a problem that affects Anatolian dialectology in general.

Despite of all these complications impairing the transparency of the picture, I am quite certain that consonant assimilations are an important parameter, based on my Trabzon dialects. In most of the province of Trabzon, *-r/-* is usually not assimilated. However, if we go to the westernmost part of the province, which is inhabited by the so-called Çepnis, an originally nomadic group distinct from the rest of the population of Trabzon, we find examples such as *talla* for *tarla* ‘field’, *tallarmıza* (Brendemoen 2002: 2, text 138/2) and also *billäşmişlär* (133/122), *davalların* (134/3), *yelläşiyorlar* (106/42), and aorist forms such as *dellerde* (135/65), *dellerd-ona* (138/6), etc. Such forms are much more rarely found further to the east in the province of Trabzon. In fact, as I have shown (2002: 1, 226), there is an important isogloss bundle dividing the westernmost parts of Trabzon, which belong to the West Anatolian group, from those spoken to the east of this line, which constitute the Eastern Black Sea Dialect group. However, where assimilated forms are found in the districts further to the east, it is significant that these districts are mountainous areas in the southern part of the province, to which e.g. text 86 belongs, where we also find the form *deller* (86/79). In fact informants such as 86 have other Çepni features in their dialect, too, which are a product of the symbiosis and linguistic convergence between the different dialect groups in the regions close to the summer pastures.

Survey of Anatolia

Before we proceed to the difficult task of giving a survey of the tendencies we have been able to establish in other parts of Turkey, based on the published studies available, one complicating factor we already have mentioned should be elaborated further, i.e. the weak articulation of syllable-final *-r* in quite a number of Anatolian dialects, which indeed causes systematic dropping of syllable-final *-r* in some dialects. A weak articulation of *-r* is not confined to Turkish of Turkey, but is also found in older stages of development of the Turkic languages, and must be the reason why the syllable-final *-r* has been dropped in forms used especially frequently, such as the Old Turkic copula verb *är-* which has developed into *i-* or *e-* in most languages, and intraterminal or present tense markers *durur* and *yorur*, which have developed into {-*DI-*} and *-yo-* and *-yu-* in various languages and dialects. Even in most spoken varieties of modern Standard Turkish, although under strong influence from the written language, the indefinite article *bir* usually has the form *bi* in front of consonants. (For the weak articulation of *-r* in Standard Turkish, see Bergsträßer 1918: 251.) In our Eastern Black Sea Coast dialects, syllable-final *-r* may be dropped in both Trabzon and Rize, but mostly in absolute auslaut position, especially in the copula suffix corresponding to ST {-*DXr*}. At least there is no *systematic* dropping of syllable-final *-r* in the Eastern Black Sea coast dialects. In East Anatolian dialects, too, dropping of syllable-final *-r* is quite rare, at least in inlaut position, but is found e.g. in Azerbaijani dialects (e.g. Terekeme dialects) in Kars, see Gemalmaz 1978 I: 201, Ercilasun 1983: 174.

Dropping of syllable-final *-r* as a *systematic* feature, however, has its nucleus in western Anatolia, i.e. in the provinces of İzmir, Manisa, Uşak, and Aydın. Sometimes the vowel preceding the *-r* is lengthened as a compensatory device, sometimes it is not, giving aorist forms such as *gideler* and *gide:ler*, *yapalar* and *yapa:lar*, besides, of course, auslaut forms such as *va* (ST *var*), plural suffix {-*IA*} (and not {-*IAr*}), etc. The resulting homonymy with optative forms for verbal stems that get a low aorist vowel does not seem to have prevented the dropping, probably because the optative 3rd person forms are marginal, almost obsolete, in most modern dialects. In her study of the Southwest Anatolian dialects, Korkmaz (1956) has shown that dropping of *-r* is less frequent going south to the southern parts of Denizli, and Muğla. She further states that dropping of *-r* is found also in the Kastamonu region and in some Eastern and Northeastern dialects, especially in the copula 3rd person form {-*DX*} and not {-*DXr*}, plural {-*IA*} and not {-*IAr*} and other auslaut positions; more rarely in inlaut position (1956: 79). For the Kastamonu region her claim is correct if we compare Ergi's little study (1991) of the dialect of Tosya in the province of Kastamonu (*alula, gelüle* p. 5). It should be added that according to my own observations, dropping of *-r* is common also to the north of İzmir all the way up to Edremit, perhaps also further north. Besides, Gülensoy (1988: 65, 105) has established the same fact for Kütahya.

A weak articulation of *-r-* is perhaps a phonetical prerequisite for both dropping it and for assimilating it to a following *-l-*. However, it should be stressed that because consonant length is a phonemic feature in practically all Turkish dialects, it would be difficult to imagine that any of the forms with a simple consonant with or without lengthening of the preceding vowel (i.e. *gideler* and *gide:ler*) could be explained as a secondary simplification of an assimilated form *gideller*. Nevertheless, judging from Korkmaz' material, there seem to be dialects in western Anatolia where forms such as *gideller* and *gide:ler* exist side by side even in one and the same speaker e.g. in the dialects in İzmir and east and south of İzmir (see Korkmaz 1956: 75).

If we disregard the area in West Anatolia where syllable-final *-r* is dropped systematically and start with an area we know especially well, i.e. the Eastern Black Sea coast, we notice that assimilation of *-rl-* > *-ll-* is found frequently, as we pointed out above, in the westernmost parts of Trabzon. This continues in the areas going westwards along the Black Sea coast. The preservation of the consonant group *-rl-*, which we find in most parts of Trabzon, however, continues into the province of Rize, which has been studied by Günay (1978). Further eastwards along the coast we mainly have to do with an East Anatolian dialect with a Kartvelian sub- or adstrate, as languages such as Laz and also Georgian are spoken in the area. Since Turkization of the whole eastern Black Sea coast seems to have taken place mostly from the inland and not from the sea, it is probable that the regions east of Rize have been Turkized relatively recently through the Çoruh and Tortum valleys, which constitute a funnel from the Erzurum area, where consonant assimilations of this kind occur frequently. Thus it is no surprise that we find assimilated forms in the easternmost parts of Rize, e.g. *kuṛal'l'aṛ*, *ḡidelleṛ* (Günay 1978: 130-131). In the dialects south of the Pontic mountain ridge, which belong to the East Anatolian dialect group, assimilation of *-rl-* > *-ll-* is a rule. Most fortunately Gemalmaz, the scholar who has prepared an unsurpassed study on the dialects of Erzurum, explicitly states that there is an isogloss running through the province of Erzurum constituted by the different realizations of the sequence *-rl-*. As illustrated on Gemalmaz' map no. 8 (1978, I: 210-211), the northern part of Erzurum has present and aorist 3rd person plural forms without assimilation, while those in the south have assimilation. It is instructive that the unassimilated forms are found in areas adjacent to (the western and central parts of) Rize, where there is no assimilation either, as we pointed out. A similar picture is found in Gümüşhane, the province immediately to the south of Trabzon, although the statements giving by San (1990) as usual are quite confusing. It seems, however, that assimilation of *-rl-* is found especially frequently in areas bordering the western parts of Trabzon, i.e. Kürtün (San 234, 260), where also a lot of other features are shared with the Çepni dialects of Trabzon. It also seems that the easternmost parts of Gümüşhane have the same feature, which would then be a continuation from Erzurum.¹

□

¹ Although San's claim (p. 153) that assimilation is especially frequent in aorist and present tense forms in the regions of Akdağ and Aydoğdu in Kelkit in the southern part of the

Whether the lack of assimilation in the remaining parts of Bayburt should be interpreted as a continuation of the situation in the Eastern Black Sea dialects or of the ambiguous situation in the Central Anatolian dialects, to which we shall return in a moment, is unclear. In East Anatolian dialects, however, assimilation is a rule. The East Anatolian dialects that have been subject to the most thorough research, Erzurum and Kars (Ercilasun 1983: 127), show this clearly (except for the northernmost part of Erzurum close to Rize which we mentioned). Assimilation is most probably typical for East Anatolian dialects further to the south, too, although lack of proper material leaves us somewhat uncertain.² At least in Erzincan, which is south of Bayburt and west of Erzurum, (for dropping of *-r-*, see Sağır 1995: 104-105), assimilation of *-rl-* to *-ll-* is very frequent, at least in the aorist 3rd person plural forms (ibid. 184), which are listed only with assimilated forms in the work by Sağır (1995: 117), e.g. *édeller*, *ğırallar* etc. (However, in the present tense conjugation, assimilated forms such as *vèrèyeller*, *yapèyèller* are, for unknown reasons, mentioned as characteristic only of the region of Kemah, while the other regions have unassimilated forms.) Assimilation is also the rule in the continuation of the East Anatolian dialects southwards into Iraq (see Bayatlı 1996: 366), and also the continuation of East Anatolian dialects into Iran and Azerbaijan. In fact, in the Azeri of the Republic of Azerbaijan, which has a rather conservative orthography, the assimilation *-rl-* > *-ll-* is not shown in writing, but is a rule in practically all kinds of spoken language (see Ergin 1971: 128). (Except for some dialects, especially in Iran, which have a wider range of assimilation possibilities such as *addar*, *koyunnar*, *karrar*, corresponding to ST *karlar*, etc., see Dehghani 2000: 47). The same kind of assimilation is a rule also in other Turkic languages of Iran, such as Khalaj (see Doerfer 1988: 164, 200),³ Kashkay, and Khorasan Turkish.

If we return to Anatolia, or more precisely to the part of Western Anatolia south of the area which constitutes the nucleus for the dropping of *-r* in syllable-final position, we find non-assimilated forms such as *bişirlä*: and *içärlä* alongside forms with *-r* being dropped such as *verilä* in one and the same text from the province of Muğla. (Korkmaz 1956: 100). From Alanya on the southern coast Demir presents some very trustworthy texts in his work on postverbal constructions in the dialect of his village (1993). In this dialect, assimilation of *-rl-* to *-ll-* is a rule, as in the east, as

□

province is not developed further in his morphological survey, it is most probably true, as this region is adjacent to Erzincan.

² For Urfa, see Edip 1991: 47 (*satallar*, with the confusing footnote that “(r - l) benzeşmesiyle *satallar* şeklinde de kullanılır”), 48 (*geliller*). For Gaziantep, see Aksoy 1945: 55 (“Geniş zaman kipinin üçüncü çoğul şahsındaki “ler” takısı kendisinden evvelki “r”yi çok defa “l”ye çevirir: *gideller*, *geliller*.” etc. It is quite unclear what “çok defa” actually signifies.) For Bitlis, cf. the following forms in the texts given by Zülfikar 1978: *yatelle* p. 311, *gelülle*, *bağelle* p. 312, *annedülle* p. 314, etc. For Mardin, no research seems to have been conducted.

³ For the dropping of final *-r* in the aorist see Doerfer 1988: 153s.

becomes obvious from forms such as *döwällär* 157/58, *gızıyörüllär* 157/81, and *del-lärmiş* 163, 5/2. The same seems to be the case with the Turkish dialect of Cyprus (see Saracoğlu 1992: 24).

In the Central Anatolian dialects further north, assimilated and non-assimilated forms are often found side by side. There is reason to believe, however, that assimilation *-r/- > -ll-* is quite common also in parts of Central Anatolia which are more or less blank on the dialect map. Especially in the western parts, such as Nevşehir, assimilations are quite frequent, as is also dropping of *-r*, giving forms like *çekeller*, *ekeller*, and *a:nadırlar* side by side (Korkmaz 1963: 128, 174). A similar picture with frequent assimilations is found also in the western parts of Konya (see Gültekin 1994: 39). We then go northwards to the Middle Black Sea coast, where Korkmaz, as mentioned above, has found numerous attestations of dropping of syllable-final *-r* in the dialect of Kastamonu. To the west of Kastamonu, this feature is found alongside assimilation of *-r/-* to *-ll-*. In Eren's 1997 study of the Western Black Sea coast dialects (Zonguldak-Bartın-Karabük), 3rd person plural aorist forms such as *dut-a:-la:* ~ *dut-al-la*, *gid-e:-le:* ~ *gid-el-le:*, *al-u:-la:* ~ *al-ul-la:* are listed as parallel forms without any comments on their geographic distribution (p. 68).⁴ In the eastern part of the Middle Black Sea coast (Ordu-Giresun), however, assimilations seem to be the rule. We may deduce this from examples in Caferoğlu's 1946 text anthology from the region, such as *asaller* (p. 9), *getürüller* (p. 10), *ğellüller* (p. 40).⁵ Assimilation is stated by Aydın (2002: 33) to be a rule in 3rd person plural present and aorist forms in the dialect of Aybastı in the southern part of Ordu, and also for the whole province of Ordu in general by Demir (2001: 90). This kind of assimilated forms go on, as stated above, until immediately east of the border to Trabzon.

If we go inland from the eastern part of the middle Black Sea coast dialects to Sivas, non-assimilated forms seem to appear more frequently, cf. Räsänen's texts from the area, where forms such as *oxurlar*, *öp'erler*, *Bayramlaşırar*, *Barışırar*, *Dêrler* occur alongside *gulallar*, *atallar*, *gêdeller*, and *giyeler* in one and the same text (1933: 50). If we go southeast, we find an area before we meet the East Anatolian dialects comprising at least Malatya, parts of Elazığ, Diyarbakır, Adana, and perhaps also Maraş, where assimilation of *-r/-* is not found to any extent. For the western parts of Elazığ, situated to the immediate east of the Euphrates river, which roughly constitutes the boundary between East and West Anatolian dialects, assimilated forms are found in the Keban dialect investigated by Buran (1997: *gidallar*, *yapallar*, *taxallar* text 1/19), but not in the dialects of the districts Baskil and Ağın: *dêrlardı* (text 40/43, but *göturullardı* text 44/38), *çağırırar* (text 47/34), *saxlırler*, *ğorxırler* (text 47/48), etc. In the same way, in the city of Elazığ, although singular assimilations such as *talla*, *söleller* do occur (Güler 1992: 29), the usual aorist 3rd

□

⁴ Of these two mechanisms in 3rd person plural forms, only dropping of *-r* is mentioned by Korkmaz in her 1965 study of the Bartın dialects (p. 21, 26).

⁵ Assimilation is common outside aorist forms, too, cf. *hatılladı* (p. 6), *veziller* (p. 10), *tal-laya* (p. 35), *zolliyalar* (p. 35), *gattelleri* (p. 36).

person plural forms show no assimilation, e.g. *dinnenürler*, *alurlar*, *açarlar* (ibid., 35-36). The lack of assimilations in Malatya immediately on the west side of the Euphrates is supported by forms such as *ederlär* in a recording I have made in the village of Korucuk, and also plural forms such as *biberler* and not **bibeller*. Gülseren's study (2000) of the dialects in Malatya—it is in fact called *Malatya İli Ağızları*, i.e. "The dialects of Malatya", but contains no attempt to draw internal dialect boundaries in the region, so it could just as well have had the title "The dialect of Malatya"—is in fact one of those studies where the grammatical part says one thing but the texts say something completely different: The phonological and morphological sections claim that assimilation *-rl-* > *-ll-* is "very common" in the aorist 3rd person plural (the present tense formation is different in these dialects), and the author gives examples such as *tikellerdi*, *yapallar*, etc. (p. 93). However, if we look at Gülseren's texts, we see at once that unassimilated forms by far outnumber the assimilated ones, e.g. *gëyindirirler*, *atarlar* (p. 258), *getirirlerdi*, *götürürlerdi* (p. 269), *danişirlär*, *dökerler* (p. 283), etc. For the lack of assimilations in Adana, I rely on oral information from my colleague in Mainz, Dr. Christiane Bulut, who knows this dialect very well. It is quite interesting to see that the dialect of Diyarbakır, too, very rarely has examples of assimilation *-rl-* > *-ll-*. This is expressed explicitly by Erten in his 1994 study (p. 16) and confirmed by his texts. He claims on the other hand that present tense forms such as *deyisiz* are examples of dropping of *-r-* (p. 17). This should, however, be taken with a grain of salt because the present tense paradigm usually has no *-r-* in this dialect, thus giving pairs such as *biçiler* 'they are cutting' (present) vs. *biçerler* (aorist).

In spite of the lack of sufficient material, it seems possible to establish a parallel between the Eastern Black Sea coast dialects and the borderland between West and East Anatolian dialects further south, partly along the Euphrates, comprising an area whose size we still do not know, but which comes down to the Mediterranean at Adana.

We should ask if this correspondence has any parallel in other parameters, i.e., if any other isoglosses follow the same path, and they do indeed. In my study on the Trabzon dialects, I have pointed out parallels in the present tense formation between approximately the same districts (2002: 1, 257-262). As the Euphrates roughly coincides with the border between the East and West Anatolian dialects, we may assume for some features the existence of nucleus areas on both sides exerting their influence from the west towards the east and from the east towards the west. Thus, as I have shown in Brendemoen 2005, in the field of Arabic loanwords, Tebriz in the east, the capital of the Akkoyunlu Turks in the latter part of the 15th century, and later of the partly Azeri-speaking Safavid dynasty, must have been a nucleus from which Arabic loanwords in a Persian phonological shape were diffused in all directions, while, at probably a somewhat later date, Arabic loanwords in a more learned shape imitating Classical Arabic were diffused from Istanbul, the capital of the Ottomans. In the case of the present tense formation, however, the picture is somewhat different, but perhaps resembles the case of assimilation *-rl-* > *-ll-* even

more: After a new present tense formation with the auxiliary verb *-yor-* came into being perhaps during the 15th century, it was diffused from the cultural centres in West Anatolia to most parts of the Balkans and eastwards into Anatolia, while the Azeri present tense realization in {-Xr} (*gâkir, yapır*, etc.), which also is developed from the same auxiliary, was diffused westwards from the Azeri cultural centres in Iran. In an area where these present tense formations met, or perhaps did not meet because the force of their diffusion was not strong enough, other present tense formations without an *-r-* exist, most probably as an archaism, i.e., as a relic of an earlier present tense formation which has been lost under the pressure of the new present tense formations elsewhere in Anatolia. In the same way, we may perhaps regard the strong tendency to assimilate *-rl-* to *-ll-* as a feature that was diffused westwards into Anatolia from Azerbaijan in the east. In the west, however, the tendency to drop syllable final *-r-* in Western Anatolia has spread eastwards, possibly preventing the tendency of assimilations to proceed very far west, and accordingly an area where the consonant group *-rl-* has been preserved as an archaism has remained in the middle. Whether this interpretation is correct or not is to some extent dependent on how the situation really is in the blank spots of the map, which partly have not been investigated, partly present a confusing picture.

The desire to avoid homonymy may also have been a factor in the preservation of the consonant group *-rl-* in the area in the southeast: As the present tense is formed with the suffix *-i*, present and aorist 3rd person plural forms of verbal stems that take the aorist vowel *-i* would almost become homonymous, being *geliler* and *geliller* respectively. It should also be mentioned that today's Istanbul dialect usually does not have assimilation of *-rl-*, but this may of course be a secondary feature due to copying from the written language. However, as I shall show, it does not seem that the Istanbul dialect at any point has had assimilation of *-rl-* to *-ll-* as a feature.

Survey of the Balkan dialects

I think the picture will gain some clarity, especially as to the diachronical aspect of the different treatments of syllable-final *-r*, if we have a look at the Turkish dialects on the Balkans.

Although the areas in the Balkans where Turkish is spoken have diminished dramatically not only because of the collapse of the Ottoman Empire, but also as a result of political events in more modern times, the Turkish Balkan dialects are quite well documented thanks to work done in the first and middle part of the 20th century. The classification of the different dialects is still disputable, but at least it is quite certain that an important dialect boundary runs through Bulgaria in a north-south direction not far east of Sofia. The dialects to the west of this line are the so-called West Rumelian dialects, which generally have preserved more archaic features than the ones further to the east. This is especially the case with the dialect spoken along the Danube in places such as Vidin and Lom, which constitute a very isolated corner of the Turkic-speaking world. These dialects, which have been the subject of a famous

study by Németh (1965), do not assimilate *-rl-* to *-ll-*. As for the dropping of syllable-final *-r-*, this does not seem to happen to any greater extent, either. This is also the case in other West Rumelian dialects, such as dialects in Macedonia like the one in Dinler (close to Ovčepole in the sub-district of Štip) and the dialect of Komanova or Komanovo, which both were studied by Eckmann (1960 and 1962); cf. aorist plural forms such as *tanarlar*, *dünerler*, *süylerler* (1962: 122). The same is the case with the dialects of Küstendil to the southeast of Sofia, and Michailovgrad (new name: Montana) between Sofia and Vidin, studied by Kakuk (1961), who does not mention assimilation *-rl-* > *-ll-* at all.

In all these West Rumelian dialects, however, other assimilations, such as *-nl-* > *-nm-* are quite common, see Kakuk 1961: 314. The same lack of assimilation *-rl-* > *-ll-* is found in Gagauz, a variety of Ottoman Turkish which has been transformed thoroughly through its symbiosis with Slavic languages, especially Russian, in a very isolated situation in the northern parts of Romania and Moldavia. In Gagauz, *-r-* is sometimes dropped in syllable final position (see Özkan 1996: 86), but assimilation *-rl-* > *-ll-* does not occur (while *-nl-* > *-nm-* is very common, cf. p. 79), and aorist 3rd person plural forms are *därlär*, *başlarlar*, *alarla*, *bilärlär*, etc. (p. 146).

Another dialect group in the Balkans is the so-called Deliorman dialect group in North East Bulgaria as exemplified by the dialect of Razgrad, which was studied by Eckmann (1950a). This dialect is similar to the ones in West Anatolia in the respect that *-r-* tends to be dropped in syllable-final position (p. 12). Assimilation *-rl-* > *-ll-* also occurs, but in the aorist, 3rd person plural forms with dropping of *-r-* and compensatory lengthening (*kalı:lar*, *geli:ler*, p. 16) seem to be the regular ones. (The present formation does not have an *-r-* in 3rd person plural, either, but this may be due to other factors.) This picture is supported by the later research performed by Dallı (1976, cf. p. 108-110 and forms such as *sürdüre:le*, *indire:le*, *yiye:le* p. 152). If we move a little south down to the wide plain that is confined by the Balkan mountains proper in the north and the Rhodopes in the south, we find dialects such as the one at Kazanlık, studied by Kakuk in 1958. Here, the tendency of *-r-* to be dropped in syllable-final position goes on, but as Kakuk remarks, assimilation, producing a double *l*, occurs quite frequently, too, giving examples such as *götürüler* besides the more frequent *ayırullar*, cf. 178.⁶ If we go further to the south, however, up into the Rhodope mountains close to the Greek border, where the dialect of the township of Kırçalı has been studied by Hazai (1959), amongst others, we see that *-r-* tends to be dropped only in absolute auslaut position, but that it is assimilated to a following *-l* as a rule (p. 218), thus giving aorist tense forms such as *gidällä*, *çıkalla* (225).

Thus we see that the three different destinies of syllable-final *-r-* we find in Anatolia, also are represented in the Balkans. The dialects where *-rl-* is preserved are

□

⁶ The loss of the *-r-* in the present tense paradigm affects all persons; still forms such as *içiyollar* do occur, *ibid.* In Kakuk's texts forms such as *alırlar* occur side by side with *takallar* and *karşılarlar* in one and the same text.

the most remote ones, and are archaic also in other respects;⁷ the ones where *-r* is usually dropped, are less remote, and are partly found along the main route of commerce and migration in Southern Bulgaria, alongside the dialects where assimilation takes place.

The Balkan dialects have had a great impact on the shaping of the Istanbul dialect, and most probably did so already at the time the city was conquered in 1453. The Ottoman court that established itself in Istanbul came there from Edirne in Thrace, which had been the Ottoman capital since 1362. By then, the Balkans, although having been in the hands of the Ottomans for less than a century, had started to become extremely important, and because there was no aristocracy in the Balkans that could prevent the Ottomans from establishing their different institutions—in contrast to the case in Anatolia—the Balkans enjoyed top priority from the Ottoman authorities. An important factor in this was the fact that the Ottoman officials to a great extent were recruited from Christian families in the Balkans. The kind of Balkan Turkish that had an impact on Istanbul Turkish most probably was the kind which is today represented by the most remote dialects, i.e. the West Rumelian dialects. It is probable that the dialects having extensive dropping of *-r* represent a later development, perhaps a later wave of immigration to the Balkans from Anatolia, and, as we have already indicated, that the tendency to prefer assimilations represents an even later stage of development. On the other hand, as soon as the Ottoman written language was established, the prestige Istanbul Turkish enjoyed as a codified language would no doubt in itself prevent assimilations and dropping of *-r-* from taking place.

Survey of older texts

Nevertheless it is an astonishing fact that older Ottoman texts written either with the Arabic or other alphabets do not give *any* examples of assimilation *-r/- > -ll-*. In the case of texts written by Turks, this could of course be explained as the effect of a graphic convention, but still it is remarkable that no example whatsoever reflects the actual pronunciation if assimilation did take place in the spoken language. Accordingly we are tempted to assume that it did not take place in the spoken language. In the so-called transcription texts, i.e. texts written mostly by foreigners, such as conversation guides, no reflexes of assimilation *-r/- > -ll-* are found either. Although most older texts represent Istanbul Turkish, texts known to have been written in Anatolia (such as *Süheyl ü Nevbahâr*, see Banguoğlu 1938: 11) do not provide us with examples of *-r/-* assimilation either. For the transcription texts this could perhaps be explained by the fact that the authors knew Ottoman orthography and transferred the principles of that to their home-made orthography in Latin, Greek, or

□

⁷ E.g., the dialects in Northwest Bulgaria have a present tense formation without an *-r* not very different from the ones found in Trabzon and in certain East Anatolian dialects, see Németh 1965: 84-86, Brendemoen 2002 1: 259-262.

Armenian script, regardless of how the actual pronunciation was in the spoken language.⁸ This is, however, contradicted by the fact that other kinds of assimilations are quite common in these texts, such as *-nl-* > *-nn-* and *-ls-* > *-ss-*, e.g. *bunmar*, *olsunnar*, *günnük*, etc. (in Karamanlidic texts, see Eckmann 1950: 196-197). This kind of assimilations, which all are at variance with Ottoman orthography, are also found in older transcription texts such as the Mühlbacher text from the 15th century (*miskinner*, *ossun*), and in the grammar by Pietro della Valle from the beginning of the 17th century (oral communication from Dr. Heidi Stein), it is stated explicitly that assimilated forms such as *ossun*, and *ossunlar* (for *olsun*, *olsunlar*) are characteristic of everyday speech. The only case where the assimilation of *-rl-* to *-ll-* is attested in older texts, is to my knowledge in Azerbaijani, but not Ottoman manuscripts (e.g. in Foy's 1903 study, where Azeri forms such as *olulla* (corresponding to ST *olurlar*) are mentioned (p. 193)). Thus, in a manuscript of the poems by the Azerbaijani poet Fuzûlî from the beginning of the 16th century, copied in Kerbelâ' in Iraq in 1576, forms such as *söyleller*, *dönderüller*, *açallar* are found (Olçay 1956: 38), showing that this kind of assimilations are quite old in Azeri.

Examples of dropping of syllable-final *-r* are, however, found in other texts, if not very frequently in older texts, e.g. in absolute auslaut position in *durla*, *dirilirle* in the so-called Mühlbacher text with Latin script from the 15th century (see Foy 1902: 241, 272-273). Evliyâ Çelebî's autograph manuscript of his monumental *Seyâhatnâme* from the 17th century has forms such as *deler*, *olular*, and *vadır*; in fact in some cases an *-r-* has been added to the manuscript, indicating perhaps that the forms without an *-r-* are oral forms (see Duman 1995: 26-27). In the same way, Evliyâ himself tells us that "tanners and rebellious artisans" used such expressions as *öldi* 'kill!' where the standard language had, as today, *öldür* (see Dankoff 1990: 89).⁹

Conclusion

Thus, *one* conclusion of this paper would be that the assimilation of the consonant group *-rl-* to *-ll-*—contrary to the assimilation of other groups—is quite recent in Western Anatolia and the Balkans, and considerably more recent than the dropping of *-r*, and also that it was imported from the east.

□

⁸ Thus, the so-called Harsány text, which is a conversation book from the 17th century (see Hazai 1973), contains no example of assimilation or of dropping of *-r*. The very few examples of other kinds of assimilations are mostly in learned words whose spelling not everybody would know in Ottoman either (p. 353-354). The case with the so-called Georgievits text from the middle of the 15th century (see Heffening 1942) is exactly the same.

⁹ It should be added that the argument that dropping of *-r* in aorist forms would cause homonymy with optative forms in the case of verbal stems that take a low aorist vowel, could be quite relevant as a possible reason why more extensive dropping did not take place in older periods, since the optative was a much more central tense in former times than it is today.

The other conclusion must be that the different destinies of syllable-final *-r-* is a relevant parameter indeed, but that—as with several other of the parameters used—certain difficulties have to be overcome before we can properly take advantage of it.

References

- Aksoy, Ömer Asım 1945-1946. *Gaziantep ağzı 1-3*. İstanbul: İbrahim Horoz Basımevi.
- Aydın, Mehmet 2002. *Aybastı ağzı (İnceleme-metin-sözlük)*. (Türk Dil Kurumu Yayınları 796.) Ankara: Türk Dil Kurumu.
- Banguoğlu, Tahsin 1938. *Altosmanische Sprachstudien zu Süheyl-ü Nevbahar*. Leipzig: August Pries.
- Bayatlı, Hidayet Kemal 1996. *Irak Türkmen Türkçesi*. (Türk Dil Kurumu Yayınları 664.) Ankara: Türk Dil Kurumu.
- Bergsträßer, Gotthelf 1918. Zur Phonetik des Türkischen nach gebildeter Konstantinopler Aussprache. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 72: 233-262.
- Boeschoten, Hendrik 1991. Aspects of language variation. In: Boeschoten, Hendrik & Verhoeven, Ludo (eds.) *Turkish linguistics today*. Leiden: Brill. 150-176.
- Brendemoen, Bernt 2002. *The Turkish dialects of Trabzon. Their phonology and historical development 1: Analysis, 2: Texts*. (Turcologica 50.) Wiesbaden: Harrassowitz.
- 2005. Ottoman or Iranian? An example of Turkic-Iranian language contact in East Anatolian dialects. In: Johanson, Lars & Bulut, Christiane (eds.) *Turkic-Iranian contact areas. Historical and linguistic aspects*. (Turcologica 62.) Wiesbaden: Harrassowitz.
- Buran, Ahmet 1997. *Keban, Baskil ve Ağın yöresi ağızları*. (Türk Dil Kurumu Yayınları 669.) Ankara: Türk Dil Kurumu.
- Caferoğlu, Ahmet 1946. *Kuzey-Doğu illerimiz ağızlarından toplamalar*. İstanbul: Bürhaneddin Erenler Matbaası.
- Dallı, Hüseyin 1976. *Kuzeydoğu Bulgaristan Türk ağızları üzerine araştırmalar*. (Türk Dil Kurumu Yayınları 450.) Ankara: Türk Dil Kurumu.
- Dankoff, Robert 1990. Turkic languages and Turkish dialects according to Evliya Çelebi. In: Brendemoen, Bernt (ed.) *Altaica Osloensia—Proceedings from The 32nd Meeting of the Permanent International Altaistic Conference*. Oslo: Universitetsforlaget. 89-102.
- Dehghani, Yavar 2000. *A grammar of Iranian Azari*. Munich: Lincom.
- Demir, Necati 2001. *Ordu ili ve yöresi ağızları. İnceleme-metinler-sözlük*. (Türk Dil Kurumu Yayınları 788.) Ankara: Türk Dil Kurumu.
- Demir, Nurettin 1993. *Postverbien im Türkeitürkischen unter besonderer Berücksichtigung eines südanatolischen Dorfdialekts*. (Turcologica 17.) Wiesbaden: Harrassowitz.
- Doerfer, Gerhard 1988. *Grammatik des Chaladsch*. (Turcologica 4.) Wiesbaden: Harrassowitz.
- Duman, Musa 1995. *Evliya Çelebi Seyahatnamesine göre 17. yüzyılda ses değişimleri*. (Türk Dil Kurumu Yayınları 616.) Ankara: Türk Dil Kurumu.
- Eckmann, János 1950a. Razgrad Türk ağzı. In: Eren, Hasan & Halasi-Kun, Tibor (eds.) *Türk dili ve tarihi hakkında araştırmalar 1*. Ankara: Türk Tarih Kurumu. 1-25.
- 1950b. Anadolu Karamalı ağızlarına ait araştırmalar. *Ankara Üniversitesi Dil ve Tarih-Coğrafya Fakültesi Dergisi* 8: 165-200.
- 1960. Dinler (Makedonya) Türk ağzı. *Türk Dili Araştırmaları Yıllığı—Belleten* 1960: 189-204.

- 1962. Kumanovo (Makedonya) Türk ağzı. In: Eckmann, Janos & Levend, Agah Sırrı & Mansuroğlu, Mecdut (eds.) *Németh armağanı*. (Türk Dil Kurumu Yayınları 191.) Ankara: Türk Dil Kurumu. 111-144.
- Edip, Kemal 1991. *Urfa ağzı*. (Türk Dil Kurumu Yayınları 25.) Ankara: Türk Dil Kurumu.
- Ercilasun, Ahmet Bican 1983. *Kars ili ağızları*. Ankara: Gazi Üniversitesi Yayınları.
- Eren, Emin 1997. *Zonguldak-Bartın-Karabük illeri ağızları*. (Türk Dil Kurumu Yayınları 686.) Ankara: Türk Dil Kurumu.
- Ergi, İsmail 1991. *Tosya ağzı*. (Tosya Kalkınma. Kültür ve Çevre Vakfı Bilim Eserleri Serisi 1.) Tosya.
- Ergin, Muharrem 1971. *Azeri Türkçesi*. (Edebiyat Fakültesi Yayınları 1633.) İstanbul: İstanbul Üniversitesi.
- Erten, Münir 1994. *Diyarbakır ağzı*. (Türk Dil Kurumu Yayınları 556.) Ankara: Türk Dil Kurumu.
- Foy, Karl 1902. Die ältesten osmanischen Transcriptionstexte in gothischen Lettern 2. *Mitteilungen des Seminars für Orientalische Sprachen. Westasiatische Studien (Berlin) 5*: 233-293.
- 1903. Azerbajjanische Studien mit einer Charakteristik des Südtürkischen. *Mitteilungen des Seminars für Orientalische Sprachen. Westasiatische Studien (Berlin) 6*: 126-193.
- Gadžanov, D. G. 1911. Vorläufiger Bericht über eine im Auftrag der Balkan-Kommission der kais. Akademie der Wissenschaften in Wien durch Nordost-Bulgarien unternommene Reise zum Zwecke von türkischen Dialektstudien. *Anzeiger der kaiserlichen Akademie der Wissenschaften. Phil.-hist. Klasse 48*: 28-42.
- Zweiter vorläufiger Bericht über die ergänzende Untersuchung der türkischen Elemente im nordöstlichen Bulgarien in sprachlicher, kultureller und ethnographischer Beziehung. *Anzeiger der kaiserlichen Akademie der Wissenschaften. Phil.-hist. Klasse 49*: 13-20.
- Gemalmaz, Efrasiyap 1978. *Erzurum ili ağızları 1-3*. (Atatürk Üniversitesi Yayınları 487.) Erzurum: Atatürk Üniversitesi.
- Golden, Peter B. 1992. *An introduction to the history of the Turkic peoples*. (Turcologica 9.) Wiesbaden: Harrassowitz.
- Gülensoy, Tuncer 1988. *Kütahya ve yöresi ağızları*. (Türk Dil Kurumu Yayınları 536.) Ankara: Türk Dil Kurumu.
- 1993. *Rumeli ağızlarının ses bilgisi üzerine bir deneme*. (Erciyes Üniversitesi Yayınları 51.) Kayseri: Erciyes Üniversitesi.
- Güler, Zülfü 1992. *Harput ağzı*. (Elazığ Belediyesi Yayınları 1.) Elazığ: Elazığ Belediyesi.
- Gülseren, Cemil 2000. *Malatya ili ağızları*. (Türk Dil Kurumu Yayınları 237.) Ankara: Türk Dil Kurumu.
- Gültekin, Mevlüt 1994. *Der türkeitürkische Dialekt von Imrenler bei Konya*. (Ph.D. dissertation). Universität Mainz.
- Günay, Turgut 1978. *Rize ili ağızları*. (Kültür Bakanlığı Milli Folklor Araştırma Dairesi Yayınları 27.) Ankara: Kültür Bakanlığı.
- Heffening, Wilhelm 1942. *Die türkischen Transkriptionstexte des Bartheolomaeus Georgievits aus den Jahren 1544-1548*. Leipzig: Deutsche Morgenländische Gesellschaft.
- Hazai, György 1959. Les dialectes turcs du Rhodope. *Acta Orientalia Academiae Scientiarum Hungaricae 9*: 205-229.
- 1973. *Das Osmanisch-Türkische im XVII. Jahrhundert. Untersuchungen an den Transkriptionstexten von Jakab Nagy de Harsányi*. The Hague-Paris: Mouton.
- Johanson, Lars 1979. Die westghusische Labialharmonie. *Orientalia Suecana 27-28*, 63-107.

- 1998. The structure of Turkic. In: Johanson, Lars & Csató, Éva Ágnes (eds.) *The Turkic languages*. London: Routledge. 30-66.
- Kakuk, Suzanne 1958. Le dialecte turc de Kazanlyk. *Acta Orientalia Academiae Scientiarum Hungaricae* 8: 169-187 and 241-311.
- 1961. Die türkische Mundart von Küstendil und Michailovgrad. *Acta Linguistica Academiae Scientiarum Hungaricae* 11: 301-386
- Kalay, Emin 1998. *Edirne ili ağızları*. (Türk Dil Kurumu Yayınları 694.) Ankara: Türk Dil Kurumu.
- Karahan, Leylâ 1996. *Anadolu ağızlarının sınıflandırılması*. (Türk Dil Kurumu Yayınları 630.) Ankara: Türk Dil Kurumu.
- Korkmaz, Zeynep 1956. *Güney-Batı Anadolu ağızları*. Ankara: Dil ve Tarih-Coğrafya Fakültesi.
- 1963. *Nevşehir ve yöresi ağızları I. Sesbilgisi (Phonétique)*. Ankara: Dil ve Tarih-Coğrafya Fakültesi.
- 1964-1965. Bartın ve yöresi ağızları. *Türkoğlu Dergisi*. [Published again in 1994 (Türk Dil Kurumu Yayınları 584.) Ankara: Türk Dil Kurumu.]
- Kowalski, Tadeusz 1929-1930. Einige Probleme der osmanisch-türkischen Dialektforschung. *Rocznik Orientalistyczny* 7: 264-280.
- 1934. Osmanisch-türkische Dialekte. In: *Enzyklopaedie des Islam*. 1913-1937. Leiden: Brill-Leipzig: Harrassowitz. 4: 991-1011.
- Kral, Piet 1980. *De in Turkije gesproken Turkse dialecten met een overzicht van de verwantschapstermen in Turkije*. (Unpublished M.A. thesis.) Leiden: Rijksuniversiteit te Leiden.
- Mollova, Mefküre 1962. Les ga-dialectes turcs dans les Balkans et leur rapport avec les autres langues turcs. *Linguistique Balkanique* 4: 107-130.
- 1970. Dimităr Gadžanov et les parlers turcs dans les Balkans. *Linguistique Balkanique* 14: 99-105.
- Németh, Julius 1956. *Zur Einteilung der türkischen Mundarten Bulgariens*. Sofia: Bulgarische Akademie der Wissenschaften.
- 1965. *Die Türken von Vidin*. Budapest: Akadémiai Kiadó.
- Olçay, Selahâttin 1956. Fuzulî'de Türkçe fiil tasrifleri. In: Korkmaz, Zeynep & Olçay, Selâhattin (eds.) *Fuzulî'nin dili hakkında notlar*. (Ankara Üniversitesi Dil ve Tarih-Coğrafya Fakültesi Yayınlarından 118.) Ankara: Ankara Üniversitesi.
- 1966. [Reprint 1995.] *Doğu Trakya yerli ağızları*. (Türk Dil Kurumu Yayınları 580.) Ankara: Türk Dil Kurumu.
- Özkan, Nevzat 1996. *Gagavuz Türkçesi grameri*. (Türk Dil Kurumu Yayınları 657.) Ankara: Türk Dil Kurumu.
- Räsänen, Martti 1933. *Türkische Sprachproben aus Mittel-Anatolien I: Sivas Vil.* (Studia Orientalia 5/2.) Helsinki.
- Sağır, Mukim 1995. *Erzincan ve yöresi ağızları (İnceleme-metin-sözlük)*. (Türk Dil Kurumu Yayınları 565.) Ankara: Türk Dil Kurumu.
- San, Sabri Özcan 1990. *Gümüşhane kültür araştırmaları ve yöre ağızları*. (Kültür Bakanlığı Yayınları 1212.) Ankara: Kültür Bakanlığı.
- Saracoğlu, Erdoğan 1992. *Kıbrıs ağızları*. (K.K.T.C. Milli Eğitim ve Kültür Bakanlığı Yayınları 21.) Ankara: K.K.T.C. Milli Eğitim ve Kültür Bakanlığı.
- Zülfikar, Hamza 1978. Van gölü çevresi ağızlarının özellikleri. In: *Ömer Asım Aksoy armağanı* (Türk Dil Kurumu Yayınları 449.) Ankara: Türk Dil Kurumu. 297-317.

On the Turkic background of two early loanwords of Turkic origin in Hungarian

Árpád Berta

Berta, Árpád 2005. On the Turkic background of two early loanwords of Turkic origin in Hungarian. *Turkic Languages* 9, 188-198.

The article discusses the Turkic background of two Hungarian words of Turkic origin: *iker* 'twin' and *ökör* 'ox'. In the first part of the article, the author discusses the differences between the Turkic etymon for the word *iker* and the forms of its base word meaning 'two' in the Turkic languages—differences which have not received sufficient attention thus far—and makes an attempt to explain these. In the second part of the article, the author offers an inner Turkic etymology for the Turkic word meaning 'ox', which is: Ancient Turkic **pökkVr-(X)z* > **pökkXz* > **pökXz* > Old Turkic *öküz*.

Árpád Berta, Department of Altaic Studies, University of Szeged, Egyetem u. 2, H-6720 Szeged, Hungary. E-mail: berta@hung.u-szeged.hu

Iker 'twin'

The editors of TESz (2: 196b-197a) reconstructed the Old Chuvash form **ikir* as the original word for the Hungarian word *iker*, forms of which have been recorded in Hungarian as of the 11th century. In their reference to earlier works (including studies by Budenz, Gombocz, Németh and Ligeti, among others), the editors of TESz indicate that the word **ikir*, an example of the common Turkic -z ~ Chuvash -r sound correspondence, derives etymologically from the Turkic numeral *iki* ~ *äki*, meaning 'two'.

Not even the slightest reference can be found in the Turkological literature that would give one grounds to question the connection between the Turkic numeral meaning 'two' and the Hungarian common noun meaning 'twin'. However, there is no explanation either for the obvious lack of sound correspondence between the base form of the Turkic word meaning 'two' and its so-called derivative, the common noun meaning 'twin'—which can also be found in a majority of the Turkic languages. Clauson's dictionary on Turkic word history and Severtjan's etymological dictionary regard the lexemes **ekkiz* 'twin' (Clauson 119b) and *ekiz* 'bliznecy, dvojnja' (Severtjan 1: 252-254) as derivatives of the Old Turkic *ekki* 'two' (Clauson 100b-101a) and *iki* 'dva' (Severtjan 1: 337-339).

This article does not question the etymological connection between the Turkic numeral meaning 'two' and the Turkic common noun meaning 'twin', but it does hope to provide an explanation for the significant differences between the two word

forms (initial vowel, medial consonant), which, although present in a number of Turkic languages, have not been examined with sufficient care.¹

In a few of the Modern Turkic languages, the phonetic differences mentioned above do not appear among the lexemes under examination. The lexemes in these languages are the following:

Oghuz languages (O): *iki* 'dva'—*ikiz* 'dvojnja, bliznecy' (Tt);
iki 'dva'—*ikiz* 'bliznecy; dvojnja' (Gag);
 Siberian Turkic languages (S): *iyi* 'dva'—*iyis* 'dvojnja' (Tuv);
iki 'dva'—*ikis* 'bliznecy, dvojnja; dvojnik' (Khak).

The data within this group indicate that the word meaning 'twin' was formed through the addition of the old final *+z (more precisely *+(X)z) formant to the base numeral. The data in the Siberian Turkic languages (Tuvan, Khakas) ending in the suffix +s—in line with the devoicing -z > -s regular sound change in these languages—present the sound one may expect as the sound resulting from the historical formant *+(X)z.

The group of languages showing phonetic difference(s) between these two words is rather larger than the one above, which included languages with proper correspondences. For a better overview of these languages and their forms, they are presented in various subgroups.

A. Difference in the correspondence of the vowels in the first syllable:

Oghuz languages: *iki* 'dva'—*äkiz* 'bliznecy, dvojnja' (Az);
iki 'dva'—*ekiz* 'dvojnja, bliznecy, dvojnjaški', *ekizäk*² 'odin iz bliznecov' (Tkm).

Only two Oghuz languages belong to this subgroup. The phonetic difference lies in the fact that the initial vowel in the numeral meaning 'two' is more closed than that in the common noun meaning 'twin'. This phenomenon—at the present state of our knowledge—cannot be explained properly. Nevertheless, it may be possible that in Azeri and Turkmen the vowel in the second open syllable had an impact on the first vowel syllable through regressive assimilation, an impact which vowels in closed syllables could not have. This is clearly on the level of a working hypothesis, and as such, calls for further investigation.

¹ The difference in sound correspondence between the two lexemes cannot serve as an argument against the etymological connection. This is so not only because the connection is borne out by semantics, but also because similar sound differences can be located in a number of languages between the base numeral and the derivative common noun, cf. e.g. English *two* and *twin*, German *zwei* and *Zwilling*, Russian *dva* and *dvojnja*.

² The Turkmen item is a derivative with the suffix +*Ak*. Cf. also below the endings of the words Tatar *igëzäk*, Bashkir *igëdäk*, Karaim of Troki *ekiz'ak*, *egiz'ak*.

B. Difference appears in the feature of the medial consonants:

Kipchak languages (K): *ikě* 'dva, dve, dvoe'—*igěz* 'bliznecy, dvojnja, dvojnjaški; (*peren.*) ljudi, blizkie i poxožie drug na druga', cf. *igězäk*: *oč igězäk* 'trojnja' (Tat);

ikě 'dva, dve, dvojka; para *prost.*'—*igěð* 'dvojnja; dvojnjaški *razg.*; bliznecy', cf. *igěðäk* id. (Bashk);

eki 'dva'—*egiz* 'bliznecy, dvojnja; para odinakovyx' (Kirg);

eki 'dva; dvu-, dvux-'—*egiz* 'bliznecy; dvojnja, dvojnjaški' (Nog);

eki 'dva, dvoe', cf. *ekiz'ak* 'dvojnjoj'—*egiz'ak* 'bliznec' (Kar T);

eki 'dva, dvoe'—*egiz* 'bliznec' (Kar C);

yekī 'dva'—*yegiz* 'bliznecy; dvojnja' (Kzk);

eki 'dva'—*egiz* 'bliznecy, dvojnja' (Kmk);

eki 'dva'—*egiz, egizle*³ 'dvojnja, bliznecy' (Krch-Blk);

Siberian Turkic languages (S): *eki* / *ekki* 'dva'—*egis* 'dvojnja, bliznecy' (Oyr Kmd).

The Kipchak languages and the Oyrot dialect are connected because the medial strong explosive in the numeral meaning 'two' is replaced by a weak explosive in the word meaning 'twin'. At this point, we can only offer an assumption as a possible explanation.⁴ It seems that as a result of the effect of the consonant in the final suffix +(X)z, the medial long consonant shortened (*-kk- > *-k-) so early that it preceded the change of intervocalic *-k- > -g-, which is a phonetic feature of the languages listed in this subgroup.

C. Alternating representations—only partly, due to internal language differences both in the case of the initial vowels and the medial consonants:

Kipchak languages (K): *eki* / *iki* 'dva, dvoe'—*egizék* 'bliznec' (Kar H);

eki / *yeki* / *yekki* / *ikki* 'dva'—*yegiz* 'bliznecy, dvojnja' (Kkalp);⁵

Siberian Turkic languages: *eki* 'dva'—*egis* / *igis* 'bliznecy, dvojnja; para' (Oyr);

Turki languages (T): *ikki* 'dva'—*egiz* 'dvojnja, bliznecy' (NUyg);

ikki 'dva'—*egiz* 'bliznecy' (Uzb);

Chuvash (Ch): *ikkě, ikě, ik* 'dva'⁶—*yěkěr* 'dvojnjoj; sdvoennyj; para', cf. *yěkěreš* 'bliznecy, dvojnja; nerazlučnyj'.

³ The Karachay-Balkar word *egizle* is a further derivative.

⁴ This is an assumption—and not a working hypothesis—as the very same phenomenon will appear in the case of *ökör*, the second Turkic loanword discussed in this article. Other examples will be listed there to strengthen this assumption.

⁵ Some of the Karakalpak data show a secondary initial *y*-. See also the similar—also secondary—*y*- in the Chuvash data.

⁶ The variety of the Chuvash data may be explained by their functional distribution, similar to the Hungarian forms meaning 'two', *két* and *kettő*.

The explanations offered as a working hypothesis and an assumption under points A and B above may also serve as an explanation for the inconsistencies in these data as well.

However, it must be emphasised that at this point, based on the present state of our knowledge, these phenomena cannot be explained with absolute certainty. Naturally, another general linguistic explanation may also be offered, yet again, as a working hypothesis: the languages in groups B and C attempted to mark a phonetic difference in order to indicate some semantic distinction.

From among the more significant modern Turkic languages, the Yakut words *ikki* ‘dva’—*igirè*, *igirèlèr* ‘bliznecy’ (RusskJakSl) were not listed in group B, since the Yakut word meaning ‘twin’ is a loanword from Mongolian.⁷ The Mongolian word, however, is of Turkic origin, and has also found its way into some of the Tungusic languages.⁸

Ökör ‘ox’

The editors of TESz (3:23a) maintain that the Hungarian word *ökör* “is a Chuvash-type Old Turkic loanword”, which possibly derived from “a Turkic form of **ökür* which entered the Hungarian language”.

As for the origin of the Turkic word, various theories have gained currency in the Turkological literature; however, within historical linguistics, two major theories can be found.⁹ Both of these regard this word as being of Indo-European origin.

One of these maintains that the Turkic word meaning ‘ox’ derived from a Tokharian A dialectal form of *okäs* (see more recently Clauson 120a),¹⁰ while the

⁷ Cf. Mongolian *ikire*, *ikere* ‘twins’ (L). For further Mongolian data, see Ligeti (1986:311). Besides Yakut, Mongolian loanwords may also be found in some of the Siberian Turkic languages: Khakas dialect *ikere* ‘dvojnja (*o životnyx*)’, Koybal *ikkärä* ‘dvojniki; die Zwillinge’ (R 1:1420).

⁸ Doerfer (1965:189-191) argues convincingly against the views held by Ramstedt (1957:113) and Poppe (1960:105), which maintain that the Turkic, Mongolian and Tungusic data derive from the Altaic base word. He also points out that the stem of the word *ekiz* ‘twin’ is the Turkic word *eki* ‘two’, which is unknown in Mongolian. Moreover, the suffix +z is also Turkic, being unknown with the same function in Mongolian. Doerfer’s entry also offers a thorough survey of the way this word has spread in other languages as well as illustrating well through a number of examples how in a number of languages throughout the world the numeral meaning ‘two’, which is part of the basic vocabulary, is a loanword like the common noun meaning ‘twin’. Doerfer lists the Hungarian word *iker* among these examples.

⁹ A good summary of the history of scholarship on the etymological change of the Turkic word meaning ‘ox’ is provided by András Róna-Tas in his unpublished dissertation, see Róna-Tas Diss. 460-467 and Sevortjan 1: 522-523.

other one (cf. Ramstedt 1957: 103-104, among others) connects the Turkic word to the well-known Indo-European word family **peku* (cf. OldInd *paśu*, Lat *pecu*, *pecus*, Goth *faihu*, Germ *Vieh* etc.).¹¹ Neither of these theories that propose an Indo-European origin would seem to be plausible.

Before suggesting an etymology which would consider an inner Turkic onomatopoeic base word for the origin,¹² let us see the data related to the Hungarian word *ökör* in the modern Turkic languages.¹³

- Oghuz languages (O): *öküz* ‘byk, vol’ (Tt); *öküz* ‘byk, vol’ (Az); *öküz* ‘vol’ (Gag); *öküz* ‘byk, vol’ (Tkm);
- Kipchak languages (K): *ügěz* ‘byk, bugaj; (*peren.*) očen' bol'šoj i sil'nyj (*o čeloveke*)’ (Tat); *ügěđ* ‘byk, vol’ (Bashk); *ögüz* ‘vol, kastrirovannyj byk’ (Kirg); *ögiz* ‘vol, byk; (*peren.*) ploho soobražajuščij čelovek’ (Nog); *égiz* ‘byk, vol’ (Kar H); *ögüz* ‘byk’ (Kar C), *ög'uz* ‘id. (Kar T); *ögiz* ‘vol’ (Kzk); *ögiz* ‘byk’ (Kkalp); *ögüz* ‘vol’ (Cr Tat); *ögüz* ‘vol’ (Kmk); *ögüz* ‘vol’ (Krč-Blk);
- Siberian Turkic languages (S): —;
- Turki languages (T): *ökuz* / *xökuz*¹⁴ ‘byk, vol’ (NUig); *öküz* ‘steer’ (NUygJarring); *oqus* ‘korova; byk, vol’ / *okus* ‘krupnyj, rogatyj skot’ / *kus* ‘byk, vol’ / *qus* ‘korova; byk, vol’ (YUygMalov 1957),¹⁵ *xükiz* ‘vol’ (Uzb);
- Yakut (Y): *oğus*, *oğos*¹⁶ ‘byk voobšče; byk, vol’ (Pekarskij 2: 1786-1787); cf. also *at oğus*¹⁷ ‘vol’ (RusskJakSI);

- ¹⁰ This is Clauson’s modified view, since he earlier (Clauson 1959) regarded this Turkic word as a derivative of a Tokharian B dialectal form *okso*. Both of Clauson’s views are criticised by Doerfer (1963:539), who points out that the Tokharian origin is highly unlikely. He argues that Clauson disregards the fact that the Mongolian word of Turkic origin is found in the form *hüker* > *üker* ‘bovine animal, ox, cow; large, big; the second of the twelve animals of the zodiac’ (L) and that the Middle Mongolian initial *h-* historically may also have derived from an Ancient Turkic **p-*.
- ¹¹ Problems with phonetics (vocalism in the first syllable!) and semantics (the Indo-European word means ‘animal; cattle, livestock’, but not ‘ox’) in this latter etymology have already been pointed out by Doerfer as well in the work mentioned above (1963:539).
- ¹² This explanation has surfaced before. A summary of the possibilities regarding an onomatopoeic base is offered in András Róna-Tas’s unpublished dissertation (Róna-Tas, Diss. 464). Róna-Tas, however, does not accept this explanation.
- ¹³ For a list of the language historical data, see Clauson 120a; Severtjan 1:521-523.
- ¹⁴ The initial *x-* in the NUyg word *xökuz* may be secondary, but may also be archaic. The same applies to the initial sound *x-* in the Uzbek word.
- ¹⁵ In the Yellow Uyghur words *kus* and *qus* the disappearance of the initial vowel is, of course, a secondary phenomenon.
- ¹⁶ The back vowel in the Yakut word may be secondary, just like in the Yellow Uyghur word.

Khalaj (Kh): —;
 Chuvash (Ch): *vākār, mākār*¹⁸ ‘byk’.

When determining the inner Turkic origin of the Turkic word meaning ‘ox’, a Turkic verb must also be included in the discussion. On the basis of its form and meaning, it seems that this verb, although not present in all the language branches, should be included in the investigation.

Oghuz languages (O): —;
 Kipchak languages (K): *ükēr-* ‘revet’; vopit’’, cf. *üksē-* ‘všlipyvat’, plakat’ všlipyvaja’ (Tat); *ükēr-* ‘revet’, ryčat’; (peren.) vyt’; šumet’’, cf. *ükhē-* ‘gromko plakat’, rydat’; prizyvno revet’ (*o životnyx*) (Bashk); *ökür-* ‘(*o byke, bugae*) revet’; (*o mužčine*) gromko plakat’ i pričitat’ (*pokačivajas’ korpusom s boku na bok, približajas’ k domu, k jurte, gde est’ ili nedavno byl pokojnik*)’, cf. *öksö-* ‘gromko plakat’, rydat’; (peren.) plakat’ sja, žalovat’ sja na sud’bu’, *öküm*¹⁹ ‘nesderžannyj, vspyl’čivij; neterpelivij, toroplivij’ (Kirg); *ökir-* ‘revet’, myčat’; vopit’’, rydat’’, cf. *öksi-* ‘rydat’; všlipyvat’ (Nog); *ékir-* ‘sto-nat’; ryčat’; revet’; myčat’ (Kar H), *ökür-* ‘plakat’, revet’; myčat’ (Kar C), *ökür-* ‘vyt’, ryčat’; zevat’ (Kar T);²⁰ *ökir-* ‘revet’ (*o korove*); (peren.) rydat’ (*o čeloveke*)’ (Kzk); *ökir-* ‘kričat’; revet’, plakat’ (Kkalp); *ökür-* ‘revet’ (Cr Tat); *ökür-* ‘gudet’; gremet’; buševat’; ryčat’; (peren.) gremet’; slavit’ sja’ (Kmk); *ökür-* ‘revet’ (Krch-Blk);
 Siberian Turkic languages (S): cf. *öksö-* ‘gor’ko plakat’, gromko plakat’, rydat’ (Oyr); *öksö-* ‘kričat’ (OyrTuba);
 Turki languages (T): *xöküri-* ‘revet’, rykat’ (*o zverjax*); rydat’, gromko plakat’; cf. *öksü-* ‘všlipyvat’, rydat’, plakat’ navzryd’ (NUyg); *ükir-* ‘revet’ (Uzb);
 Yakut (Y): —;
 Khalaj (Kh): —;
 Chuvash (Ch): *ükxēr-* ‘šumet’, gudet’; ryčat’; vyt’, kričat’.

If the Turkic noun meaning ‘ox’ is connected to the Turkic word meaning ‘to below, low’, which is justifiable,²¹ it must be decided whether the noun meaning ‘ox’ should be regarded as a derivative of the verbal base word, or the other way round.

¹⁷ The first element of the Yakut compound is the word *at* meaning ‘horse’ which is well-known in the Turkic languages (see Pekarskij 1:182).

¹⁸ The latter dialectal form with the initial *m-* (see Egorov) is secondary. At the same time, the appearance of the prothesis *v-* in the word *vākār* reflects a regularity.

¹⁹ The etymological status of the lexeme *öküm* is problematic. If it is an old derivative, it may be connected to a form with the morphological segmentation of **pök+U-(X)m*.

²⁰ KRPSI makes a mistake in suggesting that the word *aqir-* ‘revet’, stonat’ (Kar K) belongs here.

²¹ See the etymology suggested by Brockelmann (1954: 49), according to which the Turkic word *buğa* may be a derivative of the onomatopoeic base word **bu-* followed by the suf-

It seems that the common noun meaning ‘ox’ is of verbal origin, as the data fail to bear out derivation from the other direction.

The verb meaning ‘to bellow, low’ may derive from an onomatopoeic nominal base. This base may have been the form **pök* in Ancient Turkic. The base word **pök* may have been followed by the suffix *+kVr-*,²² thus, the reconstructed Ancient Turkic form must have been **pökkVr-*.

The Ancient Turkic common noun meaning ‘ox’ may be a derivative of the verb form reconstructed as **pök+kVr-* followed by the suffix *-(X)z*: **pökkVr-(X)z > *pökkXz > *pökXz*.

The possibility of the **pökkVr-(X)z > *pökkXz* development is borne out by a series of convincing morphological analogies from Erdal’s monograph on Old Turkic word formation (1991:323):

kütuz ‘a mad dog’ < **kütur-(u)z*;
ärgüz [‘snow and ice melting at the beginning of spring’] < **ärgür-(ü)z*;
munduz [‘senile, simple-minded’] < **mun-dur-(u)z*;
adüz [‘a smaller (uncultivated) piece of land’] < **adür-(i)z*;
yavüz ‘bad’ < **yavrü-z*;
sämiz ‘fat (adj., of an animal)’ < **sämri-z*.

The examples listed by Erdal reveal the expansionist behaviour of the Old Turkic suffix *-(X)z*, which has resulted in the shortening of the endings on verbal bases ending in *°r(V)-*, or, to be more precise, their elision.²³

The change of **pökkXz > *pökXz* as suggested above—that is, the shortening of the internal long consonant **-kk-*would be the same presumed change presented in

fix -gA. Severtjan (2:231-232) rejects Brockelmann’s etymology, but his discussion lacks clear argumentation.

²² The nature of the vowel in the suffix is questionable. The applicable section of Erdal’s excellent work on Old Turkic word formation (1991: 465-467) cannot be regarded as the final solution for three reasons. First, because among his examples for various onomatopoeic bases—eleven examples, to be precise, in which bases are followed by the formant which Erdal determines as *+kIr-*—there are only two bases with a labial vowel (*bü(r) + kür-*, *üş + kür-*), which may offer a possibility for determining whether the vowel in the denominal verbal formant was indeed *-X-* (that is, with four vowel variants) or *-I-* (that is, palatal and unrounded); thus, the other nine examples are not significant in this respect. Second, two of the Old Turkic examples listed by Erdal, (*bü(r) + kür-*, *üş + kür-*)—as the data reveal—vary in the nature of the second vowel (*-X-* / *-I-*). Third, the fact that Erdal examined only the Old Turkic corpus, with which—although we tend to forget this—the Ancient Turkic data may not be identified, also makes it problematic to determine the vowel in the given suffix.

²³ The regularity indicated here is borne out by a number of examples, which I plan to discuss in a separate study in the near future.

connection with the suffix $+(X)z$ in the case of the change in $*\ddot{a}kki+(X)z > *\ddot{a}kiz$ ‘twin’ discussed in the first part of this article.

The presumed change in the long consonant $-kk-$ before $-z$ and its subsequent weakening as witnessed in the Turkic language branches in which it occurred in the case of the Turkic common nouns meaning ‘twin’ and ‘ox’ is further borne out by two Turkic numerals.

Old Turkic *tokküz* (*d-*) ‘nine’ (Clouston 474b)

Oghuz languages (O): *dokuz* (Tt), *dogğuz* (Az), *dokuz* (Gag), *dokuz* (Tkm);

Kipchak languages (K): *tugız* (Tat), *tugıđ* (Bashk), *togız* (Nog), *toğuz* (Kar T, H), *dokuz* (Kar C), *toğız* (Kzk); *toğuz* / *toğız* / *toğus* / *toğıs* / *toqquz* (Kkalp), *toğuz* (Kırg), *doquz* (Cr Tat); *toğuz* (Kmk); *toğuz* (Krch-Blk);²⁴

Siberian Turkic languages (S): *toğus* (Oyr), *toğus* (OyrTuba), *toğus* (OyrKmd), *tos* (Tuv), *toğıs* (Khak);

Turki languages (T): *toqquz* (NUyg), *toqoz* / *toquz* / *toqquz* (NUygJarring), *toqış* / *to'qıs* (YUyg), *toqos* (Sal);

Yakut (Y): *toğus* (RusskJakSI);

Khalaj (Kh): *toqquz*;

Chuvash (Ch): *täxxär*, *täxär*.

Old Turkic *säkkiz* ‘eight’ (Clouston 823b)

Oghuz languages (O): *sekiz* (Tt), *säkkiz* (Az), *sekiz* (Gag), *sekiz* (Tkm);

Kipchak languages (K): *sigěz* (Tat), *higěđ* (Bashk), *segiz* (Nog), *segiz*' (Kar T), *ségiz* (Kar H), *sėkiz* (Kar C), *segiz* (Kzk), *säkkiz* / *segiz* (Kkalp), *segiz* (Kırg), *sekiz* (Cr Tat), *segiz* (Kmk), *segiz* (Krch-Blk);

Siberian Turkic languages (S): *segis* (Oyr), *segis* (OyrTuba), cf. *segizen* ‘vosem’-desjat’ (OyrKmd), *ses* (Tuv), *sigıs* (Khak);

Turki languages (T): *säkkiz* (NUyg), *sekiz* / *şekiz* (NUygJarring), *sekes* / *sak'ıs* (YUyg), *sėkis* / *sėkes* / *sekis* (Sal);

Yakut (Y): *ağıs* (RusskJakSI);

Khalaj (Kh): *säkkiz*;

Chuvash (Ch): *sakkär*.

It must also be noted that, with the exception of a few relatively new loanwords from Russian (see, e.g., *nakaz*, *ukaz*), the only word structure which can be found in

²⁴ Among the data from the Kipchak languages, the modern forms of the Old Turkic lexeme *tokkuz* with the internal $-K-$ (or, even $-KK-$) in Crimean Karaim, Crimean Tatar, Karakalpak, are exactly the same as the corresponding modern Turkic representatives of the Old Turkic numeral *säkkiz* listed below. These so-called ‘A-Kipchak’ forms may not necessarily be explained in the same way. The Crimean Tatar data—just like the whole Crimean Tatar language—may show Oghuz influence. The same may also hold true for the Crimean Karaim form. However, further data and investigation would be required to explain the variations within the Karakalpak form.

the Turkic languages mentioned above is $(C)VGVz^{25}$; no lexeme exists with the structure $(C)VKVz$. However, forms with the structures $(C)VKV(C)$ are also known to exist,²⁶ if the syllable following $-k-$ does not end in $-oz$. All of this means—it seems—that we have found (at least one of) several reasons why the old intervocalic long $-kk-$ not only shortens in some of the Turkic languages, but then also weakens the formally long strong consonant, while in other examples—in the very same languages—it merely shortens.

Abbreviations and bibliography

- ARSL = *Azerbajdžansko-russkij slovar*. Edited by Azizbekov, H. A. Baku: Azerbajdžanskoe gosudarstvennoe izdatel'stvo. 1965.
- Az = Azeri, see ARSL.
- Bashk = Bashkir, see BRH.
- Baskakov 1966 = Baskakov, N. A. *Dialekt černevyx tatar (tuba-kiži). Grammatičeskij očerk i slovar*. Moskva: Nauka.
- Baskakov 1972 = Baskakov, N. A. *Dialekt kumandincev (kumandī-kiži)*. Moskva: Nauka.
- BRH = *Bašqōrtsa-russa hüðlĕk*. Edited by Uraqsin Z. & Mäskäü, G. Näšriätä Digora. 1996.
- Ch = Chuvash, see ČVSI.
- Clauson = Clauson, G. *An etymological dictionary of pre-thirteenth-century Turkish*. Oxford: Clarendon Press. 1972.
- Clauson 1959 = Clauson, G. 1959. The earliest Turkish loan words in Mongolian. *Central Asiatic Journal* 4: 174-187.
- Cr Tat = Crimean Tatar, see KrTatRSI.
- ČVSI = *Čävašla-viräsla slovar*. Edited by Skvorcov, M. & Muskv, I. Russkij Jazyk. 1982.
- Doerfer 1963 = Doerfer, G. *Türkische und mongolische Elemente im Neupersischen*. 1. Wiesbaden: Steiner.
- Doerfer 1965 = Doerfer, G. *Türkische und mongolische Elemente im Neupersischen*. 2. Wiesbaden: Steiner.
- Doerfer-Tezcan = Doerfer, G. & Tezcan, S. 1980. *Wörterbuch des Chaladsch. (Dialekt von Charrab)*. Budapest: Akadémiai Kiadó.
- Egorov = Egorov, V. G. 1964. *Ėtimologičeskij slovar' čuvaškogo jazyka*. Čeboksary: Čuvaškoe knižnoe izdatel'stvo.
- Erdal 1991 = Erdal, M. 1991. *Old Turkic word formation. A functional approach to the lexicon* 1-2. (Turcologica 7.) Wiesbaden: Harrassowitz.
- Gag = Gagauz, see GRMSI.
- GRMSI = *Gagauzsko-russko-moldavskij slovar*. Edited by Baskakov, N. A. Moskva: Sovetskaja Ėnciklopedija. 1973.
- Judaxin = Judaxin, K. K. 1965. *Kirgizsko-russkij slovar*. Moskva: Sovetskaja Ėnciklopedija.
- K = Kipchak languages.

²⁵ See, e.g., Tatar of Kazan (Tat) *agĭz-*, *bugaz*, *igĕz*, *kigĕz-*, *mögĕz*, *nigĕz*, *sagĭz*, *sigĕz*, *tigĕz*, *tugĭz*, *ügĭz*, *ugĭz*, *ügĕz*, etc.

²⁶ See, e.g., from Tatar of Kazan (Tat) again: *akay-* and *akay*, *akĭr-*, *baka*, *bakĭr*, *bikä*, *bükän*, *yokĭ*, *kikĕr-*, etc.

- Kar H = Karaim Halich dialect, see KRPSI.
 Kar C = Karaim Crimean dialect, see KRPSI.
 Kar T = Karaim Troki dialect, see KRPSI.
 KhakRSI = *Xakassko-russkij slovar*'. Edited by Baskakov, N. A. Moskva: Gosudarstvennoe izdatel'stvo inostrannyx i nacional'nyx slovar'ej. 1953.
 Kh = Khalaj, see Doerfer-Tezcan.
 Khak = Khakas, see KhakRSI.
 Kirg = Kirghiz, see Judaxin
 Kkalp = Karakalpak, see KkalpBask.
 KkalpBask = 'Slovar' in Baskakov, N. A. 1951. *Karakalpakskij jazyk. I. Materialy po dialektologii*. Moskva: Izdatel'stvo Akademii Nauk SSSR.
 Kmk = Kumyk, see KmkRSI.
 KmkRSI = *Kumyjsko-russkij slovar*'. Edited by Bammatov, Z. Z. Moskva: Sovetskaja Ėnciklopedija. 1969.
 Koyb = Koybal, see R.
 Krch-Blk = Karachay-Balkar, see RusskKBSI.
 KRPSI = *Karaimsko-russko-pol'skij slovar*'. Edited by Baskakov, N. A. et al. Moskva: Russkij jazyk. 1974.
 CrTatRSI = *Krymskotatarsko-russkij slovar*'. Edited by Asanov, Š. A. et al. Kiev: Radjans'kaja škola. 1988.
 Kzk = Kazakh, see KzkRSI.
 KzkRSI = Maxmudov, X. & Musabaev, G. *Kazaxsko-russkij slovar*'. Alma-Ata: Izdatel'stvo Akademii Nauk Kazaxskoj SSSR. 1954.
 L = Lessing, F. D. 1973. *Mongolian-English dictionary*. Bloomington: The Mongolia Society.
 Ligeti 1986 = Ligeti Lajos 1986. *A magyar nyelv török kapcsolatai a honfoglalás előtt és az Árpád-korban*. Budapest: Akadémiai Kiadó.
 Nog 0 = Nogay, see NRSI.
 NRSI = *Nogajsko-russkij slovar*'. Edited by Baskakov, N. A. Moskva. 1963.
 NUyg = New Uighur, see UjgRSI.
 NUygJarring = Jarring, G. 1964. *An Eastern Turki-English dialect dictionary*. Lund: CWK Gleerup.
 NyK = Nyelvtudományi Közlemények. Pest [later] Budapest. 1 (1862) –
 O = Oghuz languages.
 Oyr = Oyrot, see OyrRSI.
 OyrKmd = Oyrot language, Kumandī dialect, see Baskakov 1972.
 OyrRSI = *Ojrotsko-russkij slovar*'. Edited by Baskakov, N. A. Moskva: Gosudarstvennoe izdatel'stvo inostrannyx i nacional'nyx slovar'ej. 1947.
 OyrTuba = Oyrot language. Tuba dialect, see Baskakov 1966.
 Poppe 1960 = Poppe, N. 1960. *Vergleichende Grammatik der altaischen Sprachen 1. Vergleichende Lautlehre*. Wiesbaden: Harrassowitz.
 R = Radloff, W. 1893-1911. *Versuch eines Wörterbuches der Türk-Dialecte* 1-4. Sankt-peterburg.
 Ramstedt 1957 = Ramstedt, G. J. 1957. *Einführung in die altaische Sprachwissenschaft 1. Lautlehre*. Mémoires de la Société Finno-Ougrienne 104:1. Helsinki.

- Róna-Tas Diss. = Róna-Tas, A. 1970. *Az altaji nyelvrokonság vizsgálatának alapjai. A nyelvrokonság elmélete és a csuvas-mongol nyelvviszony.* Akadémiai doktori értekezés. Budapest. Unpublished.
- RusskKBSl = *Russko-karačaevo-balkarskij slovar'*. Edited by Sujunčev, X. I. & Urusbiev, I. X. Moskva: Sovetskaja Ėnciklopedija. 1965.
- RusskJakSl = *Russko-jakutskij slovar'*. Edited by Afanas'ev, P. S. & Xaritonov, L. N. Moskva: Sovetskaja Ėnciklopedija. 1968.
- S = Siberian Turkic languages.
- Sevortjan 1-2 = Sevortjan, Ė. V. *Ėtimologičeskij slovar' tjurkskix jazykov* 1-2. Moskva. 1974-1978.
- Sal = Salar, see Tenišev.
- T = Turki languages.
- Tat = Tatar of Kazan, see TRSl.
- Tenišev = Tenišev, Ė. R. 1976. *Stroj salarskogo jazyka.* Moskva: Nauka. 1976.
- TESz = *A magyar nyelv történeti-etimológiai szótára* 1-4. Edited by Benkő, L. Budapest: Akadémiai Kiadó. 1967-1984.
- Tkm = Turkmen, see TkmRSl.
- TkmRSl = *Turkmensko-russkij slovar'*. Edited by Baskakov, N. A. et al. Moskva: Sovetskaja Ėnciklopedija. 1968.
- TRSl = *Tatarsko-russkij slovar'*. Edited by Golovkina, O.V. Moskva: Sovetskaja Ėnciklopedija. 1966.
- Tt = Turkish, see TuRSl.
- TuRSl = *Turecko-russkij slovar'*. Edited by Mustafaev, Ė. M.-Ė. et al. Moskva: Russkij Jazyk. 1977.
- Tuv = Tuvan, see TuvRSl.
- TuvRSl = *Tuvinsko-russkij slovar'*. Edited by Pal'mbax, A. A. Moskva: Gosudarstvennoe izdatel'stvo inostrannyx i nacional'nyx slovarej. 1955.
- UygRSl = *Ujgursko-russkij slovar'*. Edited by Raximov, T. R. Moskva: Sovetskaja Ėnciklopedija. 1968.
- Uzb = Uzbek, see UzbRSl.
- UzbRSl = *Uzbeksko-russkij slovar'*. Edited by Borovkov, L. K. et al. Moskva: Gosudarstvennoe izdatel'stvo inostrannyx i nacional'nyx slovarej. 1959.
- Y = Yakut, see Pekarskij; RusskJakSl.
- YUygMalov = Malov, S. E., *Jazyk želtyx ujugurov.* Alma-Ata: Nauka. 1957.

On the problems of Oghuz morphophonology

Anna V. Dybo

Dybo, Anna 2005. On the problems of Oghuz morphophonology. *Turkic Languages* 9, 199-204.

The paper deals with the alternation of forms with and without a narrow vowel in Oghuz lexical stems. Reconstructions show that Proto-Turkic possessed stems with consonant clusters which were later eliminated, i.e. simplified or resolved by vowel epenthesis. It can be demonstrated that the alternation in Proto-Oghuz was originally governed by purely phonemic rules—as still in modern Turkmen—and later overlapped by processes of morphological analogy. In the present paper, a list of native Turkish disyllabic noun stems is compared with reconstructed Proto-Turkic forms and parallels from other Turkic languages. Certain Proto-Turkic forms ended in a final cluster, while others displayed a vowel between the two last consonants. Turkish stems that exhibit the alternation go back to stems both with and without final clusters. Some stems do not display the alternation although they originate in Proto-Turkic forms with final clusters. We may assume a development from phonemically conditioned alternations to processes of paradigmatic unification of the stem forms. In the noun inflection the alternation is mostly eliminated. Almost all stems that have retained the alternation belong to semantic groups that are usually employed with possessive suffixes. We may assume a development from phonemically conditioned alternations to processes of paradigmatic unification of the stem forms.

Anna Dybo, Institute of Linguistics, Moscow, 121009, Bolshoj Kislovskij per. 1/12, Russia. Email: adybo@mail.ru

The elision/insertion of a narrow vowel is one of the patterns of alternation in Oghuz stems and in Turkic stems in general. In Modern Turkish the alternation is distributed lexically, so it must belong to the morphophonological stratum of the language system. In nominal paradigms, grammars describe it as quite parallel to the rule of loan-word adaptation, i.e. the insertion of a vowel to resolve a cluster inadmissible in the Auslaut (see, e.g., Kononov 1956: 28). A reference to the presence of an alternation is usual in Turkish lexicography. The same rule exists in Gagauz (Pokrovskaja 1964: 46-48) and probably in Azeri, but standard dictionaries of these languages contain no references to the presence of alternation in the stems. In Gagauz some “secondary” stems are noted, which have developed on the base of petrified possessive forms of the 3rd person with the alternation and later attached new possessive affixes, in the post-vocalic form: *ōlu* – *ōlu-su* (the modern standard form is *ōl* – *ōlu*) ‘son’, *burnu* – *burnu-su* ‘nose’, *anny* – *anny-sy* ‘front’, *kojnu* – *kojnu-su*

‘chest’ (cf. *kojun* – *kojunu* ‘sheep’), *bojnu* – *bojnu-su* ‘neck’, *gelin/ gelni* – *gelnisi* ‘daughter-in-law’ (recent standard form is *gelin* – *gelini*); see Dmitriev 1962: 253. Such forms are present sporadically also in other Oghuz languages. In Turkmen, the alternation is not conditioned lexically: the rule is purely phonemic: a narrow vowel in the structure VCVcl.CV¹ drops if the resulting consonant cluster belongs to the following set: RR, RS, SS, Sd, Rd, Sž, Rž (see *Grammatika turkmenskogo jazyka* 1970: 61–63; the rule is confirmed by materials in *Türkmen diliniň orfoepik sözlügi* 1978).

The full selection of two-syllable nominal stems from the Turkish-Russian dictionary edited by Mustafaev & Starostov (1977) demonstrates that the words that are genetically Turkic have such alternation only if they have a simple (from the modern point of view) disyllabic stem with phonemic structure (C)VS/RVS/R. Among the possible fricatives š is not present. But not all stems with such a structure exhibit the alternation. See, e.g., the minimal pairs *kojun*, *-jnu* ‘lap’ – *kojun*, *-jnu* ‘sheep’, *kajyn*, *-jny* ‘brother-in-law’ – *kajyn*, *-jyny* ‘beech’ or a quasi-minimal pair such as *beniz*, *-nzi* ‘face’ – *deniz*, *-nizi* ‘sea’.

According to the general principles of the nature of morphophonemic phenomena, we must suppose that historically this alternation, lexically distributed and partly defined by a phonemic rule, as found in Modern Turkish and some other Oghuz languages, developed from a purely phonemic rule that was overlapped later by processes of morphological analogy (or, more specifically, by paradigmatic unification). On the other hand, the reconstruction of Proto-Turkic lexical stems produced during the composition of a full Turkic etymological data-base has finally demonstrated the presence in Proto-Turkic of stems with consonant clusters, which were later eliminated in different ways in the various Turkic languages: they were either simplified or resolved by a narrow vowel epenthesis. Already the Turkish grammar of Jean Deny (1920: 147-148) raised the question whether the ambiguity of the behavior of Turkish nouns is related to primarily different endings of their stems. Now we can try to give a justified answer to this question.

I present here a list of native Turkic disyllabic noun stems with the alternation; they are listed with Proto-Turkic reconstructions (after the Turkic etymological data-base) and with relevant parallels from other Turkic languages.

A. Alternative data force us to reconstruct a Proto-Turkic final cluster.²

1. **bōjn* ‘neck’: *bojyn* (Old Uyghur), *bojun*, *bojyn* (MK), *bojun* (QB), Turkish *boyun*, *-ynu*, Chaghatay *bojn*, *bojun* (Sanglax, MA), Azeri *bojun*, *-jnu*, Turkmen *bojun*, Khakas *mojyn*, Chuvash *mɔj*, Yakut *mōj* (*mońno-*), Dolgan *muoj*, Tuvan *mojun*, Tofan *mōän* (*mojnu*), Gagauz *bojnu*, Salar *bojny*, Kumyk *bojun*.

2. **Kōjn* ‘lap’: *qojyn* (Yenisei Turkic, Old Uyghur), *qoj* (MK), Turkish *koyun*, *-ynu*, Middle Kipchak *qojyn* (Caferoğlu 1931). Azeri *Gojun*, *-jnu*, Turkmen *Gojun*,

¹ Here and below C is a consonant, V is a vowel, S is a fricative, R is a resonant.

² Regarding the reconstruction of clusters of the *-jn*-type in the finals of stems see Dybo 1996: 38.

Khakas *xojyn*, Chuvash *xəʷ, xü*, dialectal *xüm*, Yakut *xōj* (*xoñño-*), Dolgan *konnok*, Tuvan *xoj*, Gagauz *qojnu*, Kumyk *qojyn*.

3. **bejn* ‘brain’: *meji, meji* (Old Uyghur), *meji* (MK, QB), Turkish *bejin, -jni*, Chaghatay *miji* (MA), *mejn* (Sanglax), Azeri *bejin, -jni*, Turkmen *bejni, mejni*, Khakas *mī*, Chuvash *mimə*, Yakut *mejī*, Dolgan *meñī*, Tuvan *mē*, Tofan *mā*, Kumyk *miji*.

4. **gökr*, **Kökre* ‘chest, breast’: *kögüz* (Old Uyghur), *kögüz* (MK), Turkish *göğüs, -ğsü*, Chaghatay *kögs, kögüs* (Sanglax), *köküs* (MA), *kökre* (Velj aminov-Zernov 1868, Borovkov 1961), Azeri *köks, -ü*, Turkmen *gövüs, kükre*, Khakas *kögis*, Chuvash *kəʷgəʷr*, Yakut *köyüs*, Dolgan *köksü*, Gagauz *gūs*, Salar *göfrix*, Kumyk *kökürek*. Probably a cluster, considering the sonorization of the central consonant (clearly primarily voiceless after the Siberian and Chuvash reflexes), which can be positionally conditioned only by contact with -z (later devoiced in the Proto-Oghuz cluster?).

5. **čekn* ‘part of the shoulder between the neck and the shoulder blade’: *čikin* (Old Uyghur), Turkish *čekin, çekini/ çijin, -jni*, Chaghatay *čikin* (Velj aminov-Zernov 1868, Sanglax), Azeri *čijin, -jini*, Turkmen *čigin*, Chuvash *šan* ‘body’. Probably a cluster; otherwise we have no explanation for the variation of voiced/unvoiced consonant and for the full dropping of the guttural in Chuvash.

6. **āln* derived from **āl*, ‘front’: *alyn* (Old Uyghur), *alyn* (MK, QB), Turkish *alyn, -lny*, Chaghatay *alny-da* ‘in the presence of’ (Velj aminov-Zernov 1868), Azeri *alyn, -lny*, Turkmen *ālyn*, Khakas *alny*, Chuvash *om*, Tuvan *alyn*, Gagauz *anny*, Salar *aldy-*. Probably a cluster, cf. **gein* ‘daughter-in-law’: Chuvash *kin*, Yakut *kijit* (*plur.), Salar *kiin*, but **Kalyŋ* ‘bride-price’: Chuvash *xoləm*, Yakut *xalyŋ, xalyym*.

B. Alternative data force us to reconstruct a vowel between two final consonants.

1. **ogul* ‘son’: *oγul* (Orkhon Turkic, Old Uyghur), *oγul* (MK), Turkish *oğul, -ğlu*, Chaghatay *oγul* (Pavet de Courteille), Azeri *oγul, -γlu*, Turkmen *oγul*, Khakas *oγyl, ōl*, Chuvash *γvəl*, Yakut *uol*, Dolgan *uol*, Tuvan *ōl*, Gagauz *ōl*. Probably not a cluster, since the *γ* clearly demonstrates an intervocalic development.

2. **agyŋ* ‘mouth’: *aγyz* (Orkhon Turkic), *aγyz, aγaz* (Old Uyghur), *aγyz* (MK), Turkish *ağyz, -ğzy*, Chaghatay *aγyz*, Azeri *aγyz, -γzy*, Turkmen *aγyz*, Khakas *ās, aqsy* (3rd person), Chuvash *vəʷrəʷ, urə-lə*, Yakut *uos*, Dolgan *uos*, Tuvan *ās, aqsy* (3rd person), Tofan *ās, aqsy* (3rd person), Gagauz *ās*, Salar *aγyz*, Kumyk *awuz*. Probably not a cluster, since the *γ* clearly demonstrates an intervocalic development (cf. ***gökr*).

3. **biagyŋ* ‘liver’: *baγyr* (Old Uyghur), *baγyr* (MK), Turkish *bağyr, -ğry*, Chaghatay *baγyr* (Borovkov 1961, MA), Azeri *baγyr, -γry*, Turkmen *baγyr*, Khakas *pār*, Chuvash *pəʷver*, Yakut *byar*, Dolgan *byar*, Tuvan *bār*, Tofan *bār*, Salar *baγyr*. Probably not a cluster; cf. above.

4. **egin* ‘shoulder’: *egin* (Old Uyghur), *egin* (MK), Turkish *eğin, -ğni*, Chaghatay *egin* (Velj aminov-Zernov 1868, Pavet de Courteille), *in* (Pavet de Courteille), Azeri *äjin, -jni*, Turkmen *egin* (dialectal), Khakas *iñni*, Chuvash *avən, an*, Yakut *ien*, Tu-

van *eyin*. Primarily not a cluster, cf. the reflexes of the stem with a primary cluster, **jeġn* ‘sleeve’ (*jeġ* in the majority of languages, Chuvash *savny*, Yakut *iax*).

5. **Kādyn* ‘in-law’: *qadyñ* (Yenisei Turkic, Old Uyghur), *qadyñ* (MK, QB), *qadyñ* (QB), *qajyn* (IM), Turkish *kajyn*, *-jny*, Chaghatay, Middle Kipchak *qajyn* (Pavet de Courteille, MA, Houtsma 1894), Azeri *Gajyn*, Turkmen *Gājyn*, Khakas *xazyn*, *xasty*, Chuvash *xorɔn* (*xon* < Tatar), Tuvan *katy*, Tofan *xatty*, Kumyk *qajyn*. Probably not a cluster, which the intervocalic development in Chuvash shows, cf. the development of clusters in the stems **bydnyk* ‘moustache’: *bydyq* (MK), *byjyq* (IM), Turkish *byjyk*, Azeri *byj*, Turkmen *myjq* (dialectal), Gagauz *byjyk*, Chuvash *myjyx*; **ednye-r* ‘saddle’: *eder* (MK), Turkish *ejer*, Chaghatay *eger*, Uzbek *egar*, Azeri *jāhār*, Turkmen *ejer*, Khakas *izer*, Shor *ezer*, Chuvash *jəner*, Yakut *yŋyyr*, Dolgan *yŋyyr*, Tuvan *ezer*, Tofan *e’zer*, Gagauz *jēr*, Salar *eŋer*, Kumyk *er*; **Kadgu* ‘sorrow’: *qadyu* (Old Uyghur), *qadyu* (MK, QB), Turkish *kajġy*, Chaghatay *qajyy* (Zajaczkowski 1961, Pavet de Courteille), Uzbek *qejyy* (dialectal), Azeri *Gajyy*, Turkmen *GajGy*, Chuvash *xojya*, Kumyk *qajyy*; **edge* ‘host’: *edi* (*idi*) (Old Uyghur), *ige* (late Old Uyghur), *ije* (TT 6, TT 8), *iđi* (MK), Turkish *ije*, *ys*, *is*, Tatar *ijä*, Chaghatay *eje* (Borovkov 1961, Veljaminov-Zernov 1868), *ije* (Pavet de Courteille, MA), *ige* (Borovkov 1961, Pavet de Courteille 1820). Uzbek *ega*, (dialectal) *jigä*, *ijgä*, Azeri *jijä*, Turkmen *eje*, Khakas *ē-zi* (haplology of **eze-zi* < **ede-si*), Chuvash *ije*, *ɣja*, Tuvan *ē-zi* (haplology of **ede-si*), Kumyk *jeje*.

6. **bEñir* ‘face’: *beñiz* (Orkhon Turkic), *meñiz* (Old Uyghur), *meñiz* (MK, QB), Turkish *beniz*, *-nzi*, *benze-*, Chaghatay *beñiz*, *meñiz* (Sanglax), Azeri *bäniz*, *-nzi* (Azizbekov 1965), *-nizi* (*Azerbajdžansko-russkij slovar*), *bänzä-*, Turkmen *meñiz*, *meñze-*, Khakas *mys* (Verbickij 1884), Gagauz *beniz*, *benze-*. Probably not a cluster – cf. the development of the cluster in the stems **biñr* ‘awl’: Turkish *biz*, Azeri *biz*, Turkmen *bij/byz*, Salar *piz*; **bēñr* ‘ulcer’: *bez* MK, Chaghatay *bez*, *mez*, Turkish *bez*, Azeri *bāz*, *vāz*, Turkmen *māz*; **byñl(yk)* ‘cat’: *müš* (MK), Turkish *pyšyk* (dialectal), *pisi*, Azeri *pişik*, Turkmen *pişik*.

7. **geñir* ‘nasal cavity’: Turkish *geniz*, *-nzi*, Azeri *gäniz*, *-nzi*, *gänzik*, Turkmen *geñz-ew*, Yakut *keñeri* ‘bridge of nose’. Probably not a cluster, cf. above.

8. **göñ-il* ‘heart, mood’: *köñül* (Orkhon Turkic, Old Uyghur), *köñül* (MK, QB), Turkish *gönül*, *-nlü*; *göjün*, *göjn* (dialectal), Chaghatay *köñül* (Sanglax, MA), Azeri *könül*, *-nlü*, Turkmen *gövün*, *köñül*, Khakas *köl*, *köñn-ə*, Chuvash *kəᵐᵐəᵑ*, Yakut *köñül*, Dolgan *köñül*, Tofan *xöl*, Gagauz *gön*, Salar *göjny*, Kumyk *göñül*.

C. The presence of a cluster in protoforms is unclear.

1. **bögür*, **bögrek* ‘kidneys’: *bögür* (Old Uyghur), *bögür* (MK), Turkish *böğür*, *-ğrü*, *böjrek*, *böbrek*, Chaghatay *bögrek* (Pavet de Courteille), Azeri *böjür*, *-jrü*, *böjräk*, Turkmen *bevrek*, *bövrek*, Khakas *pügürek*, *pürek*, Chuvash *püre*, Yakut *büör*, Tuvan *büre*, Tofan *börek*, Gagauz *bür*, *büre*, *börek*, Kumyk *büjrek*.

2. **čygyr* ‘boundary, path’: *čyru-*, *čyyr* (MK), Turkish *çığır*, *-ğrı*, Chaghatay *čyyr* (Sanglax), Turkmen *čyyr*, Tuvan *šyyr*, Tofan *šyyr*.

3. **omur* ‘shoulder’: Turkish *omuz*, *-mzu*, Chaghatay *omuz* (Pavet de Courteille 1820), Turkmen *omuz*, Chuvash *ᵑᵐᵑᵑ*, Kumyk *omuz*.

4. **burun* (**buryn*) ‘nose, before’: *burun* (Old Uyghur), *burun* (MK, QB), Turkish *burun*, *-rnu*, Chaghatay *burun* (Sanglax), Azeri *burun*, *-rnu*, Turkmen *burun*, Khakas *purun*, Yakut *murun*, Dolgan *munnu*, Tuvan *murnu*, Tofan *murnu*, Gagauz *burnu*, Salar *purny*, Kumyk *burun*.

5. **Karyn* ‘belly’: *qaryn* (Old Uyghur), *qaryn* (MK, QB), Turkish *karyn*, Chaghatay *qaryn* (Sanglax, MA), Azeri *Garyn*, *-rny*, Turkmen *Garyn*, Khakas *xaryn*, Chuvash *хырм*, Yakut *xaryn*, Tuvan *xyryn*, Tofan *xyryn*, Gagauz *qaryn*, Salar *qaryn-taş* ‘relative’, Kumyk *qaryn*.

6. **ug-ur*, **ug-ra* ‘to meet; occasion; time’: *uyra-*, *uyur* (Orkhon Turkic, Old Uyghur), *uyra-*, *uyur* (MK), Turkish *uğra-*, *uğur*, *-ğru*, Chaghatay *oğur* (Radloff), *uyra-* (Borovkov 1961, Azeri *uyur*, *-yrun*, Turkmen *uGra-*, *uyur*, Gagauz *ūr*, *ūra-*, Kumyk *oğur*.

7. **iagyr* ‘heavy, pain’: *ayyr* (Orkhon Turkic, Old Uyghur), *ayyr* (MK, QB), Turkish *ağır*, *-ğri*, Chaghatay *ayyr* (Pavet de Courteille), Azeri *ayyr*, *-yry*, Turkmen *ayyr*, Khakas *ār*, Chuvash *jyvər*, Yakut *yar*, Dolgan *yara-kan*, Tuvan *ār*, Tofan *ār*, Salar *ayyr*.

Thus Turkish stems with the alternation (corresponding normally to Azeri stems with the alternation) practically could have originated with equal probability from Proto-Turkic stems with final clusters as from those without. Beside these, some stems appear that had final clusters in Proto-Turkic but resulted, in Turkish and Azeri, without alternation, e.g., **geln* ‘daughter-in-law’ (see above) > *gelin*, *gelini*, (?) **quln* ‘foal’ (cf. Chuvash *xum*) > *kulun*, *kulumu*, **kojn* ‘sheep’ (Tuvan *xoj*, Khakas *xoj*, Chaghatay *qoj* (Veljaminev-Zernov; Sanglax), Uzbek *qoj*, Kumyk *qoj*, Tatar *quj*) > *kojun*, *-junu*.

The same alternation as in nominal inflection appears in Turkic word-formation. In particular, it is widely attested in Turkish and Azeri. What has engaged our attention is that in Turkish and also in Azeri this alternation involves more nominal stems in the area of adnominal verb formation than in the nominal inflection, cf. *oju* ‘play’ – *ojunu*, but *ojna-* ‘to play’ (Proto-Turkic **oj-*), *orun* ‘place’ – *orunu*, but *orna-t-* ‘to replace’ (Proto-Turkic **or-un*). In deverbal word-formation the alternation covers practically all stems with convenient phonemic structure (cf. *čayyr-* – *čayryl*, *bujur-* – *bujruk*, *eğir-* – *eğri*, *evir-* – *evrin-*), and moreover, where variants with and without the alternation exist, words with the alternation have idiomatic meanings, which signifies that they are “older” than the ones without the alternation, cf. *čevir-* ‘to turn’ – *čevirim* ‘turn’ and *čevrim* ‘cycle, period’.

All represented facts would seem to suggest that the modern state of Turkish and Azeri originates from a state similar to the one in Turkmen (having purely phonemic conditions); and from there one can reconstruct the development by supposing processes of paradigmatic unification of the stem forms. Particularly, in the noun inflection the alternation is eliminated in the majority of cases. Among the 21 stems retaining the alternation, two stems are designations of family members, and 17 are names of body parts (see the list above), i.e. the lexemes belonging to the semantic groups that usually are employed with possessive affixes; a high incidence of such

forms clearly conditioned the conservation of the alternation. Thus, we should reconstruct for the Proto-Oghuz stage the Turkmen (or similar) situation, but this situation is neither a Proto-Turkic nor a Common Turkic one: as it was demonstrated above, in Proto-Turkic some final clusters existed in disyllabic stems that developed later in specific ways in some of the Turkic languages.

References

- Veljaminov-Zernov V. V. 1868. *Slovar' džagatajsko-tureckij*. Sankt-Peterburg: Tipografija Imperatorskoj Akademii Nauk.
- Caferoğlu A. 1931. *Abû-Hayyân. Kitâb al-Idrâk li-lisân al Atrâk*. Istanbul.
- Azerbajdžansko-russkij slovar'* 1. Baku 1986.
- Azizbekov Kh. A. (ed.) 1965. *Azerbajdžansko-russkij slovar'*. Baku.
- Borovkov A. K. 1961. "Badâ'i al-lughat": *Slovar' T'â'li Imâni Geratskogo k sočinenijam Ališera Navoi*. Moskva.
- Deny J. 1920. *Grammaire de la langue turque (dialecte osmanli)*. Paris: Ernest Leroux.
- Dmitriev N. K. 1962. *Stroj tjurkskix jazykov*. Moskva: Izdatel'stvo vostočnoj literatury.
- Dybo A. V. 1996. *Semantičeskaja rekonstrukcija v altajskoj étimologii*. Moscow: Jazyki russkoj kul'tury.
- Grammatika turkmenskogo jazyka 1. Fonetika i morfologija*. Ašxabad: Ylym, 1970.
- Houtsma M.Th. 1894. *Ein türkisch-arabisches Glossar*. Leiden.
- Kononov A. N. 1956. *Grammatika sovremenogo tureckogo literaturnogo jazyka*. Moskva: Leningrad: Izdatel'stvo Akademii nauk SSSR.
- Pavet de Courteille = Pavet de Courteille M. 1820. *Dictionnaire turc-oriental*. Paris.
- Pokrovskaja L.A. 1964. *Grammatika gagauzskogo jazyka. Fonetika i morfologija*. Moskva: Nauka.
- QB = Karakhanid Turkic according to *Qutađyu Bilig*.
- Zajączkowski A. 1961. *Najstarsza wersja turecka Husräv u Širîn Qutba 3. Słownik*. Kraków: Librairie franco-polonaise et étrangère.
- Radloff = Radlov V. V. 1899-1911. *Opyt slovarja tjurkskix narečij 4*. Sankt-Peterburg.
- Sanglax = *Sanglax. Facsimile text with an introduction and indices by Sir G. Clauson*. (E. J. W. Gibb Memorial, New Series 20.) London, 1960.
- Türkmen diliniñ orfoepik sözlügi*. Ašgabat: Ylym, 1978.
- TT 6 = Gabain A. von & Rachmati G. R. 1934. [Sitzungsberichte der Preussischen Akademie der Wissenschaften] Jahrgang 1934. Phil.-hist. Klasse, 1934.
- TT 8 = Gabain A. von 1954. *Türkische Turfan-Texte 8*. [Abhandlungen der Deutschen Akademie der Wissenschaften zu Berlin. Klasse für Sprachen, Literatur und Kunst] Berlin.
- MA = Chaghatay according to the dictionary *Muqqadimat al-Adab*.
- MK = Karakhanid Turkic according to the dictionary of Mahmüd Kāšyarī.
- Mustafaev, E. M. E. & Starostov, L. N. (eds.) 1977. *Turecko-russkij slovar'*. Moskva: Russkij Jazyk.
- Verbickij W. 1884. *Slovar' altajskogo i aladagskogo narečij tjurkskogo jazyka*. Kazan'.

Khazarica: Notes on some Khazar terms

Peter B. Golden

Golden, Peter B. 2005. Khazarica: Notes on some Khazar terms. *Turkic Languages* 9, 205-222.

This paper gives a brief overview of the question of the language(s) of the Khazar Qağanate (ca. 650 – ca. 965-969), one of the most important Turkic states of the western Eurasian steppe zone. The language is known only through transcriptions of isolated words scattered in a variety of contemporary or near contemporary sources. This study focuses on certain sound changes that appear to be characteristic of one of the principal Turkic languages of the Khazars, a language that shows a closer affinity to the Oğuro-Bulğaric languages rather than Common Turkic.

Peter B. Golden, Rutgers University, Department of History, Conklin Hall, 175 University Avenue, Newark, NJ 07102 U.S.A. E-mail: pgolden@andromeda.rutgers.edu

In 1991, Marcel Erdal, who has contributed so much to our knowledge of Old Turkic and whose many accomplishments we honor in this issue, published an important article on the name of a Khazar *ghulâm*, Îtâkh/Itaq, in the service of the ‘Abbâsîd caliphs of the first half of the ninth century (Erdal 1991). His work prompted me to review these anthroponyms (Golden 2002-2003, 2004), which I had not included in my earlier Khazar word list (Golden 1980). Erdal’s masterly overview of the current problems in the study of the sparse remnants of the Khazar language given at the First International Colloquium on the Khazars Jerusalem, 1999, (Erdal, in press) has now served as the inspiration for this brief essay. Many basic questions regarding the Khazar language (or more likely languages) remain unresolved. Largely, this is due to the lack of texts that can with certainty be identified as Khazar. It may well turn out that some of the runiform materials found in the Don-Kuban’ regions (see Kljaštornyj 1979, Kljaštornyj & Vásáry 1987, Kyzlasov 1994) and other parts of the Western Eurasian steppelands and Eastern Europe will prove to be Khazar and hence will provide the substantial texts we need to resolve some of these questions. We must await their full publication.

One of the many problems we encounter is rather basic: one cannot be sure what “Khazar” really denotes beyond its obvious political dimensions. References to the Khazars before the mid-seventh century are probably anachronistic (Golden 1980 I: 50-51, 58-59; Zuckerman 2001: 313-325). The Khazar realm was an offshoot of the Western Türk Empire and very likely ruled by a dynasty of Türk origin. Hence, the gentilics “Türk” and “Khazar” are frequently used interchangeably in the sources.

The Khazar Qağans dominated a complex union of Turkic tribal groupings (including, probably, an inner core of Türk tribes that accompanied the dynasty) speaking, one may assume, a number of Turkic languages. Was there an actual “Khazar” tribe or pre-650 AD tribal union bearing this name? Did it form in the Daghistanian-Caspian steppes, the “Berzilia” connection, prior to the sixth century as has been claimed on the basis of ethnogenetic legends recorded by Byzantine and Syriac sources (Pletněva 1976: 15)? Did it have a pre-“Berzilian” history? Indeed, is “Khazar” (*Qazar*) actually a tribal name? Róna-Tas (1982b, 1982c) connects them with the Uyğur *Qasar* tribe and the same name (or anthroponym?) noted in Uyğur runic inscriptions of the mid-eighth century, with Khazar/*Qazar* as its Bulğaric form. Was it originally a political term, denoting, perhaps, a group that broke away from the Western Türk core (cf. the later *Qazaqs* with which term *Qazar* might conceivably be related)? The ‘Abbâsid caliphs, when they created their *gulâm* army and then settled them in Samarrâ’, gave their Khazar servitors, their own land allotments, next to those they termed “Turks” (al-Ya‘qûbî 1892: 258-259, 262). As the latter were settled near the *Farâğina* ‘men of Farghâna’, who were presumably Iranians (Sogdians), but may have included some Turkic elements, these distinctions (Turk, Farghanian, Khazar) were, perhaps, geographical in origin rather than ethno-linguistic—although the latter possibility cannot be excluded. There is no doubt that the ‘Abbâsids ranked the Khazars among the “Turks” (used as a generic) and the sources often interchange the *nisbas* “at-Turkî” and “al-Xazarî” to describe *gilmân* who were Khazars or perhaps came from Khazaria.

The classical Muslim geographers, however, did not quite know where to place Khazar linguistically. Although some scholars (e.g. Gadžieva & Serebrennikov 1977: 3, following Baskakov 1969: 237, 231, among others) have declared the Khazars, on the basis of the statements of the Islamic geographers, as “ethnically close to the Bulğars” whose language, they aver, “had much in common with Khazar”, the actual reports are far more ambiguous. Al-Iştaxrî, writing in the mid-tenth century (although the first variants of his *Kitâb Masâlik al-Mamâlik*, part of the geographical school of al-Balxî, 850-934, may date to the 930s, see Kračkovskij 1957, IV: 197-198) and probably basing himself on materials from the early part of the tenth century, at the latest (Dunlop 1954: 102-104), if not considerably earlier (Zaxoder I: 49-51) reports two conflicting notices (al-Iştaxrî 1927: 222, 225): “the language of the Khazars is different than the language of the Turks and the Persians, nor does a tongue of (any) group of humanity have anything in common with it” and “the language of the Bulğâr is like the language of the Khazars, but the Burtâş have another language”. The Burtâş were located between the right bank of the Middle Volga and the Middle Don, east of the Iranian Alans, south of the Finnic Mordva and southwest of the Volga Bulğars. Their ethno-linguistic affiliations have long been the subject of debate. Finno-Ugric, Iranian (Alano-As) and Turkic connections have been suggested. They may well have been a mixed grouping (see discussion in Romašov 2002-2003: 168-179). Ibn Ĥawqal (1992: 332, 335), who was in the region after 968-9 not long after Khazaria had been overrun by the Rus’ acting in alliance with the Oğuz

Turks, see Golden (1980, I: 81-83), Konovalova (2003:171-90) basically repeats al-Iṣṭaxrī's remarks: "the language of the Khazars is different than the language of the Turks and the language of the Persians. None of the languages of humankind has anything in common with them". Al-Mas'ūdī (1894: 83), writing in the mid-tenth century, lists the Khazars among the "types of the Turks" (*ajnās at-Turk*) and adds that "they are called *Sabīr* in Turkic and *Xazar* in Persian". This is most probably a garbled reflection of the importance of *Sabir* elements, see Golden (1992: 104-106, 236), Byzantine: Σάβιροι etc. [Sabir/Savir/Sābir etc.?], Armen.: Սավիրք [Savir-k', pl.], Khazaro-Hebrew: סאַויר [Sāvīr], see Golden (1980, I: 256), among the tribes brought into the Khazar union. The Sabirs also constituted one of the tribal conglomerations of Volga Bulğaria, the سوار *Sawâr* (sometimes written: سُوَار [Suwâr]), with changes typical of Bulğar, *Sabir* > *Sawâr* (e.g. Common Turkic *il-teber* ~ Volga Bulğar *Yıl-tawâr*, see Ibn Faḍlân, Togan (1939): Arabic, 1 / Germ. Transl. 1). The Khazar and Volga Bulğar forms of this ethnonym (and shared ethnic component) appear to differ here (but, cf. the Khazar *gulâm*, Waṣīf b. Ṣawârtakīn [Sawâr Tegin] al-Xazarī, Golden 2002-2003: 25). Al-Birūnī (1923: 41-2), the Khwārazmian polymath (d. ca. 1050), in discussing the (Volga) Bulğars and Sawârs, who lived in the "most remote region of the habitable lands, near the end of the Seventh Clime", remarks that their language was "compounded (*umtazija*) of Turkic and Khazar". Understandably, al-Muqaddasī (1987: 283, writing ca. 985) termed the Khazar language "very incomprehensible" (*šadīd al-ingilâq*).

We do not know whether one or several languages were denoted under the term *Xazar* in the sources. We do not know the extent to which an "official" Khazar tongue was used in this polyglot state. Al-Iṣṭaxrī (1927: 191-2) does mention that "Khazar" was spoken by the populace of Bâb al-Abwâb (Darband) along with other "language(s) of their mountains". Hence, something that outside observers, however confused, recognized as "Khazar" must have existed, but the picture remains unclear at best. Whether traces of Khazar can be discerned in some of the northern dialects of Azeri Turkic or present day Qıpçaq languages of the North Caucasus (see Gadžieva & Serebrennikov 1977: 6-12) remains an open subject. What we can say with some assurance is that a number of Turkic languages were spoken in Khazaria, one (or more) of which probably had affiliations with Oğuro-Bulğaric. I prefer the term "Oğuro-Bulğaric" rather than "Bulğaric" as the former encompasses speakers of kindred tongues that were not, strictly speaking, "Bulğars". Róna-Tas (1982a: 119) noted that what he termed "Old Bulgarian" comprised "more dialects, and perhaps even languages" than is apparent to us today. It is important to bear this in mind. Indeed, it is not impossible that other branches of Archaic Turkic (for want of a better term), perhaps quite distinct from "Oğuro-Bulğaric" and "Common Turkic" were present, but have simply disappeared, absorbed by other Turkic groups. One should also take into account that very little of Danubo-Balkan Bulğaric, a contemporary of Khazar, has survived (Pritsak 1955, Tekin 1987, Parzymies 1994). Similarly, the Volga Bulğar inscriptions, our other major, pre-modern source for Oğuro-Bulğaric, date to a considerably later period (the late thirteenth-mid-fourteenth

century) and provide rather sparse material (Tekin 1988, Erdal 1993). The reconstruction of the history of Chuvash, the only surviving Oğuro-Bulğaric tongue, is hardly an uncontested field (see Róna-Tas 1982a, Fedotov 1996b, Tenišev 2002: 677ff.). It is probable that even if “Khazar” (or one of Khazaria’s principal languages) did belong to some branch of Oğuro-Bulğaric, it was sufficiently different that people distinguished between the two. This is very different from the relative uniformity of Common Turkic of which al-Iṣṭaxrī remarks (1927: 9) “as for the Turks, all of them, from the Toquz Oğuz, Qırğız, Kimek, Oğuz, Qarluq, their language is one. They understand one another”. The language which the Islamic geographers called “Khazar”, at best then, may have had some similarities with Bulğar, which was known to the Muslim world through trade and the Islamization of the Volga Bulğars in the early tenth century. Otherwise, Khazar seems to have been rather alien.

The remnants of the Khazar language, largely titles, names and a few toponyms, are recognizably Turkic, but complexities of interpretation abound. Thus, even the personal name discussed by Erdal (1991), Ítâx [*Itaq], while showing the *-aq/-ak* suffix found in many Oğuro-Bulğaric forms (e.g. Common Turk. *ay* ‘moon, month’ > Oğuro-Bulğ. **ayaq* > Chuv. *uyăx*, Common Turk. *qıl* ‘thick hair’ > Oğuro-Bulğ. **qılıq* > Chuv. *xělěx* ‘horse hair’ etc., see Tenišev et al. 2002: 700ff.), lacks, at least in this anthroponym, the *i-/ı-* > *yi-/yı-* shift that one finds in Oğuro-Bulğaric (cf. Common Turk. *it/ıt* ‘dog’ ~ Oğuro-Bulğ. **yutaq* > Chuv. *yută*) and is apparent in the Khazar title *yiliglyélig* (see below). We continually encounter a mixed or ambiguous picture. The name of the oft-debated Khazar fortress of Sarkel, built in ca. 840-41 (this dating has recently been demonstrated by Zuckerman 1997), is a case in point. Byzantine sources have Σάρκελ or variants of it. This can be read as *Sarkel* or *Šarkel* as Greek has no letter for *š*. The Khazar Hebrew correspondence (letter of Joseph) has שרכליל which can be read as *Šarkil* (more likely given Medieval Hebrew traditions of transcribing foreign terms) or *Sarkil*. Two epigraphs on Biblical Codex No. 51 of the Firkovič Collection in St. Petersburg, however, have סרקהל [srql] **Sarqil* = **Sarqıl* (for these forms see Golden 1980, I: 239-240, Kokovcov 1932, 105-6). In addition, a Western Circassian (Bžedux) tale has preserved the name as *Sarqahl* of the *Qazahra* (Khazars, see Colarusso 1992-1993: 63-68). Constantine Porphyrogenitus (1967: 182) writing in the mid-tenth century, but using sources dating back to the building of Sarkel, ca. 840-41 (on this date, see Zuckerman, 1997) tells us that the term means ‘white house’ (ἄσπρον ὄσπίτιον). Theophanes Continuatus (1838: 122, an anonymous collection of three or four authors covering the period 813-961 the last part of which was probably written before 963, *Prodolžatel’ Feofana* 1992: 217-219) translates this toponym with the more classical λευκὸν οἶκημα - with the same meaning. This translation is confirmed by the Rus’ name of the town, *Bela Veža* “White Fort” and the Arabic *al-Bayḏā* ‘white (town)’, although the latter was used, perhaps, for more than one town or fort in the region. Hence, an Oğuro-Bulğaric **Šarı(ğ) kil* ‘white house’ (Modern Chuvash *Šură kil* “white house”), would appear to be the best match for at least one of the forms that has come down to us. Ligeti posits the Khazar

word as *šar/šari* ‘white’ (Ligeti 1986: 18, 95, 475-457) presumably from an Oğuro-Bulgaric **šariğ*. However, its Common Turkic equivalent, *sariğ* (see Tenišev 1997: 681-682) as well as Hung. *sárga* (< Oğuro-Bulgaric, see Ligeti 1986: 18, 95) and Class. Mong. *šira*, Mod. *šar*, all denote ‘yellow’. The two Biblical Codex epigraphs coming through Firkovič may have been altered (on Firkovič and the much-debated question of his forgeries, see Vixnovič 1997), but what is one to make of the Circassian form? Colarusso (1992-1993:64), a leading specialist on the Nart Sagas and North Caucasian folklore remarks that as “Circassian lore shows very little influence, old or new, from Russian sources” the form with *s* “is unlikely to come from a European form of the name”. Thus, **Sarikel* is also not impossible. *Killgil* ‘house’ etc. may be an Iranian loanword in Turkic (Old Iran. **gr̥da* > *gil* in Southwestern Iranian [e.g. Persian] and **guli/gali* in Northwestern Iranian [e.g. Kurdish]), but this too is not without problems (Golden 1980, I: 241-242). Fedotov (1996a, I: 291-292) notes Chuvash *killkel* ‘žilišče (dvor), dom, podvor’e’ in a number of oconyms, but says nothing of a possible Iranian origin. Indeed, he compares it with Evenki *gulla* ‘žilišče, izba, xizina, zimmij dom’. Starostin et al. (2003, I: 570-571) does, indeed, derive it from Altaic **gūli* ‘dwelling, cottage’ Proto-Tung. **gūle* ‘hut dwelling-place’, Proto-Turk. **gūle* ‘house, home, hut’, Proto-Jap. *kura* ‘shed’. Yakut *kūlä* is probably a borrowing from Tungusic. Clauson (2002: 148) notes the probably unrelated Turkic suffix *-ğil/-gil* ‘apparently associated with colours’. Among the Modern Turkic languages, only Oğuz and Chuvash seem to know this word. There is no problem with a borrowing of Middle Persian *gil* etc. into Oğuz (Turkish, Azeri *gillgil*), as the latter had close relations with speakers of Persian. However, in Oğuz it denotes ‘the family of’ i.e. the household kinship unit rather than ‘house’ itself. Its Iranian origins have not really been demonstrated. A possibly better case might be made for Chuvash. Old and Middle Iranian loanwords are certainly present in a variety of languages (Turkic and Finno-Ugric) of the lower and Middle Volga region, the Ural and the Ponto-Caspian zones. Southwestern Iranian (Persian) forms could have come from merchants from Iran. The Alano-As peoples (speakers of Northeastern Iranian) are another potential source (cf. Iranian loanwords in Hungarian, Ligeti 1986: 162-174; Harmatta 1997). Nonetheless, the sparse Alanic linguistic data preclude a detailed elaboration of the process.

Another Khazar toponym of interest is سارغشن [sârgšn]: *Sariğšin* or **Sariğčîn* (Arabic ش [š] is sometimes used to represent č). The mss. (see forms in Golden 1980 I: 237-238) are unanimous in having initial س [s] not ش, hence it cannot be **šariğšin*. *Sariğšin/Sariğčîn* appears to be Common Turkic. The suffixes *-čîn/-šin*, *-ğčîn* (Ligeti 1986: 478, Clauson 2002: 149) are used with colors in Turkic and Mongol = “Yellow” or “White City”.

The well known title *beg* ‘clan or tribal chieftain’ is well-attested in Khazar. Clauson (1972: 322-3; Clauson 2002: 15) viewed *beg* as a borrowing from Chinese *bo* ‘hundred, head of a hundred men’ (Early Middle Chin. [up to ca. 600] paŋjk/pe:jk, see Pulleyblank 1991: 42 = Karlgren 1996: 206, Archaic Chin. [pre-Han, i. e. pre-third century BC] **pāk*, Ancient Chin. [ca. 600] pōk). This has been connected

with Mong. *begi* and Manchu *beile* (Sevortjan 1978: 99-100). The former is undoubtedly a borrowing from Turkic *beg*, the latter perhaps from the Old Turkic title *boyla* (Cincius 1975 I: 120). Others, however, associate *beg* with Middle Iran. *bag*, *bağ*, *bağa* < Aryan *bhaga* ‘god, lord’ etc. (Sevortjan 1978: 100, on Iranian forms, see Rastorgueva & Edel’man 2000-03 2: 48-49). There are three variants of this title associated with the Khazars in the Arabic sources and two names containing it in Arabic and Armenian sources. The earliest attestation is in Lewond (latter part of the eighth century), whose *History* (*Patmut’iwn*) covers the period 632-788 (Lewond 1982: 25). He records a Khazar invasion of Armenia, s. a. 730, sent by the *Xatun* Փարսբիթ *P’arsbit’*, the mother of the recently deceased Khazar Qağan (Lewond 1982: 107). In the name *P’arsbit’*, the final *թ* [tʰ] could be a corruption of *ք* *k’* or perhaps reflect a pronunciation of *č*, i. e. *Parsbik’/Parsbi č’*. The name consists of *pars/bars* ‘tiger, leopard, panther’ (Sevortjan 1978: 68-70; Clauson 1972: 368) + *bik’* < *beg* (on the ms. variants and texts, see Golden 1980 I: 205-06). During this same period, the Islamic sources note a very similar name, in this instance “the son of the Qağan” who leads the Khazar armies against the Arabs: Ibn A’tam al-Kûfî (d. 926) has *بارسبيك* *bârsbîk* (*Barsbik/Barsbeg*) or corruptions of it (for manuscript forms, see Golden, 1980, I: 158). Bal’amî (tenth century Persian translator of Ṭabarî’s history, with additions) has: *بارجنگ بارخیک بارحیک بارجیک* *Bârxîk, Bârĥîk, Bârĵîk, Bârjank* etc. (ms. variants in Golden 1980 I: 157-158, see there for later variants from the Turkish translation of Ṭabarî and the *Darbandnâma*). Whether the two people bearing this name are actually one and the same, but confused by our sources is also not impossible. In any event, there is clearly a form with the title *beg/big/bik’* in it. The Armenian form hints at a change in the pronunciation of Turkic *beg*. Al-Iṣṭaxrî and Ibn Ḥawqal (using a form that had already become fixed in Muslim sources) note *بک* [bk] = *beg* and *بک* (*bâk*, recte: *بلك* [ylk]) = *yilig/yélig* (see below and Golden 1980 I: 164 for ms. variants). Well before their time, however, this form had changed. Thus, their contemporaries Constantine Porphyrogenitus and Theophanes Continuatus, in connection with Sarkel, note both “the then Qağan and Beg of Khazaria”: *ὁ γὰρ χαγάνος ἐκεῖνος καὶ ὁ πῆχ Χαζαρίας* (Constantine Porphyrogenitus 1967: 182, Theophanes Continuatus 1838: 122, *Prodolžatel’ Feofana* 1992: 56 and ms. variants in Golden 1980 I: 163). Clearly, in *πῆχ* [pex = bex] we have evidence, by 840-41, of a shift *beg/bek’* [bik’] > **beğ* > *bex*. By the early tenth century, further changes had occurred. Ibn Faḍlân who journeyed to Volga Bulğaria in 921-22 notes that the deputy of the Khazar Qağan bore the title of *خاقانیه* [xâqân bäh] *Qağan Beh*. *Bex* had now become *beh*. The Hungarian historical-etymological dictionary views Hung. *bő* ‘full, rich’ as probably deriving from the Turkic *beg* via *beğ* (*MNyTESz* 1967-76 I: 356-7, see also Németh 1991: 284). Ligeti (1986), however, omits it from his analysis of the Turkic elements in Hungarian. Its development in Hungarian is usually explained as: *beğ* > **beÿ*, *böÿ* > *bő*. If *bő* does, indeed, derive from *beg*, Khazar *beh* could have been its source.

We see a similar development with the Khazar title *yilig/yélig* [*yéllig?*] ‘(junior) king, prince’ (cf. Turk. *éllig*, *élig* ‘having a realm, king, ruler, master’ < *él* ‘realm’,

Clauson 1972: 141-142, Erdal 2004: 51) with the prothetic *y-* common to Oğuro-Bulgaric (but not unknown in Common Turkic). It is unlikely that **𐰇𐰏𐰍** represents Old Turk. *yeläk* ‘banner, flag’ also found as a personal name (Tenišev 1997: 565-6) as this probably would have been **jeläg*. In addition to al-Iṣṭaxrī and Ibn Ḥawqal, as we have seen, this title is found in Yâqūt (d. 1229) and in later authors, e.g. Zakariya al-Qazwīnī (*‘Ajā’ib al-Buldān wa Axbār al-‘Ibād*, composed ca. 1273, see Qazwīnī 1969: 585) and al-Bakūwī (fifteenth century) in forms that are clearly **𐰇𐰏𐰍** [ylk]: *yilig/yelig/yeleg* (see ms. variants in Golden 1980 I: 184-185 and discussion in Golden 1975). It also appears as the name of the second son of the Hungarian ruler Árpád (late ninth century), noted by Constantine Porphyrogenitus (1967: 175): *’léλεχ* (*Yelex*) with the shift *-g > -ğ > -x* in evidence. The early Hungarian form probably reflects its Khazar source. The Hungarian place-name Üllő (in Pest County) derives from *ellig/ilig/ iläg* not *yelex* (Golden 1975; Ligeti 1986: 42, 94). Hung. *Jelő*, however, is from “Khazar-Qabar” *yelex* (Ligeti 1986: 486).

In the Long Redaction of Joseph’s letter we encounter a country north of the Black Sea (*yam Qūstandīnah*) called **בצרה** [bṣrh] undoubtedly a corruption of **בצנה** (bṣnh, Kokovcov, 1932, 31-2/102, 110n.32) = *Bacnah* = *Bačanah* or *Bäčänäh*. This is the Khazar name for the Pečenegs/Bečenegs with the *-q* or *-k/-g > -ğ > -x > -h* and ultimately *> -0* shift. In point of fact, we have several distinct forms of this name in our sources. Arabic-script accounts have **بجناك** [bjnāk] and **بجاناك** [bjânāk]: *Bäčänāk* and *Bečänāk* = *Bečenäk/Bečenäg* or *Pečeneg* (e.g. see Ibn Xurdādbih 1889: 31, Kāšgarī 1982-1985 I: 101-102 followed by Rašīd al-Dīn and Abu’l-Ġāzī, all noting a Pečeneg grouping that had been taken into the Oğuz union after the latter briefly, ca. 1036 – ca. 1050, became the dominant element in the Pontic steppes, see Golden 1992: 207-298, 264). In the Islamic sources we find another variant of this name: **بجني** [bjny] *Bačanā/Bäčänā* (or *Pačanā/Pečene*) noted by al-Mas‘ūdī in both the *Murūj* (1966-79 I: 235-236) and his *Tanbih* (1894: 180-181) among the four Turkic tribes associated with a town or place called Wulundur (a later form of the ethnonyms *Onoğundur*): *Bečenäk*, *Bajğird*, *Bečenä* and *Nogurda* [نوگردة] and *Bečenäk*, *Bečenä*, *Bajğird* and *Nogurda*. Ibn al-Aḫṭar (1965-1967 I: 339) has a truncated version of the event and notes *Bečenäk*, *Bečenä* “and two others” who had formed a union, s.a. 322/933-934 and attacked Byzantium (Knjaz’kij 2003: 15-16). As I suggested sometime ago (Golden 1975, see Arabic texts there as well), our authors may have been reporting two different traditions, each reflecting the same *two* groups of peoples, the Pečenegs (the *Bečenäk/Bäčänä*) and the Hungarian tribal union (*Bajğird* and *Nogurda* < **Onogurda*, the *Nogurda* are not the Novgorodians as Knjaz’kij, 2003: 16 and others have claimed). Rašīd al-Dīn (1373/1994 I: 60, cf. also Rašīd al-Dīn 1969: 46) has preserved **بيچنه** [bičnah] *Bäčänä*, one of the sons of Kök Xan of the Uč Oq subdivision of the Oğuz. Abu’l-Ġāzī (1603-1664) the Khivan khan and historian, in his *Šajara-yi Tarākima* (1958: Türkmen text 31) repeats Rašīd al-Dīn and later adds (1958: 41) that the **بجنه** [bjñh] *Bäjänā/Bäčänä* clan of the Türkmen was called *Īt-Bäjänä* (‘Dog-Bäčänä’) by the Salor clan, their enemies. The forms in Arabic script, as we have seen, can be read as rendering a Turkic **Bečänäk/Pečänäk* etc.

A similar ambiguity can be seen in the Greek forms Πατζινακίται, Πατζινάκοι, Πατζανάγοι etc. (Moravcsik 1958 II: 247-248) where π can render *p* or *b* producing *Bačinaq*, *Pečenek* etc. Forms with initial *p* are also found in Rus' (Печенѣгъ), Armenian Պաճինակ [Pacinak in Matthew of Edessa], Georgian Pačanikni and *Pačaniket'i* = 'Pechenegia' *K'art'lis C'xovreba* 1955-1973, I: 45, 156, 157]. Latin sources, depending on local tongues, have *Pizenaci*, *Bisseni* etc. (see Golden 1992: 264, for these forms). Attempts to connect them with the *Beiru* (Early Middle Chinese pək juawk, Late Middle Chinese puək rywk, Pulleyblank 1991: 31, 269, previously read as pək řǰjok), noted by the *Suishu* (composed 627-36) in its listing of Tiele tribes (a tribal union arcing across Eurasia), "east of Fulin" (Rome, i.e. Constantinople) near the Enqu (Liu 1958 I: 127-128, II: 569 n. 663 = Early Middle Chinese: ?ən k^hut, Late Middle Chinese: ?ən k^hyt = **Ongur*, **On[o]gur*, Pulleyblank 1991: 87, 266) and Alans (presumably in the Ponto-Caspian steppes), seem unlikely. The Hungarian term confirms a form with initial *b-*: *Báčänäg > Bäschänäg > *Besenyő* (Ligeti 1986: 268, Németh 1991: 85, 90, 172) as do Tibetan (*Be-ča-nag*), Khazar (Bačana/Báčänä) and some Latin sources. We also appear to have both velar and palatal forms. As this ethnonym is generally viewed as deriving from *baĵa(naq)* /*baĵinaq* 'brother-in-law' (Sevortjan 1978: 24-25, Németh 1991: 85, 90, 172, Ligeti 1986: 268), the Danubian Bulġaric form of which, папеногъ 'brother-in-law' (cf. Chuvash *pušana*, Fedotov 1996a I: 453) was preserved in Church Slavonic (D'jačenko 1993: 412, cf. also Serbo-Croatian *pašanac*); the velar form was probably the original. It has long been accepted that Bačana/Báčänä was a parallel form of Báčänäk/Pečeneg (Marquart 1961: 61, 63, 67, 78). I would suggest the presence of both velar and palatal forms in our sources and forms without the final *-k/-g* point to what was most probably a linguistically mixed group, including both Common Turkic and Oġuro-Bulġaric speakers each producing appropriate forms of their ethnonyms (the 'In-Law Tribe') recorded by different informants: *Baĵinaq/Baĵanaq/Bačana/Beĵänäk* ~ *Pačinaq/Pečeneg* etc. Pritsak (1975: 211ff.) suggested that the Pečenegs originally had a Tokharian (Yuezhi) base as well. It would be difficult to prove this on the basis of the surviving linguistic material (see Ligeti 1986: 506-511, Györffy 1990: 170ff). What has come down to us appears to be Common Turkic. Some terms, however, e. g. (Constantine Porphyrogenitus 1967: 166) Γερίχ (= Yäyiq or even Yeyäx, the Ural River, see below) show a shift *a > ä* explaining, perhaps the *Baĵinaq/Pečeneg* et al. variants. The early history of the Pečenegs who are first encountered in the Volga-Ural Mesopotamia and are not noted in the Orxon or other Old Turkic inscriptions, remains obscure (see overview in Golden 1992: 264-265).

Another title used by the Khazar Qaġan's deputy ruler, who by the late ninth and certainly early tenth century ran the actual affairs of state as the Qaġan became largely sacralized and tabuized, a situation clearly depicted in the Islamic sources, is: ایشا (Ibn Rusta, writing by 903): 'yšâ and ایشاد (Gardîzî, middle eleventh century, but using earlier sources): 'yšâd (see ms. variants in Golden 1980 I, 207). This is a variant of the Old Türk title *Šad* (recorded in Movsês Dasxuranc'i as Ḥωρḥ Šat', ms. variants in Golden 1980 I: 207) which in the Ashina Türk state was invariably given to

high officials of Ashina origin. In Chinese sources (Chavannes 1941, Index: 320) it is noted as *she* (EMC ʃiat), *sha* (EMC ʃəʃt /ʃɛ:t) and *cha* (EMC tʃʰəʃt/tʃʰɛ:t, Pulleyblank 1991: 279, 273, 47). The origin of this high title among the early Turkic peoples is clearly Iranian, going back to Old Iran. *xšāyaθiya* > Middle Pers. *šāh* or Old Pers. *xšāita* Avest. *xšāēta* > Sogd. 'xšēδ (see Bombaci 1974). This term is usually Arabized as *Īxšīd*/*Īxšīδ* (cf. Ibn Xurdādbih, mid-late-ninth century, for *ixšēδ*). How *xšāyaθiya* produced Türk *Šad* remains unexplained (Clauson 1975: 45). Aṭ-Ṭabarī (VI: 473, 476), however, notes the Sogdian form اخشاد [*xšād*]. This is clearly the source of the Khazar forms corrupted in the texts of Ibn Rusta and better preserved in Gardīzī: *ixšād* > *ihšad* > *išad*. Sogdian or Khwārazmian, another East Iranian tongue is the likely source. Khwārazmians were an important component of the Khazar court and personal guard (*Ors*) of the Qağans in Atil.

The Muslim geographers give a number of names of Khazar cities (or parts of the Khazar capital). One tradition, found in Ibn Xurdādbih, the *Ḥudūd al-ʿĀlam* (372/982), al-Muqaddasī and Yâqût has variants (see ms. variants in Golden I: 230-232) of the form *خملیخ [*xmlīx*] = **Xam-malix* or perhaps even *خمبلیخ [*xmblīx*] (< **Xam Balıq* < *Qam Balıq* < *Qan Balıq* 'City of the Khan'—'City of the Shaman' [*qam*] seems unlikely). Alongside of that, the al-Jaihānī tradition (early tenth century, see Göckenjan & Zimonyi 2001: 3-10) found in Ibn Rusta, *Ḥudūd al-ʿĀlam*, Gardīzī, al-Bakrī (1992 I: 446, writing in 1086) and al-Marwazī (late eleventh, early twelfth century), notes another city, *Sariğšin (noted above) mentioned along with the city of *Xutluğ (Qutluğ). Our sources clearly distinguish the two, except for Ibn Rusta (1892, 139) who says that within Sariğšin is "another city; it is called هب نلع [hb n'l] or خنبلغ [*Xnblg]". The ms. variants (see Golden 1980 I: 232) are undotted and hence open to a variety of interpretations (including Xutluğ, noted above). I think we have two variants of the toponym *Qan Balığ جنبلغ [*Xanbalıg*] and Han Balığ هنبالغ [Han Balıg < Qan Balıq]. The latter, at least in part, reflects the shift q- > x- > h-. Mixed data, such as the preceding, which may reflect two dialects, distinct Khazar languages or materials recorded by different sources at different times, make the Khazar language question so complex. The name/title of the Khazar general noted by Lewond, ca. 758-64, Xat'irlit'ber (խաթրիլիտ'բեր, see Lewond 1982: 125, ms. variants in Golden 1980 I: 197-8): Xatır il-teber < Qadır él-teber, may point to a q- > x-change already by the eighth century (although Turkic *q*, it should be noted, is often rendered by *x* in Arabic, Armenian, Georgian and other sources). Khazar *il-teber*/*él-teber* is Common Turkic rather than Oğuro-Bulğaric (cf. Volga Bulğaric *Yıl-tawâr* noted by Ibn Faḍlân, Togan, 1939: Arab. 1, Ligeti 1986: 457-8).

The Khazar title reported by Ibn Faḍlân as just below that of the Qağan Beg in the Khazar tetrarchy is *Kündü Qağan* ([kndw] كندو, the mss. of Yâqût's *Mu'jam al-Buldân* which has preserved this section of Ibn Faḍlân's *Risāla*, universally have [kndr] كندر, see Golden 1980 I: 200). The evidence, largely circumstantial, points to [kndw] كندو as the original form. Hungarian scholarly tradition has long connected this title with Hung. **Kündä* (Modern Hung. *Kende*), the title of the sacral ruler of the ninth century Magyar-led tribal union, while the *Gyula* (< Turk. *ḡulalyula*) ran the

actual affairs of state—a situation analogous to that of the Khazar *Qağan* and *Qağan Beg* (Ligeti 1986: 253-4, 368, 482, 484-5; Németh 1991: 83, 226, 236-7; Róna-Tas 1999: 342-4). Variants such as *kündü* ~ *kündä* are not unknown (cf. *ordu* ~ *orda*, Tenišev 1997: 495). It also appears in the name of the Khazar *gulâm*, Ishâq b. Kundâj [كنداج] and Kundâjîq [كنداجيق] (see forms in Golden 1980 I: 202-203) probably representing **Kündäčik*. The etymology of *kündü* is problematic in Turkic where it is found only as a Mongol loanword in Siberian Turkic (Radloff II/2: 1444-5) denoting ‘die Ehrfurcht, Höflichkeit, Ehrfurchtbezeugung, das Gastmahl’, and ‘das zweite wichtigste Amt nach dem Jaisang’. Ligeti (1986: 49) noted Sino-Korean *kuntai* ‘minister of war’, but this seems unlikely. The root may be (Starostin et al. 2003 I: 820): Altaic **k’june* ‘heavy load’: Tung. **(x)ünī-*, Mong. **kündü*, PTung. **(x)üni-* ‘to carry on the back’ PMong *kundu*, *kunule* ‘to respect’ W. Mong. *kündü* etc. > Manchu *kundu* ‘respect, honor’ (see also Cincius 1975 I: 432 *kundulê-* ‘ugoščat’ etc. Manchu *kundu* ‘čest’, *dostoinstvo*, *počet*, *uvaženie*’, etc.). Mongolic appears to be the source of this word in both Turkic and Manchu-Tungusic. Since we have no indications that Khazar was Mongolic, one can only presume that Khazar *kündü* is either an ancient loanword in Khazar (or its ancestor tongue) from Mongolic (perhaps from the ancestor tongue of Khazar to Mongolic?) or part of a much debated Altaic legacy (now under assault again, see Beckwith 2004: 184-194; Vovin 2005: 71-132). In any event, it was undoubtedly one of the Khazar terms that made their language seem “strange”. The title جارشيفر [jāwšîgr], the deputy of the *kündü*, remains unexplained as well. Perhaps, it is a garbling of جارشيفر **Jawašgir* < **javaš*, Common Turk. *yavaš* ‘gentle, mild’ + *-gir/gur* ‘the one who makes peace’ (cf. Uyğ. Buddh. *yavaš qil* ‘to make peace’ (Clouston 1972: 880).

Another obscure Khazar term is noted by Theophanes (ca. 812) who recounts the attempt by the Byzantine emperor Justinian II (reg. 685-95, 705-11) in 710/711 to punish his recalcitrant subjects in Xerson. In the course of the ensuing political maneuvers, the Khazar *tudun/tudun* (a Türk title given to administrative and fiscal officials not of royal blood, see Golden 1980 I: 215-216) died. According to Theophanes (1980: 377-379, Theophanes 1997: 527-528) εἰς δοχὴν αὐτοῦ the Khazars sacrificed the Byzantine Turmarxos and 300 soldiers. Mango translates this phrase as ‘in his honor’ (Theophanes/Mango 1997: 528), understanding *δοχή* in the Classical Greek sense of ‘reception, entertainment’. While possible, this seems unlikely. Dieter Ludwig (Ludwig 1982: 356-357) first properly identified this with *δογή* (mss. variants of Theophanes have: *δογήν*, *δογήν*, *δουγήν*). This term, in the form *δόγῃα*, is noted in Menander’s account (Menander/Blockley 1985: 178) of the funeral rites carried out in 576 for Σιλζιβούλος (**Sir Jabğu* < *Śri Yabğu*, i.e. *İštemi Qağan*, see Dobrovits 2004: 112-113), the Western Türk Qağan. The term, *doğ*, possibly *dox* in Khazar, should be compared with Old and Middle Turkic *yoğ* ‘funeral feast, wake’ (Clouston 1972: 895, cf. Qazaq *žoqta-*, Qara Qalpaq *žoqla-* ‘oplakivat’ *umeršego*’ < *yoğla-*, Sevortjan 1989: 207). Its relationship with Chuvash *šăva* ‘kladbišče’ is less certain (Fedotov 1996a, II: 89). As for the initial *δ-*, Menander (Blockley 1985: 125) also notes the Δαῖχ River (= Common Turk. *Yayıq* ‘Ural River’

< *yay-* ‘to spread’) in his account of the sixth century Turko-Byzantine embassies, which should be compared with the $\Delta\acute{\alpha}\iota\tilde{\nu}$ of Ptolemy (2nd century AD, see Pritsak 1955: 43; Moravcsik 1958 II: 116). Early Turkic-speaking populations pushed into the Kazakhstanian steppes by various movements of the Xiongnu might, possibly, be the source for this hydronym (despite the objections of Clauson 2002: 124 that this is too early for Turkic-speakers to be here). Clauson (2002: 125) further contends that there is no Turkic verb **day-*. *Yay-* < *yād-* ‘to spread out’ (Clauson 1972: 883-884, 980) *yay-* ‘to shake’ could not, in his view, have been the source for this word. He concludes that **Yadūq* = *Yayıq* “is a local pre-Turkish name”. Nonetheless, he fully accepts that Menander’s **doğ* is an earlier form of *yoğ* (Clauson 2002: 124). By this time, in Medieval Greek, the pronunciation of δ as δ [dh] was long established (by the second century AD, Browning 1983: 26-27), although it was also used to render *d* (in Modern Greek $\nu\tau$). Ibn Faḍlān 1939: Arab. 18) notes this river as $\dot{\text{ج}}\text{ا}\text{ي}\text{خ}$ [jyx] = *Ĵayix*. Ligeti (1986: 460) considers *Ĵayix* either the Bulḡar or Khazar form. Hence, it is unclear if δ of our earlier sources represents *d*, δ or even Bulḡaric *ĵ* as Németh (1991: 110-111 n.b) initially viewed it. The usual sequence is to posit an original Turkic *y-* coming from an earlier Altaic *d*, *ĵ* and *y*. It remained *y-* in Common Turkic (later changing to *ĵ*-, *ž*- in some languages) but early on in Oḡuro-Bulḡaric became *ĵ*- and subsequently *č* > *š*- and ultimately *ś* in Chuvash. Another explanation is: *y*- > *ĵ*- > *ž*- > *ś* (Tenišev 1984: 277-278). Perhaps, we should posit a δ [ð] in Proto-Turkic which became *y* > *ĵ* in Oḡuro-Bulḡaric? The question of initial *d-* is complicated by several Oḡuro-Bulḡaric loanwords in Hungarian which have *d-*, although these may be explained as a secondary development in Hungarian as Ligeti indicates (cf. *disznó* ‘swine’ < *gyisznó*, Ligeti 1986: 21, 24-25, 45, 194, 284). On the other hand, Danubo-Balkan Bulḡar may provide evidence of an initial *d-* which became *ĵ* in some Oḡuro-Bulḡar dialects, but appears to have remained *d-* in others, cf. $\text{д}\text{иломь}$ [dilom] = **dilām*, in Pritsak 1955: 73] ‘snake’ (Chuv. *šilen* < **ĵilan*, Common Turk. *ylan*). The д has also been viewed as a Slavicism (see discussion in Ligeti 1986: 474, cf. also the Bulḡar clan name Доло = *Ĵula*?). The oft-cited дохъторь [doxūtor], *doxtār* in Pritsak 1955: 73 with *-xd-* > *-xt-*] ‘pillow’ < **doxtār* < *doğtar* < **doğdar* (Chuv. *šutar/šātār* < **johtar*, according to Pritsak 1955: 43-4, Tekin 1987: 14, 67 with slightly different reconstruction) is more problematic. Pritsak (1955, 43), Räsänen (1969: 127, 204-205) and Tekin (1987: 14, 67) compared it with Mong. *joğdur* ‘long hair on the throat of a camel’ (Lessing 1982: 1067) and Turkic *yoğdu* ‘the long hair under a camel’s chin’ (Clauson 1972: 899), presumably the material from which a ‘pillow’ might be made. Fedotov (1996a, II: 158), however connects it with Common Turk. *yatur-* ‘to cause to lie down’. These all seem something of a stretch. Perhaps, this term derives from Turk. *toqu-* ‘to weave’ (cf. Middle Oḡuz *doqi-* see Tenišev 1997: 395-396): Oḡuro-Bulḡaric **doqutur* > *doxūtor*? In Chuvash, this verb has been replaced by *těrt-* < Tat. *tört-*. Thus, the initial δ of these forms (and hence our Khazar term) has yet to be explained fully. At best one may suggest an Altaic > Proto-Turkic *d-* > δ - > *y-* / *ĵ-* with some more archaic Turkic tongues perhaps

retaining *d*-. If Khazar *δουγη* is not simply a rendering of *joğ*, it could point to some very archaic features.

The Khazar ruler Joseph's response (ca. 960) to Ḥasday b. Šaprūt, the Jewish courtier of the Iberian Umayyads, contains a series of ethnonyms, hydronyms and toponyms which, notwithstanding Kokovcov's (1932) thorough analysis more than seventy years ago, merit a new and separate study. I will note only one ethnonym here. Joseph mentions the *וּנְנִיר* [wnntr], the major opponents of the Khazars in the struggle for dominion in the Ponto-Caspian steppes (Kokovcov 1932: 28). *Wnnt* denoted the Onoğundur-Bulğar or Onoğundur Empire (see Zimonyi 1990: 40-42), founded ca. 635 by the Onoğundur leader Qubrat/Kuvrat. The Onoğundur-Bulğars were defeated by the Khazars which led to their partial subjugation in the 670s. Some remained in the Pontic steppes, others eventually (early-mid eighth century) went up the Volga and founded Volga Bulğaria (Zimonyi 1990), a vassal state of the Khazars. Another grouping entered Byzantine Moesia, in 679, imposed itself on the local Slavic population giving rise to the Balkan Bulğar state with kinsmen in Pannonia. It is likely that the form of this ethnonym in Joseph's letter is Khazar of the mid-tenth century, reflecting, ultimately, the Onoğundur self-designation (see also discussions in Róna-Tas 1996: 101, 179, 219, 259). *Onoğundur* is a variant of *Onoğur* ('Ten Oğur [tribes]'). The latter form is first attested by Priskos (d. after 472) who noted, ca. 463, the chain of migrations in the steppe initiated by the Avars who pushed the Sabirs westward. The latter, then, precipitated the movement of the Šara Oğurs (Σαράγουροι), Oğurs (text: Οὔρωγοι for *'Ογούροι) and Onoğurs (Ὀνόγουροι) into the Pontic steppes and contact with Constantinople (Priskos/Blockley 1983 II: 344). The Syriac epitome of the *Ecclesiastical History* of Zacharias of Mytilene (d. after 536), "Pseudo-Zacharias Rhetor" (probably compiled ca. 569, see Pigulëvskaja 2000: 185-190), in its excursus on the northern peoples dated to 555 (Kmoskó 2004, 47-48), notes *'wngwr* (Marquart 1961: 356; Pigulëvskaja 2000: 568, Kmoskó 2004: 99) among the nomadic peoples beyond the "Caspian Gates in the land of the Huns". The *Suishu* (composed 627-636) in its listing of Tiele tribes (a tribal union arcing across Eurasia), "east of Fulin" (Rome, i.e. Constantinople) mentions the Enqu (Liu 1958 I: 127-128, II: 569 n. 663 = **On[o]gur*, see above). The nearly contemporary *Armenian Geography* (*Ašxarhač'oyc'*) compiled by Ananias Širakec'i (ca. 610-685, written prior to 636 with later additions and interpolations, Hewsens 1992: 15, 33-34) notes among the various Bulğaric peoples in the Ponto-Caspian steppes, the *Οἰκοντορ Βουλγάρων* *Okontor Bulgar* (Hewsens 1992: 55). Agathon (early eighth century) records: ἔθνος τῶν Οὐννογούρων Βουλγάρων (Moravcsik 1930: 67). Movsēs Xorenac'i (ostensibly fifth century, probably 770s with perhaps later interpolations, see Khorenatsi/Thomson 1980 I: 51, 60) notes a Bulgar *Վընդուր/ Վըրըրըր: Vtndur /Vtəndur* (Xorenac'i/Malxasyan 1961: 153, Xorenac'i/Ulubabyan 2003: 132, Khorenatsi/ Thomson 1980: 135). Theophanes calls *Qubrat/Kuvrat*, the 'lord of the Οὐνογουνδοῦροι' (Theophanes 1883 I: 356) and his contemporary, Nicephorus (who completed his *Short History* ca. 828) calls him the 'lord' of the Οὐννογουνδούρων Βουλγάρων (Nikephorus/Mango 1990: 70). More than a century later, Constantine

Porphrogenitus (1952: 85) says that the Bulğars previously called themselves 'Ονογουν-δούποι. Ibn Kalbî (ca. 820, Marquart 1924: 275) notes the الغندر [l'gndr = *Ul^hgundur?]. Al-Mas'ûdî (*Tanbih* 1894: 180) mentions, ca. 320/932, “tribes of nomadic Turks who are called الولندرية [al-wlndryh: *wulunduriyya] associated with a city called [wlndr: *Wulundur, see *Ḥudūd*/Minorsky 1970: 469-70, perhaps modern Burgas, see Knjaz'kij 2003: 15] at the eastern extremities of Rûm. He has a similar notice in his *Murûj ad-Ḍahab* (1966-1979 I: 236, see above). The *Ḥudūd al-Ālam* (*Ḥudūd*/Minorsky 1970: 160, 161) mentions the “V.n.d.r mountains” in the country of Mirvât (Great Moravia, see Göckenjan & Zimonyi 2001: 214 n. 188) and the N.nd.r (V.n.nd.r) a people north of the Khazar land on the Volga, alongside the *Burdâs (Burtâs, probably confused here with the Volga Bulğars, see Göckenjan & Zimonyi 2001: 219). The V.n.nd.r (*Vunundur*), it is generally agreed, are the Danubian Bulğars (see *Ḥudūd*/Minorsky 1970: 440-1, 465-8; Göckenjan & Zimonyi 2001: 219). The Khazaro-Hebrew form noted above, thus, should also be read as *Wonuntur/Wununtur* (*Wanuntur?* or *Vonuntur* etc.). Gardîzî (1984: 587, writing in the mid-eleventh century, has: نندر [nndr = Nandur] and نندريان [nndryân = Nanduriyân]—perhaps a corruption of وندر [wnndr] or the reflection of a later form. *Nándor* was the Hungarian name of the Danubian-Balkan Bulğars, surviving in Hungarian place-names and as *Nándorfehérvár* ‘Nándor White City’ = Belgrad (Kiss 1978: 397, 455, Ligeti 1986: 268-9; Róna-Tas 1996: 219, 259). Thus, we have several variants of this ethnonym. By the early seventh century, *Onoğundur* had, apparently, for reasons still unexplained, developed another form *Ol[u]x/gundur recorded in Armenian and Muslim sources. Although in Arabic mss., the confusion of medial *nûn* and *lam* is not unknown, the Armenian form clearly indicates the presence of an *-l-*. This form, perhaps from *Uluğ Onoğundur (‘the Great Onoğunders’) > *Ulux Onowundur (with medial *ğ* > *w*) > *Uluh Onowundur and was then conflated into *Wulunundur* > *Wulundur*. The Khazar form **Wonuntur/Wununtur/Wanuntur* (< *Onoğundur*) shows the prothetic *w-*, known to Oğuro-Bulğaric (whether the same shift, *o-* > *wo-* (or *u-* **wu-*, later in Chuv. *wă/vă-*, cf. Common Turk. *uzun* ‘long’ ~ Chuv. *vărăm*), occurred in Khazar cannot be determined as the Khazar form of this name may simply reflect the *Wonundur/Onoğundur* self-designation.

The shift *-q/ -ğ* > *-x* > *-h* > 0, typical of Oğuro-Bulğaric (Tenišev 2002: 693-698), seems to be a feature of Khazar as well. Clearly, an Oğuro-Bulğaric tongue, or something close to it, was one of the languages identified with the Khazars and elements of it were refracted in the royal Khazar Hebrew correspondence.

Bibliography

- Abu'l-Gâzi Bahâdur Khan 1958. *Šajara-yi Tarâkima. Rodoslovnaja Turkmen*. Moskva-Leningrad: Izdatel'stvo Akademii Nauk SSSR.
- The Armenian geography. Transl. by R. H. Hewsen 1992. *The geography of Ananias of Širak (Ašxarhaç' oyc')*. *Beihefte zum Tübinger Atlas des Vorderen Orients*. (Reihe B, Geisteswissenschaften Nr. 77.) Wiesbaden: Reichert.

- Al-Bakrî, 1992. *Kitâb al-Masâlik wa'l-Mamâlik 1-2*. Edited by A. P. van Leewen & A. Ferre. Tunis: Dâr al-Ġarb al-Islâmî.
- Baskakov, N. A. 1969. *Vvedenie v izučenie tjurkskix jazykov*. Moskva: Vysšaja Škola.
- Beckwith, C. I. 2004. *Koguryo. The language of Japan's continental relatives*. Leiden: Brill.
- Al-Birûnî 1923. *Kitâb al-Âtâr al-Bâqiyya 'an al-Qurûn al-Xâliyya*. Edited by C. E. Sachau. 1878. Reprint: Leipzig: Harrassowitz.
- Bombaci, A. 1974. On the ancient Turkish title *Şad*. In: *Gururâjamañjarikâ. Studi in onore di G. Tucci*. Napoli. 176-193.
- Browning, R. 1983. *Medieval and Modern Greek*. Cambridge: Cambridge University Press.
- Chavannes, E. 1941. *Documents sur les Tou-kiue (turcs) occidentaux*. Paris: Adrien Maisonneuve.
- Cincius, V. I. et al. 1975. *Sravnitel'nyj slovar' tunguso-man'čžurskix jazykov*. Leningrad: Nauka. 2 vols.
- Clauson, Sir Gerard 1972. *An etymological dictionary of pre-thirteenth century Turkish*. Oxford: Oxford University Press.
- Clauson, Sir Gerard 1975. The foreign elements in early Turkish. In: L. Ligeti (ed.) *Researches in Altaic languages*. Fourteenth PIAC Meeting, Szeged, August 22-28, 1971. (Bibliotheca Orientalis Hungarica 20.) Budapest: Akadémiai Kiadó. 43-49.
- Clauson, Sir Gerard 2002. *Studies in Turkic and Mongolic linguistics*. Royal Asiatic Society 1962. Reprint: London: Routledge Curzon.
- Colarusso, J. 1992-1993. Two Circassian tales of Huns and Khazars. *The Annual of the Society for the Study of Caucasia* 4-5: 63-75.
- Constantine Porphyrogenitus 1952. *De thematibus*. Edited by A. Pertusi. Città del Vaticano: Biblioteca apostolica vaticana.
- Constantine Porphyrogenitus 1967. *De administrando imperio*. Edited by J. Moravcsik, transl. by R. Jenkins. Washington, D. C.: Dumbarton Oaks.
- D'jačenko, G. 1993. *Polnyj cerkovno-slavjanskij slovar'*. 1900. Reprint: Moskva: Izdatel'skij otdel moskovskogo patriarxata.
- Dobrovits, M. 2004. A nyugati türkök első uralkodójáról. *Antik Tanulmányok* 48: 111-114.
- DTS: Nadeljaev, V. M. et al. 1969. *Drevnetjurkskij slovar'*. Leningrad: Nauka.
- Dunlop, D. M. 1954. *The history of the Jewish Khazars*. Princeton: Princeton University Press.
- Erdal, M. 1991. Ein unbemerkter chasarischer Eigenname. *Türk Dilleri Araştırmaları* 1991: 31-36.
- Erdal, M. 1993. *Die Sprache der wolgabolgarischen Inschriften*. (Turcologica 13.) Wiesbaden: Harrassowitz.
- Erdal, M. 2004. *A grammar of Old Turkic*. Leiden: Brill.
- Erdal, M. in press. The Khazar language. In: Ben-Shammai, H. & Golden, P. B. & Róna-Tas, A. (eds.) *The Khazars*. Leiden: Brill.
- Fedotov, M. R. 1996a. *Ėtimologičeskij slovar' čuvaškogo jazyka 1-2*. Čeboksary: Čuvaškij Gosudarstvennyj Institut Gumanitarnyx Nauk.
- Fedotov, M. R. 1996b. *Čuvaškij jazyk*. Čeboksary: Izdatel'stvo Čuvaškogo Universiteta.
- Gadžieva, N. Z. & Serebrennikov, B. A. 1977. Areal'naja lingvistika i problema vosstanovlenija nekotoryx čert isčeznuvšix jazykov. *Sovetskaja Tjurkologija* 3: 3-12.
- Gardîzî 1363/1984. *Ta'rix-i Gardîzî*. Edited by 'A. Ĥabîbî. Tehran: Dunyâ-yi Kitâb.
- Göckenjan, H. & Zimonyi, I. 2001. *Orientalische Berichte über die Völker Osteuropas und Zentralasiens im Mittelalter. Die Ġaihânî-Tradition*. Wiesbaden: Harrassowitz.

- Golden, P. B. 1975. The Q'azaro-Hungarian title/personal name $\text{𐌆𐌿𐌹𐌸𐌰} - \text{𐌰𐌺𐌹𐌸}$. *Archivum Eurasiae Medii Aevi* 1: 37-43.
- Golden, P. B. 1980. *Khazar studies 1-2*. (Bibliotheca Orientalis Hungarica 25/1-2.) Budapest: Akadémiai Kiadó.
- Golden, P. B. 1992. *An introduction to the history of the Turkic peoples*. (Turcologica 9.) Wiesbaden: Harrassowitz.
- Golden P. B. 2002-2003. Khazar Ghulâms in caliphal service: Onomastic notes. *Archivum Eurasiae Medii Aevi* 12: 15-27
- Golden P. B. 2004. Khazar Ghulâms in Caliphal service. *Journal Asiatique* 292/1-2: 279-309.
- Györfly, Gy. 1990. *A magyarság keleti elemei*. Budapest: Gondolat.
- Harmatta, J. 1997. Iráni nyelvek hatása az ősmagyar nyelvre. In: Kovács, L. & Veszprémy, L. (eds.) *Honfoglalás és nyelvészet*. Budapest: Balassi Kiadó.
- Hewsen 1992. See *The Armenian geography*.
- Ḥudūd al-'Ālam. 1970. Transl. by V. Minorsky. (Gibb Memorial Series, n.s. 11.) London. 1937, Rev. ed. London: Luzac.
- Ibn al-Aḡrī 1965-1967. *Al-Kāmil fī'l-Ta'rīx. Chronicon quod perfectissimum inscribitur*. Edited by C. J. Tornberg. Leiden: Brill 1851-1876. Reprint: Beirut: Dār Ṣādir
- Ibn Faḍlān 1939. *Ibn Faḍlān's Reisebericht*. Edited and German transl. by A. Z. V. Togan. *Abhandlungen für die Kunde des Morgenlandes* 24/3. Leipzig: Deutsche Morgenländische Gesellschaft.
- Ibn Faḍlān 1956. *Kniga Axmeda ibn Fadlana o ego putešestvii na Volgu v 921-922 gg.* Edited and translated by A. P. Kovalevskij. Xar'kov: Gosudarstvennyy Universitet imeni Gor'kogo. [This contains a facsimile of the Mašhad ms.]
- Ibn Ḥawqal 1992. *Kitāb Ṣūrat al-Arḍ*. Beirut: Ḥayāt.
- Ibn Xurdādbih 1889. *Kitāb Masālik wa'l-Mamālik*. Edited by M. J. de Goeje. Leiden: Brill.
- al-Iṣṭaxrī, 1927². *Kitāb Masālik al-Mamālik*. Edited by M. J. de Goeje, Leiden: Brill
- Karlgren, B. 1996. *Grammata Serica recensa*. Stockholm, 1957. Reprint: Taipei: SMC Publishing.
- K'art'lis C'xovreba* 1955-1973. *K'art'lis C'xovreba*. Edited by S. Qauxč'išvili. T'bilisi: Saxelgami-Sabčot'a Sak'art'velo. 4 vols.
- Kāšġarī 1982-1985. Maḥmūd al-Kāšġarī *Compendium of the Turkic dialects (Dīwān Luyāt al-Turk)*. Edited and transl. by R. Dankoff in collaboration with J. Kelly. (Sources of Oriental Languages and Literatures 7.) Cambridge, Mass.: Harvard University Press. 3 vols. Khorenatsi/Thomson: Moses Khorenats' i 1978. See under Xorenac'i, Movsēs.
- Kiss, L. 1978. *Földrajzi nevek etimológiai szótára*. Budapest: Akadémiai Kiadó.
- Kljaštornyj, S. G. 1979. Xazarskaja nadpis' na amfore s gorodišča Majaki. *Sovetskaja Arxeologija* XIX: 270-275.
- Klyashtorny, S. G. 1997. About one Khazar title in Ibn Faḍlān. *Manuscripta Orientalia* 3/3: 22-23.
- Kljaštornyj, S. G. & Vásáry, I. 1987. A Runic inscription on a bull-skull from the Volga region. In Kara, Gy. (ed.) *Between the Danube and the Caucasus*. Budapest: Akadémiai Kiadó: 171-179.
- Kmoskó, M. 2004. *Szír írók a steppe népeiről*. Edited by Sz. Felföldi. Budapest: Balassi Kiadó.
- Knjaz'kij, I. O. 2003. *Vizantija i kočevniki južnorusskix stepej*. Moskva: Aletejja.
- Kokovcov, P. K. 1932. *Evrejsko-Xazarskaja perepiska v X veke*. Leningrad: Izdatel'stvo akademii nauk SSSR.

- Konovalova, I. G. 2003. Padenie Xazarii v istoričeskoj pamjati raznyx narodov. In: *Drevnejšie gosudarstva Vostočnoj Evropy 2001*. Moskva: 171-190.
- Kračkovskij, I. Ju. 1957. *Arabskaja geografičeskaja literature*. In: Kračkovskij, I. Ju. 1957. *Izbrannye sočinenija*. Moskva-Leningrad: Izdatel'stvo akademii nauk SSSR. Vol. 4.
- Kyzlasov, I. L. 1994. *Runičeskie pis'mennosti evrazijskix stepej*. Moskva: Vostočnaja Literatura.
- Lessing, F. et al. 1982. *Mongolian-English dictionary*. Bloomington: The Mongolia Society. Corrected reprint of Berkeley-Los Angeles, 1960.
- Lewond 1982. *History of Lewond. The eminent vardapet of the Armenians*. Transl. by Z. Arzoumanian. Philadelphia: St. Sahag & St. Mesrob Armenian Church.
- Ligeti, L. 1986. *A magyar nyelv török kapcsolatai a honfoglalás előtt és az Árpád-korban*. Budapest: Akadémiai Kiadó.
- Liu Mau-tsai 1958. *Die chinesischen Nachrichten zur Geschichte der Ost-Türken (T'u-Küe)*. Wiesbaden: Harrassowitz. 2 vols.
- Ludwig, D. 1982. *Struktur und Gesellschaft des Chazaren-Reiches im Licht der schriftlichen Quellen*. (PhD dissertation.) Münster: Philosophische Fakultät der Westfälischen Wilhelms-Universität.
- Marquart, J. 1924. Ein arabischer Bericht über die arktischen (uralischen) Länder aus dem 10. Jahrhundert. *Ungarische Jahrbücher* 4: 261-238.
- Marquart, J. 1961. *Osteuropäische und ostasiatische Streifzüge*. Leipzig. 1903. Reprint Hildesheim: Georg Olms.
- Al-Mas'ūdī 1894. *Kitāb at-Tanbīh wa 'l-Isrāf*. Edited by M.J. de Goeje. Leiden: Brill.
- Al-Mas'ūdī 1966-79. *Murūj al-Dahab wa Ma'din al-Jawhar*. Edited by C. Pellat. Beirut: Manshūrāt al-Jāmi'at al-Lubnāniyya. 7 vols.
- Menander 1985. *The History of Menander the Guardsman*. Edited and transl. by R. C. Blockley. Liverpool: F. Cairns.
- Al-Muqaddasī 1408/1987. *Aḥsan at-Taqāsīm fī Ma'rīfat al-Aqālīm*. Edited by M. Maxzūm. Beirut: Dār Iḥyā at-Turaṯ al-'Arabī.
- MNyTESz: *A magyar nyelv történeti-etimológiai szótára*. Benkő, L. et al. (eds.). 1967-1976. Budapest: Akadémiai Kiadó. 3 vols.
- Moravcsik, J. 1930. Zur Geschichte der Onoguren. *Ungarische Jahrbücher* 10: 53-90.
- Moravcsik, J. (ed.) 1958². *Byzantinoturcica*. Berlin: Akademie Verlag. 2 vols.
- Németh, Gy. 1991. *A honfoglaló magyarság kialakulása*. Budapest 1930. Rev. ed. Budapest: Akadémiai Kiadó.
- Nikephoros Patriarch of Constantinople 1990. *Short history*. Edited and transl. by C. Mango. Washington, D.C.: Dumbarton Oaks.
- Parzymies, A. 1994. *Język protobułgarski*. Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.
- Pigulëvskaja, N. V. 2000. *Sirijskaja srednevekovaja istorija*. St. Petersburg: Dmitrij Bulanin.
- Pletněva, S. A. 1976. *Xazary*. Moskva: Nauka.
- Priskos 1981. 1983. *The fragmentary classicizing historians of the later Roman Empire. Eunapius, Olympiodorus, Priscus and Malchus*. Edited and transl. by R. C. Blockley. Liverpool: F. Cairns. 2 vols.
- Pritsak, O. 1955. *Die bulgarische Fürstenliste und die Sprache der Protobulgaren*. Wiesbaden: Harrassowitz.

- Pritsak, O. 1975. The Pečenegs: A case of social and economic transformation. *Archivum Eurasiae Medii Aevi* 1: 211-235.
- Prodolžatel' Feofana* 1992. Edited by Ja. I. Ljubarskij. St. Petersburg: Nauka. [See also *Theophanus Continuatus*.]
- Pulleyblank, E. 1991. *Lexicon of reconstructed pronunciation in early Middle Chinese, late Middle Chinese, and early Mandarin*. Vancouver: University of British Columbia Press.
- al-Qazwīnī 1389/1969. *'Ajā'ib al-Buldān wa Axbār al-'Ibād*. Beirut: Dār Šādir.
- Radloff (Radlov), W. (V.) 1963. *Opyt slovarja tjurkskix narečij*. St. Petersburg 1893-1911. Reprint: Moskva: Izdatel'stvo Vostočnoj Literatury. 4 vols. in 8 parts.
- Räsänen, M. 1969. *Versuch eines etymologischen Wörterbuchs der Türk Sprachen*. Helsinki: Suomalais-Ugrilainen Seura.
- Rastorgueva, V. S & Edel'man, D. I. 2000, 2003. *Ėtimologičeskij slovar' iranskix jazykov*, Moscow: Vostočnaja Literatura. [2 vols. published thus far.]
- Rašīd al-Dīn 1969. *Die Geschichte der Öguzen des Rašīd al-Dīn*. Edited and transl. K. Jahn. (Österreichische Akademie der Wissenschaften, Philosophisch-historische Klasse, Denkschriften 100.) Wien: Hermann Böhlau Nachf.
- Rašīd al-Dīn. 1373/1994. *Jāmi' al-Tawārix*, Edited by M. Rawšan & M. Mūsawī. Tehran: Alburz.
- Romašov, K. 2000-2003. Istoričeskaja geografija Xazarskogo Kaganata (V-XIII vv.) [Parts I and II]. *Archivum Eurasiae Medii Aevi* 11 (2000-2001), 219-338; [Part III], 12 (2002-2003), 81-221, [Parts IV and V] 13 (2004), 185-264.
- Róna-Tas, A. 1982a. The periodization and sources of Chuvash linguistic history. In: Róna-Tas, A. (ed.) *Chuvash studies*. Wiesbaden: Harrassowitz. 113-169.
- Róna-Tas, A. 1982b. A kazár népnévről. *Nyelvtudományi Közlemények* 84: 349-379.
- Róna-Tas, A. 1982c. Ujabb adatok a kazár népnév történetéhez. *Nyelvtudományi Közlemények* 85: 126-133.
- Róna-Tas, A. 1996. *A honfoglaló magyar nép*. Budapest: Balassi Kiadó. English transl.: Róna-Tas, A. 1999. *Hungarians and Europe in the Early Middle Ages*. Budapest, 1999. Central European University Press. Revised English translation of Róna-Tas, A. 1996.
- Sevortjan, É. V. 1978. *Ėtimologičeskij slovar' tjurkskix jazykov. Obščetjurkskie i mežtjurkskie osnovy na bukvu "B"*. Moskva: Nauka.
- Sevortjan, É. V. et al. 1974 (ongoing). *Ėtimologičeskij slovar' tjurkskix jazykov* Moskva: Nauka, 7 vols. [Volumes are noted by year of publication.]
- Starostin, S. & Dybo, A. & Mudrak, O. 2003. *Etymological dictionary of the Altaic languages*, Leiden: Brill. 3 vols.
- aṭ-Ṭabarī 1967-1969. *Ta'rix aṭ-Ṭabarī. Ta'rix ar-rasūl wa 'l-mulūk*. Edited by Abu'l-Faḍl Ibrāhīm. Cairo: Dār al-Ma'ārif. 10 vols.
- Tekin, T. 1987. *Tuna Bulgarları ve dilleri*. Ankara: Türk Dil Kurumu.
- Tekin, T. 1988. *Volga Bulgar kitabeleri ve Volga Bulgarcası*. Ankara: Türk Dil Kurumu.
- Tenišev, É. R. et al. 1984. *Sravnitel'no-istoričeskaja grammatika tjurkskix jazykov. Fonetika*. Moskva: Nauka.
- Tenišev, É. R. et al. 1997. *Sravnitel'no-istoričeskaja grammatika tjurkskix jazykov. Leksika*. Moskva: Nauka.
- Tenišev, É. R. et al. 2002. *Sravnitel'no-istoričeskaja grammatika tjurkskix jazykov. Regional'nye rekonstrukcii*. Moskva: Nauka.
- Theophanes 1980. *Theophanis chronographia*. Edited by C. de Boor, Leipzig. 1883. Reprint: Hildesheim: Georg Olms. 2 vols.

- Theophanes 1997. *The chronicle of Theophanes Confessor*. Edited and transl. by C. Mango & R. Scott with the assistance of G. Greatrex. Oxford: Oxford University Press.
- Theophanes Continuatus 1838. *Historiae*. Edited by I. Bekker. Bonn: Weber. [See also the Russian transl. *Prodolžatel' Feofana*. 1992.]
- Vixnovič, V. L. 1997. *Karaim Avraam Firkovič evrejskie rukopisi, istorija, puteštvoja*. St. Peterburg: Peterburgskoe vostokovedenie.
- Vovin, A. 2005. The end of the Altaic controversy. *Central Asiatic Journal* 49: 71-132.
- Xorenac'i, Movsês 1961. *Patmu'iwn Hayoc'*. Edited by St. Malxasyan. Yerevan: Haypethrat.
- Xorenac'i, Movsês 2003. *Patmu'iwn Hayoc'*. Edited by B. Ulubabyan. Yerevan: Gasprint.
- Xorenac'i, Movsês: Moses Khorenats'i 1978. *History of the Armenians*. Transl. by R. W. Thomson. Cambridge, Mass.: Harvard University Press.
- Al-Ya'qûbî 1892. *Kitâb al-Buldân*. Edited by M. J. de Goeje. Leiden: Brill.
- Zaxoder, B. N. 1962, 1967. *Kaspijskij svod svedenij o Vostočnoj Evrope*. Moskva: Izdatel'stvo Vostočnoj Literatury.
- Zimonyi, I. 1990. *The origins of the Volga Bulghars*. (Studia Uralo-Altaica 32.) Szeged: Universitas Szegediensis de Attila József Nominata.
- Zuckerman, C. [K.] 1997. Two notes on the early history of the *thema* of Cherson. *Byzantine and Modern Greek Studies* 21: 210-222.
- Zuckerman, K. 2001. Xazary i Vizantija: pervye kontakty. *Materialy po arxeologii, istorii i étnografii Tavrii* 8: 312-333.

Turkic–Persian bilateral code copying

Lars Johanson

Johanson, Lars 2005. Turkic-Persian bilateral code copying. *Turkic Languages* 9, 223-228.

The paper deals with the convergent development of southeastern Turkic on eastern Persian in Central Asia due to intensive bilateral copying. One aim is to discuss a relative chronology of the copying processes involved, some of which may be of high age, while others are more recent. Certain shared features are due to Turkic influence, e.g. the use of forms of specific auxiliary verbs as markers of focal intraterminality. Iranian influence may be assumed in the use of preterit-presents in Turkic varieties. Further Turkic influence lies behind the formation of various actional periphrases in eastern (Tajik) Persian.

Lars Johanson, Department of Oriental Studies, Johannes Gutenberg-University Mainz, DE-55099 Mainz, Germany. E-mail: johanson@uni-mainz.de

Turkic–Persian contacts in Central Asia

Turkic and Persian varieties have a long history of symbiosis in Central Asia, resulting in considerable influence in both directions. The convergent development began many centuries before Uzbek and Tajik were established as modern standard languages, intensive bilateral copying already taking place from the 11th century. Certain shared features are due to south-eastern Turkic influence on pre-Tajik eastern Persian. Others are due to eastern Persian influence on pre-Uzbek southeastern Turkic. Since the developments are highly complex, it is difficult to determine the direction of copying and to pinpoint the developmental stages of the languages involved at the time of copying. The following discussion of a complex case of bilateral influence will illustrate these difficulties. (For questions of code copying, see Johanson 1992, 1993, 1996, 1997, 1998a, 1999a, 1999b, 2000a, 2001, 2002a, 2002b.)

Focal intraterminals

Many innovations observed in Tajik Persian are instances of selective copying of semantic and combinational properties from Turkic. Some of them concern grammatical markers of the verb system (Johanson 1992: 245-246, 2000: 99-101). The issue to be discussed here is a case of renewal of focal intraterminality.

The viewpoint aspect of intraterminality, typical of present/imperfect categories, means envisaging an event between its outer limits. Intraterminals show various degrees of focality. Focal intraterminals (“progressives”) express actional concentration, focus, on the event at a given intraterminal viewpoint (Johanson 1971: 100-101, 133-134, 2000b: 76-78, 85-93).

Persian forms intraterminals by means of the prefix *mi-/me-*, e.g. *mikærd/mekærd* ‘was doing’. Eastern Persian displays focal intraterminal markers derived from the verbs *istvdæn* ‘to stand up, to stand’ and *xvræftæn* ‘to lie down, to lie’. These devices seem to have been copied from Turkic markers derived from the corresponding verbs *tur-* ‘to stand up, to stand’ and *jat-* ‘to lie down, to lie’.

Types based on ‘to stand’

The focal intraterminal periphrases in Tajik Persian are mostly based on *istvd-æ* plus a copula (‘to be’). The non-past item is *-æ istvd-æ æst*, e.g. *ræftæ istvdæ æst* ‘is going’. The corresponding past item is *-æ istvd-æ bud*, e.g. *ræft-æ istvdæ bud* ‘was going’ (Kerimova 1966: 225, Windfuhr 1990: 544). Northern dialects exhibit contracted forms such as the non-past items *ræfsvdæs* or *ræfsvs* ‘is going’ and the past items *ræfsvde bud* or *ræfsvdut* ‘was going’ (Rastorgueva 1964: 108-109).

The form *istvd-æ* is a postterminal (“perfect”) participle formed from the stem *istvd-* of *istvdæn* ‘to stand up, to stand’. The viewpoint aspect of postterminality, typical of resultative/perfect categories, means envisaging an event after its relevant outer limit (Johanson 2000b: 104-106). The postterminal participle in *-æ* is also used in finite forms such as the present perfect, e.g. *ræftæ (æst) ~ ræftæs* ‘has gone’, and the past perfect (pluperfect), e.g. *ræftæ budæ (æst)* ‘had gone’. The Turkic postterminal converb in *-(I)b* is a close functional equivalent of the postterminal participle in *-æ*. Thus, non-past postterminals in *-(I)b*, e.g. Uzbek *barib*, Uyghur *berip* ‘has gone’, correspond to *ræftæ (æst)*. Past postterminals in *-(I)b ærdi* such as *-(i)b edi*, *-(u)widi*, *-(i)w-idi*, e.g. Uzbek *barib edi*, Uyghur *beriwidi* ‘had gone’, correspond to *ræftæ bud ~ ræftæs bud*.

It has often been claimed that Tajik Persian focal intraterminals of the type *-æ istvdæ* + copula are copies of Turkic items. Nothing speaks against this assumption. Soper, however, remarks that the Uzbek focal intraterminality marker is derived from *jat-* ‘to lie, to lie down’, and not from *tur-* ‘to stand up, to stand’. If an Uzbek marker had been copied, he argues, it should have been the one derived from *jat-*. The Uzbek marker would have been expected to be derived from *xvræftæn* and not from *istvdæn* (1987: 86-87, 1996:67-68). This is obvious: Tajik Persian items such as *ræfsv(dæ)s* ‘is going’ could not possibly have been modeled on Uzbek items such as *barjæpti*. They may, however, have developed from an older type of Turkic focal intraterminals copied into eastern Persian at a much earlier time.

Renewal of focal intraterminals

Renewals of the expression of focal intraterminality are known from the history of many languages. Most Turkic languages have undergone at least one renewal during their known history (Johanson 1976, 1989, 1995, 1998b: 113-116, 1999c, 1999d). The formal starting-point has been a postverb construction consisting of a lexical verb + a converb marker + an auxiliary verb. The auxiliary is a desemanticized verb that modifies the actional content of the lexical verb, describing the way the action is

performed, e.g. *jaz-a tur-* ‘to write continuously’. Here *jaz-* means ‘to write’, *-a* is the converb marker, and *tur-* conveys the notion of durativity.

Further grammaticalization of certain postverb constructions has led to the creation of aspect-tense markers. The expression of focal intraterminality has been renewed by means of periphrases of this kind, the starting-point being a postverb construction indicating durativity, habituality, nontransformativity, etc. (Johanson 1995). Southwestern Turkic has used *jori-* ‘to move’ for this task, e.g. Turkish *yazıyor* ‘writes, is writing’. Most Turkic languages, however, have employed *tur-* ‘to stand up, to stand’. The pattern for this first known renewal of focal intraterminality was *-A turur*, with the auxiliary (‘stands’) in the old intraterminal form in *-(V)r*, producing items such as *jaz-a turur* ‘stands writing’ > ‘is writing’, later on with a reduced the material shape, e.g. *jazadi*, *jazat*, *jaza*. In modern southeastern Turkic, these items are represented by Uzbek *-æ-di* and Uyghur *-i-du*, which are now defocalized so-called “present-future” forms, e.g. Uzbek *jæædi*, Uyghur *jazidu* ‘writes, will write’.

The Tajik Persian focal intraterminal type *-æ istod-æ æst/bud*, which is based on ‘to stand’, may well be a copy of this earlier Turkic type. The type *ræftæ istodæ æst* is analogous to the southeastern Turkic type *bara turur* ‘is going’ < ‘stands going’.

It is, however, impossible to pinpoint the stages of grammaticalization of the relevant items at the time of copying. We do not know at what stage of development of the model code the item was copied, whether the item was still an actional marker or already a viewpoint aspect marker, what stage its material shape represented, at what developmental stage of the basic code the copy was acquired, etc. The model for *-æ istodæ æst* may have been the above-mentioned form *-A turur*. But it may also, as we will see, have been a more recent, less grammaticalized item of the type *-(i)b turub* (*turur*).

Products of later renewals

In many Turkic languages further renewals of the expression of focal intraterminality have taken place. Southeastern Turkic has employed periphrases with *jat-* ‘to lie down, to lie’, consisting of a lexical verb + a converb marker + the auxiliary *jat-* + a converb marker + a copula verb. Patterns of this type have produced a variety of forms. **-A jat-ib turur* has yielded Uzbek *-(æ)jotibt(i)* and *-jæp(ti)*, e.g. *jææjotibt(i)*, *jææjæpt(i)* ‘is writing’. The type **(I)b jat-ib turur* has yielded Uyghur *-iwati(du)*, e.g. *jeziwatidu* ‘is writing’. Even this type may have influenced Tajik Persian. In the dialects displaying the strongest Uzbek impact, the marker of focal intraterminality can also be based on *xɔræftæn* ‘to lie down, to lie, to sleep’ (in the literary register: *xɔb ræftæn*), which is the equivalent of Uzbek *jot-*. Given the formal and semantic analogies, these markers actually seem to be selective copies from Uzbek (Ras-torgueva 1952b: 230; 1964: 111-113). Nevertheless, it is impossible to define exactly which developmental stages and variant forms they go back to.

Southeastern Turkic has produced another type of similar forms, consisting of the lexical verb + a converb marker *-(I)b* + an auxiliary verb + the converb marker *-(I)b*

(+ **turur*). The auxiliary verb is derived from verbs meaning ‘to stand’, ‘to sit’, ‘to lie’ or ‘to move, to go’. Uzbek thus exhibits forms with *tur-* ‘to stand’, *otir-* ‘to sit’, *jot-* ‘to lie’, *jür-* ‘to go, to walk’, e.g. *jözib turibdi*, *jözib otiribdi*, *jözib jotibdi*, *jözib jüribdi*. Modern Uyghur has similar forms with *tur-* ‘to stand’, *oltur-* ‘to sit’ and *žür-* ‘to go, to walk’, e.g. *jezip turup*, *jezip olturup*, *jezip žürüp*. These items represent relatively young stages of grammaticalization. The material reduction is minimal, and the desemanticization of the auxiliaries, three postural verbs and one motion verb, has not been completed. The constructions still imply shades of meaning of the corresponding lexical verbs. Thus, Uyghur *jezip turup* means ‘writes/is writing in a standing position’, *jezip olturup* ‘writes/is writing in a sitting position’, *jezip žürüp* ‘writes continuously, regularly, periodically’. The constructions serve to express actional modification rather than forming aspect-tense items. At the beginning of their development, new aspectual items may, however, be difficult to distinguish from actional items (Johanson 1991, 1995, 1999c, 199d, 2000b: 95-97).

Tajik Persian displays similar constructions, consisting of the postterminal participle in *-æ*, which corresponds to the Turkic converb in *-(I)b*, plus one of the verbs *istvdæn* ‘to stand’, *xvræftæn* ‘to lie’, *šišæn* ‘to sit’ and *gæštæn* ‘to go’. We have already mentioned the types *-æ istvd-æ æst/bud* and *-æ xvræft-æ æst/bud*, which are similar to Turkic constructions with *tur-* ‘to stand up, to stand’ and *jat-* ‘to lie down, to lie’.

Turkic patterns have certainly served as models for the corresponding Persian Tajik periphrases. But at what developmental stages of the model code elements did the copying take place? Was it one of the items *-A turur*, *-(I)b turur*, *-(I)b tur-ub turur* and *-(I)b jot-ib turur* that served as the model? Or was it some corresponding item at a later stage of grammaticalization? The item *-æ istvdæ æst* is most probably a selective copy from Uzbek (Rastorgueva 1952b: 230, 1964: 132-133), but it is impossible to determine whether it ultimately goes back to *-A turur*, to *-(I)b turur* or to *-(I)b tur-ub turur*. Even if *xvndæ istvdæs* ‘is reading’ corresponds directly to Uzbek *oqup turibdi*, it cannot be excluded that it goes back to a different pattern, maybe an older pattern such as *oqub turur*.

Initiotransformative markers

Southeastern Turkic exhibits postterminal forms with *turub*, *jatib* etc. serving as intraterminal markers, e.g. Uzbek *jözib turibdi*, *jözib jotibdi* and Uyghur *jezip turup*, *jezip olturup*. The postterminal forms *turub*, *jatib*, etc. correspond to intraterminals such as *turur* and *tura* ‘stands’ in comparable periphrases of other Turkic languages. How can postterminal forms be used for forming focal intraterminals in Uzbek and Uyghur? The reason is that they are postterminal forms of initiotransformative verbs, a kind of preterit-present items (Johanson 2000b: 161-163). Initiotransformatives are verbs such as *tur-* and *jat-*, each covering two phases of an action: (i) ‘to stand up, to stop’ and (ii) ‘to stand’; (i) ‘to lie down’ and (ii) ‘to lie’. The postterminal aspect of initiotransformatives envisages an event after its initial limit. Items such as *turub* and

jatib thus convey the meanings 'having stood up' = 'standing' and 'having lain down' = 'lying'. They may refer to the same objective situation as described by corresponding intraterminals, cf. English *is hidden* = *has hidden* = *is hiding*.

Southeastern Turkic may thus use the preterit-present items of the type *turub*, *jatib*, *olturub*, *jürüb* in the conjugation instead of *-(V)r* forms. The lexical verbs meaning 'to stand', 'to lie', 'to sit' and 'to go' do not need the normal auxiliary-based markers to express focal intraterminality; postterminals such as *turub* and *jatib* are used instead. This usage, which is typical of the southeastern Turkic conjugation, is most probably the result of a very old Iranian influence; cf. the Persian type *nišastah* 'seated, sitting'. The corresponding eastern Persian postterminals *istvdæ*, *xvræftæ*, *šišṭæ* and *gæštæ* are used in the same way, e.g. *istvdæ* (*æst*) = *turub* '(has stood up)' 'is standing'.

Conclusion

The copying processes dealt with above are obviously bilateral. We find a clear southeastern Turkic impact on eastern Persian verb forms, but also traces of a reverse influence. On the one hand, Turkic influence is observed in the use of forms of certain auxiliary verbs as markers of focal intraterminality. On the other hand, Iranian influence may be observed in the use of preterit-presents in Turkic varieties. Further Turkic influence may be assumed behind the formation of various actional periphrases in Tajik Persian. Some of the copying processes leading to innovations in the verb systems may be of high age, whereas others may be more recent. In order to establish a relative chronology of the copying processes involved, we would need more precise historical data concerning the developmental stages of the elements copied.

References

- Johanson, Lars 1971. *Aspekt im Türkischen. Vorstudien zu einer Beschreibung des türkisch-türkischen Aspektsystems*. (Studia Turcica Upsaliensia 1.) Uppsala: Almqvist & Wiksell.
- Johanson, Lars 1976. Zum Präsens der nordwestlichen und mittelasiatischen Türksprachen. *Acta Orientalia* 37: 57-74. [Also in Johanson 1991: 99-116.]
- Johanson, Lars 1989. Aorist and present tense in West Oghuz Turkic. *Journal of Turkish Studies* 13: 99-105.
- Johanson, Lars 1991. *Linguistische Beiträge zur Gesamtturkologie*. (Bibliotheca Orientalis Hungarica 37.) Budapest: Akadémiai Kiadó.
- Johanson, Lars 1992. *Strukturelle Faktoren in türkischen Sprachkontakten*. (Sitzungsberichte der Wissenschaftlichen Gesellschaft an der J. W. Goethe-Universität Frankfurt am Main 29: 5.) Stuttgart: Steiner.
- Johanson, Lars 1993. Code-copying in immigrant Turkish. In: Extra, G. & Verhoeven, L. (eds.) *Immigrant languages in Europe*. Clevedon & Philadelphia & Adelaide: Multilingual Matters.
- Johanson, Lars 1995. Mehrdeutigkeit in der türkischen Verbalkomposition. In: Erdal, M. & Tezcan, S. (eds.) *Beläk bitig. Sprachstudien für Gerhard Doerfer zum 75. Geburtstag*. (Turcologica 23.) Wiesbaden: Harrassowitz. 81-101.

- Johanson, Lars 1996. Kопierte Satzjunktoren im Türkischen. *Sprachtypologie und Universalienforschung* 49: 1-11.
- Johanson, Lars 1997. Kopien russischer Konjunktionen in türkischen Sprachen. In: Huber, D. & Worbs, E. (eds.) *Ars transferendi. Sprache, Sbersetzung, Interkulturalität*. Frankfurt: Peter Lang. 115-121.
- Johanson, Lars 1998a. Code-copying in Irano-Turkic. *Language Sciences* 20: 325-337.
- Johanson, Lars 1998b. The history of Turkic. In: Johanson, L. & Csató, É. Á. (eds.) *The Turkic languages*. London: Routledge. 81-125.
- Johanson, Lars 1999a. Frame-changing code-copying in immigrant varieties. In: Extra, G. & Verhoeven, L. (eds.) *Bilingualism and migration*. (Studies on Language Acquisition 14.) Berlin & New York: Mouton de Gruyter. 247-260.
- Johanson, Lars 1999b. The dynamics of code-copying in language encounters. In: Brendemoen, B. & Lanza, E. & Ryen, E. (eds.) *Language encounters across time and space*. Oslo: Novus Press. 37-62.
- Johanson, Lars 1999c. Grenzbezogenheit in Aspekt und Lexik am Beispiel türkischer Postverbiaalkonstruktionen. In: Breu, W. (ed.) *Probleme der Interaktion von Lexik und Aspekt (ILA)*. Tübingen: Max Niemeyer. 129-139.
- Johanson, Lars 1999d. Typological notes on aspect and actionality in Kipchak Turkic. In: Abraham, W. & Kulikov, L. (eds.) *Tense-aspect, transitivity and causativity. Essays in honour of Vladimir Nedjalkov*. (Studies in Language Companion series 50.) Amsterdam & Philadelphia: Benjamins. 171-184.
- Johanson, Lars 2000a. Linguistic convergence in the Volga area. In: Gilbers, D., Nerbonne, J. & Schaeken, J. (eds.) *Languages in contact*. (Studies in Slavic and General Linguistics 28.) Amsterdam & Atlanta: Rodopi. 165-178.
- Johanson, Lars 2000b. Viewpoint operators in European languages. In: Dahl, O. (ed.) *Tense and aspect in the languages of Europe*. Berlin & New York: Mouton de Gruyter. 27-187.
- Johanson, Lars 2001. On Bulgarian copies of Turkish suffixes. In: Iglá, B. & Stolz, T. (eds.) *“Was ich noch sagen wollte...”. A multilingual Festschrift for Norbert Boretzky on occasion of his 65th birthday*. (Studia Typologica 2.) Berlin: Akademie-Verlag. 177-180.
- Johanson, Lars 2002a. *Structural factors in Turkic language contacts*. [With an introduction by Bernard Comrie.] London: Curzon.
- Johanson, Lars 2002b. Contact-induced linguistic change in a code-copying framework. In: M. Jones & E. Esch (eds.) *Language change: The interplay of internal, external and extralinguistic factors*. (Contributions to the Sociology of Language 86.) Berlin: Mouton de Gruyter. 285-313.
- Kerimova, Aza A. 1966. Tadžikskij jazyk. In: Vinogradov, V. V. (ed.) *Jazyki narodov SSSR I*. Moskva: Nauka. 212-236.
- Rastorgueva, Vera S. 1952. Očerki po tadžikskoj dialektologii 1. *Varzobskij govor tadžikskogo jazyka*. Moskva: Nauka.
- Rastorgueva, Vera S. 1964. *Opyt sravnitel'nogo izučenija tadžikskix govorov*. Moskva: Nauka.
- Soper, John D. 1987. Loan syntax in Turkic and Iranian: The verb systems of Tajik, Uzbek and Qashqay. Los Angeles.
- Soper, John D. 1996. *Loan syntax in Turkic and Iranian*. Revised and edited by A. J. E. Bodrogligeti. (Eurasian Language Archives 2.) Bloomington, Indiana: Eurolingua.
- Windfuhr, Gernot L. 1990. Persian. In: B. Comrie (ed.) *The world's major languages*. New York & Oxford: Oxford University Press. 523-546.

Notes on *-QALAQ* in Shor

Dmitrij M. Nasilov

Nasilov, Dmitrij M. 2005. Notes on *-QALAQ* in Shor. *Turkic Languages* 9, 229-233.

The affix *-qalaq-* is found in Khakas, Tuva, Altay, Shor, Chulym-Turkic, Bačat-Teleut, Baraba and Tomsk Tatar Turkic. The analytical markers *-a + elek* and *-a + ilik* are considered to be its analogues in Kirghiz and Yakut along with Dolgan. The described forms of all Siberian Turkic languages are characterized by the similar meaning of an action which has not yet taken place; the secondary nature of the origin of their markers is obvious. The author assumes that these forms go back to a combination of the intensifying particle (*e*)*le* and the negative predicative noun *joq* with either the adverb *-a/-y/-u* or the verbal noun *-ig/-gi/-yq/-qy/-qu*.

Dmitrij M. Nasilov, Moscow State University, Institute of Asian and African Studies. Ul. Oranžerejnaja, 20. 141200, Puškino, Russia.

The Shor language is a part of the area of Turkic languages where the verbal marker *-QALAQ* functions as a participle or as a tense stem, e.g. *tüle-gelek at* ‘a horse has not been saddled yet’, *šiq-qalaq ay* ‘a moon that has not risen yet’, *at-qalaq kiyik* ‘a roe that has not been killed yet’; *aŋči kel-gelek* ‘the hunter has not come yet’, *at-qalaq-pyn* ‘I have not shot yet’, *taŋ čari-ŋalaq pol-ŋan* ‘the dawn has not broken yet’.

This suffix is also observed in Khakas, Tuvan, Altay Turkic, Chulym Turkic, Bačat Teleut, Baraba and the Tomsk variety (Esipova 1993: 17). Tofan astonishingly stands out against these languages. Accepted counterparts of this affix are the markers *-A + elek* and *-A + ilik* in Kirghiz and Yakut (also Dolgan) respectively; they also function as predicates (Korkina 1970: 239; Ubrjatova 1985: 31, 1988: 471). All these forms share a similar meaning, the expression of a quality or a condition that has not yet appeared; in other words, the action has not occurred yet, but is expected (the Altay grammar of 1869, Dyrenkova 1941), not occurred yet at the time being (O. Böhtlingk). Unrealized action is meant here. However, some authors give a different evaluation of the predicative use of the form as *verbum finitum*: Korkina defines Yakut *-A ilik* as “the mood of the unfulfilled action”, e.g. present tense *En miigin iitt-ahata ilik-kin* ‘You have not given me either food or drink yet’, and past tense *Utuya ilik ete* ‘(S)he had not slept yet (at that time)’ (1970: 247-249). According to Čispijakov, the respective Shor form is “the past imperfect verb tense”, e.g. *Ayaš pürülen-gelek* ‘The tree has not lost its leaves’ (1992 : 116). Tadykin believes that in Altay Turkic the participle *-GALAQ* only conveys (negative) past tense meaning relative to the moment of speech, e.g. *Ol bar-ŋalaq* ‘He had to go, but he had not gone

yet' (1971: 90). In Čulym Turkic, the main meaning is said to be an action which has not taken place in the past, e.g. *Qoy soq-qalaq-pis* 'We have not slaughtered the sheep yet' (Birjukovič 1981: 67). Isxakov & Pal'mbax consider the same form in Tuvan as the future of the expected time (1961: 391).

Comparing the semantics of the mentioned forms with Yakut *-A ilik* and Kirghiz *-A elek*, Korkina concludes that they are not only similar to each other at present, but probably have a common origin (1970: 246). Ubrjatova recognizes the common origin of the given affixes too, relating them, however, to the secondary participle forms of a later origin (1985: 32). As for the origin of these forms, the authors of the comparative-historical grammar of the Turkic languages adhere to the same opinion (1988: 471). Schönig considers a common source of the forms although he notes the difficulties of explaining the final results of a development from a common archetype, "if this archetype existed" (1998: 135).

It follows from the explanations quoted above that the Yakut and Kirghiz forms represent an earlier stage of a development that led to the replacement of the analytical forms by the suffix *-QALAQ*. This raises some questions about the original form of the first and second components of the analytical construction. Though there are different opinions on this, it is common to suppose an original ambiguity of the component *ilik / elek*. E. I. Korkina says that it is difficult to solve the problem of the etymology of *ilik / elek / kalak* and that this question has to be studied specially (1970: 246). Ubrjatova defines *ilik* as a word meaning 'which not... yet' (1976: 54). Ščerbak leaves *elek* without comments (1977: 160, 1981: 96). Böhlingk (1989: 471) and Monguš & Sat (1968: 611) assume a separate word, Yakut *ilik* 'not happening yet', 'not having occurred yet', Tuvan *elek* 'it is too early' (*e-+-lek?*), 'till, until'. Judaxin interprets Kirghiz *elek* as a verbal negation 'still not yet' (1965: 947).

Most researchers probably agree that the second component of the construction is *ilik / elek*. The next question concerns the first component of the construction. It seems to be clear for Yakut that this is a form of the common Turkic geconverb *-A/-y*. Böhlingk first expressed this opinion (1989#: 415), and his view is supported by modern Yakut linguists. The similar problem concerning Kirghiz is solved in the same way (Judaxin 1965: 947).

As for the affix *-QALAQ*, its initial velar is derived from an ancient verbal noun *-GU/-GI* (Ščerbak 1981: 178; Birjukovič 1981: 66), *-IG/-IK* (Sat 1959: 79), or *-GAN* (Castrén 1857: 74; Benzing 1959: 4). A number of specialists in Turkic philology assume a variant of *-IIQ/-IUQ* in this affix (Birjukovič 1981, Jusupov 1985: 83). Jusupov takes this suffix to go back to *elek*, cf. the participle suffix *-GIIIQ* expressing a potentiality in Tatar dialects and in some other Turkic languages (1985: 83). This comparison with the form marker *-GIIIQ* < *-GUIIU*, well known from Old Uyghur documents, with the particle *elek* is semantically weak (and probably also phonetically).

In Old Turkic, *-GUIIUQ*, derived from *-GU*, was used to express future action accompanied by the same modality nuances typical of *-GU*, i.e. the necessity and obligation, e.g. Old Uyghur *bilgülük ol* 'should be known', *bilgä biliginä sizik köñül*

turyur-ma-yuluq ol 'it is impossible to doubt his wise competence', *üzük sözlä-gülük ärmäz men* 'I shall not tell a lie', *darnini säkiz yüz qata oqı-yuluq ol* 'he must recite the formula eight hundred times'.

The etymologies just quoted do not explain the suffix *-LAQ*, if it is not a rest of *elek*. This suffix does not seem to mark verbal forms and can hardly be correlated with *-GUIUQ*.

As mentioned, the forms under discussion have in all Siberian languages the meaning of an action which has not taken place. According to Ubrjatova, it should be expedient to try to explain its negative semantic. Attention should also be drawn to Benzing's etymology, which is not mentioned in the works cited. Benzing derives the Altay-Sayan suffix *-GALAQ*, as well as the Kirghiz and Yakut forms from the analytical construction **-yan + ele + yoq* (1959: 4), which probably includes the intensifying particle *ele*, characterized by a wide range of functions, e.g. in Kirghiz (Batmanov 1940: 64-68). Batmanov refers Kirghiz *ele* to the category of "syntactic words" like *eken* and *emes*. The particle *ele* can be compared also with the Common Turkic expressive-emphatic particle *la* (*ele* < /e/ + *le*). In Old Turkic and Middle Turkic this particle is often added to the verb, intensifying the action or condition expressed by it, conveying a categorical shade of meaning, e.g. *öldüm-le* 'I have died!' (Ščerbak 1987: 99-100). In Kirghiz, the combination of the past tense marker *-DI*, the perfect marker *-GAN* or the past iterative marker *-čU* with *ale* allocates the action as a separate episode after which the further narration will follow. In combination with the present-future marker *-A* or with the presumable future marker *-Ar* gives them the meaning of the opportunity to perform an action in the future (Batmanov 1940 : 66). The intensifying role of *ale* is very significant in this case.

Benzing's construction makes the negative semantics of the considered form transparent. The use of *yoq* for verbal nouns negative forms creation is a widespread Turkic phenomenon (Ščerbak 1981: 97). This negative predicative *yoq* is added to *-GAN* and the nouns of action in *-IG* and *-GU*, which have played a considerable role in development of tense system in the Turkic languages (Blagova 1958). Such forms are found in Old Uyghur, e.g. *baliqtaqı ig toya ketgüsi yoq* 'there is pestilence in the city, and the infection will not leave'. Tadykin remarks that in Altay Turkic, the participle *-GALAQ* and the negative form of the participle *-GAN* are often inter-changeable. The meaning of *-GALAQ* is close to the negative form of *-GAN*; it consists of combination of this participle with the word *yoq* 'not available' (1971: 88). The same semantic conformity is observed in Fu-yü Kirghiz (Schönig 1998: 136⁶).

As regards the use of *yoq* in tense paradigms, specific attention should be paid to the negative forms of the present tense in Yellow Uigur *-u + yoq-tır*. In this form the negative *yoq* is added to the common Turkic converb in *-U/-A/-y*, e.g. *men kel-ü yoq-tır* 'I do not go, I am not going' (Tenišev 1976a: 86); cf. Salar *piser var-yox-tır* 'we do not go, we are not going' (Tenišev 1976b: 140). The use of *yoq* in tenses is observed in Čulym Turkic too (Birjukovič 1981: 46-58). These examples show that *yoq* can be added to adverbial forms.

Thus, the origin of Yakut *ilik* and Kirghiz *elek* may be a combination of the intensifying particle and the negative word *yoq*, i.e. *(e)le + yoq > elek/ilik*. This complex component was combined either with the converb in *-U/-A/-y* or with the verbal noun *-IG/-GU/-GI*, marking the analytical form with a negative value of an action that has not happened. The analytical form was used in attributive and predicative functions, and it could be substantivized, as in Yakut. Benzing's suggestion of *-GAn* as the first component is more problematic since it assumes complicated phonetic changes. Consequently, Ubrjatova's opinion to the effect that the Yakut and Sayan forms may be secondary formations of the participle can be accepted. The Shor participle and tense marker *-QALAQ* may thus be the result of a phonetic development of an ancient analytical construction.

References

- Batmanov I. A. 1940. *Grammatika kirgizskogo jazyka 3: Tipy otglagol'nyx obrazovanij i ix funkcii*. Frunze: Kirgizizdat.
- Benzing, J. 1959. Classification of the Turkic languages. In: Deny, Jean & Grønbech, Kaare & Scheel, Helmuth & Togan, Zeki Velidi (eds.) *Philologiae turcicae fundamenta* 1. Aquis Mattiacis: Steiner. 1-5.
- Birjukovič R. M. 1981. *Morfologija čulymsko-tjurkskogo jazyka*. Saratov: Knižnoe izdatel'stvo.
- Blagova G. F. 1958. Sootnositel'nye glagol'nye formy i ix razvitie v uzbekskom jazyke. *Vo-prosy jazykoznanija*, 1958: 4.
- Böhtlingk O. N. 1889. *O jazyke jakutov*. Novosibirsk: Nauka.
- Castrén M. A. 1857. *Versuch einer koibalischen und karagassischen Sprachlehre*. St.-Petersburg: Kaiserliche Akademie der Wissenschaften.
- Čispiyakov E. F. 1992. *Učebnik šorskogo jazyka*. Kemerovo: Kemerovskoe knižnoe izdatel'stvo.
- Dyrenkova N. P. 1941. *Grammatika šorskogo jazyka*. Moskva & Leningrad: Izdatel'stvo AN SSSR.
- Esipova A. V. 1993. *Opredelitel'naja funkcija pričastij v šorskom jazyke*. Novosibirsk.
- Grammatika 1869. *Grammatika altajskogo jazyka*. Kazan': Izdatel'stvo Kazanskogo universiteta.
- Isxakov F. G. & Pal'mbax A. A. 1961. *Grammatika tuvinskogo jazyka. Fonetika i morfologija*. Moskva: Nauka.
- Judaxin K. K. 1965. *Kirgizsko-russkij slovar'*. Moskva: Slovarnoe izdatel'stvo.
- Jusupov F. J. 1985. *Neličnyje formy glagola v dialektax tatarskogo jazyka*. Kazan': Tatknigizdat.
- Korkina E. I. 1970. *Naklonenija glagola v jakutskom jazyke*. Moskva: Nauka.
- Morfologija 1988. *Sravnitel'no-istoričeskaja grammatika tjurkskix jazykov: Morfologija*. Moskva: Nauka.
- Sat Š. Č. 1959. Pričastie v tuvinskom jazyke. In: *Učenyje zapiski TuvNIJALI* 3. Kyzyl: Kyzylskoe knižnoe izdatel'stvo.
- Schönig C. 1998. A new attempt to classify the Turkic languages 3. *Turkic Languages* 130-151.

- Ščerbak A. M. 1977. *Očerki po sravnitel'noj morfologii tjurkskix jazykov: Imja*. Leningrad: Nauka.
- Ščerbak A. M. 1981. *Očerki po sravnitel'noj morfologii tjurkskix jazykov: Glagol*. Leningrad: Nauka.
- Ščerbak A. M. 1987. *Očerki po sravnitel'noj morfologii tjurkskix jazykov: Narečije, služebnye časti reči, izobrazitel'nye slova*. Leningrad: Nauka.
- Tadykin V. N. 1971. *Pričastija v altajskom jazyke*. Gorno-Altajsk: Altajskoe knižnoe izdatel'stvo.
- Tenišev Ė. R. 1976a. *Stroj saryg-južurskogo jazyka*. Moskva: Nauka.
- Tenišev Ė. R. 1976b. *Stroj salarskogo jazyka*. Moskva: Nauka.
- Tuvinsko-russkij slovar'*. 1968. Moskva: Slovarnoe izdatel'stvo.
- Ubrjatova E. I. 1976. *Issledovanija po sintaksisu jakutskogo jazyka 2: Složnoe predloženie 1*. Novosibirsk: Nauka.
- Ubrjatova E. I. 1985. *Istoričeskaja grammatika jakutskogo jazyka*. Jakutsk: Jakutskoe knižnoe izdatel'stvo.

Concessive and adversative constructions in Siberian Turkic

Irina Nevskaya

Nevskaya, Irina 2005. Concessive and adversative constructions in Siberian Turkic. *Turkic Languages* 9, 234-251.

In this paper, the semantic and structural types of concessive and adversative constructions in Siberian Turkic languages are investigated. In the semantic domain, we distinguish concessive proper and conditional-concessive constructions as well as real and unreal ones, the latter falling into hypothetical and counterfactive constructions. Generalised concessive constructions versus non-generalised ones represent another classificatory criterion. As for their structure, concessive constructions with the conditional form of the dependent predicate in combination with the particle *DA* represent the core of such constructions in all Turkic languages. Additionally, there exist language specific means of expressing concessive and adversative relations: constructions with imperative forms of the predicate in the concessive clause, constructions with various participial forms of the dependent predicate and some contextual means of expressing concession.

Adversative constructions are a more recent means of expressing concessive-adversative relations. In Siberian Turkic, we find only a few adversative conjunctions, most of them are copied structurally or fully from Russian. In addition, a number of modal introductory phrases are on the way to being grammaticalised as adversative conjunctions.

Irina Nevskaya, Institute of Philology of the Siberian Division of the Russian Academy of Sciences, Novosibirsk, Russia. Email: Nevskaya@em.uni-frankfurt.de

1. Introductory remarks

Concessive and adversative relations alongside causative, consequential, final, resultative and conditional ones express determination of one situation by another and belong to the group of inter-propositional relations. Concession is a relation opposite to cause. A situation which is seen as a cause brings about another situation which is its consequence: *Because he had not done his homework, he got a bad mark*. A concessive situation brings about an anti-consequential situation (or an adversative one) which takes place in spite of the first situation, contrary to our expectations and contrary to a normal, i.e. causal, succession of events: *Although he had not done his homework, he got a good mark*.

	1 st situation	2 nd situation
<i>Causative-consequential relations:</i>	cause	consequence
<i>Concessive-adversative relations:</i>	concession	anti-consequence

Causative-consequential relations are normally expressed by means of specialised causative (1a) or consequential (1b) constructions. In a causative construction, the causative situation is marked by a causative connecting element (a conjunction, a particle or an adverb in the connector function grammaticalised to different degrees) (1a); in (1b), the consequential situation is marked by a grammaticalised consequential element:

- (1) a. *Because he had not done his homework, he got a bad mark.*
 b. *He had not done his homework; therefore, he got a bad mark.*

Similarly, we can distinguish concessive (2a) and adversative (2b) constructions:

- (2) a. *Although he had not done his homework, he got a good mark.*
 b. *He had not done his homework, but he got a good mark.*

Concessive constructions have been recently studied cross-linguistically by Bondarko 1996, Haspelmath & König 1998 and Xrakovskij 2004c; the latter edition is a monograph written by a group of authors. It contains an overview of Turkic concessive constructions (Isxakova, Nasilov & Nevskaya 2004). In this article, we describe concessive and adversative constructions in South Siberian Turkic in more detail. First, we distinguish their semantic types (Section 2). Then, we describe the structural and semantic types of concessive constructions in South Siberian Turkic (Section 3). Most of them are mono-finite, i.e. they are of the synthetic type (Čeremisina, Skribnik 1986) and consist of a matrix clause with the predicate in a finite verb form and of a dependent clause with the predicate in a nonfinite form (e.g. a converb or a participle). These are the most typical means of expressing concessive-adversative relations in Turkic. Concessive constructions with an imperative form of the dependent predicate are a rare exception. Adversative constructions with adversative conjunctions are a more recent means of expressing concessive-adversative relations in Siberian Turkic. They are bi-finite and contain adversative conjunctions, copied (Johanson 1992) from Russian either fully or structurally (Section 4). Ways of expressing concessive-adversative relations at the textual level are described in Section 5.

For illustration, we use the language material found in grammar descriptions (Anonymous 1884, Baskakov 1958, 1966, 1972, 1975, 1985, Čeremisina 1995, Dyrenkova 1941, Isxakov & Pal'mbax 1961, Nevskaya 1993, Radloff 1966, Ubrjatova 1982 etc.) as well as our Shor field data and experimental material collected according to Xrakovskij's questionnaire on concessive constructions (Xrakovskij 2004b).

<i>bol-zo,</i>	<i>üy-in-e</i>	<i>süy-di</i>
be-COND	wife-POSS3-DAT	present-ACC
<i>al-ıp</i>	<i>ber-e</i>	<i>e-d-i.</i>
buy-CONV	give-FUT.PART	be-PAST-3

‘Even if Akchabay had not got his salary, he would have bought a present for his wife.’

2.2. Real, hypothetical and counterfactive constructions

In real constructions, both correlated situations are either factive or are seen by the speaker as quite plausible (3a and 3b). The predicate of the matrix clause is typically in the indicative mood. Counterfactive constructions present these situations as unreal, i.e. the speaker knows that the described situations have not taken place (3c). Only conditional-concessive interpretation of the construction is possible. The predicate of the matrix clause is in the conjunctive mood in this case (the future participle of the lexical verb plus the preterit of the auxiliary verb *e-* ‘be’) and the predicate of the concessive clause bears an analytical conditional marker consisting of a perfect participle of the lexical verb and of the conditional form of the auxiliary verb *pol-* ‘be’ in combination with the particle *DA*.

Hypothetical concessive constructions present both situations as problematic, but not really impossible (4). The presupposition is that the speaker considers this possibility.

Altay					
(4)	<i>Poezd</i>	<i>öyinde</i>	<i>de</i>	<i>kel-er</i>	<i>bol-zo,</i>
	train	on.time	PTL	come-PRF.PART	be-COND
	<i>baştaḡı</i>	<i>d’uum-ğa</i>	<i>oroyt-ıp</i>	<i>qal-ar</i>	<i>e-d-is.</i>
	opening	session-DAT	be.late-CONV	stay:AUX-FUT.PART	be-PST-1PL

‘Suppose the train arrived/had arrived on time, we would still be/have been late for the opening session.’

In (4), the predicate of the matrix clause is also in the conjunctive mood, but this form has the meaning of supposition here. The analytical conditional form consists of the future participle of the lexical verb and of the conditional form of the auxiliary verb. Only conditional-concessive interpretation of the construction is possible. The time reference is determined only by tense adverbs. The situations expressed in both counterfactive and hypothetical constructions are unreal ones.

2.3. Generalised concessive semantics

Following Xrakovskij 2000, we distinguish the following two types of generalised concessive semantics: non-iterative (4) and iterative (5) expressed by means of specialised generalised constructions. Non-iterative generalised constructions are emphatic. They contain indefinite pronouns combined with the concessive particle *DA* as their structural markers: *Shor kem de* ‘no matter who’, *qandıy da* ‘no matter

which', etc. In iterative constructions, the action performed by the same agent takes place many times, or different agents perform the same action.

Altay

- (5) *Qanayda* *da* *qapšağayla-za-η,* *qamčī-la*
 how PTL hurry-COND-2SG whip-INST
ad-īη-dī *soq-po!*
 horse-POSS2-ACC hit-NEG

'(No matter) how much you are in a hurry, do not hit your horse with the whip!'

Shor

- (6) *Parčīn kiži* *pil-ze* *de,* *ayt-paan-ča.*
 every person know-COND PTL say-NEG-PRS
 'Although everyone knows (that), they do not tell.'

Generalised concessive constructions can be either real or unreal as well as either concessive proper or conditional-concessive. Combining all the semantic criteria, we get the following semantic types of concessive constructions. They are illustrated with Tuvan examples here.

I. Non-generalised constructions

I.1. Non-generalised concessive proper constructions

- (7) *Ča's* *čap* *tur-za* *daa,* *Petrov* *öön-den*
 rain fall:CONV stand:AUX-COND PTL P. house-ABL
zontik *čoq* *ün-üp* *kel-gen.*
 umbrella without go.out-CONV come:AUX-PRF
 'Although it was raining, Petrov left the house without his umbrella.'

I.2. Non-generalised conditional-concessive constructions

I.2.1. Non-generalised real conditional-concessive constructions

- (8) *Ča's* *čay-za* *daa,* *Petrov* *öön-den*
 rain fall-COND PTL P. house-ABL
zontik *čoq* *ün-üp* *kel-ir.*
 umbrella without go.out-CONV come:AUX-AOR/FUT
 'Even if it rains, Petrov (always) leaves the house without his umbrella.'
- (9) *Ča's* *čap* *kel-ze* *daa,* *Petrov* *öön-den*
 rain fall:CONV come-COND PTL P. house-ABL
zontik *čoq* *ün-üp* *kel-ir.*
 umbrella without go.out-CONV come:AUX-AOR/FUT
 'Even if it is raining, Petrov will leave the house without his umbrella.'

I.2.2. Non-generalised unreal conditional-concessive constructions

I.2.2.1. Non-generalised hypothetical constructions

- (10) *Ča's* *čap* *kel-ir-daa* *bol-za,*
 rain fall:CONV come:AUX-FUT.PART-PTL be-COND
öön-den *zontik* *čoq* *ün-üp*
 house-ABL umbrella without go.out-CONV
Petrov *kel-ir* *iyik.*
 P. come:AUX-FUT.PART PTL
 'Suppose it were raining, Petrov would still leave the house without his umbrella.'

I.2.2.2. Non-generalised counterfactual constructions

- (11) *Ča's* *čap* *tur-yan-daa* *bol-za,*
 rain fall:CONV stand:AUX-PRF.PART-PTL be-COND
Petrov *zontik* *čoq*
 P. house-ABL umbrella without
ün-üp *kel-ir* *iyik.*
 go.out:CONV come:AUX-FUT PTL
 'Even if it had been raining, Petrov would have left the house without his umbrella.'

Or

- (12) *Ča's* *čap* *tur-yan-daa* *bol-za,*
 rain fall:CONV stand:AUX-PRF.PART-PTL be-COND
Petrov *zontik* *čoq*
 P. house-ABL umbrella without
ün-üp *kel-gey* *ertik.*
 go.out-CONV come:AUX-OPT PTL
 'Even if it had been raining, Petrov would have left the house without his umbrella.'

The subjunctive mood in Tuvan is built with either the future participle or the optative form of the lexical verb plus the modal particle *iyik* or *ertik*. These particles are structural analogues of the auxiliary *edi*: all of them go back to the Old Turkic auxiliary verb *är-* 'be' in the preterit *-D* (*edi*), or in the form of the perfect participle *-DOk* (*ertik*), or in the form of the evidential past *-yOk* (*iyik*).¹

II. Generalised concessive constructions

II.1. Generalised concessive proper constructions

- (13) *Petrov* *qim-dan-da* *ayür-za,* *qim-daa* *aŋaa*
 P. who-ABL-PTL ask-COND who-PTL he:DAT

¹ Concerning the Old Turkic formants, see Erdal 2004.

χarīn ber-ip šīda-vaan.
 answer give-CONV be.able-PRF.NEG
 ‘No matter whom Petrov asked [about it] (Although Petrov asked everyone about it), nobody could give him an answer.’

II.2. Generalised conditional-concessive constructions

II.2.1. Generalised real generalised conditional-concessive constructions

- (14) *Petrov qīm-dan-da aytīr-za, qīm-daa aḡaa*
 P. who-ABL-PTL ask-COND who-PTL he:DAT
χarīn ber-ip šīda-vas.
 answer give-CONV be.able-AOR/FUT.NEG
 ‘No matter whom Petrov asks [about it] (Even if Petrov asks everyone about it), nobody will be able to give him an answer.’

II.2.2. Generalised unreal conditional-concessive constructions

II.2.2.1. Generalised hypothetical conditional-concessive constructions

- (15) *Petrov qīm-dan-da aytīr-ar bol-za, qīm-daa*
 P. who-ABL-PTL ask-FUR.PART be-COND who-PTL
aḡaa χarīn ber-ip šīda-vas iyik.
 he:DAT answer give-CONV be.able-FUT.PART.NEG PTL
 ‘Suppose Petrov asked everyone [about it], nobody would still be able to give him an answer.’

II.2.2.2. Generalised counterfactual conditional-concessive constructions

- (16) *Petrov qīm-dan-da aytīr-ğan bol-za, qīm-daa*
 P. who-ABL-PTL ask-PRF.PART be-COND who-PTL
aḡaa χarīn ber-ip šīda-vas iyik.
 he:DAT answer give-CONV be.able-FUT.PART.NEG PTL
 ‘No matter whom Petrov had asked [about it] (even if Petrov had asked everyone about it), nobody could have given him an answer.’

3. Concessive constructions in Siberian Turkic

3.1. Concessive constructions with the conditional form

Concessive constructions with the conditional form of the dependent predicate are the most widespread means to express concessive semantics in Turkic. As we have seen, they can express both concessive proper and conditional-concessive semantics, denote real, hypothetical and unreal situations and render generalised and non-generalised concessive relations. Further we describe some of their structural features. Special attention is paid to the temporal localisation of the correlated situations in such constructions. It is worth noting that in contrast to English, German or Russian

concessive constructions where the anti-consequence can precede the concessive situation, Turkic concessive constructions mirror the temporal correlation of these situations iconically: the concessive situation always precedes the anti-consequence.

3.1.1. Structural features

Along with the form *-SA* such constructions normally contain concessive particles: *TAGI/TAA/DA*, *CI* or *LA*, etc. However, the conditional construction with the “plain” form *-SA* can also occasionally express concession (6).

Shor

- (17) *Quday-γa ižen-ze-η, poy-uη čanil-ba!*
 god-DAT trust-COND-2SG self-POSS2SG err-NEG
 ‘Although you trust in God, do not make mistakes yourself!’

The main clause can contain adversative conjunctions and particles or modal adverbs (Altay: *tünej le*, *d’ani la* ‘nevertheless, still’). They additionally mark the situation of anti-consequence.

The temporal localisation of the concessive and anti-consequential situations in such constructions depends on many factors: the tense/mood marker of the finite predicate (i.e. whether indicative or non-indicative), the structure of the conditional form itself (whether the conditional marker is added to the lexical verb itself or is added to the auxiliary verb *pol-* while the lexical verb takes a participial form; which participial form the lexical verb gets in the latter case, etc.) and on the presence of certain time adverbs.

3.1.2. Temporal localisation as expressed in concessive constructions with the simple conditional form *-SA*

In the concessive constructions with the simple conditional form of the dependent predicate, the temporal localisation of the concessive situation depends on the temporal localisation of the situation expressed by the matrix clause. Thus, the form *-SA* displays a feature here of a typical converb: absence of an independent temporal semantics, i.e. relative tense.

3.1.2.1. Both the concessive situation and that of anti-consequence can be temporally non-localised. These are repeatedly occurring situations.

Tuvan

- (18) *Al-za daa, “al-d-ïm” di-ves,*
 take-COND PTL take-PST-1SG say-NEG.AOR/FUT
či-ze daa, “či-d-ïm” di-ves.
 eat-COND PTL eat-PST-1SG say-NEG.AOR/FUT
 ‘Although he (always, repeatedly) takes, he does not say that he took; although he (always, repeatedly) eats (something), he does not say that he ate.’

3.1.2.2. Both situations are present ones:

Shor

- (19) *Qayizi kiži pil-ze de, ayt-paanča.*
 which person know-COND PTL say-NEG.PRS
 ‘Although somebody knows (that), he does not say (it).’

3.1.2.3. Both situations refer to the past:

Khakas

- (20) *Uzi-rya sayin-ip, xaraan čap-sa daa,*
 sleep-INF think-CONV eye close-COND PTL
noya da kirbik-ter čara la
 why PTL eyelash-PL separately PTL
oylas tar-yan-nar.
 run spread-PRF-PL
 ‘Although she tried to close her eyes, she could not (lit.: her eyelashes were separated).’

3.1.2.4. Both situations are located in the future:

Tofan

- (21) *Ol gel-se tä, sooda-vas.*
 he come-COND PTL say-NEG.AOR/FUT
 ‘Although he comes, he will not say (anything about this).’

3.1.3. Temporal localisation as expressed in concessive constructions with analytical conditional forms

3.1.3.1. Both situations are located in the future:

Altay

- (22) *Erten ada-m tura-niň üst-in de*
 tomorrow father-POSS1SG house-GEN roof-POSS3ACC PTL
d’aza-r bol-zo, men ogo
 repair-FUT.PART be-COND I he:DAT
boluř-paz-ım.
 help-AOR/FUT.NEG-1SG
 ‘Although my father repairs the roof tomorrow, I will not help him.’

3.1.3.2. Both situations refer to the past:

Altay

- (23) *Men köömöy dö işte-gen bol-zo-m,*
 I badly PTL work-PRF.PART be-COND-1SG
d’e d’aqši işte-p al-ata-m.
 but well work-CONV take-IMPF-1SG
 ‘Although I worked badly, I earned much.’

3.1.3.3. The concessive situation is situated in the past, the anti-consequence is situated in the present:

Altay

(24)	<i>Aqčabay</i>	<i>d'uunda-p</i>	<i>bar-ba-ğan</i>	<i>da</i>	<i>bol-zo,</i>
	A.	meet-CONV	go-NEG-PRF.PART	PTL	be-COND
	<i>ol</i>	<i>emdi</i>	<i>kem-di</i>	<i>de</i>	<i>kemdir-beyt.</i>
	he	now	who-ACC	PTL	receive-NEG.PRS

'Although A. has not gone to the meeting, he does not receive anybody now.'

3.1.3.4. The concessive situation refers to the past, the anti-consequence to the future:

Altay

(25)	<i>Aqčabaj</i>	<i>d'ed-ip</i>	<i>te</i>	<i>kel-gen</i>	<i>bol-zo,</i>
	A.	reach-CONV	PTL	come-PRF.PART	be-COND
	<i>men</i>	<i>oğo</i>	<i>telefon</i>	<i>soq-poz-ım.</i>	
	I	he:DAT	telefon	beat-NEG.AOR/FUT-1SG	

'Although Akchabay has already returned (home), I will not call him.'

We see that the concessive situation can be localised in a different period of time than the situation expressed in the main clause when the conditional form is an analytical one and the lexical verb takes a participial formant determining the temporal localisation of its action.

All the examples in the sections 3.1.2 and 3.1.3 can also have a conditional-concessive interpretation. All of them are real. The situations can be either factive (consequently the construction gets a concessive proper interpretation) or non-factive (consequently the construction gets a conditional-concessive interpretation).

3.1.4. Unreal conditional-concessive constructions

In unreal conditional-concessive constructions i.e. in hypothetical (15) and counterfactual (16) ones, the temporal localisation is neutralised. We need additional contextual markers (like temporal adverbs) to refer the situation to a certain time period.

Altay:

(26)	<i>Poezd</i>	<i>(bügün/keče/ertene)</i>	<i>öyinde</i>
	train	today/yesterday/tomorrow	on.time
	<i>kel-er</i>	<i>bol-zo,</i>	
	come-FUT.PART	be-COND	
	<i>baştapqı</i>	<i>d'uun-ğa</i>	<i>oroyt-ıp</i>
	first	session-DAT	be.late-CONV
	<i>de</i>	<i>qal-ar</i>	<i>e-d-is.</i>
	PTL	stay-FUT.PART	be-PST-1PL

'Suppose that the train had arrived on time (yesterday), we would have been late for the opening session anyway. / Suppose the train arrived on time (today/tomorrow), we would be late for the opening session anyway.'

Altay

(27) *Poezd* (bügün/keče/ertene) öyinde *de*
 train (today/yesterday/tomorrow) on.time PTL
kel-gen *bol-zo,*
 come-PRF.PART be-COND
baštapaqı *d'uun-ğa* *oroyt-ıp* *qal-ar*
 first session-DAT be.late-CONV stay-FUT.PART
e-d-is.
 be-PST-1PL
 'Even if the train had arrived on time (yesterday), we would have been late for the opening session. / Even if the train arrived on time (today/tomorrow), we would be late for the opening session.'

3.2. Concessive constructions formed by other forms

Such constructions are very diverse and language specific. Most of them are not specialised on expressing concessive relations and render a concessive meaning only in certain contexts. Concessive particles can be contextual markers of such uses.

3.2.1. Concessive constructions with the imperative forms

Imperative forms are found in concessive constructions quite often cross-linguistically: compare Russian *Bud' ja xot' volšebnikom, ja by i togda ne smog by vypolnil' tvoego želanija* 'Even if I were a magician, I would not be able to make your wish come true.' Such constructions always represent a non-factive concessive situation:

Tuvan

(28) *Day* *bedik* *daa* *bol,* *buura-ar,*
 mountain high PTL be get.ruined-FUT
dalay *terey* *daa* *bol,* *qurya-ar.*
 sea deep PTL be dry.out-FUT
 'Even if a mountain is high (lit.: be a mountain high), it gets ruined, even if a sea is deep (lit.: be a sea deep), it dries out.'

Tofan

(29) *Sen* *čor-iy* *tä* *ber,* *men* *ilya-vas-men*
 you go.away-CONV PTL give:AUX I cry-NEG.FUT-1SG
 'Even if you go away (lit.: you go away), I will not cry.'

3.2.2. Participles in case forms with postpositions

In Shor, there exists a highly specialised concessive construction with the meaning 'in spite of the situation A (concession), the situation B (anti-consequence) happens'. The dependent predicate is expressed by the participle *-GAn* with the postposition *üstüne* [*üst-ün-e* upper.part-POSS3-DAT]. The construction always renders factive

situations and is more emphatic than the standard construction with the conditional form.

Shor

(30)	<i>Men</i>	<i>elči</i>	<i>is-qan</i>	<i>üstüne,</i>
	I	ambassador	send-PRF.PART	POSTP
	<i>meenj</i>	<i>kel-er-im-ni</i>	<i>pil-be-d-iñ</i>	<i>či!</i>
	I:GEN	come-FUT.PART-POSS1SG-ACC	know-NEG-PST-2SG	PTL

'In spite of the fact that I had sent an ambassador, you did not know that I would come?!'

3.3. Contextual means of expressing concessive semantics

A number of constructions with the core meaning of a different type can express concession in certain contexts. These are mostly temporal constructions that may convey different types of causal semantics (cause, condition, purpose or concession) that can be induced by a correlation of the lexical meanings of the verbs in the matrix and the dependent clauses, by some structural markers (like certain verb forms or the presence of certain particles), by a pragmatic reading of a situation alone.

3.3.1. The form *-GAndA*

This form is made up of the perfect participle *-GAn* in the Locative case form. In Siberian Turkic, it is functionally close to gerunds. The prime function of the complex constructions with the dependent predicate in this form is temporal: the dependent situation determines the time of the matrix one. If the content of the matrix situation contradicts our expectations based on the content of the dependent situation, we may have concessive correlation of these situations. The concessive semantics is that of the factive type.

Shor

(31)	<i>Sen</i>	<i>alıñ</i>	<i>pol-yan-da,</i>
	you:2SG	fool	be-PRF.PART-LOC
	<i>pis-tiñ</i>	<i>tin-ıbis-ti</i>	<i>al-d-iñ</i>
	we-GEN	soul-POSS.1PL-ACC	save-PST-2SG

'Although you are a fool/Being a fool, (but) you have saved our souls.'

3.3.2. The form *-ArGA*

The Tuvan form *-ArGA* has also primarily temporal functions, but it can express concession contextually. It is the aorist-future participle in the Dative case form, but it functions as a converb. The concessive situation is always factive. This form can serve as a periphrastic equivalent of the conditional form *-SA* with the particle *DA* in concessive proper constructions. Thus, (13) can be transformed into (32) without any semantic losses.

- (32) *Petrov qim-dan-da aytir-ar-ga, qim-daa aḡaa*
 P. who-ABL-PTL ask-FUT.PART-LOC who-PTL he:DAT
χariñ ber-ip šida-vaan.
 answer give-CONV be.able-PREF.NEG
 ‘No matter whom Petrov asked [about it] (Although/When Petrov asked everyone), nobody could give him an answer.’

3.3.3. The converb $-(X)p$

This converb is of the contextual type (Nevskaja 1993), i.e. its semantics is always determined by the lexical or structural factors.

Shor

- (33) *Ol čiiš-ti as či-p, ebire köp qaraqta-pča.*
 he meal-ACC little eat-CONV around much look-PRS
 ‘Although he eats little (eating little), he is looking around a lot.’

3.3.4. The negative converb

The negative counterpart of the converb $-(X)p$ tends to express causal semantics of different types even more often than the positive converb.

Shor

- (34) *Paliq qoştan-maan tabıraq, qaranyı qoştan-d-ı*
 fish approach-NEG.CONV fast gradually approach-PST-3SG
 ‘Although the fish did not approach fast (Not approaching fast), it approached gradually.’

4. Adversative constructions

Such constructions are bi-finite ones, with adversative conjunctions as their structural markers. Here, two clauses – one representing a concessive situation and another an anti-consequential one – are joined by adversative conjunctions or modal phrases in the process of grammaticalising to become adversative conjunctions. In Siberian Turkic, we find only a few adversative conjunctions of Turkic origin. In some cases we may suppose that their adversative functions are structural copies of the corresponding Russian conjunctions. Siberian Turkic also has some conjunctions copied directly from Russian.

4.1. Constructions with Turkic conjunctions

One of the most striking examples of structural copying of adversative functions from Russian is the Altay conjunction *d'e*. This is also an interjection expressing consent: ‘yes!’ and a particle having adversative functions expressing the meaning ‘however’ (among other functions). We suppose that this element has developed the adversative functions following the Russian word *da*, which is also an interjection of consent, an adversative particle (‘however’, ‘nevertheless’) and an adversative

conjunction in Russian. In fact, both in Russian and in Altay, it is often very difficult to draw a line between the particle and the adversative conjunction.

Altay

(35) *Oni kem de d'araš emes de-p ayt-pas,*
 he:ACC who PTL handsome no say-CONV say-NEG.FUT
d'e onıñ ċıray-ı kem-ge de
 however he:GEN face-POSS3SG who-DAT PTL
d'ara-bayt.
 be.liked-NEG.PRS
 '(Although) nobody would say that he is not handsome, but nobody likes his face.'

The Shor conjunctions *añtebe* has developed from a postpositional phrase *aay tebe* [DEM3:GEN towards]. It renders a very specific adversative meaning of compensation 'instead of the situation A (concession), the situation B (anti-consequence) has happened'.

Shor

(36) *Čulat paž-ın-da pir da*
 river upper.reaches-POSS3-LOC one PTL
paliq tart-pa-d-i-lar, añtebe čulat
 fish catch-NEG-PST-3-PL instead river
pel-tir-i-neñ küsküş-ter tart-t-ı-lar.
 lower.reaches-POSS3-ABL kind.of.fish-PL catch-PST-3-PL
 '(Although) they have not caught a single fish in the upper reaches of the river, but they have caught much fish in the lower reaches of the river instead.'

4.2. Adversative conjunctions borrowed from Russian

The conjunction *no* 'but' serves as an example of a conjunction copied from Russian. It is widely used in Shor and in oral varieties of other South Siberian languages (Nevskaja 1999, Nevskaja 2000). Language purists usually proscribe its use in the written varieties of the well-established literary South Siberian Turkic languages like Altay or Khakas, but this is not the case for a young literary language like Shor where we find this conjunction also in published literary works.

Shor

(37) *Sın-ın-a ċet par-ıp ebire kör-d-i,*
 trunk-POSS3-DAT reach go:AUX-CONV around look-PST-3
no pir da torum körün-meen-ča.
 but one PTL cone appear-NEG-PRS
 'He has climbed (the tree) and looked around, but no cones are seen.'

5. Ways of expressing concessive-adversative relations on the textual level

5.1. Adversative constructions with introductory modal phrases

A great diversity of introductory modal phrases with adversative meaning serves as a reservoir for developing adversative conjunctions. Structurally, they are dependent clauses formed according to the concessive patterns with the conditional form *-SA* and the concessive particle *DA*. They act at the textual level as structural elements and usually appear in the beginning of a sentence anaphorically referring to the previous sentence, which represents a concessive situation. They are normally punctuated like dependent clauses, i.e. they are separated from the rest of the sentence by a comma.

Shor

(38)	<i>Tegri</i>	<i>sooqtan</i>	<i>par-tir,</i>	<i>čay-arya</i>	<i>pelen</i>	<i>pol</i>
	sky	get.cold	go:AUX-IND	rain-INF	ready	be
	<i>par-tir.</i>	<i>Endig</i>	<i>da</i>	<i>pol-za,</i>	<i>Alexey</i>	<i>ayradrom-ya</i>
	go:AUX-IND	so	PTL	be-COND	A.	airport-DAT
	<i>par-arya,</i>	<i>te-p,</i>	<i>em-nej</i>	<i>šiq-t-i.</i>		
	go-INF	say-CONV	house-ABL	go.out-PST-3		

‘It got cold, and it was going to rain. However, Alexey left the house in order to go to the airport.’

The modal phrase can also appear without comma, which is evidence that it has been integrated into the intonational structure of the sentence:

Shor

(39)	<i>Služba</i>	<i>aar</i>	<i>iriz-i</i>	<i>čoq</i>	<i>kerek</i>	<i>čilep</i>	<i>pildir-d-i.</i>
	service	hard	joy-POSS3	without	matter	like	appear-PST-3
	<i>Endig</i>	<i>da</i>	<i>pol-za</i>	<i>talaš-čan</i>	<i>ebes</i>	<i>pol-yan!</i>	
	so	PTL	be-COND	retreat-IMP.F.PART	no	be-PRF	

‘My army service appeared to be a hard and joyless matter. However, (but) I could not retreat!’

Finally, appearing in the same complex structure together with the concessive situation and being integrated intonationally, this modal phrase can fulfil the function of an adversative conjunctive element:

Shor

(40)	<i>Ol</i>	<i>čaqšī</i>	<i>kiži</i>	<i>pol-yan</i>	<i>endig</i>
	he	good	person	be-PRF.PART	so
	<i>da</i>	<i>pol-za</i>	<i>už-i</i>		
	PTL	be-COND	end-POSS3		
	<i>paž-i</i>	<i>čoq,</i>	<i>salyin</i>	<i>uškaš</i>	<i>sayiš-tiy</i>
	head-POSS3	without	wind	like	thought-WITH

kiži *pol-γan.*
 person be-PRF
 ‘He was a good man, but a very light-headed and thoughtless one (lit.: without the beginning and the end, with thoughts like the wind).’

5.2. Asyndetical constructions

Concessive-adversative relations can be expressed on the textual level by mere juxtaposition of the sentences expressing the concessive and the adversative situations respectively:

Shor
 (41) *Apšiy* *sīraŋay* *tün-ma* *sal-īp*
 old.man totally breath-INST put-CONV
 ertiš-t-i.
 drink-PST-3
 Apšiy-dīŋ *erbekte-rge* *aqs-ī* *emen*
 old.man-GEN speak-INF mouth-POSS3 nice
 pol-d-u.
 become-PST-3
 Ayd-arya *söz-ün* *tabin-mas* *pol-d-u.*
 say-INF word-POSS3ACC find-PART.AOR.NEG become-PST-3
 ‘The old man swallowed [the alcohol] in a gulp. He felt like talking (lit.: his mouth became nice to speak). [However,] he could not find what to say (Lit.: words to say).’

Abbreviations

AOR	Aorist	IND	Indirective
ACC	Accusative	INF	Infinitive
ABL	Ablative	INST	Instrumental
AUX	Auxiliary (element)	LOC	Locative
CONV	Converb	NEG	Negation
COND	Conditional	PART	Participle
DAT	Dative	PL	Plural
DEM	Demonstrative	POSS	Possessive
DIR	Directive	PRF	Perfect
FUT	Future	PST	Past
GEN	Genitive	PRS	Present
IMP	Imperative	PTL	Particle
IMPF	Imperfect	SG	Singular

Acknowledgements

I would like to express my gratitude to my colleagues Ajana Ozonova and Bajlak Orzhak, who helped me to collect and to evaluate the Altay and Tuvan material used in this paper.

Literature

- Anononymous. 1884. *Grammatika altajskogo jazyka, sostavlenaja členami altajskoj missii*. Kazan': Izdatel'stvo Kazanskogo universiteta.
- Baskakov, N. A. 1958. *Altajskij jazyk (vvedenie v izučenie altajskogo jazyka i ego dialektov)*. Moskva-Leningrad: Izdatel'stvo Akademii nauk SSSR.
- Baskakov, N. A. 1966. *Severnye dialekty altajskogo jazyka. Dialekt černevyx tatar (tyba-kiži)*. Moskva: Nauka.
- Baskakov, N. A. 1972. *Severnye dialekty altajskogo jazyka. Dialekt kumandincev (kumandykiži)*. Moskva: Nauka.
- Baskakov, N. A. (ed.). 1975. *Grammatika xakasskogo jazyka*. Moskva: Nauka.
- Baskakov, N. A. 1985. *Dialekt lebedinskix tatar-čalkancev (kuu-kiži)*. Moskva: Nauka.
- Bondarko, A. V. (ed.) 1996. *Teorija funkcional'noj grammatiki. Lokativnost'. Bytijnost'. Possessivnost'. Obyslovennost'*. Leningrad: Nauka. Leningradskoe otdelenie.
- Böhtlingk, O. 1851. *Über die Sprache der Jakuten*. Sankt Peterburg: Imperatorskaja Akademiya nauk.
- Čeremisina, M. I. (ed.). 1995. *Grammatika sovremennogo jakutskogo literaturnogo jazyka. Sintaxis*. Novosibirsk: Nauka.
- Čeremisina, M. I. & Skribnik, E. K. et al. 1986. *Strukturnye tipy sintetičeskix polipredikativnyx konstrukcij v jazykax raznyx sistem*. Novosibirsk: Nauka.
- Dyrenkova, N. D. 1941. *Grammatika šorskogo jazyka*. Moskva-Leningrad: Izdatel'stvo Akademii nauk SSSR.
- Erdal, M. 2004. *A grammar of Old Turkic*. Leiden & Boston: Brill.
- Haspelmath, M. & König, E. 1998. Concessive conditionals in the languages of Europe. In: Van der Auwera, J. (ed.). *Adverbial constructions in the languages of Europe*. Berlin: Mouton de Gruyter.
- Isxakov, F. G. & Pal'mbax, A. A. 1961. *Grammatika tuvinskogo jazyka*. Moskva: Izdatel'stvo vostočnoj literatury.
- Isxakova, X. F. & Nasilov, D. M. & Nevskaja, I. A. 2004. Chapter 15. Ustupitel'nye konstrukcii v tjurkskix jazykax. In: Xrakovskij, V. S. (ed.). *Tipologija ustupitel'nyx konstrukcij*. Sankt-Peterburg: Nauka. 425–452.
- Johanson, L. 1998. The structure of Turkic. In: Johanson, L. & Csató, É. (eds.) *The Turkic languages*. London: Routledge. 30-67.
- Johanson, L. 1992. *Strukturelle Faktoren in türkischen Sprachkontakten*. (Sitzungsberichte der Wissenschaftlichen Gesellschaft an der J. W. Goethe-Universität, Frankfurt am Main 29, 5.) Stuttgart.
- Nevskaja, I. A. 1993. *Formy deepričastnogo tipa v šorskom jazyke*. Novosibirsk: Izdatel'stvo NGU.
- Nevskaja, I. A. 1998. The revival of literary Shor. *Turkic Languages* 2: 253–270.
- Nevskaya, I. A. 2000. Shor-Russian contact features. In: Gilbers, D. & Nerbonne, J. & Schaeken, J. (eds.) *Languages in contact*. (Studies in Slavic and General Linguistics 28.) Amsterdam-Atlanta: Rodopi. 283–298.
- Radloff, W. 1866. *Proben der Volksliteratur der türkischen Stämme Süd-Sibiriens. I*. Sankt Peterburg: Imperatorskaja Akademiya nauk.
- Rassadin, V. I. 1978. *Morfologija tofalarskogo jazyka v sravnitel'nom osveščanii*. Moskva: Nauka.

- Ubrjatova, E. I. 1950. *Issledovanija po sintaxisu jakutskogo jazyka. Čast' I. Prostoje predloženie*. Moskva–Leningrad: Nauka.
- Ubrjatova, E. I. 1976. *Issledovanija po sintaxisu jakutskogo jazyka. Čast' II. Složnoe predloženie*. Moskva–Leningrad: Nauka.
- Ubrjatova, E. I. (ed.) 1982. *Grammatika sovremennogo jakutskogo literaturnogo jazyka*. Moskva: Nauka.
- Xrakovskij, V. S. 2000. Opyt analizy universal'nyx ustupitel'nyx konstrukcij. In: Bondarko, A. V. & Šubik, S. A. (eds.) *Problemy funkcional'noj grammatiki. Kategorii morfolologii i sintaxisa v vyskazyvanii*. Sankt Peterburg: Nauka.
- Xrakovskij, V. S. 2004a. Ustupitel'nye konstrukcii: semantika, sintaxis, tipologija. In: Xrakovskij, V. S. (ed.) *Tipologija ustupitel'nyx konstrukcij*. Sankt-Peterburg: Nauka. 9–91.
- Xrakovskij, V. S. 2004b. Anketa dlja opisanija ustupitel'nyx konstrukcij. In: Xrakovskij, V. S. (ed.) *Tipologija ustupitel'nyx konstrukcij*. Sankt-Peterburg: Nauka. 559–591.
- Xrakovskij, V. S. (ed.) 2004c. *Tipologija ustupitel'nyx konstrukcij*. Sankt-Peterburg: Nauka.

Between Cooperative and Plural: Kirghiz type “Cooperative suffixes” in modern literary Uyghur

Julian Rentzsch

Rentzsch, Julian 2005. Between Cooperative and Plural: Kirghiz type “Cooperative suffixes” in modern literary Uyghur. *Turkic Languages* 9, 252-261.

In Kirghiz, the Turkic Cooperative suffix regularly functions as a Plural morpheme in the third person. Similar uses are also found in neighboring languages, one of which is Uyghur. Although these usages are recorded in native publications, they are rarely mentioned in Western ones. This contribution aims at pointing out the range of uses of the Cooperative suffix in Standard Uyghur. It shows that this morpheme displays functions resembling those in Kirghiz, albeit in a less regularized way.

Julian Rentzsch, Seminar für Orientkunde, Johannes Gutenberg-Universität Mainz, DE-55099 Mainz, Germany. E-mail: rentzsch@mail.uni-mainz.de

1. Introduction

It is a well-known fact among Turcologists that the Cooperative Suffix (*V*)*š* regularly functions as a third person plural marker in the standard Kirghiz verb paradigms, though with approximately the same restrictions on obligatoriness as *IEr* in Turkish (e.g. Imart 1981: 803-807, 2196-2198). Western Turcology largely seems to assume tacitly that this phenomenon is restricted to Kirghiz (e.g. Johanson 1998: 43; Kirchner 1998: 349). Nonetheless, it is a recorded fact that the same pattern can be recognized in some variety or other of at least Kazakh (e.g. Begaliyev & Sawranbaev 1944: 102; *Qazaq tili encyklopediyası*: 121), Uzbek (e.g. Reshetov & Shoabdurahmonov 1978: 152-157; Muhamadjonov 1983: 116-118) and Uyghur (e.g. Kaydarov et al. 1966: 207-209), but this fact is rarely formulated in the literature, and we cannot always be sure from the data presented about the exact use of the items in question. Still, Omeljan Pritsak in his highly dialect-oriented description of modern Uyghur (1959) observes that “Im Neuuigurischen gebraucht man oft für den Plural des Verbums simplex den reziproken Stamm (vgl. das Kirgisische)” [In modern Uyghur, the reciprocal stem is often used for the plural of the simple verb (cf. Kirghiz)] (1959: 553). In standard Uzbek, plural-like uses of the Cooperative suffix do occur at least in certain types of texts, as the following examples¹ from the *Afandi latifalari* (1989) show:

¹ I use here the new Uzbek orthography for transcription.

- (1) *Xotin, tanish mullavachchalar kelishdi,*
 wife acquainted son of a mulla:P come:COOP.PAST.3
oshni katta qil.
 meal:ACC big make:IMP.2S
 ‘Dear wife, some acquainted students have arrived, prepare an ample meal!’ (p. 14)
- (2) *Afandidan odamlar: “Siz kattami, akangiz kattami?”*
 efendi:ABL man:P You big:Q brother:POSS.2s big:Q
deb so’rashdi.
 QUOT ask:COOP.PAST.3
 ‘The people asked the Efendi: Which one is elder, you or your brother?’ (p. 55)
- (3) *Dengizdagi baliqlar sasib qolmasin*
 sea:LOC.REL fish:P smell:CV ACTION:NEG.IMP.3
deb tuzlab qo’yishgan.
 QUOT salt:CV ACTION:COOP.POST.3
 ‘They salted the sea, so that the fishes in it might not smell.’ (p. 69)

This article does not deal with the situation in *dialects* of Central Asian Turkic. It is concerned with the functional distribution of the Cooperative suffix in modern written standard Uyghur and aims at showing that the “Kirghiz type” of plural marking observed by Pritsak for the dialects also diffuses into written standard Uyghur as an optional marker for plurality.

2. The “official” situation in standard Uyghur

As opposed to Uzbek, standard Uyghur lacks a third person plural marking device of the type **keldiler* ‘they came’. The form *keldi* is normally used for both third person singular and plural. As modern Uyghur more regularly than many other Turkic languages indicates the subject of a sentence overtly (cf. Uyg. *u keldi* vs. Turkish *?o geldi* ‘s/he came’), ambiguities rarely occur. Still, the distinctiveness gap between [\pm PLUR] for the third person in the verb paradigms leaves a blank for semantic extension of the original Cooperative suffix (uyg. *ömlük deriye*, literally ‘communion degree’). The Uyghurs themselves are usually quite conscious of some surplus semantics of this suffix in addition to simple cooperative meaning.² For example, the entry *ömlük deriye* in the large six-volume Uyghur dictionary UTIL defines this item as “a unit that indicates that a grammatical subject consisting of two or more human beings performs the main action directed against or rivaling one another, and that the main action is performed by a grammatical subject consisting of two or more human

² Even the grammatical term hints at this fact.

beings. As, e.g., ‘they went’, ‘they worked’” (UTIL, 5: 855).³ The latter part of the definition indeed comes very close to the notion ‘plural’, although the exact meaning of the examples given cannot be assessed due to the lack of situational context. Strikingly enough, ‘classical’ cooperative verbs like *riqabetleş-* ‘to compete’ and *sözleş-* ‘to talk’ are not even mentioned in the examples. We will now cast a look on how the suffix *(V)ş* is actually employed in modern Uyghur literary texts.

3. Towards the notion of plural: Examples of *(V)ş* in Uyghur

The specter of meanings covered by Uyghur *(V)ş* comprises a semantic continuum from reciprocal in its most restricted form to general plural.⁴ This continuum can be roughly divided into four sub-fields:

1. Restricted Cooperative (reciprocal): An action performed mutually.
2. Cooperative Proper: A coordinated action performed interdependently, but not necessarily mutually or reciprocally.
3. Cooperative Plural (Open Cooperativeness): An action performed jointly, but partly independently, not necessarily coordinated.
4. Genuine Plural: An action performed totally independently, no notion of cooperativeness at all.

1 and 2 are considered the common Turkic core meanings of the cooperative suffix *Iş*,⁵ whereas 3 and 4 are excentric in that they are not typical for large parts of the Turkic world and represent a diachronic extension of the core meanings. Naturally, the cooperative core meanings are fully covered by the Uyghur item as in all other Turkic languages.

Sub-meaning 1: Restricted Cooperativeness (Reciprocal)

The Restricted Cooperative sub-meaning frequently occurs in verbs like *riqabetleş-* ‘to compete’, *sözleş-* ‘to talk’, *muñdaş-* ‘to chat’, *uruş-* ‘to beat one another’, *öltürüş-* ‘to kill one another’, *öpüş-* ‘to kiss one another’, etc. Quite a few combinations like these can be considered lexicalized; still *(V)ş* is fully productive in this sub-meaning, and in futile contexts verbs can be freely reciprocalized with this suffix.

As the reciprocal meaning of *(V)ş* is common place in Turkic, one example for this use will suffice:

³ *İkkidin artuq ademdin terkip tapqan gramatik iginin esliy heriketni bir-birige qaritip yaki musabiqilişip elip bərişini ve esliy heriketnin ikkidin artuq ademdin terkip tapqan gramatik ige teripidin orunlinişini bildüridiyan derije. Mesilen, bərişti, işlişip berdi ge oxşaş.*

⁴ In certain lexemes, *Iş*—usually in the combined denominal verb suffix *IEş*—seems to reflect other meanings that do not require more than one participant. This combination frequently conveys a meaning of self-centred processual development, e.g. *jiddiyleş-* ‘to become earnest’, *xitaylaş-* ‘to Sinisize oneself’. A related but slightly different example is *yéqinlaş-* ‘to approach’. This use, which can be considered common Turkic, will not be dealt with in this essay.

⁵ Cf. e.g. Erdal 1991: 578-583.

- (4) *Avazıñiz biz bilen qalidu, siz*
 voice:POSS.2S we with remain:[-PAST].3S you
ketsıñizmu biz siz bilen
 go:COND.2S.EMP we you with
sözlişiverimiz.
 talk:COOP.ACTION:[-PAST].1P
 ‘Your voice will remain with us; even if you go we will talk to you.’ (Sabir, Qerzdar: 204)

Sub-meaning 2: Cooperative Proper

This shade of meaning, encountered frequently in Turkic, also occurs both in lexicalized combinations and productively, e.g. *qatnaş-* ‘to participate’, *jidelleş-* ‘to revolt’, *üğiniş-* ‘to learn collectively’, *oquş-* ‘to read/study collectively’. Although Cooperative Proper is a regular pattern in Turkic, the two examples given here represent a rather atypical use as they exceed the use encountered e.g. in Turkish:

- (5) *Qalğan gep-sözlerni keyin deyişermiz.*
 remain:VA speech-word:P.ACC later say:COOP.MOD.1P
 ‘We will discuss the rest later.’ (Asim, Yıylima insan: 60)
- (6) *Eger u yerge bərişni xalimisañ,*
 if DET place:DAT go:VN:ACC want:NEG.COND.2S
hazırqi ornuda işlevérisen,
 now:REL place:POSS.2S.LOC work:ACTION:[-PAST].2S
séniñ telipini keyin oylišimiz.
 you:GEN claim:POSS.2S.ACC later think:COOP.[-PAST].1P
 ‘If you do not want to go there, simply work at your present place; we will consider your claim later.’ (Asim, Yıylima insan: 365)

Attention has to be paid with respect to the verb *oylaş-*, which can carry two meanings:

1. Cooperative of *oyla-* ‘to think’ and 2. lexicalized verb *oylaş-* ‘to think’. Example 6 is thus a case of ambiguity.

Sub-meaning 3: Cooperative Plural (Open Cooperativeness)

The following examples represent cases where an action is performed together but at least partly independently and not clearly coordinated:

- (7) *[Harvikeşler]niñ beziliri nahayiti muñluq avazda*
 cart driver:P.GEN some:P.POSS.3 extremely mournful voice:LOC
naxşa éytşa beziliri çüje
 song sing:COND.3 some:P.POSS.3 chick

xorazniy *avazidek* *inçike* *avazda*
 cock:GEN voice:POSS.3.EQU shrill voice:LOC
naxša *ëytišatti.*
 song sing:COOP.INTRA^{LF}.PAST.3

‘While some of the cart drivers were singing in a very sad voice, others were singing in a shrill, chicken-like voice.’ (Asim, *Yiylima insan*: 32)

- (8) *Köpçilik* *sizniy* *mu'ellim* *bolyanliqijizni*
 majority you:GEN teacher be:VN.POSS.2S.ACC
hörmet qilišti.
 credit:COOP.3

‘The majority credited you with the fact that you are a teacher.’ (Sabir, *Qerzdar*: 79)

- (9) *Başqilarmu* *hezretke* *egişip* *başlirini*
 other:P.too excellency:DAT according to head:POSS.3P.ACC
sel-pel *ëgip* *hörmet* *bildürüšti.*
 slightly bow:CV esteem display:COOP.PAST.3

‘Also the others showed their respect by slightly bowing their heads according to his excellency.’ (Ilyas, *Ëyir tiniqlar*: 51)

Note that in this example the participants do not pay respect reciprocally, rather a group of people bows to one person of high esteem.

- (10) *Zakirmiş* *ayiniliri* *uni* *Zakir* *gilem* *dep*
 Zakir:GEN friend:P.POSS.3 PPR:ACC Zakir carpet QUOT
atişatti.
 name:COOP.INTRA^{LF}.PAST.3

‘Zakir’s friends called him Zakir Carpet.’ (Asim, *Yiylima insan*: 1)

- (11) *Keçte* *işçılar* *bazarya* *yişilip* *tamaq*
 night:LOC worker:P market:DAT assemble:CV food
yëyişti.
 eat:COOP.PAST

Tamaqtin *këyin* *işçılar* *özliriniy*
 meal:ABL after worker:P RFL:P.POSS.3.GEN
aram *alidiyan* *orniya* *këtişti.*
 rest take:VA place:POSS.3.DAT go:COOP.PAST

‘At night, the workers assembled in the market and had their meal. After the meal, they went to their resting places.’ (Asim, *Yiylima insan*: 61)

While the first action (*yëyişti*) could well be interpreted as a cooperative action of type 2 (not as type 1 Restricted Cooperativeness though, as the workers do not eat one another), action 2 (*këtişti*) clearly is an example of Open Cooperativeness, as everybody goes to his individual resting place. This action is thus performed individually, though not totally independently.

- (12) *Ular qaqaqlıšıp külüšti.*
 PPR:P make ha ha:COOP.CV laugh:COOP.PAST.3
 ‘They laughed heartily.’ (Sabir, Qerzdar: 176)

The persons do not laugh at one another, but they laugh together.

- (13) *Ular [...] hemrahi bilen muñdaşqač çaykılaryya*
 PPR:P companion:POSS.3 with chat:CV gull:P.DAT
bolka, mëve čėčip, çaykılarnıñ ozuq üçün
 roll fruit share:CV gull:P.GEN fodder for
qilyan küreš’lirini tamaşa qılıšidiken.
 make:VA fight:P.POSS.3.ACC watch:COOP.[-PAST].3.IND
 ‘They chat with their companions, feed the gulls on rolls and fruit, and watch the gulls’ fighting for the feed.’ (Sabir, Qerzdar: 183)
- (14) *Ular mėniñ tonušturuşumni aňlap tolimu xoşal*
 PPR:P I:GEN introduction:POSS.1S.ACC hear:CV very glad
bolušti.
 become:COOP.PAST.3
 ‘When I introduced myself, they became very happy.’ (Sabir, Qerzdar: 205)
- (15) *U bu geplerni qilip qaqaqlap*
 PPR DET speech:P.ACC make:CV make ha ha:CV
küldi, bizmu külüштуq.
 laugh:PAST.3 we:too laugh:COOP.PAST.1P
 ‘Speaking like that he laughed heartily, and we laughed, too.’ (Sabir, Qerzdar: 216)

Sub-meaning 4: Genuine Plural

In these examples the action is performed by more than one person, and clearly totally independently, possibly even not synchronically:

- (16) *Pakar ėdirliq qaptalliriya jaylaşqan bu*
 low hill:DNN slope:P.POSS.3.DAT be situated:VA DET
şehir yolliri qişliq pelto, ötük,
 city road:P.POSS.3 winter:DNN coat boot
qulaqça kiyişken ademliri bilen bizge
 fur hat wear:COOP.VA man:P.POSS.3 with we:DAT
yėyi tuyuldi.
 new feel:PASS.PAST.3
 ‘These city roads on the slopes of low hills with its people who had put on winter coats, boots and fur hats felt new for us.’ (Sabir, Qerzdar: 131)

The action of putting on clothes is performed totally independently and without any temporal coincidence.

- (17) *Yillar ötti, şəherdin kömürge çıqqan*
 year:P pass:PAST.3 town:ABL coal:DAT come out:VA
harvikeşler yəni hökümet qurulyanliqi,
 cart driver:P new government establish:PASS.VN.POSS.3
kompartiye heqqide yəni xeverlerni sözlep
 communist party about new news:P.ACC tell:CV
yürüşti.
 march:COOP.PAST.3
 ‘Years passed, and the cart drivers coming from the city for coal brought fresh news about the formation of a new government and the communist party.’
 (Asim, Yiylima insan: 62)
- (18) *Qışliq kiyim kiygen ademler aldirişip*
 winter:DNN clothes put on:VA man:P hurry:COOP.CV
yürüşmekte.
 march:COOP.INTRA^{HF}.3
 ‘People wearing winter clothes were running in haste.’ (Sabir, Qerzdar: 132)
- (19) *Yol çétidiki bendiñlerde olturuşqan*
 way outside:POSS.3.LOC.REL bench:P.LOC sit:COOP.VA
adamlarni tamşa qilip mañmaqtimen.
 man:P.ACC watch:CV go:INTRA^{HF}.1S
 ‘I am traveling ahead, watching people sitting on benches on the roadside.’
 (Sabir, Qerzdar: 138)
- (20) *Napolé'on ve Gitler armiyisi [...] muşu*
 Napoleon and Hitler army:POSS.3 DET
yerlerde özliriniñ miñliyan
 place:P.LOC RFL:P.POSS.3.GEN thousand:DNV.VA
jesetlirini qaldurup, haryın, alaqqade,
 corpse:P.POSS.3.ACC bury:CV fatigued frightened
vehime içide öz xoşayinlirini
 fear inside:POSS.3.LOC RFL leader:P.POSS.3.ACC
qarışip yerbke qarap
 curse:COOP.CV west:DAT look:CV
qéçişqan.
 flee:COOP.POST.3
 ‘Napoleon’s and Hitler’s armies buried their thousands of corpses right here and fled fatigued, in fright and fear to the west, cursing their leaders.’
 (Sabir, Qerzdar: 144)

Here, two totally independent actions are referred to.

- (21) *Ular* *mèni* *öylirige* *teklip qilišti.*
 PPR:P I:ACC house:P.POSS.3.DAT invite:COOP.PAST.3
 ‘They invited me to their homes.’ (Sabir, Qerzdar: 153)

4. Conclusion and prospects

Standard Uyghur, like many Central Asian Turkic varieties, lacks a genuine third person plural marker in the verb conjugation paradigm. To fill the morphological gap, the semantics of the original cooperative suffix has been considerably extended, such that *(V)š* synchronically covers the whole semantic field from the most restricted cooperativeness as represented by the Reciprocal to the most general, individual plural.

The Cooperative shows a high affinity towards plurality by nature, as cooperative actions are rarely performed by a single actor. In search of a way to mark verbs for plurality in a certain language, broadening the semantic specter of the Cooperative suffix is quite a logical choice. As the Cooperative is actually just a special case of the notion “plural”, we need not be astonished at this type of semantic extension. It should be noted, though, that the function of the Turkic Plural is not just to designate plurality but also to individualize (cf. Johanson 1991). Cooperativeness being a rather homogeneous notion—and thus quite contrary to individuality—, the development from Cooperative to Plural is therefore a remarkably large step. *(V)š* items of the Kirghiz type therefore cover a semantic field larger than superficial consideration might suggest.

As example 15 shows, this use of *(V)š* is not restricted to the third person, so that Pritsak’s observation for the dialects that the use of *(V)š* as a plural marker partly extends to the first and second person plural can also be applied to the written standard language. Combinations with the first and second person are much less frequent, though. This may be due to the economic imperative to avoid redundancy.

As examples 9, 16, 18, 19 show, *(V)š* can also pluralize non-final verbs, which is in many cases impossible with the other verbal plural marker *IEr*. *(V)š* therefore offers a highly flexible applicability. The distribution of the Turkish Plural, for example, appears to be more restricted.

From examples 7, 13, 14, 20, on the other hand, it becomes obvious that the scope of *(V)š*- plurals may extend to non-final predicates, as is the case with *IEr*- plurals, too.

As shown so far, the meaning of *(V)š* in Uyghur represents a continuum between Restricted Cooperativeness (reciprocal) and Genuine Plural. It is a continuum of decreasing strictness with respect to cooperativeness and comprises both more homogeneous and more individual notions. The unit does not have several meanings: All shades of meaning are derivable from one basic meaning (*Grundbedeutung*). The Open Cooperative and Plural sub-meanings represent a semantic and functional extension derived from the Cooperative core meaning.

Clearly, this article raises more questions than it answers. I am not able to present a frequency analysis here about the presence or absence of *(V)š* in third person plu-

rals, nor am I able to comment on whether the distribution of (*V*)*š*-plurals varies according to the specific dialectal background of the writers. From my reading I get the impression that the use for third person plural is very common, whereas it is quite exceptional in the other persons. But what exactly is the distribution of (*V*)*š*-plurals? What characterizes the combinability with different TAM-markers? In how far does the functional extension of (*V*)*š* provoke disambiguation strategies for the Cooperative Proper domain? I am also not totally sure yet whether the (*V*)*š*-plural is restricted to human beings. Most desirable would be research on the areal distribution of (*V*)*š*-plurals in the Central Asian Turkic dialects in order to establish isoglosses. It is obvious that useful results can only be achieved through extensive fieldwork with rather substantial text corpuses.

What I hope to have shown is that Kirghiz-style plurals are well represented in Uyghur, too, although not necessarily recognized as such by Uyghur and non-Uyghur grammarians. Taking into consideration that in Kazakh and Uzbek dialects—at least partly even in standard Uzbek—similar tendencies can be found, we conclude that it is a widely spread plural marking type in Central Asian Turkic, which accidentally has only become fully standardized in Kirghiz.

Glosses

ABL	ablative	LOC	locative
ACC	accusative	MOD	modal unit
ACTION	actionality operator	NEG	negation
COOP	cooperative	NP	noun proper
CV	converb	P	plural
DAT	dative	PASS	passive
DEM	demonstrative pronoun	PAST	past
DET	determinator	POSS	possessive
DNN	denominal noun	POST	postterminal
DNV	denominal verb	PPR	third person personal pronoun
DVN	deverbal noun	Q	question
DVV	deverbal verb	QUOT	quotation particle
EMP	emphasis	REL	relational particle
EQU	equative	RFL	reflexive pronoun
GEN	genitive	S	singular
HF	high focal	VA	verbal adjective, participle
IMP	imperative	VN	verbal noun
IND	indirective	1	first person
INTRA	intraterminal	2	second person
LF	low focal	3	third person

References

- Afandi latifalari* 1989. Toshkent: Ghafur Ghulom nomidagi adabiyot v san'at nashriyoti.
 Asim, Feyret 2000. *Yiylima insan*. Ürümçi: Šinjañ xelq nešriyati.

- Begaliyev, G. & Sawranbaev, N. 1944. *Qazaq tiliniñ grammatiykası*. Almatı: Qazaqtıñ memleketlik birikken baspası.
- Deny, Jean & Grønbech, Kaare & Scheel, Helmuth & Togan, Zeki Velidi (eds.) 1959. *Philologiae turcicae fundamenta* 1. Aquis Mattiacis: Steiner.
- Erdal, Marcel 1991. *Old Turkic word formation. A functional approach to the lexicon*. Turcologica 7. Wiesbaden: Harrassowitz.
- Friederich, Michael 2002. *Uyghurisch Lehrbuch*. In Zusammenarbeit mit Abdurishid Yakup. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Hahn, Reinhard F. 1991. *Spoken Uyghur*. Seattle, London: University of Washington Press.
- Hahn, Reinhard F. 1998. Uyghur. In: Johanson & Csató (eds.) 1998, 379-396.
- Ilyas, Yüsüp 1998². *Èyir tiniqlar*. Ürümçi: Şinjañ xelq neşriyati.
- Imart, Guy 1981. *Le Kirghiz (Turk d'Asie Centrale Soviétique). Description d'une langue de littérisation récente. Avec une étude sur: Le dialecte kirghiz du Pamir afghan par Rémy Dor*. Aix-en-Provence: Publications de l'Université de Provence.
- Johanson, Lars 1991². Bestimmtheit und Mitteilungsperspektive im türkischen Satz. In: Johanson, Lars 1991. *Linguistische Beiträge zur Gesamturkologie*. Bibliotheca Orientalis Hungarica 37, Budapest: Akadémiai Kiadó, S. 225-242. First published *Zeitschrift der Deutschen Morgenländischen Gesellschaft*, Suppl. III: 2, 1977, 1186-1203.
- Johanson, Lars 1998. The structure of Turkic. In: Johanson & Csató (eds.) 1998, 30-66.
- Johanson, Lars & Csató, Éva Á. (eds.) 1998. *The Turkic languages*. London & New York: Routledge.
- Kaydarov, A. T. & Talipov, T. T. & Cunvazov, Yü. & Isma'ilov, I. 1966. *Hazırqı zaman uyğur tili* 2. Almuta: Qazaq SSR Nauka neşriyati.
- Kirchner, Mark 1998. Kirghiz. In: Johanson & Csató (eds.) 1998, 344-356.
- Muhamadjonov, Qasim 1983. *Janubiy Qozoghistondagi O'zbek shevalari morfologiyasi*. Toshkent: O'zbekiston SSR Fan nashriyoti.
- Pritsak, Omeljan 1959. Das Neuuigurische. In: Deny, Jean & Grønbech, Kaare & Scheel, Helmuth & Togan, Zeki Velidi (eds.) 1959, 525-563.
- Qazaq tili encyklopediyası*. Qazaqstan damuw iynstiytuwıfı. Almatı: IDK-TIPO redakciyalıq-baspa ortalıyı.
- Reshetov, V.V. & Shoabdurahmonov, Sh. 1978. *O'zbek dialektologiyasi*. Toshkent: O'qituvchi nashriyoti.
- Sabir, Zordun 1999. *Qerzdar*. Ürümçi: Şinjañ yaşlar – ösmürler neşriyati.
- Tömür, Xemit 1987. *Hazırqı zaman uyğur tili gramatikisi (Morfologiyası)*. Beijing: Milletler neşriyati.
- Tömür, Hämit 2003. *Modern Uyghur Grammar (Morphology)*. Translated by Anne Lee. Dil ve Edebiyat Dizisi 3. İstanbul: Yıldız. (Übersetzung von Tömür, Xemit 1987.)
- UTIL: *Uyğur tiliniñ izahlıq luyıti* 1990-1999. Şinjañ uyğur aptonom rayonluq milletler til-yéziq xizmiti komitétı luyet bölümi tüzgen. (6 vols.). Beijing: Milletler neşriyati.

First language dominance and language shift of Turkish youngsters in Western Europe

Kutlay Yağmur

Yağmur, Kutlay 2005. First language dominance and language shift of Turkish youngsters in Western Europe. *Turkic Languages* 9, 262-280.

In this paper, the language use, choice, dominance and preference of Turkish youngsters in five multicultural Western European cities are presented. Reported language proficiency of Turkish students between the ages of 6-11 is documented for Göteborg, Hamburg, The Hague, Brussels and Lyon. Also, the informants' language choice, language dominance, and language preference are presented cumulatively. Derived from this database, a (pseudo)-longitudinal profile and a language vitality index for Turkish youngsters of different age groups and, ultimately, of different generations will be construed. The data presented here is derived from the Multilingual Cities Project, which has been carried out among primary school students in six large European cities.

Kutlay Yağmur, Babylon, Center for Studies of Multilingualism in The Multicultural Society, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands.

Introduction

Language maintenance and shift have been researched in many different contexts. There are a number of models developed for the investigation of language maintenance and shift as documented by Clyne (1991). Those developed by Kloss (1966), Giles et al. (1977), Smolicz (1981), and Bourdieu (1982) identify different factors that are important in language maintenance (or shift). The factors involved are generally divided into two categories: those affecting a speech community and those affecting individuals within a speech community (Kipp, Clyne, & Pauwels 1995). Group factors include size and distribution of an ethnic group, the policy of the host community towards minority languages, the position of the language within the cultural value system of the group, and proximity or distance of the minority language to or from majority language, while birthplace, age, period of residence, gender, education/qualifications, marriage patterns, prior knowledge of majority language, reason for migration, and language variety are considered to be individual factors (Kipp et al. 1995: 123). However, it is not always easy to draw the line between individual and societal factors, as there is an ongoing interaction between an individual and the speech community. In most cases, these factors are interrelated both on the individual and on the group level. In language contact situations one's native language is not a fixed and stable system but rather a changeable one.

Possible relationships between ethnolinguistic vitality perceptions and language maintenance/shift have been investigated extensively. In order to provide empirical evidence into language shift and maintenance, in this study, a vitality index for Turkish groups in five cities is calculated. Rather than employing the original subjective ethnolinguistic vitality questionnaire (Bourhis et al. 1981), a language vitality index on the basis of language proficiency, choice, dominance and preference is developed (for details see Extra & Yağmur 2002). As our database is rich in scope and number of informants, results obtained are highly representative for the given cities.

Language shift or bilingualism

Sociolinguists in a variety of national contexts have investigated some aspects of first and second language use of Turkish speakers in a second language environment (Akıncı 1999; Aarsen 1996; Aarts 1994; Backus 1996; Boeschoten 1990; Johanson 1993; Kurtböke 1998; Necef 1996; Pfaff 1991, 1994; Schaufeli 1991; Türker 2000; Verhoeven 1987; Yağmur 1997; Yağmur & Akıncı 2003). Some of these studies concentrated on synchronic variation and diachronic language change. On the basis of lexical borrowing, semantic transfer, and codeswitching data, researchers claimed language change and, ultimately, language attrition. For instance, interpreting the *utterance* by a single informant presented in Türker (2000: 172):

- (1) *bisiklet* *sürdük*
bicycle drive-PST-1PL

instead of

- (2) *bisiklet-e* *bindik*
bicycle-DAT mount-PST-1PL

Myers-Scotton (2002: 199-200) came to the conclusion that (1) is triggered by the lexical-conceptual structure of the Norwegian term. Reflecting on Türker's interpretation, Myers-Scotton suggests that the informant, being influenced by the Norwegian concept 'to drive' a bicycle, has used the verb for 'drive' (*sür-*) instead of the expected verb (*bin-*). Furthermore, it is suggested that the omission of dative case (-e) on *bisiklet* indicates convergence at the level of morphological realization patterns. Considering the standard Turkish form (*binmek*) the first interpretation might seem relevant; however, given the fact that in most central Anatolian towns, the verb (*bisiklet sürmek*) is the most common form, the above interpretation might not be plausible. The divergence from standard Turkish should not be interpreted as semantic triggering. In the same vein, Myers-Scotton (2002), on the basis of Türker's interpretation, suggests that the Turkish dative case (-e) on *bisiklet* is omitted. As a matter of fact, the word *bisiklet* in '*bisiklet sürdük*' does not take the dative case.

The findings of codeswitching and language acquisition studies provided valuable evidence on the development of Turkish and the changes it is undergoing in the

immigration context. Single or multiple case studies enable researchers to gain deeper insight into the language acquisition and also the language shift process. However, we need complementary data derived from large-scale studies to avoid broad generalizations formed on the basis of data derived from a limited number of informants.

The present linguistic situation in the immigration context has been shown to be a transitional one, with processes of language loss and shift in the second and third generation (Johanson 1993). Nevertheless, investigation of language loss and shift requires a different methodology than codeswitching studies. Claiming Turkish language loss might be problematic if it is based solely on codeswitching and code copying data derived from very few informants. Whether such copying is an indication of language change or attrition is not yet definite. In the same vein, whether these copies are made because the speakers cannot access these words in their mental lexicon is not certain. It might simply be the case that these lexical items from the second language are more active in the speakers' mental lexicon and the speakers just insert these L1 items in their L2 or L2 words in their L1. In any case, it is not certain if these bilingual Turkish speakers would speak in the same way to monolingual Turkish speakers. Without having such data, it would be premature to suggest language attrition on the basis of code-mixing data. (For a comprehensive treatment of the methodological issues concerning language loss and shift, see Yağmur 2004.)

Another problem is the issue of *language change*. If we take the Turkish case, it is easy to see that most of the Turkish speakers immigrated to other countries around the 1960s and what they mostly maintain is the Turkish spoken in the 60s. Given the varying distances between Turkey and the countries of immigration, the amount of contact with the homeland varies. For instance, Turkish immigrants in Western Europe have more contacts with the homeland, which brings their language use closer to Turkey, but Turkish speakers in Australia have limited contact with Turkish spoken in Turkey. It is certainly possible to claim that in the meantime, in Turkey itself considerable change has taken place in the Turkish lexicon due to the purification movement of language reform. As a result of their frequent contacts with Turkey, Turkish immigrants in Western Europe are more up to date with such changes than the Turkish immigrants in Australia. Moreover, language change is a natural process for all living languages, and during that process, the change experienced in a monolingual context will definitely be different from a language contact situation, where two or more languages are in interaction. Therefore, when we measure the differences between the two contexts, what we claim as lost might simply be an artifact of the varying degree of change in two different contexts.

In general, in talking about first language attrition in a second language environment, it is mostly presumed that L1 attrition occurs because of the dominant and invasive role of the mainstream language. In an immigration context, the second language environment definitely limits the use of the first language. However, even though linguists pay considerable attention to language contact as a major factor in altering languages through interference and borrowing, the sociological setting in which language contact and bilingual behavior occurs has not received sufficient

study. In order to be able to find the effects of this domain limitation on first language proficiency, not only linguistic factors but also extralinguistic factors need to be taken into consideration.

The research reported in this paper will present evidence of generational language shift of Turkish youngsters. The extralinguistic factors causing language shift and loss are discussed in great detail elsewhere (See Yağmur forthcoming).

Present study

The findings reported in this study are derived from the Multilingual Cities Project (MCP). The aims of the MCP are to gather, analyze and compare multiple data on the status of immigrant minority languages at home and at school. The project is carried out in six cities, in which Germanic and/or Romance languages have a dominant status in private and public life. In alphabetical order, these cities are Brussels, Göteborg, Hamburg, Lyon, Madrid, and The Hague. In each city, representatives of the local educational authority and researchers from a local university took part in the study. Apart from the Scandinavian countries, there is no European tradition of collecting home language statistics on multicultural (school) population groups. Our method of carrying out home language surveys amongst primary school children in each of these cities partly derives from experiences abroad with nationwide or at least large-scale population surveys in which commonly single questions on home language use were asked. In contrast to such questionnaires, our survey is based on multiple rather than single home language questions and on cross-nationally equivalent questions.

Purpose

As a consequence of socio-economically or politically determined processes of migration, the traditional patterns of language variation across Western Europe have changed considerably over the past several decades (Extra & Verhoeven 1993, Extra & Gorter 2001). Industrialized Western European countries have a growing number of immigrant minority populations, which differ widely, both from a cultural and from a linguistic point of view, from the indigenous populations. Also the mainstream indigenous populations have different views and attitudes towards these new groups. In spite of more stringent immigration policies in most European Union countries, the prognosis is that non-indigenous populations will continue to grow as a consequence of the increasing number of political refugees, the opening of the internal European borders, and political and economic developments in Central and Eastern Europe and in other regions of the world. It has been estimated that in the year 2000 at least one third of the population under the age of 35 in urbanized Western Europe had an immigrant background.

Given the decreasing significance of nationality and birth country criteria in the European context, the combined criterion of self-categorization and home language use is a potentially promising long-term alternative for obtaining basic information

on the multicultural composition of European cities. The added value of home language statistics is that they can offer valuable insights into the distribution and vitality of home languages across cultures and can thus raise the awareness of multilingualism. Empirically collected data on home language use can also play a crucial role in the context of education. Such data will not only raise the awareness of multilingualism in multicultural schools; they are in fact indispensable tools for educational policies on the teaching of both the national majority language as a first or second language and the teaching of immigrant minority languages.

In sum, the rationale for collecting, analyzing and comparing multiple home language data on multicultural school populations derives from three different perspectives:

Taken from a *demographic* perspective, home language data can play a crucial role in the definition and identification of multicultural school populations;

Taken from a *sociolinguistic* perspective, home language data can offer insights into both the distribution and vitality of home languages across cultures, and can thus raise the public awareness of multilingualism;

Taken from an *educational* perspective, home language data are indispensable tools for educational planning and policies.

Finally, identification of multilingual populations in schools plays an important role in raising awareness about the multiculturalism in society and accordingly contributes to positive interaction between groups.

In the next sections, after presenting the research methodology (the design, data collection instruments and research population), relevant findings with regard to Turkish students, who took part in this research, will be presented.

Research Method

The data have been collected by means of a specially designed survey instrument for students in primary schools. In this section, we will go into the details of questions in the questionnaire and the processing details of the data.

Design of the survey instrument

The questionnaire for data collection was designed after an ample study and evaluation of language-related questions in nation-wide or large-scale population research in a variety of countries with a history of migration and minorization processes (see Broeder & Extra 1998). Table 1 gives an outline of the questionnaire. As far as the design of the questionnaire is concerned, a number of conditions need to be met. In the first place, the questionnaires are intended for young students, which is why the format and language need to be suitable for such young pupils. (Very young students filled out questionnaires either with the help of trained adults or with their teacher). Also, the students who speak a language other than or next to the mainstream languages in their homes should answer the screening question. As mentioned earlier, in developing the screening question, extensive research had been conducted

into similar types of home language surveys in multicultural contexts. On the basis of the experiences of other countries, a screening question with utmost coverage was developed.

Table 1. Outline of the MCP questionnaire

<i>Questions</i>	<i>Focus</i>
1-3	Personal information (name, age, gender)
4-8	School information (city, district, name, type, grade)
9-11	Birth country of the pupil, father and mother
12	Selective screening question (Are any other languages than X ever used in your home? If yes, complete all the questions; if no, continue with questions 18-20)
13-17	Language repertoire, language proficiency, language choice, language dominance, and language preference
18-20	Languages learnt at/outside school and demanded from school

The first 8 questions provide background information about the informant and the school. In accordance with the privacy legislation in most of the European cities, the *name* variable was either not included or not processed. The answers to questions 9-12 make it possible to compare the status of birth country data and home language data as demographic criteria. The countries and languages explicitly mentioned in questions 9-12 are determined on the basis of the most recent municipal statistics about immigrant children at primary schools; thus, the list of languages for, e.g., Hamburg is quite different from the one used in Madrid. The language profile, specified by questions 13-17, consists of the following five dimensions:

Language repertoire: the number and type of (co-) occurring home languages;

Language proficiency: the extent to which the pupil can understand / speak / read and write the home language;

Language choice: the extent to which the home language is commonly spoken with the mother, the father, younger and older brothers/sisters;

Language dominance: the extent to which the home language is spoken best;

Language preference: the extent to which the home language is preferred to be spoken.

Taken together, the four dimensions of language proficiency, choice, dominance and preference result in a language vitality index. On the basis of questions 18-20, a school language profile can be produced. This profile provides information about the language education in and outside school, as well as the need for instruction in a given language.

Informants

The number of informants for all MCP cities exceeds 160,000, but we will only present the findings on Turkish background students. In total, the data from 10,258 Turkish students, age 4-13, from five cities are presented. Table 2 includes the total population in participating cities and details of the study.

Table 2. Overview of the MCP database (* Flemish-medium schools only; ** *Réseau d'Éducation Prioritaire* only)

City	Total number of schools	Age range	Total number of schools in the survey	Total number of pupils in the schools	Total number of pupils in the survey
Brussels	117 *	6-12	110 *	11,500	10,300
Hamburg	231 public	6-10	218 public	54,900	46,000
	17 Catholic		14 Catholic		
Lyon	173 **	6-11	42 **	60,000	11,650
Madrid	708 public	5-12	133 public	202,000	30,000
	411 Catholic		21 Catholic	99,000	
The Hague	142 primary	4-12	109 primary	41,170	27,900
	30 secondary	12-17	26 secondary	17,000	13,700
Göteborg	170	6-12	122	36,100	21,300

In this large-scale project, only the data on the Turkish population from five of the cities were analyzed for this study. There are 659 students (6%) from Brussels, 454 students (4%) from Göteborg, 4996 students (49%) from Hamburg, 480 students (5%) from Lyon, and 3666 students (36%) from The Hague. There are only 3 students in Madrid, which is why Madrid is not taken separately in the construction of language profiles. The number of female students 4,859 (47%) was less than the number of male students 5,223 (51%), while 176 students (2%) did not specify any gender. In Table 3, the distribution of Turkish students across age groups is presented.

Table 3. The distribution of Turkish students across age groups

Age	Frequency	%
4	404	3.9
5	437	4.3
6	947	9.2
7	1765	17.2
8	1883	18.4
9	1855	18.1
10	1678	16.4
11	815	7.9
12	314	3.1
13	57	0.6
Missing	103	1.0
Total	10258	100.0

Results

As indicated earlier, data were collected among primary school students. However, only in The Hague, were data collected both in primary and secondary schools to gain a deeper understanding of the issues of language proficiency, choice, dominance, preference, and ultimately the extent of bilingualism. Firstly, the overall findings on primary school students in 5 cities (Brussels, Göteborg, Hamburg, The Hague, and Lyon) are presented. Next, language profiles for the whole population are given, and finally, the pseudo-longitudinal profiles on language dominance and preference of The Hague are presented so that we obtain a complete picture of the bilingual language competence of Turkish immigrant youngsters. Presenting all the findings is beyond the scope of this article; thus, only the most essential results are presented here (for a comprehensive treatment of the Turkish student population, see Yağmur, forthcoming; and for the whole project, see Extra & Yağmur 2004). Also, the findings for each city are published in local languages or in English. These publications provide rich insight into various aspects of multilingualism in the given cities. (See Akıncı & de Ruiter & Sanagustin 2004; Extra & Aarts & Avoird & Broeder & Yağmur 2001; Fürstenau & Gogolin & Yağmur 2003; Nygren-Junkin & Extra 2003; Verlot et al. 2003 for overall findings per city). In Table 4, on the basis of descriptive statistics, the global findings of the survey are presented. 10,258 students indicated that next to or other than the main-stream language Turkish is spoken in their homes but, as seen in Table 4, not all of these students can understand or speak the language.

Table 4. Cross tabulation on home language versus other dimensions in the survey, all (4-13) age groups (N = 10258)

Turkish spoken at home versus ...	N	Percent (%)
Understanding Turkish	9866	96.2
Speaking Turkish	9632	93.9
Reading Turkish	6000	58.5
Writing Turkish	5448	53.1
Speaking Turkish with the mother	8260	80.5
Speaking Turkish with the father	7626	74.3
Speaking Turkish with younger brothers/sisters	4254	41.5
Speaking Turkish with older brothers/sisters	3389	33
Speaking Turkish with the best friend	3626	35.3
Turkish as the best spoken language	4852	47.3
Turkish as the most preferred language	4822	47
Want to learn Turkish at school	2666	26
Receive Turkish instruction outside school	1624	15.8
Receive Turkish instruction at school	1163	11.3

In Table 5, birth countries of both the informant and the parents are presented (Ages 4-5 and 12-13 are not included in the table).

Table 5. Distribution of birth countries of Turkish pupils, mothers and fathers (Ages 6-11; N = 8942)

Birth Country	Student		Mother		Father	
	N	%	N	%	N	%
Germany	4167	47%	697	8%	524	6%
The Netherlands	1959	22%	182	2%	149	2%
Turkey	1360	15%	7132	80%	7358	82%
Belgium	519	6%	150	2%	96	1%
France	379	4%	55	1%	55	1%
Sweden	329	4%	46	1%	29	0%
Iraq	6	0%	9	0%	19	0%
Bulgaria	3	0%	9	0%	12	0%
Macedonia	6	0%	29	0%	31	0%
Morocco	4	0%	27	0%	17	0%
Russia	5	0%	6	0%	5	0%
Other countries	43	0%	158	2%	194	2%
Unknown	162	2%	449	5%	453	5%
Total	8942	100%	8942	100%	8942	100%

As seen from Table 5, most of the students are born in Germany (47%), followed by The Netherlands (22%). Comparatively, smaller numbers of the students are born in Turkey (15%). On the other hand, the majority of the mothers (80%) and fathers

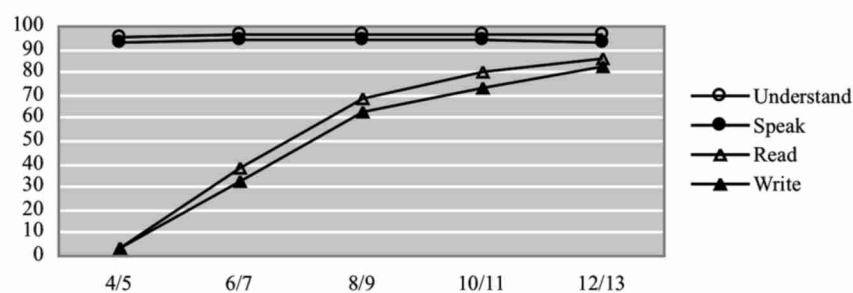
(82%) are born in Turkey. The birth country criteria show that most of the students are second-generation. On the basis of the home language repertoire question, it is possible to see which other languages are spoken next to Turkish in the domestic domain. The languages reported in addition to Turkish show that the linguistic situation in the immigration context is changing. Some of the students speak more than two languages and in some homes Arabic or Kurdish is spoken next to Turkish. From Table 6, it is easy to see that students coming from places where Turkish has been in contact with other languages report that they also speak Albanian, Russian, or Serbian/Croatian. A research project carried out in Hamburg among adolescents empirically demonstrated that the Turkish language is so important for communication that it is also used by adolescents of non-Turkish descent (Auer & Dirim, 2000, cited in Fürstenau et. al. 2003). In Table 6, languages other than Turkish spoken in the homes of Turkish students are reported.

Table 6. Languages other than Turkish used at home by the Turkish language group (Ages 6-11; N = 8942)

French	381	Russian	16	Chinese	7
Kurdish	375	Serbian/Croatian/Bosnian	14	Roman/Sinhi	7
English	133	Berber	13	Urdu	7
Arabic	96	Farsi	13	Aramese/Assyrian	6
Swedish	29	Greek	12	Dari/Pashto	6
Albanian	17	Portuguese	12	Hind(ustan)i	6
Italian	17	Polish	11	23 other languages	41
Spanish	17	Armenian	7		

In the following figures, language proficiency (Figure 1), choice (Figure 2), dominance (Figure 3), and preference (Figure 4) of the whole population are presented.

Figure 1. Turkish language proficiency



When we closely examine Figure 1, it is apparent that students, in all age groups, have very high speaking and understanding rates (over 90%), while reading and writing scores are comparatively lower. However, after age 10, writing and reading skills exceed 80%. In Figure 1, the importance of literacy skills acquisition in schools is quite apparent. If students do not receive instruction in Turkish, such high scores might not be obtained. In any case, given the history of migration, Turkish proficiency levels of second and third generation students are quite high.

In the literature on first language maintenance, availability of various domains for first language use has been shown to be vital for mother tongue maintenance. As documented by Yağmur (2004) and Yağmur & Akıncı (2003) Turkish immigrants in Western Europe have rich social networks in which they can use Turkish for their day-to-day communication in a number of domains. In a way, Turkish is not only limited to the domestic domain. As shown in Figure 2, in the home domain, Turkish is regularly used with the mother (77-87%) and with the father (70-85%). With younger and older siblings, Turkish language use is much lower (between 38% and 50%). Nevertheless, the informant might not have a younger or older brother/sister, in which case, the above findings should not be treated as absolute differences. Even so, as children get older, Turkish language use with brothers and sisters is also increasing. In terms of language use with the best friend, Figure 2 provides interesting information. First of all, it is a well-researched sociolinguistic phenomenon that second and third generation immigrants speak a mix of Turkish and mainstream language with one another. In spite of that, Turkish language use with the best friend shows that Turkish students mostly have Turkish-speaking friends, and the pattern of Turkish language use shows a strong increase as the children get older. As shown with The Hague data, when children are in the 10/11 age group, Turkish use with the best friend is 35%, but when they are 16/17, Turkish use is 65%, which shows a strong in-group orientation. The prevalent social and cultural attitudes in the receiving societies might have a certain role in the shaping of this in-group orientation.

Figure 2. Language choice for Turkish

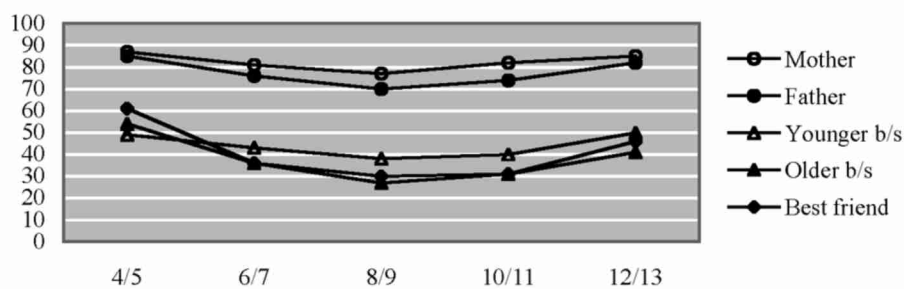


Figure 3 highlights another interesting language use phenomenon. When the language dominance pattern of Turkish students is examined, a rather intriguing picture emerges. The use of mainstream language and Turkish shows unexpected patterns. Because Turkish is the home language of many immigrant Turkish families, it is normal that at the age of 4/5, Turkish is highly dominant. After children start learning the mainstream language at school, a certain amount of decline in Turkish language skills is observed. Nevertheless, after age 12/13, students report a decline in their skills in the mainstream language and an increase in their Turkish skills. As it is, shown with the Dutch data below, Turkish skills are reported to be higher than the mainstream language around the ages 16/17, a circumstance which definitely requires further research.

Figure 3. Language dominance in Turkish and in the mainstream language

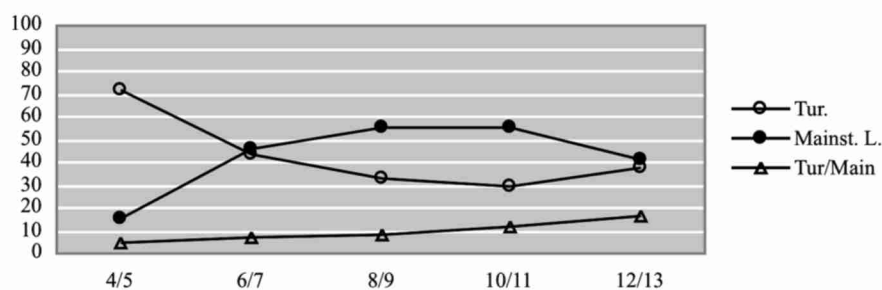
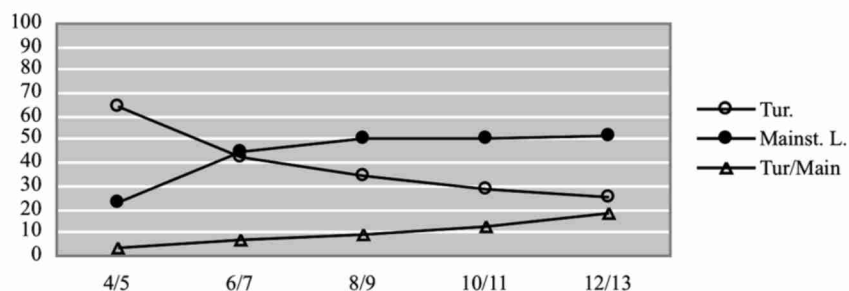


Figure 4. Language preference in Turkish and in the mainstream language



As seen in Figure 4, language preference, however, shows a different pattern than that of language dominance. Turkish youngsters are dominant in their first languages when they begin school and accordingly they prefer Turkish in communication. As Turkish children receive instruction mainly in the mainstream language, one would

expect them to be fully dominant in the mainstream language. Assuming a link between dominance and preference, students would be expected to have a higher preference for the mainstream language; accordingly, the emerging pattern in Figure 4 shows that students increasingly prefer the mainstream language to Turkish. However, as shown by the Dutch data, when children are 14/15 years old, their language preference begins to change.

The above figures (1-4) present the overall findings in the given age groups for the five cities involved in the study. Varying extralinguistic and demographic factors in the receiving societies might influence the language proficiency of Turkish youngsters. In all cities, understanding Turkish is reported to be above 88% for all age groups. Turkish speaking skills show a highly comparable pattern to that of comprehension skills. The reading skills in Turkish increase as children get older. We observe a highly congruent pattern in all the cities. When students are 10/11 years old, reading skills are 83% for Brussels, 75% for Göteborg and Hamburg, 73% for Lyon and 86% (highest percentage) for The Hague. Similar to reading skills, Turkish writing skills are congruent in the five cities but, as expected, the percentages observed are lower compared to reading skills. When students are 10/11 years old, writing skills are 77% for Brussels, 71% for Göteborg, 66% for Hamburg, 67% for Lyon and 82% (highest percentage) for The Hague.

Vitality of Turkish

Derived from the home language profiles for different cities, presented in the above section, a cross-city and pseudolongitudinal comparison of the four dimensions of language proficiency, language choice, language dominance, and language preference are made. For this calculation, these four dimensions have been operationalised as follows:

- Language proficiency:* the extent to which Turkish is understood by the informants;
- Language choice:* the extent to which Turkish is spoken with the mother;
- Language dominance:* the extent to which Turkish is spoken best;
- Language preference:* the extent to which this home language is preferred to be spoken.

The operationalisation of the first and second dimension (proficiency and choice respectively) is aimed at a maximal coverage. Language understanding is commonly the least demanding of the four language skills, and the general trend is that the mother acts as a major gatekeeper for intergenerational language transmission, which is also confirmed by the data presented above (see Figure 2 above). From the analyses on the basis of four language dimensions mentioned above, we ultimately construct a cumulative language vitality index (LVI) for Turkish youngsters in each city.

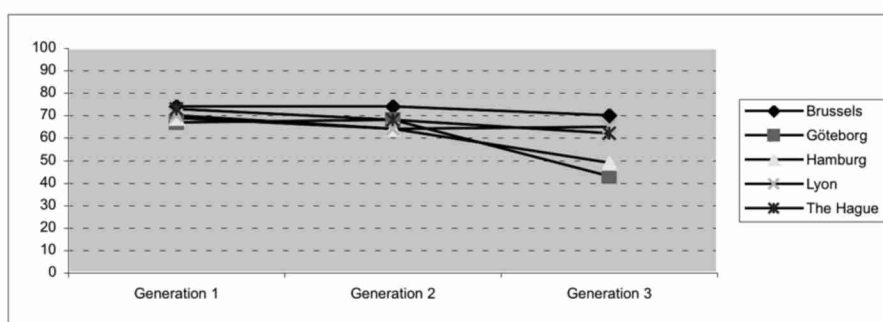
The LVI is based on the mean value of the presented scores for the four obtained language domains. This LVI is by definition an arbitrary index, in the sense that the chosen dimensions with the chosen operationalisations are equally weighted. Findings on the language vitality index (in cumulative %) for each city (Table 7) and language shift across generations (Figure 5) are presented. Generations are operationalised as follows: the informant is categorized as *first generation* if both the informant and the parents are born in Turkey; *second generation*, if the informant is born in the country of residence and the parents are born in Turkey; *third generation*, if both the informant and the parents are born in the country of residence. On the basis of this categorization, intergenerational language shift can be estimated. As mentioned earlier, because students start and finish primary school at varying ages in these five different cities, the following findings are based on the analyses on the common age groups: 6-11.

Table 7. Language vitality per age group and city (in cumulative %)

Ages	Brussels	Göteborg	Hamburg	Lyon	The Hague	Average
6/7	73	69	66	65	75	70
8/9	75	67	62	63	68	67
10/11	71	66	65	68	65	67
Average	73	67	64	65	69	68

On the basis of the findings reported in Table 7, it is seen that Turkish students in Brussels have the highest vitality measure, followed by The Hague. The lowest scores are observed in Hamburg and Lyon. Nonetheless, it is important to point out that of all the minority groups in the given cities, the Turkish language group has the highest vitality in Brussels, The Hague, and Lyon. In Hamburg Turkish occupies the fourth place in the vitality index; while in Göteborg, it holds the fifth place among the top-20 most spoken languages.

Figure 5. Language shift across generations

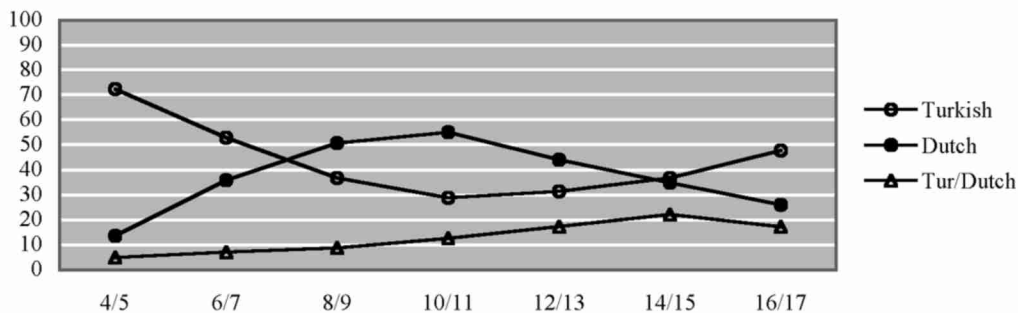


The decreasing language vitality index of Turkish across generations gives us the rate of shift to the dominant language. As seen in Figure 5, the highest intergenerational shift to the dominant language is observed in Göteborg and Hamburg. The least shift is seen in Brussels, followed by Lyon and The Hague. The high vitality of Turkish in Brussels seemed intriguing, for which reason, a number of parents have been contacted. In the French-dominant city of Brussels, mostly high SES (socio-economic status) Turkish parents send their children to Flemish-medium schools so that their children can learn Flemish in the best possible way. Apparently, practical purposes play a role in the school choice of Turkish parents. As the families mostly have high socio-economic status, their school choice is conscious; moreover, the children in such homes might be exposed to rich Turkish language use. Accordingly, such families might deliberately be supporting Turkish language maintenance. Interestingly, it is not only the high SES Turkish people but also the French elite who send their children to Flemish-medium schools. The fact that both Flemish and French are required for employment in government offices, plays a major role in parents' choice of Flemish-medium schools (see Verlot et. al. 2003).

Dominance versus Preference

Dominance in a language and preference for that language do not always go hand in hand. The evidence obtained in the Dutch context across different ages is highly interesting. As indicated earlier, only in the Dutch context was data collected both in primary and secondary schools (ages 4-17). The following two figures on language dominance and preference reveal highly intriguing results concerning linguistic competence versus attitudes.

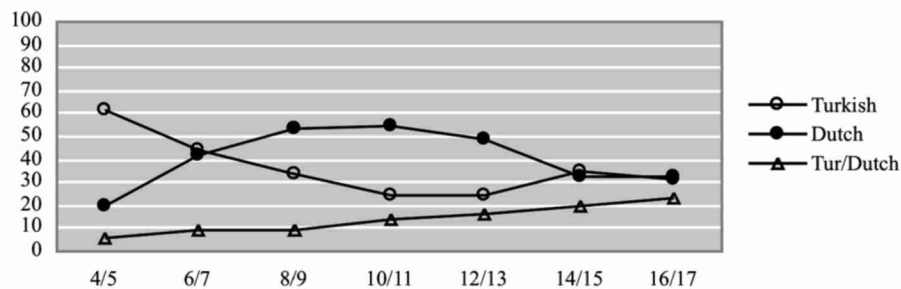
Figure 6. Language dominance in Turkish and Dutch



When the language dominance of Turkish students is examined, a rather intriguing pattern emerges. The use of Dutch and Turkish shows unexpected patterns. Because Turkish is the home language of many immigrant Turkish families, it is normal that

at the age of 4/5 Turkish is highly dominant. After children start learning Dutch at school, a certain amount of decline in Turkish language skills might be considered normal because children receive very limited hours of instruction in Turkish (1.5 hours per week). If there were bilingual programs in schools, proficiency levels in Turkish and Dutch might be similar. Until age 10/11, children become increasingly dominant in Dutch while dominance in Turkish declines. This can be explained as the result of the children's immersion in Dutch. However, what we cannot easily explain is the pattern after age 11: children start becoming less dominant in Dutch and more dominant in Turkish. At the age of 16/17, students report that they are more dominant in Turkish than Dutch. In spite of 12 years of Dutch-only schooling and in spite of all the societal Dutch input, Turkish students report that they are more dominant in Turkish than Dutch. Apparently, attitudinal rather than linguistic factors play a role in the above pattern (for details see Yağmur, forthcoming).

Figure 7. Language preference for Turkish and Dutch



When Figure 7 is examined, language preference shows a similar pattern to that of language dominance. Turkish children are dominant in their first language when they begin school and accordingly they prefer Turkish in communication. Like language dominance, Turkish youngsters' language preference changes over the years. Therefore, from age 14/15 onwards, Turkish speakers have a balanced preference for Turkish and Dutch. Also, their interchangeable use of both languages increases over the years.

Conclusions

The findings presented here provide overall patterns for language proficiency, choice, dominance, and preference of Turkish youngsters in five multicultural European cities. Because these results are derived from a large sample, the findings have high representation for the population. The findings in the Dutch context (pseudo-longitudinal perspective) clearly show that Turkish youngsters' language preference

and dominance change over time, and that there is not always a one-to-one correlation between dominance and preference.

Overall findings of the Multilingual Cities Project, and this study in particular, show that Turkish is one of the most vital immigrant languages in the European context. It is generally accepted that immigrant languages are used by the first generation of immigrants extensively, are less prevalent among the second generation, and are almost not to be found in the third generation. However, the findings of this study clearly show that Turkish is the home language even for third-generation Turkish immigrant children. First language proficiency among all children is considerably high. Due to local circumstances and the structure of the Turkish population, there are some differences between the countries with respect to language proficiency, choice, and preference.

Even though the findings are based on self-reported data, the emerging patterns for language shift are extremely interesting. On the basis of 'birth country' criteria, a declining vitality of Turkish across generations is observed. This decline is more pronounced in Göteborg and Hamburg; while, a more or less, stable situation is observed in Brussels, Lyon and The Hague. Such differences should be further investigated.

References

- Aarssen, J. 1996. *Relating events in two languages*. (Studies in Multilingualism 2.) Tilburg: Tilburg University Press.
- Aarts, R. 1994. *Functionele geletterdheid van Turkse kinderen in Turkije en in Nederland*. De Lier: Academisch Boeken Centrum.
- Akinci, M. A. & de Ruiter, J. J. & Sanagustin, F. 2004. *Plurilinguisme à Lyon: Le statut des langues à la maison et à l'école*. Paris: L'Harmattan.
- Akinci, M. A. 1999. *Développement des compétences narratives des enfants bilingues turc-français en France âgés de 5 à 10 ans*. [Ph.D. dissertation, Université Lumière Lyon 2, Lyon.]
- Backus, A. 1996. *Two in one: Bilingual speech of Turkish immigrants in the Netherlands*. (Studies in Multilingualism 1.) Tilburg: Tilburg University Press.
- Boeschoten, H. 1990. *Acquisition of Turkish by immigrant children. A multiple case study of Turkish children in the Netherlands aged 4 to 6*. (Turcologica 6.) Wiesbaden: Harrassowitz.
- Bourhis, R. Y. & Giles, H. & Rosenthal, D. 1981. Notes on the construction of a 'subjective vitality questionnaire' for ethnolinguistic groups. *Journal of Multilingual and Multicultural Development* 2, 145-155.
- Bourdieu, P. 1982. The economics of linguistic exchanges. *Social Science Information* 16, 645-668.
- Broeder, P. & Extra, G. 1998. *Language, ethnicity and education: Case studies on immigrant minority groups and immigrant minority languages*. Clevedon: Multilingual Matters.
- Clyne, M. 1991. *Community languages: The Australian experience*. Cambridge: Cambridge University Press.

- Extra, G. & Verhoeven, L. (eds.) 1993. *Immigrant languages in Europe*. Clevedon: Multilingual Matters.
- Extra, G. & Aarts, R. & Avoird, v. d. T. & Broeder, P. & Yağmur, K. 2001. *Meertaligheid in Den Haag: de status van allochtone talen thuis en op school*. Amsterdam: European Cultural Foundation.
- Extra, G. & Gorter, D. (eds.) 2001. *The other languages of Europe*. Clevedon: Multilingual Matters.
- Extra, G. & Yağmur, K. 2002. Immigrant minority languages in multicultural Europe. In: Stromqvist, S. (ed.) *The diversity of languages and language learning*. Lund: Lund University Press. 57-72.
- Extra, G. & Yağmur, K. (eds.) 2004. *Urban multilingualism in Europe: Immigrant minority languages at home and school*. Clevedon: Multilingual Matters.
- Fürstenau, S. & Gogolin, I. & Yağmur, K. (eds.) 2003. *Mehrsprachigkeit in Hamburg. Ergebnisse einer Sprachenerhebung an den Grundschulen*. Münster: Waxmann.
- Giles, H. & Bourhis, R. Y. & Taylor, D. M. 1977. Towards a theory of language in ethnic group relations. In: Giles, H. (ed.) *Language, ethnicity and intergroup relations*. London: Academic Press. 307-348.
- Johanson, L. 1993. Code-copying in immigrant Turkish. In: Extra, G. & Verhoeven, L. (eds.) *Immigrant languages in Europe*. Clevedon: Multilingual Matters. 197-221.
- Kipp, S. & Clyne, M. & Pauwels, A. 1995. *Immigration and Australia's language resources*. Canberra: Australian Government Publishing Service.
- Kloss, H. 1978 [1966]. German American language maintenance efforts. In: Fishman, J. (ed.) *Language loyalty in the United States*. New York: Arno Press. 206-252.
- Kurtböke, P. 1998. *A corpus-driven study of Turkish-English language contact in Australia*. [Ph. D. dissertation, Monash University, Melbourne.]
- Myers-Scotton, C. 2002. *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
- Necef, M. Ü. 1996. *Ethnic identity and language shift among young Turks in Denmark*. [Ph.D. dissertation, University of Odense, Odense].
- Nygren-Junkin, L. & Extra, G. 2003. *Multilingualism in Göteborg: The status of immigrant minority languages at home and at school*. Amsterdam: European Cultural Foundation.
- Pfaff, C. 1994. Early bilingual development of Turkish children in Berlin. In: Extra, G. & Verhoeven, L. (eds.) *The crosslinguistic study of bilingual development*. Amsterdam: Royal Netherlands Academy of Arts and Sciences. 75-97.
- Pfaff, C. 1991. Turkish in contact with German: Language maintenance and language loss among immigrant children in Berlin (West). *International Journal of the Sociology of Language* 80, 97-129.
- Schaufeli, A. 1991. *Turkish in an immigrant setting: A comparative study of the first language of monolingual and bilingual Turkish children*. [Ph.D. dissertation, Amsterdam University, Amsterdam.]
- Smolicz, J. J. 1981. Core values and cultural identity. *Ethnic and Racial Studies*, 4, 75-90.
- Türker, E. 2000. Turkish-Norwegian Codeswitching: Evidence for intermediate and second generation Turkish immigrants in Norway. [Ph. D. dissertation, University of Oslo, Oslo.]
- Verhoeven, L. 1987. *Ethnic minority children acquiring literacy*. Dordrecht: Foris Publications.
- Verlot, M., Delrue, K., Extra, G., & Yağmur, K. 2003. *Meertaligheid in Brussel: De status van allochtone talen thuis en op school*. Amsterdam: European Cultural Foundation.

- Yağmur, K. 1997. *First language attrition among Turkish speakers in Sydney*. (Studies in Multilingualism 7.) Tilburg: Tilburg University Press.
- Yağmur, K. 2004. Language maintenance patterns of Turkish immigrant communities in Australia, and Western Europe: The impact of majority attitudes on ethnolinguistic vitality perceptions. *International Journal of the Sociology of Language* 165, 121-142.
- Yağmur, K. 2004. Issues in finding the appropriate methodology in language attrition research. In: Schmid, M. & Köpcke, B. & Keijzer, M. & Weilemar, L. (eds.) *First Language Attrition: Interdisciplinary Perspectives on Methodological Issues*. Amsterdam & Philadelphia: John Benjamins. 133-164.
- Yağmur, K. & Akıncı, M. A. 2003. Language use, choice, maintenance and ethnolinguistic vitality of Turkish speakers in France: Intergenerational differences. *International Journal of the Sociology of Language* 164, 107-128.
- Yağmur, K. forthcoming. *Turkish abroad: Educational and sociolinguistic issues*. Amsterdam: Aksant.