

Werk

Titel: Articles

Ort: Wiesbaden

Jahr: 1997

PURL: https://resolver.sub.uni-goettingen.de/purl?666048797_0001 | LOG_0030

Kontakt/Contact

Digizeitschriften e.V.
SUB Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen

✉ info@digizeitschriften.de

Some remarks on the *-mİş* past in the Eastern Black Sea coast dialects

Bernt Brendemoen

Brendemoen, Bernt 1997. Some remarks on the *-mİş* past in the Eastern Black Sea coast dialects. *Turkic Languages* 1, 161-183.

In the Turkish dialects of Trabzon and Rize, the *-mİş* past tense (and also the inferential forms of the copula) are used very little, except by individual speakers and groups of speakers whose language is especially heavily influenced from Standard Turkish. It may be discussed if this could be linked to the Greek substrate or adstrate found in part of these dialects, since Greek does not have inferential as a separate verbal category. However, in the same areas the *-mİş* past tense is used sporadically in another meaning, i.e. postterminal. The article discusses whether this also may be due to Greek influence or whether it should be interpreted as an archaism from Old Anatolian Turkish.

Bernt Brendemoen, Institutt for østeuropeiske og orientalske studier, Universitetet i Oslo, P. O. Box 1030 Blindern, N-0315 Oslo, Norway.

1. Introduction

In his excellent book on the Turkish dialects in the province of Rize, Turgut Günay (1978) noted that what he calls the inferential past tense (“öğrenilen geçmiş zaman”) has a limited use, because the inferential notion can also be expressed with the regular past tense in *-DI*, and that the inferential past may be used for “events and actions which take place for a certain period and under certain conditions. Inferential past conveys a notion equal to the notion of certainty conveyed by the regular past tense,” Günay states.¹ At the same time, the inferential forms of the

¹ “... fiillerin öğrenilen geçmiş zaman anlatımlarını, görülen geçmiş zaman kipi ile karşılamak bölgede yaygın bir eğilime bağlı olduğundan, gerçek fonksiyonuyla geniş bir kullanış alanı bulamayan bu kip, bir çok örnekte, belli bir süre ve belli

copula (“rivayet”) are used quite rarely (and only in the 3rd person) both in its regular copula function (corresponding to Standard Turkish (ST) e.g. *ev güzelmiş*) and together with participles to form compound verbal forms corresponding to ST e.g. *gidiyor-muş-um*, *gidecek-miş-im*, etc. (cf. Günay 1978: 177 and 184). Based on my own extensive material from Trabzon, I shall try to develop and redefine the rather vague notion expressed by Günay and to explain it within a diachronic frame.

2.1. Lack of inferential *-miş* past in the Eastern Black Sea dialects

As is well-known to everybody who has studied Lars Johanson’s works on Turkish aspectology, the notions of the *-miş* past constitute one of the most complicated fields in Turkish grammar. Without reproducing the details of this complex feature, it should be pointed out that besides its inferential and / or postterminal function, i.e. what Lars Johanson calls its “diagnostic” use, it may also be used “historically”, i.e. to express the action itself in a narrative of past events.² This is especially common in popular style, e.g. in tales and stories. A narrative may also be *-DI* based, i.e., *-DI* past is the unmarked tense used to express the action as it proceeds. When singular *-miş* past forms are used within this *-DI* based frame, they convey a strong notion of inferentiality (Johanson 1971: 79). I would even go so far as to say that *-miş* forms are compulsory in a *-DI* past based narrative where inferentiality is to be expressed. There are also other narrative techniques that may be used; cf. the special use of either present tense or the aorist in jokes and anecdotes. As pointed out by Johanson (1971: 83) based on Räsänen’s Middle Anatolian texts, in this kind of popular language the use of the aorist as a narrative base is confined to Nasreddin Hoca tales, while the present tense in *-yor* is much more common in other tales and anecdotes. Both in literary and popular language, narrative styles are often mixed.

As in Rize, in most districts in Trabzon, too, the most striking aspect of the use of the *-miş* past (or rather *-miş* without vowel harmony)³—

şartlara bağlı olarak cereyan eden oluş ve kılışlar için de kullanılabilmekte; yani görülen geçmiş zaman kipinin kesinlik anlatımına eş bir anlatım taşımaktadır” (Günay 1978: 153).

² This is perhaps most clearly expressed in Johanson (1993: 123).

³ This is the most common form of this suffix in Trabzon (and Rize, cf. Günay 1978: 177) (and furthermore the one that should be expected, taking the archaic character of the dialects into account).

when compared to ST—is the fact that it quite seldom occurs in the most frequent functions it has in ST, i.e., it is neither the basic tense for a narrative as a historical tense nor used (within or outside *-DI* past based narratives) to convey inferentiality. In short, *-mİş* is used extremely rarely.

For example, in the following *-DI* and *-yor* based narrative from Trabzon it is evident that the notion expressed by the boldface verbal *-DI* past form necessarily would have had to be expressed with *-mİş* past in ST: The informant (from the municipality of Uzungöl in Çaykara), relating an accident he had during his military service many years ago tells of awakening from anesthesia in the hospital.⁴

*Ayıldum / baktum nerede yatıyorum / hastahane / ışıklar yanıyor / ondan sora ayvax / dedum ayağuma bakayım var mı yok mu / baktum ki yok / kes-tiler oni //*⁵

‘I woke up. I looked to see where I was lying. The lights are on. Then, oh my God! I thought I would look at my leg, if it was there or not. I saw that it was not. They had cut it off.’

In ST, *kesmişler* would be the compulsory form; *kesmişlerdi* would perhaps also be possible, but as we see, the last two VPs are introduced by *baktum ki*, which in ST as well can be used to determine the tense of the verbs expressing the contents of the observation, like in direct speech (cf. Johanson 1971: 72).

⁴ Examples from various dialect texts are quoted in their original transcription, except for some examples from the texts published by Räsänen and Gemalmaz, which have been somewhat adjusted to my own system for typographic reasons. In the examples from Trabzon, which I have recorded and transcribed myself, the following transcription symbols may not be self-explanatory: *ü* denoting a half-front rounded high vowel; *ı* a half-front unrounded high vowel; *ı°* a slightly rounded back high vowel; *ç* a dental affricate; *B, D, G* unaspirated unvoiced stops; *ʼ* denotes palatalization. Thus, *gʼ* and *kʼ* are close to the Standard Turkish affricates *c* and *ç*; *ð* is a dental spirant; */* indicates a pause.

⁵ The story goes on like this: *Ondan sora geldi bir kadun / Bundan aşai Goli kesik / Bu da fransuz harbinda bizum askere ma:(l)zeme taşırken attiler / fransuzlar da kolini kestiler* / ‘Then a woman came. Her arm was cut off from here and down. When she had been carrying goods to our soldiers during the war with the French, they had hit it. And the French had cut her arm off.’

Quite a number of informants such as the one just cited do not use the suffix *-mİş* at all, neither the *-mİş* past, the past participle in *-mİş* nor the inferential form *imİş* of the copula. When an anecdote or story is related only in *-DI* past without any instances of *-mİş* past, this may of course be interpreted as a result of there being no need to express any inferential notion in the discourse—after all, in ST, too, there is a great degree of optionality as to how the speaker should present an event. Therefore, in some cases it may be difficult to decide whether the lack of *-mİş* forms in an informant is caused by a different choice of narrative technique or by the fact that the informant's verbal system lacks the inferential category. To decide this, more extensive material from each informant would be needed than that which I have at my disposal.

2.2. *-DI* past also expressing relative anteriority

It is equally difficult to decide whether the lack of pluperfect tense in *-mİştİ* in a narrative should perhaps likewise be considered as caused by a tendency of not marking past tense in a vivid narrative where relative anteriority is of no importance—a tendency also found in other dialects, cf. the following example from Räsänen's *Sprachproben* from Konya quoted by Johanson (1971: 72):

Babam beni horaza bindirdi. Horaz yürümedi. Horazın boynuna inne batırdım, yürüttüm. Horazın boynu iğneden yara olmuş. Babam horazı baytarlara gösterdi.

'Mein Vater hob mich auf den Hahn. Der Hahn ging nicht. Ich steckte eine Nadel in den Hals des Hahns und zwang ihn so zu gehen. Der Hahnenhals wurde von der Nadel verwundet [*richtig*: war ... verwundet worden; *L. J.*]. Mein Vater zeigte den Hahn den Tierärzten.'

Thus, in the following example from Trabzon the boldface *-DI* past forms probably would correspond to ST *-mİş* past forms expressing inferentiality; however it is also possible to interpret them as corresponding to ST pluperfect expressing relative anteriority, or what ST would express by *-mİştİ*. As pointed out above, however, it seems that this distinction quite frequently has no relevance. (68 year-old Emine Bülbül from the village of Çamlıdüz in Maçka tells about her former Greek neighbours who had come from Greece to visit her:)

*Yuurt mayaladım verdım ulara / Da kudilan⁶ aldi gittiler / demek Gelen (u)-
rumlar / onı aldı oyle gittiler / etmek boyle bi parça aldi Gittiler k'a⁷ bizim et-
meumuzden / o şeye / o yäna / misir etmeü he // aldi gittiler / oyle demek bu-
ralari **Gariplandile⁸ Buralari köreslediler⁹ buralari //***

'I made yoghurt and gave it to them. They took it with them in a pot. The Greeks that came, I mean. They took it with them when they left. They took with them a piece of bread this size because it was our (or 'their') bread, to that place. Over there. Yes, corn bread. They took it with them when they left. They missed these places, you see, they felt the need to see these places again.' (Or 'They had missed these places, they had felt the need to see these places again.')

2.3. Use of *-mlş* past due to influence from ST

It is significant that the Çepnis, who live in the westernmost part of the province, and whose dialect is quite different from the rest of Trabzon, use *-mlş* (and *-mlştI*) as extensively as is the case in ST and with the same function as there.¹⁰ The most common narrative strategy employed by my Çepni informants is to use *-mlş* throughout an anecdote or tale or to start with *-mlş* and after a while switch to present tense. The ST use of the *-mlş* past is sporadically also found in Trabzon informants outside the Çepni area, but this seems to be an idiolectic feature, due to influence from ST through the mass media or schools. In the passages where the ST use of *-mlş* is found, other "urbanisms" or elements taken directly from ST commonly appear, too.

The notion that both the inferential category and *-mlş* as a narrative base are not really inherent in the dialect but a feature copied from ST becomes especially clear when we look again at the informant Emine Bül-bül mentioned above, who uses the suffix *-mlş* only three times¹¹

⁶ *Kudi* "ağzı dar çömlek, güveç" (DS 2990, cf. 48/41).

⁷ Corresponds to ST *güya*, with the meaning "diye".

⁸ Same meaning as *garipsemek* 'to feel lonely and homesick' (Redhouse 1968: 385), or rather, being transitive, *garipsemek* "özmek" (DS 1927).

⁹ Cf. *göresimek* (*göreslemek*) "Göreceği gelmek, özmek" (DS 2159) (a Çepni word, found in Beşikdüzü).

¹⁰ For the Çepni dialect in general, cf. Brendemoen (1989).

¹¹ All of them have a postterminal notion (cf. below), such as:

throughout 5 pages of text. However, as an answer to the question “Ma-sal, hikâye biliyor musunuz?” she gives the following interesting answer:

*Biliyorum*¹² ama yarî yarîyâ unuttum olarî şîndî hatîrîmda kalmayler Gi / un-larî oGuma oGuduk biz / unlari ani mektepte biz ma:cir hacan gelduk muha-cirluktan hacan gelduk biz Gücu:duk / Gardeşimlan / Bizi mektebe verdiler olari okudu / oGuma oGuduk ki oni // Bi kücük çobban **varmiş** / yalancılık **yaBar** (i)mîş / yalanci yalanci / sana kimse inanmas sözuna kimse kanmaz / yalanci bi çoban **imîş** / bi çoban **imîş** / günlerde **çıkmiş** kırlara çoban **Basmîş Baarmîş** ki / kurt geldi imdaD edîn / bu sefer bi Geren Goptiler ya-lan dedi / bi daha koptiler yalan dedi / bi daha esseh geldi / dediler gi'tme-diler / esseh yalanci: kurt yedi //

‘I do, but I have half-way forgotten them. I don’t remember them now. We read them ... in school. When we came back as refugees, when we came back from the *muhacirlik*,¹³ we were small. We—me and my brother—were sent to school, and we read them. We read it in reading: There was a little shepherd. He used to lie. “Liar, liar, nobody believes you, nobody is fooled by what you say.” There was a shepherd who was a liar. There was a shepherd. Once the shepherd went out into the countryside and cried out: “A wolf has come, help!” Once they came running, and he said, “It is a lie!” Once more they came running, and he said “It is a lie!” And once (the wolf) came for real. They said (...?), and did not go. The wolf ate the liar for real.’

Evler isüz / evler hep şen ... / hani Donanmış / eşyâ hep içinde / insan yokolarda. ‘The houses were empty. The houses were popu.. / I mean, they were furnished. There were things inside them. But there was nobody there.’

Habu karşılardan gideyGân da a:ca bi hartuma çaGılmış / yufka bi harDima haboyle çaGılmış bi a:ça. ‘Walking on the other side, (I saw that) a board had been nailed to a tree. A thin board had been nailed to a tree like this.’

Asas türkleri asas burda do:müş beyîmiş. ‘The real Turks (in this village) have been born and brought up here.’

¹² The ST form she uses (instead of expected *biluy(u)rum* or *biliyrum*) is conditioned by the ST form used in the question posed by myself.

¹³ In connection with the Russian occupation of North-Eastern Turkey during the First World War, a great part of the Turkish population in the occupied areas on the Eastern Black Sea coast fled to areas further west and returned to their villages after the Russians left in connection with the revolution in 1917. This status and period of self-imposed exile is called *muhacirlik*.

In the first part of the small tale, not only the use of *-mlş* and *imiş* (partly as a suffix *-(y)mlş* as in ST),¹⁴ but also other elements show that this is a reproduction of the text in the textbook the informant read at school in her childhood, such as the form *yalancılık*, where the dialect would have *yalancılık*, the form *sana* where the dialect form would be *saa* or *sa:*, and the word *küçük* where the dialect form would have a back vowel in the second syllable. We also notice that other elements become genuine dialect forms as soon as the informant breaks the convention learned at school and starts using *-DI* past, such as *Geren* for ST *kere*, the verb *kop-* in the sense ‘to run’,¹⁵ and the word *esseh* ‘true’.¹⁶ It is not impossible that the switch to *-DI* past is triggered by the direct speech *kurt geldi*.

The influence of ST on the individual speakers constitutes the main obstacle against establishing a clear picture of the geographical distribution of *-mlş* inferential in Trabzon. In general, however, it seems that outside the Çepni areas, the valleys Araklı, Arsin and Yomra constitute an area where the use is closer to what is found in Central Anatolian dialects or ST than elsewhere. With respect to other linguistic features, too,¹⁷ this area is somewhat closer to Central and East Anatolian dialects than the areas to the east (Of, Çaykara, and to some extent Sürmene¹⁸) and west (Maçka, Trabzon Merkez, Akçaabat, Tonya, and parts of Vak-

¹⁴ In the Trabzon dialects the forms of the copula stem *i-* (especially of course *idi*, but also the much rarer *imiş*) are usually not used as suffixes.

¹⁵ The impression created by DS (p. 2921) that this verb is used in Anatolian dialects in general in the meaning “koşmak, hızlı gitmek”, but in Trabzon in the meaning “kalkmak, ayrılmak, çıkmak, gitmek” does not seem to be supported by Emiroğlu (1989: 158), who states that it can have the meaning “koşmak” also in Trabzon.

¹⁶ Cf. *esseh* “sahih, doğru” (Günay 1978: 316). The word no doubt has a connection with the Arabic root in *sahih*.

¹⁷ E.g. in the use of zero anaphora contrary to other parts of Trabzon, where the anaphoric pronoun usually is expressed, cf. Brendemoen (1993), or in the area of word order, cf. Brendemoen (forthcoming, chapter 2,5).

¹⁸ I.e. the southernmost parts of the Monahos Deresi basin such as the villages Arpalı, Yılmazlar, Dağardı, Fidanlı. In villages situated closer to the sea and along the Gorgor Deresi, the picture is identical with Araklı, Arsin, and Yomra. In Caferoğlu’s material from Trabzon, the *-mlş* past based tale “Pisig” (1946: 153-155) is related by a person from Sürmene.

fikebir) of it are. An interesting example of tense mixing is found in a long fairytale related by a woman in Kalecik/Araklı, where the tense used throughout the tale is *-mİş*, except for the middle of the following passage:

O k'öpek gene gelmiş / köpek ama demek dev / dev da çok ... / şey hallında / k'öpek şekline girmiş / g'elmiş / ill'a Gızı vercesın Ba: / Demiş işte biraz da ha sevelım onu biz da havesluumuzu alalım bi_ssene daa olsun bizüm da / bu nı vercek saa / e pek'i g'itmiş k'öpek gitmiş soralardan bir gün beun bir gün-e işte bir zaman öyle geşti herhal_arası / ondan sora k'öpek De gelmiş / geldi G'i kak dedi işte kızı Giyindurdiler ufak Gızı ama oldi beş on yaşında / gi:ndırdiler Guşattiler / kakti verdiler k'öpe: / köpek / ali gittiler oni / bi evä al'ı gitti beyu:G i'ste bilmem kırk oDali kırkbir oDali evi var / bilmem hank'i Da:un üstünde / vermişler ona / o da aşmış bir odai: aşmış bir odai: ki işte haburalar işte hep bizımdur şöyledir böyledir kimi / ya: doli kimi piriş Doli kimi şeker doli kimi ... / işte dünyadä her ne yiyecek var_ısä o kırGBır oda ama kıkbirinçi odai sakın aşmıcasin / demiş / kız gene sabredememiş demiş kı / ben bu oda da açarım bakaim ne var burda / bakmış ki şe ...

'That dog came again. A dog, but in fact the giant. And the giant was very ... It had been transformed into a dog. It came. "You shall absolutely give me the girl," it said. (They answered:) "Let us fondle her a bit more; let us too have our desire fulfilled. Let us have one more year, and then we shall give her to you." All right, the dog left, and later one day, after time had elapsed, the dog came again. He came and said "Get up". They dressed the girl, the little girl; she had become between five and ten years old. They dressed her and adorned her, and then they gave her to the dog. The dog took her away (?). It took her to a house, it had a big house with approximately forty or forty-one rooms. On the top of I don't know which mountain. They gave her to him. And he opened one room after the other (and said): "These belong to us, this and that. Some are full of butter, some are full of rice, some are full of sugar, some ..." All the foodstuff there is in the world was in those forty rooms, but "the forty-first room you must not open," he said. But the girl could not resist (the temptation) and said (to herself), "I shall open this room, too, and see what there is inside." She saw that it was ...'

In this passage, the *-mİş* past chain is suddenly interrupted by *-DI* pasts. One may ask if this possibly could be triggered by the sentence *Bir zaman öyle geşti herhal_arası*, which could be perceived as a comment

made by the narrator, not actually belonging to the narrative itself. Although narrative techniques may be mixed in ST, insertion of *-DI* past forms in *-mlş* based narratives do not seem to occur. The example above should consequently be interpreted as a result of the tendency of semantic merge between the *-DI* past and *-mlş* past in the area.

Another example illustrating the same merge—or rather confusion—is from Tonya/Ağır Köyü, where the informant gives the following answer to the question “Bir hikâye anlatabilir misiniz?”:

Hikâyälärdän bent’ım haBerı’ım yok / be(n) mektepli Deilim Gi // eski zamanlarda bir var_idi bir yo:mu’ş / anam Ba’bam munteşeuni sallamışım / Ben / DuGur DuGur ederdi.

‘I don’t know any stories. I have never gone to school. Once upon a time ... I rocked the hinge of my parents.’¹⁹ It used to clatter.’

Here, the informant’s effort to imitate ST has resulted in a strange mixture of tenses.²⁰

2.4. Influence from Pontic Greek?

In general, the characteristics of the Eastern Black Sea dialects may be grouped under either of the two labels archaisms or innovations. The archaisms may be dated back to Old Anatolian Turkish, while most of the

¹⁹ *Munteşe* (ST *mentşe*) ‘hinge’ is used instead of *beşik* ‘cradle’. There is also an inexplicable contrast between the form *sallamışım* and the past aorist *ederdi*.

²⁰ A similar picture is found in the tale reproduced in Caferoğlu (1946: 162-165), where the informant mixes *-DI* past and *-mlştIr* as narrative base. No doubt the informant, who comes from the village of Zelego (today Taşören) in Of (today Çaykara), and who (considering the position of his village) would not be expected to use the *-mlş* past at all (cf. below) has adopted *-mlştIr* from ST. Thus, it is significant that the instances of *-mlştIr* decrease as the tale proceeds (and as the informant, feeling more at ease with the interview situation, tends to drop ST expressions). A similar tense mixing (but with fewer instances of *-mlş* and *-mlştIr* forms) is found in another tale narrated by the same informant on p. 250-253. In this connection it should not be forgotten that most of Caferoğlu’s informants were prison inmates who must have had extensive contact with fellow-inmates speaking if not ST, then at least different dialects. Thus, e.g. the tale found on p. 242-246 has very little value as a dialect text.

innovations are caused by the existence of substrate or adstrate languages in the area. In Trabzon, and especially in the areas mentioned above (Of and Çaykara, and to some extent Sürmene, Maçka, Akçaabat, Tonya, and Vakfikebir), the only substrate or adstrate language that has left any systematic traces on the Turkish dialect—or, to put it somewhat differently, whose codes have been copied into the Turkish dialect—is Greek, or rather, Pontic Greek, which was spoken extensively in the area until the population exchange in 1923, and which is to some extent still spoken in some villages in the *ilçes* of Çaykara and Tonya. It is interesting to remark that the lack of *-mlş* is especially noticeable in areas where Greek is still being spoken, although the above-mentioned difficulty created by the influence of ST in modern times makes it impossible to establish any clear geographic distribution. As Greek—both Standard Greek and Pontic Greek—lacks the inferential as a separate verbal category,²¹ an explanation of the lack of *-mlş* in these Turkish dialects as due to Greek influence would seem quite convincing at first glance: The lack of the inferential category has been copied into Turkish, and consequently the (secondary) use of the *-mlş* past as a historical tense synonymous with the *-DI* past in stories and tales became superfluous, too, and was discontinued. That this, however, is only a simplified half-truth becomes obvious if we consider the fact that the main semantic notion conveyed by the *-mlş* past at the time when Turkish penetrated the area (after which point it became largely isolated from the language development in Central Anatolia) was not inferentiality but postterminality, a point I shall develop below.

3.1. The postterminal use of the *-mlş* past

As mentioned in the introductory quotation from Günay, there is in some confined districts a quite widespread use of *-mlş* in the Eastern Black Sea dialects where ST would prefer *-DI* past. This is especially conspicuous in the first (and second) person singular and plural, and in-

²¹ In Standard Modern Greek, a kind of inferential may be expressed by the so-called subjunctive with the particle *θα* (cf. Mackridge 1985: 274-275), but the notion conveyed by this form is not directly parallel to Turkish, being rather what we could characterize as “emphatic inferentiality” (or better, “prägnante Inferentialität”) corresponding e.g. to Turkish expressions with participle + *olacak* or *olsa gerek*, e.g. *θα της μίλησε* ‘Onunla konuşmuş olacak’ (‘He must have talked with her’).

icates that the past tense system in the Eastern Black Sea dialects is not entirely copied from Greek. This use is found in Rize and especially the eastern parts of Trabzon. Examples from Rize:

(Kalkandere/Kayabaşı Köyü:) *Otuz sene bu işe ç'alışmışım.* (Günay 1978: 251).

'I have worked in this job for thirty years.'

(Kalkandere/Yokuşlu:) *Biz geç'urmuşık 'dertler, hastaluklar.* (Günay 1978: 253).

'We have had problems and diseases.'

Examples from Trabzon:

(Of/Keler, as an answer to the question "Bu köyde mi büyüdüünüz?") *Bu köyde habu k'öyün aşā:sında be'yündüm habu çāmeni'n yan'ina beyinmişim g'etmişim buriya yirmi sene otüz senede buriya g'älmışım.*

'I grew up in this village, in the lower part of this village; I have grown up beside this mosque, and I came here, I came here twenty or thirty years ago.'

(Of/Korucuk:) *DokküzyüzoDüseGis senesinde buna kuloumilän va:Gif olmuşım.*

'I heard this with my own ears in the year 1939.'

(Çaykara/Tüfekçiler:) *Şeye gitünüs / oGeneye / iki köyi da gezdi'nı^oz / ben orıya da bir sene çalıştı^om / o aşā:Gİ^oogeneni^on / altına bi Deyermen yapmışı^om olara.*

'You have been to Ogene,²² and you have visited both villages. I worked there for one year, too; I have built a mill for them below Aşağı Ogene.'

(Of/Cumapazarı:) *Ben çocuGDum / on oniki yaşında var_ıdım / ruslar burya geldi / ben / alano / ve ços Daina etmeG_GeDürmişum askere / ben / etmeG_GeDürmişum da: ben(d)en büyü:G_adamlarla bara:bar eD-meG_GeDİ^ordı^om.*

²² The village Karaçam in the upper part of Çaykara.

‘I was a child. I was ten or twelve years old. The Russians came here. I brought bread to the soldiers on the mountains Alano and Ços. I brought bread to the mountain. I carried bread together with men older than myself.’

(Sürmene/Fidanlı:) *Bunlarlan bi_yerdeydi evumuz aşada Biz yanı bunlan bi_ara:da beyi°mişük.*

‘Our house was in the same place down there as these people; I grew up together with him.’

(Sürmene/Dirlik:) *Arakli çarş(ıs)ında benim ma:zam var_ıdı / hätta iki dane o:lum vardı bi da bila:derim / bi_ssoba al'mişuk uş dört kilo et al'mişuk / eve ben gelincäs aksu hanlarına g'eldim.*

‘I had a shop in the center of Araklı. I was with two of my sons and my brother. We had bought an oven and three, four kilos of meat. On my way home I came to the *hans* at Aksu ...’

(Yomra/Çamlıyurt:) *Banka müdürü bu parai tutti / ... / unci° / gäl buray bu parai nerden almi°şün para saxtadu.*

‘The bank director held up the note. “Flour-seller, come here! From where have you got this money? It is fake,” (he said).’

(Arsın/Işıklı:) *Ben Gaymakla büyümi°şüm ha.*

‘I have grown up on *kaymak*, you see!’

(Akaçaabat/Baltacı; as an answer to our suggestion that the informant should tell us about his pilgrimage to Mecca:) *A:natmışım ondan bi nâzet²³ alamayük_ki.*

‘I have told it before; we won’t get any pleasure from that!’

(Akçaabat/Çatalzeytin:) *Sen hı°ristiyan ol' ne olusan ol / sen olmuşün hıristiyan hıristiyan olabilüsün ama canlı adam yaGamasın.*

‘Be a Christian, be whatever you like. You have become a Christian, you may be a Christian, but (still) you can’t burn people alive!’

²³ Corresponds to ST *lezzet*.

(Tonya/Ağır:) *Biz دنیا: karişmamışık / Boyle: / kocaDuG_Gideyruk* (the informant is talking about herself and her sister).

‘We have not got married. That’s how it is. We have grown old.’

(Vakfikebir/Ağaçlı:) *Ha_şuriya geldim ay_şığı:dı baktım nã_adam va ne bişe kayboldu / ecinli:ydi dediler oni_işte ha_uni görmüşüm burda.*

‘I came up there; there was moonlight; I looked, but neither the man was there nor anything else; they had disappeared. He was possessed by djinns, that’s what they said; him I have seen here.’²⁴

3.2. Postterminal use of *-mİş* past as a Turkic phenomenon

Without exaggeration I think we can say that ST most probably would use the *-DI* past in all these examples. These are all examples of the *postterminal* use of the *-mİş* past, a use where what Lars Johanson calls the “diagnostic dimension” of the narrative perspective is prevalent. This is quite similar to the notion conveyed by the perfect tense in English, Scandinavian languages and Northern German. This postterminal use of the tense is different not only from the use of the *-mİş* past as a basis for the discourse but also from the inferential use; in fact, it seems that the inferential use of the *-mİş* past historically is a secondary development of the postterminal use. This original postterminal use of *-mİş* is found in ST, too; for some reason, however, it is not used in the first person without an additional strong notion of inferentiality. This use of the *-mİş* past is well known from Azeri, where *-mİş*, to quote Johanson (1971: 289), “findet allgemeinere Anwendung zur Hervorhebung der diagnostischen Dimension.”—“Allgemeinere” of course in comparison with ST,

²⁴ Examples from Caferoğlu’s material (1946) are quite scarce, most probably due to the fact that the texts either are tales or manis:

Of: Zêlêgo köyünden [i.e. today Caykara/Taşören]: *Çeçuğun piri bir ruya kormışdür, hocasında dedi ki: ben bir ruya kormüşüm.* (p. 162) ‘A boy had a dream, and told his teacher: “I have had a dream.”’

Pazar: Horti köyünden’ [i.e. Pazar/Kocaköprü]: *Cəlin-da hədcaya dediki: köpek mopek bağlamışım. Sana neler saklamışım, nani zekiyem nani.* (p. 268) ‘And the bride said to the imam: “I have tied the dog and so on. I have saved some nice things for you, nani zekiyem nani.”’

as in the following example cited by Grunina (1976: 17): *Suleymanı älä²⁵ bu saat Jerevana yola saldım. Beş min manat da pul vermişäm.* ‘I just now sent Süleyman to Erivan. And I have given him a sum of five-thousand rubles.’ Whether there is a geographical continuum between Azeri and the Black Sea dialects in this respect is difficult to say; in the dialects of Erzurum, which have quite a number of characteristics in common with Azeri, the postterminal use of *-mİş* is found indeed, although only sporadically;²⁶ in the Gümüşhane (and Bayburt) texts collected by San (1990), however, there do not seem to be any examples of clearly post-terminal use of the *-mİş* past at all (or rather, any examples of the *-mİş* past in the first or second persons) in spite of the fact that San mentions quite a number of examples of such first and second person forms—taken out of their context—in the morphology chapters of his book (San 1990: 241-247).²⁷ It is important that the first person *-mİş* past forms are not found in the Çepni areas in Trabzon, which in many respects—as

²⁵ *Älä* must be an error for *elä*, cf. Grunina’s translation: “Вот только что проводил Сулеймана в Ереван. И пять тысяч рублей дал ему.”

²⁶ Cf. the following examples from Gemalmaz (1978, vol. 2):

(The informant is talking about having lost his way in the fog:) *Yoli şaşır-tmışam yol sağdan gelirken, ben sola doğru yanbegi getmişem.* (p. 33, l. 20) ‘I lost my way; when the road was coming from the right, I turned to the left’. (However, the use of *-mİş* in this example could perhaps also be interpreted as inferential.)

(Praying to God during a storm:) *Heş deyil hayvanata bişey olmasın. Ele ağ-lamışam. Elece düşmüşem yuhiya.* (p. 33, l. 27) “‘At least do not let anything happen to the animals.’ That is how I cried, and then I fell asleep.’

Ahen ben onnañ söylememeşüm da:. (p. 37, l. 16; the meaning is unclear)

Bilmeyrım ben; gocalmışım. (p. 55, l. 17) ‘I do not know; I have grown old.’

(Talking about an encounter with a bear:) *Biz de böyle daşların arhasında siper almışeş.* (p. 73, l. 15) ‘And we have / had taken shelter behind the stones.’

Ama heş de ordan geçmemeşem. (p. 73, l. 33) ‘I have / had not passed through that place.’

Especially the form *unutmuşum / unutmuşam* etc. seems to be quite frequent (e.g. p. 56 l. 44-45, p. 67 l. 16 etc.), but this, which is also found in ST, could also be explained as a kind of inferential.

²⁷ An example of postterminal use of the *-mİş* past in the third person from Gümüşhane could be the following example from Kelkit: *Babamın on çocuğu varıdı. Beş de ölmüş.* ‘My father had ten children. Five are dead.’ (San 1990: 337)

mentioned above—are closer to the Central Anatolian dialects. It should be pointed out that the use of the *-mlş* past in Khalaj is quite similar to what we find in Azeri and in the Black Sea dialects; as Doerfer most aptly remarks, stating that the use of the Khalaj form in *-mlş* differs from ST, this tense “... ist im Deutschen immer einfach mit dem Perfekt zu übersetzen. Daher sind auch (im Gegensatz zum Ttü) Belege mit 1. und 2. Person recht häufig ...” (1988: 176). Thus, within an Anatolian context it would seem reasonable to characterize the postterminal use of *-mlş* past in parts of Trabzon and Rize as an archaism.

3.3. Postterminal use of *-mlş* past as a calque on Greek?

However, a priori the possibility should not be disregarded that this postterminal use of the *-mlş* past has come into existence as a result of influence from Indo-European languages, i.e. Persian for Azeri and Khalaj and Greek for the Black Sea dialects. The Persian compound perfect of the type *xarīde-am* ‘I have bought’ does indeed have a postterminal aspect, and so does the Modern Greek compound perfect of the kind *έχω γράψει* ‘I have written’. Although the Greek perfect of this type most probably developed “after 1453” (Schwytzer 1959, 1: 130), Common Greek has had the means of expressing the postterminal aspect since its earliest stages of development, partly by regular tenses, partly through periphrastic and compound forms expressing the stative, or what Johanson calls “prägnante Postterminalität” or, in English, high-focal postterminality (cf. Schwytzer 1959, 1: 812). However, it should be kept in mind that the Modern Greek tense generally used to cover postterminality is the aorist and *not* the compound perfect (cf. Thumb 1912: 123). Since the aorist also covers terminality (corresponding to Turkish *-DI* past), this means that the aspectual distinction between terminality and postterminality is of secondary importance in Modern Greek. In Modern Greek, the compound perfect is a characteristic of written style and—to quote Thumb (1912: 163)—“not frequently used ... The use of the perfect *έχω δέσει* [‘I have bound’] is least common, the aor. indic. quite frequently having the force of our perfect ...” In most Greek dialects, perfect tense seems to express the stative, or “prägnante Postterminalität”. Even if the static connotation and the “Prägnanz” (high focality) partly (i.e., more or less as an idiosyncratic feature) seem to be lost in contemporary Standard Greek today (perfect tense (and a corresponding pluperfect tense) more and more frequently being used to express postterminality, semantically equal to the aorist, cf. Mackridge

1985: 116-117),²⁸ what is important in this connection is that the compound Greek perfect (and pluperfect) is unknown to Pontic Greek (cf. Mackridge 1987: 127), or, to put it another way, that the distinction terminality—postterminality is unknown to Pontic Greek.

3.4. Postterminal use of *-mİs* past as an archaism within a Turkic frame

As pointed out by Johanson (1993: 118), postterminality is a fundamental aspectual notion not only of Turkish of Turkey, but of Turkic languages in general, as postterminal formants are found both in Southeast and Northwest Turkic languages (i.e. the suffix *-GAn*), and also in Yakut (i.e. the suffix *-BIT*). Consequently, at least the above-mentioned instances of *-mİs* in the Black Sea dialects in the first and second person cannot represent a code-copying from Greek,²⁹ but something genuinely Turkish.³⁰

²⁸ It has been suggested that this increased use of the perfect has been caused by the frequent use of *passé composé* in French, a language (allegedly) extensively spoken in Athenian high society. It would, however, probably be more proper to consider this development within the framework devised by Johanson (1993: 122), where he postulates a development “Stativ → Perfekt → Konstativ → Terminale”, a development which most probably could be postulated as a frequent development in general language typology (cf. the similar development in Romance languages).

²⁹ The Black Sea dialects frequently also use another genuinely Turkish device to express postterminality, or rather “prägnante” or “statische Postterminalität” i.e. the verbal noun in *-mA*, e.g. (Of/Cumapazarı):

Bu saral da ovadan gelmedir. ‘This man called Saral, too, has come from the plain.’

Especially (Vakfikebir/Ağaçlı):

Buların adamı nerden geldi° / kâfkasyadan mı geldi / şeyden mi Geldi / burayın adamı bu ... / bura hep Gelme zaten / biz hep gelmeyi°k buraya / ... / heBisi gelmedir / siz gelme deyil misii°z / ordan Brakıp buraya Gelme deyil misii°z / hebisi gelmedir burda / G’ökyüzünden buräyä enmedi. ‘From where have the people in these places come? Have they come from the Caucasus? Have they come from ... ? The people here—this place is all (full of) newcomers. We all have come here from somewhere. ... All are newcomers. Haven’t you come from somewhere? Haven’t you left that place and come here? Everybody here has come from somewhere else. They haven’t descended here from heaven.’

As somewhat vaguely indicated above, the postterminal 1st and 2nd person *-mlş* forms should definitely be considered archaisms: In an article from 1993, Johanson clearly states his view that both the inferential use of the originally stative or postterminal tenses and their use as bases for narratives (without any diagnostic connotation) are secondary developments. The same view is—*mutatis mutandis*—expressed by Adamović (1985, especially p. 201), who, however, does not mention the work done by Johanson in this field 14 years earlier or refer to his terminology. Adamović's survey of the different functions of the verbal noun *-mlş* in older texts is, however, very useful. One of the most common uses in Old Anatolian Turkish (except for the use as a finite verb where a full paradigm seems to be developed in the first half of the 14th century) is what Adamović calls the "attributive-adverbial use" (I would call it a "converbial" or "gerundial" use), e.g. in the sentence *ilan sovuq-dan buymiş yatur* 'die Schlange liegt erstarrt vor Kälte' (1985: 187); another function is the use as an adjective attribute (e.g. *üşimiş ilan* 'unterkühlte Schlange') or as a noun. In all these uses the notion is post-terminality. It is quite interesting to observe that these non-finite uses are also sporadically found in Trabzon and notably in the same districts of Trabzon and Rize (especially in the eastern part of Trabzon) as where examples of the postterminal use of the first person forms are found, thus representing other kinds of the original uses of *-mlş*. The converbial use, which is found very rarely and only in the easternmost parts of Trabzon,³¹ is especially interesting because it differs somewhat from the use in ST: In ST, a sentence like *Kanepede uyumuş yatıyorum* 'I am lying sleeping on the sofa' is not necessarily perceived as containing any converbial element; the form *uyumuş* is perceived as short for *uyumuşum* in the same way as the much more common *Kanepede uyumuş yatıyordum* is considered a shortened form of *Kanepede uyumuştum*,

³⁰ However, it should be added that the lack of pluperfect in Pontic Greek perhaps could be one of the reasons why anterior postterminality is also expressed by *-DI* past in the Eastern Black Sea Turkish dialects (and not by *-mlştI*). But as mentioned initially (2.2. above), in spoken narratives in Turkish dialects in general, there seems to be a tendency of relative anteriority not necessarily being expressed, which might be a more probable explanation.

³¹ In the same group of villages close to the İyidere river where also other very special features are found, cf. Brendemoen (1996a and 1996b).

yatıyordum, through “Suffixabwurf” (Johanson 1971: 71).³² Such examples are also found in the easternmost parts of Trabzon.³³ However, since *-mlş* forms within a narrative frame with a *-DI* past base are automatically perceived as inferential—or ungrammatical (cf. Johanson 1971: 71, 79)—a sentence such as **Şişeyi kadehe boşaltmış içti* will make no sense to an ST speaker, whereas in older language, and in the Eastern Black Sea dialects, it most probably would have the meaning ‘Having emptied the bottle into the glass, he drank (it).’ The two following examples represent this type, where the converbial use of *-mlş* is found within a *-DI* based frame:

³² However, modern ST examples such as *Kadın, ... yanağını avucuna dayamış uyuyan küçük kızına bakarak yeniden bir ağlama tutturdu*. (from the short story “Ümit Fakirin Ekmeği” by Nezihe Meriç, reproduced in A. Tietze: *A Turkish Literary Reader*, Bloomington 1963, p. 97) betrays a converbial use also in ST. A “converbial” or “gerundial” interpretation of the use of the *-mlş* participles (i.e. as parallel to *-(y)Ip* and *-(y)ArAk* gerunds) in ST in examples similar to those above has been suggested by Kononov (1956: 424) and by Drimba (1976: 58). A more apt name than “converbial”, however, would be “participium coniunctum”.

³³ For example the following:

(Of/Balaban:) *Durduk Dışariya çıGmaa / sordum yaşlı adama boyle / kafasını dikmiş / dışarı çıkayı*. (8 / 27) ‘We started to go out. I asked the old man. He is walking out hanging his head (i.e. broodingly).’ (In ST, *kafasını (başını) dikmek* naturally means ‘to be conceited’, but in the eastern part of Trabzon it obviously means the opposite, i.e. ‘to hang one’s head’ (occurs 4 times in my material); see the following examples:)

(Of/Yanıktas:) *Bu adam çok Dertli:di / Gitti evine / Başını d’ikmiş Dürüy / o GaBlumbaa da dedi_on_n’e Başını d’ikmiş Dürüyüsün / o da d’edi oa*. ‘This man was very sorry. He went home. He is standing hanging his head. Then the turtle said, “Why are you standing hanging your head?” He told it why.’

(Sürmene/Arpalı:) *Bir div Garısi yatayr / uzanmış / ki memesi işlemiş*. ‘A giant woman is lying (there) stretched out because her breast has festered.’

(Yomra/Tandırılı:) *Bundan beş ay öncä / burdän aşā: bi orman arabasi / o:dun yüklenmiş gide:di* (corresponds to ST *gidiyordu*). ‘Five months ago a forest truck had loaded firewood and was going down from here.’

(Maçka/Şimşirli:) *Orada işkü orada fuçularlan yulmuş dürüdi orda (dürüdi)* corresponds to ST *dururdu*). ‘Liquor used to stand there heaped up in barrels.’

(Of/Balaban; about a woman who went to the priest to have him read over her sick child and who then understood that the priest was trying to seduce her:) *Daa okutmamış bunı aldı çecuunu koydı bastı yola geldi / okutmadi_ni.*

‘Without letting him read over him anymore she took her child and set off and came back (to her village). She did not let him read (over it).’³⁴

(Ibid.) *İsa Beyamberu ressim var kilisenin icerisinde / çıkmış pa'Baz ona yöneldi ellerin kôşsine vuruyi.*

‘There is a picture of the prophet Jesus inside the church. The priest went up and turned towards it and strikes his hands against his breast.’

Another most interesting example where the converbial use is quite clear is the following:

(Sürmene/Arpalı:) *Abu gülizar dedi nişanını Brakmış ben gülizarı almamış almam.*³⁵

‘Abu Gülizar, he said, has left her token. Not having taken Gülizar, I will take nobody.’ (= ‘If I cannot have Gülizar, I won’t have anybody.’)

A couple of sentences later with the same meaning:

Ben almam ben dedi bülbülü_güliza:rımı dedi almamış Dedi ben dedi kimseyi almam.

For this converbial use of *-mİş* and especially the examples with *almamış alamam* above, the following remark by Emiroğlu in his etymological study of the Trabzon Maçka dialect (1989: 13) is interesting, although the example he gives is not very illuminating: “Madan / meden zarf fiil yerine mamış / memiş biçimi kullanılır: *yemek lememiş* (misprint for *yememiş* ?) *olu mi*”, which must mean ‘is it possible without having eaten’.

My texts also contain a few examples of the attributive use of the *-mİş* participle:

³⁴ For the use of *okumak / okutmak* with this specialized meaning cf. Brendemoen (1996a: 42).

³⁵ The existence of such constructions has been asserted by Prof. İlhan Başgöz, Bloomington, and Mr. Mehmet Bilgin, Sürmene/İstanbul.

(Çaykara/Taşlıgedik:) *Bir gün xoca / okumuş ađam idi / bakti bir ađam Barçalanyor.*

‘One day the imam—he was a learned man—saw that a man was being torn to pieces.’

(Sürmene/Yemişli:) *Be(n) yi(r)mibeş kuruşa akşama kada bāt çekmiş adamı^om.*

‘I am a man who has dug with a digging fork (from morning) to evening for twenty-five kuruş.’

Perhaps the following two interesting examples also belong here:

(Çaykara/Taşlıgedik:) *Köyün xocaları okumuş varıdı.*

‘There were imams in the village who were learned.’

(Maçka/Yazlık:) *Orda bir iki masa kurtı^olmış varıdı.*

‘There were a couple of tables that had been set up.’

4. The inferential copula form *imiş*

As mentioned initially, some of my informants, especially in the westernmost districts and in the Sürmene-Araklı-Arsın-Yomra region, have been influenced by ST and use *-miş* both as discourse base in narratives and to some extent as inferential. In this connection it is interesting to observe that the *-miş* form of the *copula* is extremely rare outside the districts or informants whose verbal system shows a strong impact from ST: I have found only some 15 examples of *imiş* in the 250 pages of transcribed text I have from outside the Çepni area. In most of these 15 cases, *imiş* is used together with participial stems, creating a notion of inferentiality. As the inferential notion conveyed by the copula form *imiş* is most probably one of the origins of the development postterminality → inferentiality in the *-miş* past (cf. Adamović 1985: 201), the very low frequency of this form in Trabzon most probably should be interpreted as due to the fact that inferentiality basically is a notion foreign to the Eastern Black Sea dialects. The lack of *imiş* forms outside the third person, too, shows that the inferential has not developed properly in the area.³⁶ A good example of the inconsistent use of *imiş* is the following,

³⁶ For the same feature in Rize, see Günay (1978: 177-178, 184).

where it becomes obvious that the informant is trying to copy an ST code without being quite aware of what he is copying:

(Akçaabat/Şinik:) *Änvär Başsa / ruslardan bizim bu çepei rüslara kendi sat-miş şöyle / lapor yaparmış da / ... / kormiş şışe içerisine / mantarlar-mış atarmış süya / aşā:dän iki üçyüz kilometre aşā'idän ruslar alıdı oni / bakardı işte filan yerden Geç /*

'Enver Paşa himself sold this front to the Russians in the following way: He used to make reports ... He used to put them into bottles which he corked and threw into the water. Two or three hundred kilometres further down, the Russians would take them up and look: "Pass through at this or that point."

5. Conclusion

To sum up: *-mlş* past does not traditionally convey a notion of inferentiality in the Eastern Black Sea dialects, partly because this notion was not developed when Turkish started to be spoken in the area and became isolated from the general development of Anatolian dialects, partly because the development of such a notion was prevented by the Greek sub- or adstrate, to which inferential is a foreign verbal category. The sporadic use of *-mlş* past to express inferentiality or as an unmarked narrative tense, alongside the use of *imiş* forms of the copula, is a quite recent innovation due to the influence of ST. The sporadic use of *-mlş* past to convey the postterminal aspect, especially found in the Eastern part of Trabzon, is an archaism that has survived in spite of the Greek sub- or adstrate.

References

- Adamović, Milan 1985. *Konjugationsgeschichte der türkischen Sprache*. Leiden: Brill.
- Brendemoen, Bernt 1989. Trabzon Çepni ağzı ve Tepegöz hikâyesinin bir Çepni varyantı. In: Kayaoğlu, İ. G. et al. (eds.) *Trabzon Kültür-Sanat Yıllığı* 88-89. (Trabzonlular Kültür ve Yardımlaşma Derneği yayınları 2.) İstanbul. 13-22.
- Brendemoen, Bernt 1993. Pronominalsyntax in den türkischen Schwarzmeerdialekten – syntaktische Innovation oder Archaismus? In: Laut, Jens Peter & Röhrborn, Klaus (eds.) *Sprach- und Kulturkontakte der türkischen Völker – Materialien der zweiten Deutschen Turkologen-Konferenz, Rauischholzhausen, 13.-16. Juli 1990*.

- (Veröffentlichungen der Societas Uralo-Altaica 37.) Wiesbaden: Harrassowitz. 51-73.
- Brendemoen, Bernt 1996a. Case merge in the Black Sea dialects—A Kartvelian substrate feature? In: Berta, Árpád & Brendemoen, Bernt & Schönig, Claus (eds.) *Symbolae Turcologicae—Studies in honour of Lars Johanson*. Stockholm: Swedish Research Institute in Istanbul. 41-59.
- Brendemoen, Bernt 1996b. Some Remarks on the copula in a “micro-dialect” on the Eastern Black Sea Coast. In: *Türk Kültürü Araştırmaları (Prof. Dr. Zeynep Korkmaz’a armağan)* 32/1-2, 1994 (1996). 107-115.
- Brendemoen, Bernt forthcoming. Einige Bemerkungen über die Wortstellung in anatolischen Dialekten. To appear in: *Akten der 3. Deutschen Turkologenkonzferenz, 4.-7. Oktober 1994, Leipzig*.
- Caferoğlu, Ahmet 1946. *Kuzey-doğu illerimiz ağızlarından toplamalar*. İstanbul.
- Doerfer, Gerhard 1988. *Grammatik des Chaladsch*. (Turcologica 4.) Wiesbaden: Harrassowitz.
- Drimba, Vladimir 1976. La détermination prédicative en turc de Turquie. *Ural-Altaische Jahrbücher* 48, 49-64.
- DS = *Türkiye’de halk ağzından derleme sözlüğü* 1-12. Ankara: Türk Dil Kurumu. 1963-1976.
- Emiroğlu, Kudret 1989. *Trabzon-Maçka etimoloji sözlüğü*. Ankara.
- Gemalmaz, Efrasiyap 1978. *Erzurum ili ağızları* 1-3. (Atatürk Üniversitesi yayınları 487.) Erzurum: Atatürk Üniversitesi basımevi.
- Grunina, Elvira A. 1976. K istorii semantičeskogo razvitiia perfekta -miş. *Sovetskaja Tjurkologija* 1976: 1, 12-26.
- Günay, Turgut 1978. *Rize ili ağızları*. (Kültür Bakanlığı Folklor Araştırma Dairesi yayınları 27.) Ankara.
- Johanson, Lars 1971. *Aspekt im Türkischen*. (Studia Turcica Upsalien-sia 1.) Uppsala: Almqvist & Wiksell.
- Johanson, Lars 1993. Zur Entstehung historischer Präterita im Türkischen. *Türk Dilleri Araştırmaları* 3, 119-127.
- Kononov, A. N. 1956. *Grammatika sovremennogo tureckogo literaturnogo jazyka*. Moscow, Leningrad: Nauka.
- Mackridge, Peter 1985. *The modern Greek language*. Oxford: Oxford University Press.

- Mackridge, Peter 1987. Greek-speaking Moslems of north-east Turkey: Prolegomena to a study of the Ophitic sub-dialect of Pontic. *Byzantine and Modern Greek Studies* 11, 115-137.
- Redhouse yeni Türkçe-İngilizce sözlük. İstanbul: Redhouse. 1968.
- San, Sabri Özcan 1990. *Gümüşhane kültür araştırmaları ve yöre ağızları*. (Kültür Bakanlığı yayınları 1212.) Ankara.
- Schwytzer, Eduard 1959. *Griechische Grammatik* I-III. (Handbuch der Altertumswissenschaft, 2. Abt. 1. Teil.) München: C. H. Becks Verlag.
- Thumb, Albert 1912. *Handbook of the Modern Greek vernacular*. Edinburgh: T. & T. Clark.

Turkish relative clauses: A tale of two participles

Geoffrey Haig

Haig, Geoffrey 1997. Turkish relative clauses: A tale of two participles. *Turkic Languages* 1, 184-209.

Hankamer and Knecht's 1976 account of the conditions governing the choice of participle in Turkish relative clauses has, with some minor modifications, remained the most widely accepted one up to the present. Their account rests on three assumptions: (a) Participle choice is primarily determined by the syntactic function of the target of relativization (TR); (b) the rules governing participle choice can be stated in positive terms for the free participle (i.e. the *-An* participle); (c) it is necessary to invoke a subject incorporation process to account for certain types of relative clauses.

In this paper, a condensed version of chapter 6 from Haig (forthcoming), I claim that all three assumptions are misguided. Some alternative proposals are then suggested which (a) are couched in positive terms for the possessed participle, rather than the free participle; (b) do not necessitate an incorporation process; and (c) ultimately relegate participle choice to an epiphenomenon of a language-specific constraint on subject expression in nominalizations.

Geoffrey Haig, Seminar für Allgemeine Sprachwissenschaft, Universität Kiel, Olshausenstr. 40, 24098 Kiel, Germany. E-mail: haig@anglistik.uni-kiel.de.

0. Introduction

It is well known that there are two types of participle used in Turkish relative clauses: Possessed participles (PP's), and free participles (FP's). PP's obligatorily carry possessive marking indicating the person of the subject of the relative clause. FP's on the other hand may not carry such possessive morphology. A large part of recent research on

relative clauses has concentrated on formulating rules to account for the distribution of FP's and PP's in different types of relative clause.¹

The most widely accepted account of participle choice in Turkish was first formulated by Hankamer & Knecht (1976) and recently reiterated in Barker, Hankamer & Moore (1990). I will refer to these proposals as the "standard account". The standard account is based on three assumptions: (1) It assumes that the syntactic function of the target of relativization (TR) is the primary factor in determining participle choice; (2) it is framed in positive terms for the FP, defining the use of the PP in negative terms; and (3) it rests crucially on the existence of a subject incorporation process.

In this article, I will be arguing that all three assumptions are incorrect, and be suggesting an alternative account. My claims are discussed and motivated in more detail, and with extensive authentic examples, in Haig (forthcoming, ch. 6). Here I will be presenting, in a revised and condensed form, the major conclusions of that discussion. In section 1 of this paper, I will review the standard account and point out several drawbacks inherent to it. In section 2 I will develop some alternative proposals. More specifically, I will be (a) developing an idea of Dede (1978), according to which participle choice is not primarily linked to the form of the relative clause subject; (b) pointing out the inadequacies of the "subject incorporation" hypothesis; and (c) justifying why the rules for participle choice should be formulated in terms of the PP, rather than the FP. In section 3 I will recapitulate the main points of the argument and present some more general conclusions.

1. The standard account of participle choice

The standard account of participle choice is based on three principles, which I will refer to as the Primary Principle, the No-Subject Principle, and the Mother-Node Principle. In this section I will first briefly present

¹ See for example Underhill (1972), Hovdhaugen (1975), Hankamer & Knecht (1976), Dede (1978), Knecht (1979), Erdal (1981), Csató (1985), Nilsson (1985), Zimmer (1987), Johanson (1990: 204-206), Barker, Hankamer & Moore (1990), Kornfilt (1991: 72-78), Sezer (1991: 92-154), Erkman-Akerson & Ozil (1996), Zimmer (1996) and Haig (forthcoming, ch. 6). For an informed discussion of the views of earlier grammarians on this matter, see Erdal (1981).

and exemplify these three principles before going on to point out their major drawbacks.

1.1. The Primary Principle

The apparently fundamental insight on participle choice was formulated by Underhill as follows:

- (1) "The most obvious generalization is that when the head noun is the subject of the underlying sentence, a construction of the *-En* type (= FP) appears, while if the head noun is not the subject, a construction of the *-Dig* type (= PP) appears." (Underhill 1972: 88)

We may term the observation made in 1 the Primary Principle. Notice that the tacit assumption behind it is that participle choice is primarily determined by the syntactic function of the TR.² Let us briefly examine how 1 applies to straightforward examples:

- (2) (Çiçekoğlu 1992⁴: 135)
 (Kuşfür-den yeni çık-mış) iki kadın
 hairdresser-abl just go out-perfFP two woman
 'Two women (who had just come from the hairdresser).'

The head noun, *iki kadın*, is the underlying subject of the verb *çık-* 'go out'. 1 would predict the FP, and indeed this is what we find. Consider now 3:

² In his 1972 paper, Underhill himself actually abandoned this generalization in favor of another, based on word order and case in the syntactic structures from which relative clauses are allegedly derived. His claims, based on Standard Theory (Chomsky [1965] 1988), rest crucially on the assumption that the basic word order in the syntactic structures underlying the relative clause is the same as that in pragmatically neutral surface strings under the condition that all NP's are definite. As it is difficult to justify this assumption on independent grounds, we may safely ignore Underhill (1972) henceforth—see Sezer (1991: 92-104) for an enlightened discussion. In Underhill's grammar ([1976] 1987: 276), the Primary Principle regains preeminence.

- (3) (Pamuk 1993¹⁶: 199)
 (Rüya'nın oku-duğ-u) polisiye roman-lar
 Rüya-gen read-PP-poss3s detective novel-pl
 'The detective novels (which Rüya reads).'

In 3, *polisiye romanlar* 'detective novels' is the underlying direct object of *oku-* 'read', and hence the PP is used, again in accordance with 1.

1.2. The No-Subject Principle and subject incorporation

The Primary Principle fails to correctly account for the data when the relative clause is subjectless, i.e. when the verb is a passivized intransitive. Consider the following example, in which the relative clause verb is the passive form of *gir-* 'enter':

- (4) (sokak-tan içeri gir-il-en) kapı
 street-abl inside enter-pass-FP door
 'door (through which one enters from the street)'

1 would predict the PP rather than the FP in 4, because the TR *kapı* 'door' is a non-subject, a local argument. In fact, it turns out that the FP is almost invariably used with relativization out of a subjectless clause, regardless of the syntactic function of the TR. In other words, 4 is representative of a group of regular exceptions to the Primary Principle.

This fact was noted by Hankamer & Knecht (1976),³ who formulated the following additional principle to account for it:

- (5) The No-Subject Principle:
 "If there is no subject in the RC at the time of RC formation, the (PP)-construction is impossible, and only the (FP)-construction is chosen."
 (Hankamer & Knecht 1976: 132)

A second group of exceptions to 1 is illustrated by 6, taken from Barker, Hankamer & Moore (1990: 26):

³ In fact Lewis (1967: 262) had already noted that the FP is used in subjectless relative clauses. As Underhill (1972) draws heavily on Lewis, it is doubly odd that Underhill's account simply ignores examples such as 4.

- (6) (*bacağın-ı an sok-an*) *kız*
 leg-poss3s-acc bee sting-FP girl
 'the girl (whose leg a bee / some bees stung) [original translation]'⁴

Here again the TR has a non-subject syntactic function in the relative clause, yet, against the prediction made by 1, the FP is used. Hankamer & Knecht (1976) proposed the following solution: They noted that the subjects in relative clauses like 6 were "indefinite". They then suggested that indefinite subjects "undergo a demotion which has the effect of rendering the sentence subjectless" (Hankamer & Knecht 1976: 133). Thus according to Hankamer and Knecht, subjects in examples such as 6 are in fact not subjects at all, and the use of the FP can felicitously be accounted for by the No-Subject Condition 5.

In more recent studies (e.g. Barker, Hankamer & Moore 1990), the terms "indefinite" and "subject demotion" have been replaced by "non-specific and generic" and "subject incorporation" respectively, but the basic insight remains unaltered. In section 2.2 I shall be examining the validity of this analysis in more detail.

1.3. The Mother-Node Principle

The final group of exceptions to 1, also noted by Underhill (1972), concerns relativization of genitive attributes of the subject. We may illustrate this with the following example:

- (7) (*kız-ı ağla-yan*) *kadın*
 girl-poss3s cry-FP woman
 'woman (whose daughter is crying)'
 lit.: '(her-daughter-crying) the woman'

The finite clause corresponding to 7 is as follows:

- (8) *Kadın-ın kız-ı ağlı-yor*
 woman-gen girl-poss3s cry-prog(3s)
 'The woman's daughter is crying.'

⁴ I use this example solely because it is so widely quoted in the literature. It should be emphasized that it is in fact a totally atypical example of the process it is intended to illustrate (cf. detailed discussion in Haig, forthcoming, § 6.3.2.5).

The TR in 7, *kadın* ‘woman’, is not the underlying subject of the relative clause, but the possessor, i.e. a genitive attribute, of the relative clause subject (*kız-ı*). Therefore the Primary Principle 1 would predict the PP, but in 8 the FP is used.

It turns out that the FP is used whenever the TR is a genitive attribute of the relative clause subject. Furthermore, not only do genitive attributes of the subject relativize with the FP, but *any* type of subconstituent of the subject. Consider the following example of relativization out of a clausal subject:

- (9) (Ağaoğlu 1992: 106)
(daha önce çöz-me-m gerek-en) bir şey
 still earlier solve-inf-poss1s be necessary-FP a matter
 ‘a matter / problem (which (I) must solve / sort out first)’

The finite sentence corresponding to 9 has a nominalization as its subject, headed by the possessed infinitive *çöz-me-m*:

- (10) (*Daha önce bir şey çöz-me-m gerek-iyor*)
 still earlier a matter solve-inf-poss1s be necessary-prog(3s)
 ‘It is necessary (that (I) first solve a matter / problem).’
 lit.: ‘(My solving a matter first) is necessary.’

The TR, *bir şey* ‘something’, is not a subject, but the direct object of the possessed infinitive *çözmem* ‘that I solve’. Nevertheless, the FP is still used in 9.

Consider now a more complicated example:

- (11) (Yetiş 1993: 129)
(hangi devir-de yaz-ıl-dığ-ın-ı bil-me-miz)
 which era-loc write-pass-PP-poss3s-acc know-inf-poss1pl

mümkün ol-ma-yan) bu not
 possible be-neg-FP this note
 ‘this note, (of which our knowing in which era it was written is not possible)’

The finite clause corresponding to 11, with the subject NP in brackets, is given in 12:

- (12) *(Bu not-un hangi devir-de yaz-ıl-dı-ğ-ın-ı*
 this note-gen which era-loc write-pass.PP.poss3s.acc
bil-me-miz) mümkün değil
 know-inf.poss1pl possible not(3s)
 ‘(Our knowing in which era this note was written) is not possible.’

Here again the TR (*bu not* ‘this note’) is not a subject, but the genitive attribute of one of the subconstituents of the subject NP. What 7, 9 and 11 illustrate is that, as far as participle choice is concerned, it appears to be irrelevant how deeply embedded in the relative clause subject the TR actually is. In other words, participle choice seems to treat both simple genitive attributes, as in 7, and all other types of subconstituents of the subject alike. Hankamer & Knecht (1976) proposed the following generalization to account for these facts:

- (13) The Mother-Node Principle:
 “If a subconstituent of a major constituent of the RC is relativized, the participle is chosen which would be appropriate for relativization of the major constituent itself.” (Hankamer & Knecht 1976: 127)

Participle choice for the “major constituent itself” is of course determined by the Primary Principle 1. Therefore, because the relativized constituents in the examples 7, 9 and 11 are subconstituents of the relative clause subject, it follows from the Mother-Node Principle that the FP will be used.

Csató (1985) points out some counter-examples to the Mother-Node Principle, and it has since been the subject of considerable debate (see discussion and references in Haig (forthcoming, § 6.2.3.1). Nevertheless, I maintain that the Mother-Node Principle, or something very much like it (an alternative is given in 35 below), is necessary in any account of participle choice. In what follows, I will be assuming that it is valid, and ignoring those types of relative clause which are accounted for by the Mother-Node Principle, more specifically, relative clauses where the TR is some subconstituent of the relative clause subject.

1.4. Summary and critique of the standard account

The standard account for participle choice in relative clauses may be summed up as follows (cf. Barker, Hankamer & Moore 1990: 23 and Kornfilt 1991: 74 for similar summaries):

1. The Primary Principle: The FP is used for relativization of subjects, the PP for non-subjects.
2. The No-Subject Principle: If the relative clause is subjectless at the time of relative clause formation, use the FP.
3. The Mother-Node Principle: When subconstituents of a major clause constituent are relativized, the choice of participle is in accordance with what would be predicted by the Primary Principle for the head of that constituent.

As Kornfilt notes (1994: 74), these three statements have remained largely unchallenged as an explanation for the choice of participles up to the present. She adds however that, “while being accurate generalizations, (they) are not explanatory”. Nor is it clear from what, if any, underlying principle the three conditions can be derived, for they are based on completely heterogeneous criteria: The Primary Principle rests on the syntactic function of the TR; the No-Subject Condition is based purely on the presence or absence of a subject in the relative clause. How the Mother-Node Principle relates to either of the other two is unclear.

On closer inspection, the second two principles appear suspiciously like arbitrary additions, serving the sole purpose of patching up the gaps not covered by the Primary Principle. Given the extent, and the systematic nature of those gaps, one might have expected that the Primary Principle itself be reconsidered, but oddly enough, with the sole exception of Dede (1978), to whom I return below, this option has scarcely been considered.

A further odd characteristic of the standard account is the following: There is no attempt to clearly state the conditions under which the PP is used. Rather, the conditions are stated in positive terms for the FP, and negative terms for the PP (this feature is more pronounced in the presentation of Barker, Hankamer & Moore 1990: 23). Yet, intuitively, one would consider the PP to be the more marked member of the opposition, an assumption I will justify in section 2.4. Therefore, one would expect the conditions to be stated in the reverse manner, i.e. in positive terms for the marked member, the PP, while the unmarked one is considered

the elsewhere case, used when the special conditions for the use of the marked member are not met.

Finally, the No-Subject Condition necessitates that certain types of errant subject, e.g. the subject in 6, be explained away via an incorporation process. But, as I will point out in section 2.2, there are no independent grounds for assuming such a process in Turkish.

2. Some alternative proposals

2.1. Collapsing the Primary Principle and the No-Subject Principle

Let us first consider two simple examples of FP-constructions:

- (= 4) *(sokak-tan içeri gir-il-en) kapı*
 street-abl inside enter-pass-FP door
 'door (through which one enters from the street)'

- (14) *(ev-e dön-en) kız*
 house-dat return-FP girl
 'girl (who is returning home)'

The standard account invokes two separate principles to account for the use of the FP in 4 and 14: The FP in 4 is accounted for by the No-Subject Principle 5, which states that the FP is used when the relative clause is subjectless "at the time of relative clause formation". As the verb is an intransitive passive, the clause is of course subjectless. The use of the FP in 14 on the other hand is explained via the Primary Principle, which states that the FP is used when the TR is subject. This way of looking at things suggests that the two examples above have nothing in common, and the use of the FP in each case is motivated by two quite unrelated principles.

But there is another, and to my mind simpler approach: Neither of the relative clauses in 4 and 14 contains a surface subject, and indeed, none could be supplied without impairing grammaticality. Of course the respective *sources* of the subjectlessness in 4 and 14 are quite different: The relative clause in 4 is genuinely subjectless, because its predicate is a passivized intransitive, which is always subjectless in Turkish. We might term this *deep subjectlessness*. 14 on the other hand is subjectless because the TR is subject, and is hence deleted from the relative clause in the process of relative clause formation. But suppose participle choice

were not sensitive to the source of any subjectlessness, i.e. suppose participle choice were not a deep-level phenomenon in relative clause formation at all, but a fairly trivial surface process. If that were the case, we would not need to distinguish between the deep subjectlessness of 4 and the subjectlessness of 14, and we need not indulge in any conjecture about subjectlessness “at the time of RC-formation”, but simply rely on the criterion of surface subjectlessness. If we make that assumption, we can account for both with one and the same principle, which we may provisionally formulate as follows:

- (15) Use the FP when the relative clause is, for whatever reason, subjectless, i.e. a subject cannot be supplied without impairing grammaticality.

The immediate advantage of this approach is that we cover both 4 and 14 with a single principle, whereas the standard account needs two principles, the Primary Principle and the No-Subject Principle. 15 effectively makes the same statement as the No-Subject Principle, except that we dispense with any reference to subjectlessness “at the time of relative clause formation”. Notice that 15 makes no reference whatsoever to the syntactic function of the TR. In other words, we have rendered the Primary Principle redundant.

It should, however, be noted that we would need to modify 15 to cover examples where the TR is a sub-constituent of the subject, which I discussed in 1.3. Such an additional stipulation is perfectly feasible, and is briefly discussed in section 3. I will not be considering such cases further here.

2.2. Subjects, no subjects and semi-subjects

Our rule 15 binds participle choice to a single factor, namely the presence or absence of a subject in the relative clause. Now we have already encountered examples where the relative clause does contain a subject, but the FP is found, thereby violating 15. Consider for example 6:

- (= 6) (*bacağ-in-ı arı sok-an*) *kız*
 leg-poss3s-acc bee sting-FP girl
 ‘the girl (whose leg a bee / some bees stung) [original translation]’

The standard account offers a neat solution to this dilemma: Subjects such as *arı* ‘bee’ in 6 are said to be incorporated, hence the relative

clause contains no subject. This solution is unquestionably elegant, and as long as the examples contain the kind of bare noun subject found in 6, it appears to be quite plausible. But once the full range of data is considered, it appears considerably less so. Consider the following two authentic examples:

- (16) (Nesin 1995: 89)

(*Arkeolojik kazı-lar yap-ıl-an*)
 archaeological excavation-pl make-pass-FP

bir bölge-ye gel-di-k.

an area-dat come-pst-1pl

‘(We) arrived at an area (in which archaeological excavations were being carried out).’

- (17) (Pamuk 1995⁹: 524)

Ömer (iç-in-de kocaman bir soba yan-an)
 Ömer inside-poss3s-loc huge an oven burn-FP

geniş bir oda-da bir satranç sorunu çöz-üyor-du.

spacious a room-loc a chess problem solve-prog-pst(3s)

‘Ömer was in a spacious room (in which a huge stove was burning), solving a chess problem’

We notice that the FP is also used in the relative clauses in 16 and 17. There is only one possible explanation for this in terms of the standard account: The subject in 16, *arkeolojik kazı-lar* ‘archaeological excavation+pl’, and in 17, *kocaman bir soba* ‘a huge stove’, must have undergone incorporation.

This is a conclusion that few syntacticians would feel comfortable with. First of all, the subject in 16 has plural marking, while that in 17 has an indefinite article. Hopper & Thompson (1984: 711) state quite clearly that an incorporated noun “invariably loses the ability to take determiners and inflections”. I would like to note that this pattern is by no means unusual—in Appendix 3 of Haig (forthcoming) there are 23 authentic examples of such subjects with plural marking or articles. Nor is it possible to demonstrate by any type of independent syntactic test known to me that the subjects in these sentences have lost their subject status, i.e. that the clauses concerned are genuinely subjectless. Finally,

even in Baker's (1988) broader conception of noun incorporation, the possibility of incorporating a transitive subject is categorically excluded (1988: 81). Yet that is what proponents of the incorporation analysis are claiming for 6. Baker himself (1988: 452, fn. 8) is doubtful whether subject incorporation in Turkish qualifies as noun incorporation at all.

The only conclusion I can draw from these facts is that we must reject the incorporation analysis. One might, if one wished to salvage it at all costs, postulate two different processes: Noun incorporation, which only affects bare nouns, and some other type of looser "compounding by juxtaposition" (Reuse 1994: 2844). This would, I think, unnecessarily complicate matters: Surely we are dealing with one and the same process, namely a gradual and subtle loss of syntactic autonomy, the most prominent reflection of which is loss of genitive case marking. But the affected NP's (not just nouns) lose neither their argument status, nor their status as phonetic words. The process is not restricted to bare, generic and nonspecific nouns, but permeates to affect referential, quite elaborated, but usually indefinite NP's. The term "noun incorporation" is misleading when applied to Turkish because it implies that the affected entities are noun roots, whereas in Turkish they are NP's, and because it implies a greater degree of phonetic coalescence and loss of argument status than is justified by the Turkish data. I suggest that a more appropriate term would be *case stripping*.⁵

Rejecting the incorporation analysis leaves us with a terminological problem: What are we to call the genitiveless subjects in examples like 16 and 17? They are subjects, but not subjects enough to take genitive marking. "Genitiveless subjects" would be possible, but is confusing if we wish to talk about finite clauses, where subjects are normally nominative. I suggest therefore the term *semi-subjects* for those subjects of relative clauses (and of course of other types of nominalization, for example, complement clauses) which do not take genitive marking. "Semi-subject" is merely a convenient and relatively innocuous label, which avoids the unfortunate connotations of the term "incorporated subject". The term "semi-subject" is intended to convey the fact that such "subjects", while being ignored in terms of genitive marking, retain sufficient

⁵ The same term can of course be applied to the loss of accusative marking on direct objects, for which I also feel that the term "noun incorporation" is misguided (cf. Haig forthcoming, § 6.2.5 for a justification of this view).

syntactic substance to enable them to take plural marking, articles and nominal modifiers. Semi-subjects may of course occur in finite clauses as well, but in that environment, the difference between a semi-subject and a subject is not expressed by any segmental means, but by a loss in word order freedom and a shift in stress pattern.⁶ I will not be going into the semantic and pragmatic factors which trigger case stripping of relative clause subjects here—a detailed investigation of those factors may be found in Haig (forthcoming, § 6.3).

2.3. The Genitive Subject Condition

So far we have concentrated on identifying the conditions under which the FP is used. We have established that it is used when the relative clause is subjectless, or when it contains a semi-subject. These two environments are given in 18 and 19 respectively:

- (18) (... FP) head noun
 (19) (subj.+Ø FP) head noun

Now, interestingly, it turns out that in both of these environments, the PP is also possible. Lewis (1967: 262) quotes an example of a subjectless relative clause, i.e. corresponding to 18, with a PP:

- (20) (*normal-e dönül-düğ-ü*) *bir sıra-da*
 normal-dat return-pass-PP-poss3s a time-dat
 ‘at a time (when things were returning to normal)’

Johanson (1990: 213-214) also suggests that in subjectless relative clauses, both the FP and the PP are “in principle” possible.

The PP also cooccurs with semi-subjects, a fact that is pointed out in Nilsson (1985: 79) and in Erdal (1981: 33). Erdal quotes the following example with the semi-subject *dondurma* ‘icecream’:

- (21) (*dondurma sat-ıl-dığ-ı*) *yer*
 icecream sell-pass-PP-poss3s place
 ‘place (where icecream is sold)’

⁶ See Dede (1986: 153-154) on the shift in stress patterns (she, however, talks of “subject incorporation”).

There is no doubt that constructions such as 20 and 21 are marginal in modern Turkish—in the corpus of over 1,000 relative clauses used in Haig (forthcoming), there was nothing corresponding to them. They do however illustrate that the possibility of using the PP in these environments cannot be excluded completely. It may be then that the use of the FP in these environments should not be considered the result of a strict syntactic constraint, but simply evidence of an extremely strong tendency. I will explore this possibility further below.

Let us now turn to the PP. The single environment in which the PP must occur is quite straightforward: It is always required when the relative clause contains a genitive-marked subject, or when such a subject is suppliable without impairing grammaticality. We may represent this schematically as follows:

- (22) (subj.+gen ... PP) head noun

In the environment 22 the FP never occurs. As this is the only type of environment where the rule for participle choice admits no exceptions, it would make more sense to base our account of participle choice on this environment, and frame it in terms of the PP rather than the FP. Thus we reformulate 15 in terms of the PP rather than the FP to yield the following rule:

- (23) The Genitive Subject Condition:

When the subject of the relative clause takes genitive marking, the PP is used and the FP is impossible. Elsewhere, the FP is always possible and vastly preferred.

23 accounts for the same set of data that the Primary Principle and the No-Subject Principle of the standard account cover. Apart from being a more economical account, it is also preferable in other respects: Firstly, it obviates the necessity for assuming an incorporation process, because it is based on the criterion of presence or absence of genitive marking, not of a subject. Secondly, it admits the possibility of the PP occurring in environments such as 20. Thirdly, it is formulated in positive terms for the PP, whereas the FP is considered the elsewhere or default case. This is certainly correct, for reasons I will elaborate on in section 2.4.

I should hasten to add that a similar conclusion was reached by Dede (1978). She also rejected the standard account, and proposed instead a

“Genitive Suffix Attachment Rule” (GSAR), which conditioned the form of the subject. If the subject is genitive marked, the PP occurs, otherwise the FP. Thus for her the problem was reduced to the question “When does the GSAR apply?” (1978: 69). Up to this point, I share her conclusions entirely.

Unfortunately, Dede proceeded to explain the use of the genitive on subjects of nominalizations by appealing to its function as an indication of grammatical relations. This led her to the claim that the genitive suffix “does not apply unless it is needed to indicate the grammatical relation of the subject in the embedded sentence” (1978: 73). This claim is not tenable, for reasons that cannot be discussed here. Furthermore, Dede provided no viable alternative to Hankamer and Knecht’s Mother-Node Principle 13, which is, as I have stated, essential in any account of participle choice. The flaws in Dede’s argument were promptly pointed out in Knecht (1979), and as a result, Dede’s basically correct insight was ignored in subsequent research, which continued to work in the framework of the standard account.

2.4. The possessed participle as the marked member of the participle opposition

In Haig (forthcoming, ch. 6), I argue that the PP is the marked member of the participle opposition in relative clauses. I will briefly sum up the arguments in favor of that position here:

On purely formal grounds, the PP is the more marked member of the opposition, as it is the morphologically more complex, involving two morphemes, participle and possessive, rather than one (cf. also Zimmer 1996: 162-163). Further facts from language usage, and from related Turkic languages, point in the same direction:

1. PP’s are as tokens in texts (in attributive function), the rarer construction: Of the total participles in the corpus of over 1,000 relative clauses used in Haig (forthcoming), only 29% were PP’s.⁷

⁷ This may of course simply reflect the rather trivial fact that subjects are more commonly relativized on than non-subjects, as claimed for example in Keenan (1975) on the basis of written English data. However, the total number of subject relativizations does not exceed 50% in Keenan’s data, and in material from spoken data (e.g. Fox 1987 and Slobin 1986), it decreases.

2. In L1 acquisition, PP's are acquired later and with greater amounts of errors (Slobin 1986).⁸
3. The PP in relative clauses is seldom found in the Turkic languages outside of Turkish, and emerged later in Turkish itself.

It is important to note that my claim that the PP is the marked member of the opposition is restricted to the occurrence of the two forms in one particular syntactic function, namely as the verb of a relative clause. PP's of course occur in another syntactic function, namely as complement clauses. I would suggest that this is in fact their primary function, and their usage in relative clauses (which are, in terms of syntactic function, adnominal attributes) is an intrusion into a functional domain for which they are not suited. The FP on the other hand is a genuine verbal adjective, predestined to occur in attributive function—in fact the sole function of the FP is to create relative clauses. In many Turkic languages the etymological equivalent of the FP is found in all types of relative clause. Only in Turkish and Azerbaijani do we find the PP used in relative clauses at all.⁹

The claim that the PP is the marked member of the opposition in relative clauses leads us to the question of why the PP should be used in Turkish in relative clauses at all. Why cannot Turkish, like many other Turkic languages, get by with one and the same participle, the FP, for all types of relativization? In the next section I propose an answer to this question.

2.5. The function of the possessed participle in relative clauses

First of all, let us compare an FP and a PP construction. We can display the two types of relative clause schematically as follows:

- | | | |
|------|------------------------------------|-----------------|
| (24) | ((semi-subject) FP) head noun | FP-construction |
| (25) | (subj.+gen ... PP+poss.) head noun | PP-construction |

⁸ According to Zimmer (1987: 59-60) the tendency to use the FP in contexts where the PP would be expected is even more widespread in the spoken language, particularly of younger speakers, i.e. there may be a diachronic shift towards increasing the domain of the less marked FP at the cost of the more marked PP.

⁹ Cf. Schönig (1997) and Csató (1996).

What is it that the PP-construction achieves which the FP-construction does not? The obvious answer is that the PP carries possessive marking, thereby enabling the person of the subject to be identified.

Assuming then that the principle difference lies in the presence of possessive marking on the participle, the question arises as to why Turkish should need an additional participle, the sole purpose of which is to act as a platform for that possessive morphology. In other words, why is it not possible for one and the same participle to be used both with and without possessive morphology? We could reasonably expect to find, instead of 25, something like 26:

- (26) *(subj+gen ... FP+poss.) head noun

But 26 is, as Sezer (1991: 120) puts it, “ruled out”. The question as to why 26 is impossible in Turkish is by no means trivial, especially in view of the fact that this type of construction is attested in other Turkic languages. Consider the following examples from East Middle Turkic (both examples from Eckmann 1959: 126):

- (27) (*oltur-yan-ım*) *yär-dä* *monçuk-lar-ım-nı* *bul-du-m*
 sit-FP-poss1s place-loc pearl-pl-poss1s-acc find-pst-1s
 ‘(I) found my pearls at the place ((I) had been sitting)’

- (28) (*äſit-gän-im*)-ni *unut-ma-dı-m*
 hear-FP-poss1s-acc forget-neg-pst-1s
 ‘(I) did not forget ((what) (I) had heard)’

In East Middle Turkic the same participle, the etymological equivalent of the Turkish FP, is also used for all types of relativization. In other words, there is no alternation between different types of participle, but simply an alternation [\pm possessive] on one and the same participle.

The reason that Turkish does not permit structures such as 26 is not in fact directly related to relative clause-formation at all, but lies in more general constraints on nominalizations in Turkish. In Turkish, all verbal nominals, i.e. infinitives and participles, must be characterizable as either possessed or non-possessed, but not both (see Haig, forthcoming, ch. 3 for detailed discussion). A possessed verbal nominal permits expression of its underlying subject via possessive morphology, a non-possessed one does not. For example, the infinitives in *-mA* and *-(y)İſ* are pos-

sessed verbal nominals, the infinitive in *-mAk* is a non-possessed verbal nominal.

The FP is a non-possessed verbal nominal, a fact which can be illustrated with examples such as 29:

- (29) *ilk gel-en-imiz*
 first come-FP-poss1pl
 'our first visitors (lit. our first comers)'

The possessive marking on the FP in 29 does not, and *cannot*, refer to the underlying subject of *gel-* 'come'. Sezer (1991: 120) reaches the same conclusion: With the FP "AGR [agreement] cannot govern anything inside the RC".

The PP on the other hand is a possessed verbal nominal. The fundamental difference between the two types of participle is illustrated in 30, which contains a PP and an FP, both with 1s possessive marking:

- (30) *sev-diğ-im başka, sev-en-im başka*¹⁰
 love-PP-poss1s different(3s) love-FP-1s different(3s)
 'he / she that I love is one person, he / she that loves me is another'

The FP+1s possessive marking *sev-en-im* means literally 'my lover', i.e. 'the one that loves me'. Thus the 1s possessive marking indicates the underlying object of the verb *sev-* 'love'. Crucially, possessive marking on an FP can *never* refer to the underlying subject of the participle. The 1s possessive marking on the possessed participle *sev-diğ-im* 'he / she that I love' on the other hand can *only* refer to the person of the underlying subject.

The reason why structures such as 26 are not possible in Turkish is simply that the FP, as a non-possessed verbal nominal, is incapable of supporting possessive morphology as an indication of its underlying subject. Thus the answer to the question: Why is the PP necessary in Turkish relative clauses at all? is: The PP is necessary as a platform for possessive marking identifying the subject of the relative clause, some-

¹⁰ Taken from the song *Düşler Sokağı* from the cassette entitled *Oyun* by the group Ezginin Günlüğü. I am grateful to Friederike Braun for drawing my attention to this example.

thing which the FP is simply incapable of supporting. This feature is rooted in a broader characteristic of Turkish verbal nominals generally, and has nothing to do with the syntactic function of the TR in relative clauses.

We have now established why the PP is necessary in some relative clauses: It is necessary as a platform for possessive morphology as an indication of the subject of the nominalized relative clause. Whenever it is not necessary to indicate the person of the relative clause subject (because there is no genitive-marked subject in the relative clause), the PP does not normally occur. In relative clauses, then, there appears to be a constraint which prevents PP's from occurring when the relative clause does not contain a genitive-marked subject. This is rather an odd fact, for if we look further afield at other types of nominalizations where PP's also occur, for example complement clauses, we discover that there the PP occurs freely with and without a genitive marked subject. Consider the following examples:

- (31) (Riemann 1990: 108)

(yatak-tan biraz önce çık-ıl-mış
 bed-abl little before get out-pass-perf

ol-duğ-u) belli-ydi
 be-PP-poss3s clear-pst(3s)

'It was clear (that someone had got up out of the bed a short time ago).'
 lit.: 'It was clear (that out-of-the-bed a little earlier had been got out).'

In 31, the PP heads a genuinely subjectless complement clause. In 32, the PP occurs with a semi-subject, *bomba*:

- (32) (Pamuk 1995⁹: 55)

(Abdülhamit'e bomba at-ıl-acağ-ı)
 Abdülhamit-dat bomb throw-pass-futPP-poss3s

kim-in akl-ın-a gel-ir-di?
 who-gen mind-poss3s-dat come-aor-pst(3s)

'Who could have imagined (lit.: 'to whose mind would have come ...') (that a bomb would be thrown at Abdülhamit)?'

31 and 32 illustrate that there is no fundamental constraint in the grammar that prevents a PP from heading a nominalized clause which does not contain a genitive-marked subject. Yet in relative clauses, just such a constraint is operative: Lack of a genitive-marked subject in the nominalized clause (overwhelmingly) results in a shift from one type of verbal nominal, the PP, to another, the FP.

My explanation for this fact is related to my earlier claim that in relative clauses (but nowhere else in the grammar), the PP is the more marked verbal nominal. Therefore, when the strict conditions necessitating its use are relaxed, the unmarked FP regains its position. Elsewhere, for example in complementization, the PP is the unmarked verbal nominal, and is not subject to any restrictions. It therefore occurs freely in subjectless nominalizations, such as 31. These facts also help to explain why the PP is occasionally tolerated in examples such as 20 and 21: The use of the FP in this environment is not the result of a strict syntactic constraint, but of a markedness condition, which may occasionally be violated.

3. Conclusions

My account differs from the standard account primarily in that I abandon the Primary Principle, according to which the syntactic function of the TR is the main factor in determining participle choice. To be sure, there is a rough correlation between subject and non-subject roles of the TR on the one hand and the use of the FP and PP on the other. But this does not necessarily justify assuming a causal relationship between the syntactic function of the TR and the form of the participle.

What participle choice in Turkish is all about is an alternation between constructions with and without possessive morphology. This emerges most clearly when we turn our attention to other Turkic languages. In many Turkic languages, for example Uzbek, one and the same participle, the etymological equivalent of the FP, can be used for all types of relativization. Consider the following two examples from Uzbek:

- (33) (Jarring 1938: 54)
 (aqajāt-kan) su (...) az waχt toχta-p qal-di
 flow-FP water little time stop-ger stay-pst(3s)
 ‘the (flowing) water stopped briefly’

The TR *su* 'water' is clearly the underlying subject of the relative clause verb. We find, as we would in Turkish, the FP (or rather its etymological equivalent). 34 on the other hand is an example of non-subject relativization, but again we find the same participle used as in 33:

- (34) (Jarring 1938: 145)
- | | | | |
|------------|----------------|----------------|-------------------|
| <i>sen</i> | <i>agar'da</i> | <i>(uşa-ni</i> | <i>bujur-gan)</i> |
| you | and if | he-gen | order-FP |
-
- | | | |
|---------------|--------------|-----------------|
| <i>jol- i</i> | <i>bilen</i> | <i>ket-sa-ŋ</i> |
| way-poss3s | with | go-cond.2s |
- 'And if you go on the road (which he ordered)'

The sole structural difference between 33 and 34 is the presence of the possessive marking on the head noun *jol-i* 'road+poss3s' in 34. That in itself is motivated by the genitive-marked subject of the relative clause, the pronoun *uşa-ni* 'he+gen'. One could of course proceed to formulate a rule such that the head noun takes possessive marking when it is in some non-subject role in the relative clause, but that would surely be a clumsy and indirect way of describing the structure. The more direct way is to link the possessive marking on the head noun to the presence of a genitive-marked subject in the relative clause.

The fact that there is a participle choice in Turkish at all can be attributed to the fact that (a) Turkish verbal nominals (of which the participles are a subset) are strictly specified according to whether or not they can express their underlying subject via possessive morphology; and (b) in Turkish, the possessive marking indicating the subject of the relative clause affixes to the participle, not the head noun. Thus what surfaces in other Turkic languages as an alternation [\pm possessive morphology] somewhere in the relative construction (usually on the head noun itself) turns up in Turkish as an alternation in the type of participle.

My account also differs from the standard account in that I frame the rule for participle choice in positive terms for the PP, while I consider the FP to be the default participle in relative clauses. If the FP is the default case, then we would expect to find it used in a variety of seemingly disconnected functions, and indeed this is the case (e.g. with subjectless relative clauses, when the head noun is a subconstituent of the relative clause subject etc.) Thus the standard account, which is framed in terms of the FP, is faced with the difficult task of finding a common denomi-

nator among quite diverse functions. It makes more sense to frame one's account in terms of the marked member of the opposition, the PP, because we would expect that its usage is bound to a more restricted set of conditions.

Finally, my proposals differ from the standard account in that I dispense with the incorporation hypothesis, for reasons outlined in section 2.2. Rather, I consider that there is a broader process at work, which I termed case stripping. Under case stripping, the subject of a nominalization (not just of a relative clause) loses its genitive marking, but it retains its syntactic status as a subject, and its status as a phonetic word. I termed such subjects semi-subjects.

My proposals may be summed up as follows: The FP and the PP are in complementary distribution in the function "verb of a relative clause". The FP is the unmarked, the PP the marked member of the opposition.

The basic principle governing participle choice is expressed in the Genitive Subject Condition:

When the subject of the relative clause takes genitive marking, the PP is used and the FP is impossible. Elsewhere, the FP is always possible and vastly preferred.

The Genitive Subject Condition itself is motivated by a general feature of Turkish nominalizations: Only certain types of verbal nominals are licensed to allow subject expression via possessive morphology. The FP does not allow this. Therefore, when subject identification via possessive morphology is necessary, a possessed verbal nominal, in this case the PP, steps in to do the job. But the use of the PP in relative clauses remains an intrusion into a domain for which it is not specialized. Whenever the conditions requiring its presence are relaxed, the unmarked FP regains its position.

The Genitive Subject Condition covers the same data as the Primary Principle and the No-Subject Principle of the standard account. But it is not a complete account of participle choice, as it does not account for those cases covered by the Mother-Node Principle of the standard account. We could augment the Genitive Subject Condition with a constraint along the lines of 35, which would enable the present account to cover the same set of data as the standard account:

(35) The Subconstituent of Subject Condition:

When the TR is any subconstituent of the relative clause subject, then the relative clause subject does not take genitive marking.¹¹

As I have stated above, there is no ready explanation for the Subconstituent of Subject Condition. Until one is forthcoming, we must simply accept it as an empirically adequate description.

In my view, progress in the analysis of Turkish relative clauses has been severely hampered by the following assumption, which is so firmly entrenched in traditional lore on Turkish relative clauses that it has virtually never been seriously questioned:

“In Turkish, as in English, the form of the relative clause is determined by the grammatical role of the head noun in the included sentence; in particular, this determines the choice between the subject and the object participles.” (Underhill 1987: 276)

The alternative proposals outlined here are framed in terms of readily observable surface features; I have made no attempt to postulate deeper syntactic structures from which the relative clause is derived. Nor do I assume that the Turkish relative clause is amenable to the same kind of structural analysis as, say, the English relative clause. The only independent evidence for such an assumption appears to me to be the rough translational equivalence of the two structures.

Sources used in the examples:

- Ağaoğlu, A. 1992. *Ölmeye yatmak*. İstanbul: Simavi Yayınları. [1973].
 Çiçekoğlu, F. 1992⁴. *Sizin hiç babanız öldü mü? (Öyküler)*. İstanbul: Can. 132-135
 Jarring, G. 1938. *Uzbek texts from Afghan Turkestan*. Leipzig: Harrassowitz, Lund: Gleerup.
 Nesin, A. 1995. *Sizin memlekette eşek yok mu? (Aziz Nesin'in Aziz Nesin'den seçtikleri)*. İstanbul: Doğan Yayın Holding.
 Pamuk, O. 1993¹⁶. *Kara Kitap*. İstanbul: Can.
 Pamuk, O. 1995⁹. *Cevdet Bey ve Oğulları*. İstanbul: İletişim.

¹¹ See Csató (1985) and Zimmer (1996) on some marginal exceptions to this rule.

- Riemann, W. (ed.) 1990. *Çağdaş Türk Öyküleri – Zeitgenössische türkische Erzählungen*. München: Deutscher Taschenbuchverlag.
- Yetiş, K. 1993. Edebiyat nazariyesi kitaplarında Namık Kemal'in eserlerinin örnek olarak değerlendirilişi. In: *Doğumunun yüzellinci yılında Namık Kemal* (= Atatürk Kültür Merkezi Yayını 67; Türk Fikir ve Sanat Adamları Dizisi 8). Ankara: Türk Tarih Kurumu basımevi. 117-136.

References

- Baker, M. 1988. *Incorporation: A study of grammatical relations changing*. Chicago, London: University of Chicago.
- Barker, C. & Hankamer, J. & Moore, J. 1990. Wa and Ga in Turkish. In: Dziwirek, K. & Farell, P. & Mejías-Bikandi, E. (eds.) *Grammatical relations: A cross-theoretical perspective*. Stanford, California: Center for Study of Language and Information. 21-43.
- Chomsky, N. 1988. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press. [1965].
- Csató, É. Á. 1985. A syntactic analysis of participial constructions in modern Turkish. In: *Beşinci Milletler Arası Türkoloji Kongresi*, 23-28.09.1985. Tebliğler 1: Türk Dili, Cilt 1. İstanbul: Edebiyat Fakültesi Basımevi. 39-56.
- Csató, É. Á. 1996. A typological review of relative clause constructions in some Turkic languages. In: Rona, B. (ed.) 1996, 28-32.
- Dede, M. 1978. Why should Turkish relativization distinguish between subject and non-subject head nouns? *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 4. 67-78.
- Dede, M. 1986. Definiteness and referentiality in Turkish verbal sentences. In: Slobin, Dan I. & Zimmer, K. (eds.), 147-163.
- Eckmann, J. 1959. Das Chwarezmtürkische. In: Deny, J. & Grønbech, K. & Scheel, H. & Togan, Z. V. (eds.) *Philologiae turcicae fundamenta* 1. Aquis Mattiacis: Steiner. 113-137.
- Erdal, M. 1981. Turkish participles and the absence of reference. In: Young, D. (ed.) *Studies presented to H. J. Polotsky*. East Gloucester, Mass.: Pirtle & Polson. 21-49.
- Erkman-Akerson, F. & Ozil, Ş. 1996. /en/ and /dıǵı/ crossings—genitival NP's and sentential clauses as relative clause subjects. In: Konrot, A. (ed.) *Modern studies in Turkish linguistics. Proceedings of the Sixth International Conference on Turkish Linguistics*. Eskişehir, 12-14 August, 1992. Eskişehir: Anadolu University. 1-12.

- Fox, B. 1987. The noun phrase accessibility hierarchy reinterpreted: Subject primacy or the absolutive hypothesis. *Language* 63 (4), 856-870.
- Haig, G. forthcoming. *Relative constructions in Turkish*. Wiesbaden: Harrassowitz.
- Hankamer, J. & Knecht, L. 1976. The role of the subject / non-subject distinction in determining the choice of relative clause particle in Turkish. In: *Papers from the Sixth Meeting of the North Eastern Linguistics Society*, 31.10-2.11.1975. (= Montreal Working Papers in Linguistics Vol. 6, May 1976) McGill University. 123-135. [Also published in: Aissen, J. & Hankamer, J. (eds.), *Harvard Studies in Syntax and Semantics* 2 (1976).]
- Hovdhaugen, E. 1975. Relative clauses in Turkish. 1. *Türk Dili Bilimsel Kurultayına sunulan bildiriler* 1972. (Türk Dil Kurumu yayınları 413). Ankara: Türk Dil Kurumu basımevi. 551-554.
- Johanson, L. 1990. Subjektlose Sätze im Türkischen. In: Brendemoen, B. (ed.) *Altaica Osloensia. Proceedings of the 32nd Meeting of the Permanent International Altaistic Conference. 12-16 June, Oslo. 1989*. Oslo: Universitetsforlaget. 193-218.
- Keenan, E. 1975. Variation in universal grammar. In Fasold, R. & Shuy, R. (eds.) *Analyzing variation in language*. Washington: Georgetown University. 136-148.
- Knecht, L. 1979. The role of the genitive suffix in relative clauses: A reply to Dede. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 5. 180-197.
- Kornfilt, J. 1991. Some current issues in Turkish syntax. In: Boeschoten, H. & Verhoeven, L. (eds.) *Studies on modern Turkish. Proceedings of the Third Conference on Turkish Linguistics*. Tilburg: Tilburg University. 60-92.
- Lewis, G. 1967. *Turkish grammar*. Oxford: Oxford University Press.
- Nilsson, B. 1985. *Case marking semantics in Turkish*. [Ph.D. dissertation, Stockholms Universitet.]
- Reuse, W. de 1994. Noun incorporation. In: Asher, R. & Simpson, J. (eds.) *The encyclopaedia of language and linguistics*. (10 vols.) Oxford: Pergamon. 2842-2847.
- Rona, B. (ed.) 1996. *Current issues in Turkish linguistics. Proceedings of the Fifth Annual Conference on Turkish Linguistics*. School of Oriental and African Studies, 15-17 August, 1990. Ankara: Hitit Yayınevi.

- Schönig, C. 1992 / 1993. Relativsatzbautypen in den sogenannten altaischen Sprachen. *Acta Orientalia Academiae Scientiarum Hungaricae* 46, 327-338.
- Schönig, C. 1993. Finitprädizierte Relativsätze in Sprachen des altaischen Typs. *Türk Dilleri Araştırmaları* 3, 181-191.
- Schönig, C. 1997. A new attempt to classify the Turkic languages 2. *Turkic Languages* 1, 263-278.
- Sezer, E. 1991. *Issues in Turkish syntax*. [Ph.D. dissertation, Harvard University.]
- Slobin, D. & Zimmer, K. (eds.) 1986. *Studies in Turkish linguistics*. Amsterdam: Benjamins.
- Slobin, D. 1986. The acquisition and use of relative clauses in Turkic and Indo-European languages. In: Slobin, D. & Zimmer, K. (eds.) 1986, 273-294.
- Underhill, R. 1972. Turkish participles. *Linguistic Inquiry* 3, 87-99.
- Underhill, R. 1987. *Turkish grammar*. Cambridge, Mass.: MIT Press. [1976].
- Zimmer, K. 1987. Turkish relativization revisited. In: Boeschoten, H. & Verhoeven, (eds.) *Studies on Modern Turkish. Proceedings of the Third Conference on Turkish Linguistics*. Tilburg: Tilburg Press. 57-61.
- Zimmer, K. 1996. Overlapping strategies in Turkish relativization. In: Rona, B. (ed.) 1996, 159-164.

The first *Altınkök* inscription

Talât Tekin

Tekin, Talât 1997. The first *Altınkök* inscription. *Turkic Languages* 1, 210-226.

To date, the first *Altınkök* inscription has been studied and edited by Radloff, Orkun, Malov, Kljaštornyj, Vasil'ev and Mori. In their editions some misread words and phrases have been corrected. Yet, they still contain a number of misread and misconstrued words and phrases. In this paper, many new readings and interpretations are offered: *yirildim* 'I was separated', *qaymatın* 'without turning away', *ilâz* 'pain, sorrow', *atsar alp ârtiğiz i tutsar küç ârtiğiz ä* 'O, you were tough in throwing (arrows) and you were powerful in capturing (the enemies)', *inilig bört* 'young wolf with younger brothers', *oça bars* 'little tiger', *botomuz* 'our camel colt', *sâçlinmä* 'don't be parted', *bars täğim ä* 'O, my tiger-like one!', *altun soğa yış* 'the Altay and Songa mountains', *art oğul taş oğul* '(my) last son, my son abroad', *tad çına barsım* 'my young and little tiger', etc.

Talât Tekin, Gülden Sokak 16/10, Kavaklıdere, Ankara, TR-06690 Turkey.

The two *Altınkök* monuments were discovered in 1878 by Korčakov and Markov in an old graveyard near the small *Altınkök* Lake which is located on the right of the Abakan River, ten km away from Bondarevo. In 1881, the two monuments were brought to the Minusinsk Museum by Mart'janov and registered under the numbers 27 and 28.

The steles are of brown sandstone, rectangular in shape, the upper parts of which are curved.

The measurements of the first monument are 136.54 x 43.5 x 25 cm. It contains nine lines written vertically on the three sides of the monument, the lines being separated from one another by raised furrows. On the front side, the lines run in a U-shape (Radloff 1895: 332; Kljaštornyj 1976: 258-259; Vasil'ev 1983: 25).

The first *Altınkök* inscription was studied and published by Radloff, Orkun, Malov, Kljaštornyj, Vasil'ev and Mori.

It has not been easy to determine the right order of the sides and lines of the inscription. Radloff, Orkun and Malov wrongly followed the front-right-left sides sequence. According to Kljaštornyj and Vasil'ev the appropriate order of the sides is left-front-right, which, I believe, is correct. Mori, on the other hand, preferred the order right-front-left side.

As for the order of the lines on the right side, Radloff, Orkun, Malov and Mori read them in reverse order, i.e. from the bottom to the top. Kljaštornyj and Vasil'ev, on the other hand, read these lines correctly, i.e. from the top to the bottom.

Furthermore, Vasil'ev and Kljaštornyj read the lines on the front side in the following order: The top outer line + the bottom outer line + the top inner line. In my opinion, the true order of the lines on the front side is as follows: The bottom outer line + the top inner line + the top outer line.

Transcription of the text

- L1. on (a)y : iltđi : ög(ü)m ä : k(ä)l(ü)rti : il(i)mkä : (ä)rd(ä)m üč(ü)n : m[(ä)n] y(ı)r(i)ld(i)m [ä]
 L2. (e)l(i)m ök(ü)nč(i)ñä : q(a)l(ı)n : y(a)γ(ı)qa : q(a)ym(a)t(i)n : t(ä)g(i)p(ä)n : (a)dr(ı)ld(i)m a : y(ı)ta
 L3. in(i)ñ(i)zkä : ič(i)ñ(i)zkä : ing(ä)n yūki : il(ä)z : tüş(ü)rt(ü)ñ(ü)z
 F1. (a)ts(a)r (a)lp : (ä)rt(i)ñ(i)z i tuts(a)r küč : (ä)rt(i)ñ(i)z ä : (i)n(i)l(i)g bört oča b(a)rs : (a)dr(ı)lm(a) yıtu
 F2. bot(o)m(u)z um(a)y b(ä)g(i)m(i)z : biz uya : (a)lp (ä)r : özin : (a)l(ı)tı qılm(a)d(i)ñ : özl(ü)k (a)t : öz:(i)n : üc (ä)r(i)g (a)lm(a)d(i)ñ : yıta : (e)z(ä)nč(ü)m ä : küz(ä)nč(ü)m ä : (a)dr(ı)lma : s(ä)čl(i)nm[ä] : ögürd(i)m
 F3. y(e)rd(ä)ki b(a)rs t(ä)g(i)m ä : (ä)rd(ä)ml(i)g(i)m ä : bökm[(ä)d(i)m]
 R1. (a)ltun soña y(ı)š k(e)y(i)ki : (a)rt (o)γ(u)l t(a)š (o)γ(u)l t(a)d čına b(a)rs(ı)m (a)dr(ı)lu b(a)rd[ı] : y(ı)ta
 R2. tört (i)n(i)l(i)gü (ä)rt(i)m(i)z : b(i)zni : (ä)rkl(i)g : (a)d(ı)rtı : y(ı)ta
 R3. (ä)r (ä)r<d>(ä)m (ü)č(ü)n : in(i)m (e)č(i)m : uy(u)r(ı)n üč(ü)n : b(ä)ñ:güm(i)n : tikä : b(e)rti

Translation

- L1. 'O, my mother! She carried me for ten months (in her womb), (and then) brought (me) forth. In order to (display my) manly qualities to my people, I have been separated (from this world)!'

- L2. 'To the regret of my people, I attacked the numerous enemy without turning away from it and (thus) was parted (from this world)! Alas!'
- L3. '(With your death) you placed a female camel's load of grief on (the shoulders of) your younger and elder brothers.'
- F1. 'O, you were tough in throwing arrows! O, you were powerful in capturing (the enemy)! (O, you,) wolf cub with younger brothers! (O, you,) little tiger! Do not be parted (from us)! Alas!'
- F2. '(O,) our camel colt! (O,) our Umay(-like) beg! You did not let our lives, the lives of (your) brave kinsmen, be taken. You did not take (with you) the life of (your) favorite horse and the three (kins)men. Alas! O, my advance guard! O, my protector! Do not be parted and separated (from us)! (O), my joy (in life)!'
- F3. 'O, my tiger-like one on the earth! O, my virtuous one! (I did) not have enough of (you)!'
- R1. 'The wild animal of the Altay and Songa mountains, (my) last son, (my) son abroad, my young and little tiger, having been parted (from us), he went away. Alas!'
- R2. 'We were four younger brothers all together. Erklig, (the god of the underworld,) parted us. Alas!'
- R3. 'Because of (my) manly qualities and because of their capability, my younger and elder brothers erected (this) everlasting monument of mine.'

Explanations

1. L 1. *iltdi*: The final letter is /I/ in the Finnish and in Radloff's atlases, but it is /A/ in Vasil'ev (1983: 65), and it is transliterated as /ä/ on p. 25. But the letter is clearly /I/ in the photograph on p. 103.

2. L 1. *ög(ü)m ä*: Radloff, Orkun and Malov regarded the final letter /A/ as a separation mark. But in this inscription the words are separated from one another by a colon. Kljaštornyj and Mori rightly regarded the final /A/ here as the Old Turkic interjection of address.

3. L 1. *k(ä)l(ü)rti*: The labial vowel sign of the second syllable is lacking.

4. L 1. *m(ä)n*: This word is found in Radloff's runic text, and following him, in Orkun's, Malov's, Kljaštornyj's and Mori's editions. It is not found in Vasil'ev's copy and transliteration, however. But a letter resembling /m/ is clearly visible in the photograph on p. 102.

5. L.1. *y(i)r(i)ld(i)m ä*: In Radloff's runic text this word is spelt *YrldmA*. Radloff (1895: 334), Orkun (III: 103), Malov (1952: 25) and Kljaštornyj (1976: 261) read this word *yerledim* and translated it as 'I settled down'. Mori read the same word as *yerildim* and translated it as 'I have been separated (from my homeland)' (1986: 5); Vasil'ev has a lacuna in his transliteration (1983: 25), but in the photograph on p. 102 we find the group of letters /rSdmA/ which can be read (*ä*)r (*a*)s (*ä*)d(i)m *ä* and this can be understood as 'O, my men, women and properties!'. But such a phrase does not complete the sentence which begins as *il(i)mkä (ä)rd(ä)m üč(ü)n ...*

Furthermore, Vasil'ev put the group of letters /rSdmA/ at the end of line 8 (1983: 25). But in the photograph on p. 102 this group of letters can be seen at the end of the first line of the inscription.

If this group of letters is really /Yrldm/, it can be read *yirildim* or *yirildim* and be understood as 'I have been separated'. The verb *yiril-* 'to be split apart, be parted' occurs in the Uyghur texts and in some cases it forms a binary with its synonym *adril-*, e.g. *adrılıyahı yirilgäli* 'since we parted and separated' (Ht VII 2064) etc.

The same verb form occurs as /yrldm/, i.e. *y(i)r(i)ld(i)m* (wrongly read *yärltim* by Malov) on the Xemčik-Čirgaki inscription (back side, line 1).

The verb *yiril-* survives in modern languages: Uzb. *yiril-*, NUyg. *žiril-*, Kzk. *žiril-*, Kirg. *žiril-*, etc. Its base *yir-*, too, survives in some modern languages, e.g. Yak. *si:r-* 'to split, crack, break off, break into two pieces' < **yi:r-*, Trk. *yir-* id., *yirik* 'split, cracked', etc. The New Uyghur, Kazakh and Kirghiz forms of the verb suggest that it was back-vocalic originally (cf. OAT *ırıl-* 'ayrılmak, uzaklaşmak') and the Yakut form shows that the vowel /i/ of this word was long. The forms *yer-*, *yeril-*, *yerük* / *yerök* in EDPT (955, 965) and OTWF (256, 295, 686, 815) should therefore be corrected.

6. L.2. *(e)l(i)m*. In the two atlases and in Orkun, Malov and Mori the word is spelt /llm/, i.e. *il(i)m*; but in the photograph on p. 103 the letters under discussion are clearly /lm/. In Vasil'ev's copy and transliteration, too, the word is spelt /lm/.

7. L.2. *q(a)ym(a)t(i)n*. The word is spelt /KYmTN/ in the two atlases and in Vasil'ev. But Radloff read this word *kıymitu*, and Orkun, Malov and Kljaštornyj read the word as *qıymatın* and translated it as 'with courage'. All these readings and interpretations are wrong.

It was Mori (1986: 5) who found the appropriate solution: *qaymatın* 'without turning away'. As is known, the verb *qay-* with this meaning is

attested in the Uyğur texts. It is also found in Karakhanid Turkic: *qay-* 'to turn away or back', *qay-a kör-* 'to look back, look behind someone', *qaytar-* 'to turn back' (tr.) < **qay-t-t-ar-* (EDPT: 674, 675).

8. L 3. *ing(ä)n* 'female camel'. Radloff read this word *ingän (?)* with a question mark, but he could not explain it. Orkun read it *in?ğın*, but he did not give a translation. It was Malov who first read and explained the word correctly: *ingän yüki* 'verbljužij v'juk' (1952: 55). Kljaštornyj and Mori read and interpreted the phrase in the same way.

9. L 3. *il(ä)z*. FAtlas /Ilz/; RAtlas /Iŋz/; Vasil'ev: /İld/.

Radloff wrongly regarded this word as the 2nd p. possessive suffix of the plural, i.e. *-iŋiz*, and added it to the preceding *yüki*, reading this as *yükiŋiz* 'eure gewaltige Last' (p. 334). Orkun: *iŋiz* 'idiniz' (?); Malov: *iŋäz* 'v nerešitel'nosti, v zamešatel'stve'; Kljaštornyj: *ild* 'spuskat', snimat'', Mori: *eš[s]iz* (together with *ingän yüki*) 'deve yükü (kadar) yalnızlığı'.

These readings and interpretations cannot be accepted.

The runic letters /d/ and /z/ closely resemble each other, especially when they are small. Accepting the spelling /Ilz/ as true, I read this word *iläz* and connect it with Khak. *iles* 'pain, sorrow, distress'. *ing(ä)n yüki il(ä)z* 'a female camel's load of grief' is a good metaphor expressing the grief suffered by the brothers of the deceased warrior upon his death. Khak. *iles* seems to have been derived from Khak. *ile-* 'to suffer pain' < **elä-*.

10. F 1. *(a)ts(a)r (a)lp (ä)rt(i)ŋ(i)z i*. Radloff: /TSRLp:rtŋzI/; Vasil'ev: /TSzLp : rtŋz/.

In the Finnish Atlas and in Radloff the first three letters are /TSR/. But Radloff read the clause as *at ašar alp (atsar alp) ärtiŋiz* and translated it as 'ihr waret ein Held, der Pferde verzehrte (ein schiessender Held)' (p. 333). Orkun read this passage in the same way, i.e. as *at ašar alp ertiŋiz* and translated it into Turkish as 'At aşan kahraman idiniz'.

These readings and translations are wrong. Since there is a special sign for /š/ in this inscription, the back-vocalic sign /S/ cannot be read /š/.

Malov read this passage correctly as *atsar alp ärtiŋiz* and translated it as 'Vy byli geroem-strel'kom' (1952: 53). Mori read this group of letters in the same way and translated it as 'Ok atarsanız cesur idiniz' (1986: 6).

According to Vasil'ev and Kljaštornyj, however, the first three letters are /TSz/, and not /TSR/. It is for this reason that Kljaštornyj read the first word as *atasız* and translated the whole sentence as 'Bez otca Vy

geroem byli' (1976: 261). But if this were the case, the word would have been spelt /TASz/. Besides, the word for 'father' is *qaŋ*, and not *ata* in these inscriptions.

Not only for this reason but for contextual reasons as well, I believe that the first three letters in this line are /TSR/ and I read the whole clause as *atsar alp ärtiŋiz i*. As can clearly be seen, there is an obvious parallelism between this and the following clause. (For the *i* at the end of this clause see below.)

11. F 1. *tuts(a)r küč (ä)rt(i)ñ(i)z ä* FATlas: /tUTSRkÜç:rtLzA/; Radloff, Vasil'ev: (together with the preceding /I/): /ItUTSRkÜçrtŋz/.

Radloff: *ät ut ašar (ät utsar) küč ärtiñiz* 'Ihr waret ein Mächtiger, der Fleisch und Rinder verzehrte (der Habe gewann)' (p. 333); Orkun: *it ut ašar küč ertiñiz* 'et (?) öküz ašan güç idiniz' (1936-1941, 3: 101-102). Malov, disregarding the first two letters, read the whole passage as *utsar küč ärtiñiz* and translated it as 'Vy byli sil'ny, priobretaja bogatstvo' (1952: 53); Kljaštornyj read the same passage as *it utsar köč ert(t)iñiz a* 'Kogda psy presledovali (dič'), Vy pronosilis' mimo kočevij!' (1976: 261). Mori: *it utsar küč ärtiñiz a* 'Köpeği kovalarsanız güçlü idiniz' (1982: 6).

These readings and interpretations cannot be accepted for both grammatical and contextual reasons. Clauson's *utsar küç ertiñiz* 'you were strong in conquering' (EDPT: 693), too, cannot be accepted, for it disregards the preceding two letters, i.e. /It/.

In my opinion, putting the first letter /I/ aside for the time being, we may read the rest of it as *tutsar küč ärtiŋiz ä*, thus having a clause parallel to the preceding *atsar alp ärtiŋiz*.

The use of the front-vocalic sign /t/ instead of the back-vocalic /T/ in /tUTSR/ causes no problem since there are other examples of this kind of practise in the inscription, e.g. /YtA/ (line 2), /YItU/ (line 5), /LtI/ (line 6), etc.

As for the first letter of the second clause, it is /A/ in the Finnish Atlas. If it is really /I/ and not /A/, it is in all likelihood an exclamation expressing praise, tenderness and endearment (cf. Yak. *i*: 'oh!', Kirg. *i*., *i*:y id., etc.).

Thus we may read the two clauses as *atsar alp ärtiŋiz i tutsar küč ärtiŋiz ä*.

Another alternative reading would be *atsar alp ärtiŋiz itu*, *atsar küč ärtiŋiz ä*, regarding *itu* at the end of the first clause as a variant of the exclamation /YItU/, i.e. *yitu*, occurring at the end of this line. But I pre-

fer the first solution because of the rhymed opposition between *atsar* and *tutsar* in the two parallel clauses: *atsar alp ärtiñiz i, tutsar küč ärtiñiz ä*.

12. F 1. *(i)n(i)l(i)g*. FATlas: /nrg/; Radloff: /Inlg/; Orkun, Malov: /Inlg/; Vasil'ev: /nlg/.

Radloff (together with the following /BÜrl/): *inilig böri* 'der jüngere Brüder habende Wolf'; Orkun: *iniliğ böri* 'yavru kurt'; Malov: *inilig böri* 'Volk ... imejuščij mladšix brat'ev'; Kljaštornyj: *inilig bürt* '(dux) smerti so svoej mladšej brat'ej'; Mori: *inilig bürt* 'Küçük erkek kardeş sahibi ölüm tanrısı'.

There can be no doubt that the group of letters /nlg/ is a scribal error for /Inlg/ or /Inllg/, i.e. the adjective *inilig* 'having younger brothers'. The initial and medial vowels are also absent in /nlgÜ/, i.e. *(i)n(i)l(i)gü* in line 8.

13. F 1. *bört*. FATlas: /bÜrt/; Radloff: /bÜrl/; Vasil'ev: /bÜrt/.

Radloff, Orkun and Malov: *böri* 'wolf'. But in Korpus the word is spelt /bÜrt/. Kljaštornyj read this as *bürt*, and translated it into Russian as '(dux) smerti', i.e. 'the spirit of death', seeing a relation between this word and *bürt* 'nightmare' in MK. Mori, too, accepted this reading and interpretation.

In my opinion, this word, which is quite similar to OT *böri* 'wolf', could etymologically be related to it. But it is a *hapax legomenon*. However, it reminds me of the noun *bört* or *börtü* in the well-known Modern Turkish phrase *börtü böcek* 'insects and the like' (< ?*bört ü böcek*). As is well known, the wolf has long been regarded as a sacred animal by the Turks and its name has always been taboo. As a result the Common Turkic *bö:ri* has been replaced by a word meaning 'worm' or 'insect' in the Oghuz group, e.g. Trk. *kurt*, Az. *gurd*. We know that in Turkmen, alongside the old word *bö:ri*, the word *mö:jek* 'insect' is also used to denote 'wolf'. I would like to point out here that in some Anatolian dialects the phrase *böcü börtü* is used to denote wild animals like wolves, jackals and pigs: *böcü börtü* 'kurt, çakal, domuz gibi zararlı sayılan hayvanlar' (Ödemiş, Kiraz köyleri; Eşme-Uşak), etc. (DS II: 755-756). If Turkish *bört* or *börtü* in *börtü böcek* is cognate to *böri*, it is very probable that *böri*, too, was not the original name given to this animal by the ancient Turks.

14. F 1. *oča*. Radloff: *uča* 'entflieht (fliegt)'; Orkun: *uça* '[uça?]'; Malov: *böri uça* 'Volk bežal' (p. 53); Kljaštornyj: *inilig bürt uç a* 'Sgin' (dux) smerti so svoej mladšej brat'ej!' (1976: 261), Mori: *inilig bürt uç a* 'Ah! Küçük erkek kardeş sahibi ölüm tanrısı, def ol!' (1986: 6).

All these readings and interpretations are wrong. In my opinion what we have here is the adjective *oča* ‘young, little, youngest’ modifying *bars*. The word *oča* has not yet been attested in Old and Middle Turkic, but it appears in some northeastern languages as *oča* and *oči*: Shor *oča* ‘jüngster, kleinst, letzter’, Sag. *oči* id. (Radloff, Wb. I: 1135, 1137), Khak. *oči* id.

15. F 1. (a)dr(ı)lm(a) yitu . FATlas, Radloff, Vasil’ev: /DRLmYItU/.

Radloff: *adırılma itü* ‘so trennt sich doch der Tiger nicht (von den Seinen)’ (pp. 332-333); Orkun: *adırılma itu* ‘ayrılma ey!’ (1936-1941, 3: 101-102); Malov: *bars adırılma itu* (?) ‘bars ne otdelilsja’ (1952: 53); Kljaštornyj: *adırılma yitu* ‘ne pokidaj (naš)’ (p. 261); Mori: *adrılma yitu* ‘Bars, ayrılma (ölme). Üzgünüm!’ (1986: 6).

Radloff, Orkun and Malov thought that the letter /Y/ of *yitu* belonged to the preceding verb. But a verbal form like *adrılma* is impossible in Old Turkic. The first word spelt /DRLm/ must definitely be *adrılma*. It is obvious that the scribe simply forgot to write the final /A/ here. Kljaštornyj and Mori read the word correctly.

As for the second word, which is spelt /YItU/ instead of /YITU/, it must be an exclamation of pity, perhaps a variant of the commonly used *yita*.

16. F 2. bot(o)m(u)z. FATlas: /UKmz/; Radloff, Vasil’ev: /BUTmz/.

Radloff: *bu atımız* ‘dieser unser Name’, Orkun: *bu atımız* ‘bu bizim adımız’, Malov: *bu atımız* ‘Èto naše imja’, *bu atımız*. Kljaštornyj: [bu?] *atımız* ‘naše zvanie’, Mori: *bu atımız* ‘Bu bizim adımız (or adımızdır)’.

All these readings and interpretations are wrong. Since the deceased is one person he would not say *bu atımız* in the sense of ‘my name’. Besides, the word *bu* would be out of place in such an utterance.

What we have here is the word *botu* or *boto*, i.e. ‘camel colt’, having the 1st p. poss. suffix of the plural and used here figuratively for the deceased younger brother.

17. F 2. um(a)y b(ä)g(i)m(i)z ‘Our Umay(-like) beg’. FATlas: /UmY : bgmd/; Radloff: /UmYbgbz/; Vasil’ev: /UmYbgmz/.

Radloff: *umay bæg biz* (together with the preceding *bu atımız*) ‘(Dieser unser Name ist) Umay Beg, (der sind) wir’; Orkun: *umay beg biz* ‘(bu bizim adımız) Umay beg’dir’; Malov: *Umay bæg biz* ‘(Èto naše imja-Umay beg’, Kljaštornyj: *umay begimiz (begmiz)* ‘(Naše zvanie...), naš beg-Umaj (var.: naše zvanie takovo-my umaj-begi)’; Mori: *umay bay biz* ‘Umay boyuyuz biz’.

These readings and interpretations cannot be correct, for the group of letters in question is certainly /UmYbgmz/, i.e. *umay bāgimiz* which can literally be understood only as 'our Umay beg'. But *Umay*, being the name of the goddess who is believed to look after women and children, could hardly be a name given to a male child as Clauson remarked (EDPT: 165).

How could and should the phrase *Umay bāgimiz* be understood? The only solution I can find for the time being is to construe it as 'our Umay(-like) beg', i.e. 'our beg (who protects us as) Umay (does)'.

18. F 2. *biz uya (a)lp (ä)r özin* 'Our lives, the lives of (your) brave kinsmen (acc.) ...'.

The word *uya* 'relative, kinsman' in this phrase was misunderstood by Radloff who translated the passage as 'Uns folgend, hast den Heldenmann selbst nicht erniedrigt (?)' (1895: 333). Orkun: *biz uya alp er* 'Biz kahraman kardeş'; Malov: *biz uya alp är* 'My-nasledstvennyj muž geroj'; Kljaštornyj: *biz uya alp er* 'my xrabrye vojny (naševo) roda-plemeni (var.: my rodiči, xrabrye vojny)'; Mori: *biz uya* 'Biz bir boyuz'.

All these interpretations are wrong. The verb meaning 'folgen' is *ud-*, not *uy-* in Old Turkic. Secondly, OT *uya* means 'brother, blood brother, kinsman', and not 'kahraman' or 'tribe, clan'.

Furthermore, the word *öz* here means 'life', not 'self', as in the examples given below: *özüg uzun bolzun* 'May your life be long!' (Irk Bitig: 47), *sansız tümän özlüg ölüür* 'They kill countless myriads of living beings', *qısyā özlüg yaşıg tınlıylar* 'short-lived mortals', *uzun özlüg yaşıy* 'long-lived', etc. (EDPT: 286), etc.

19. F 2. *(a)l(ı)tı qılm(a)d(ı)ñ*. Radloff: /LtIKILmDñ/ *alti kılmađıñ* 'nicht erniedrigt (?)'; Orkun: *eleti kılmađıñ* 'göndermedin'; Malov: *äläti (alti?) kılmađıñ*; Kljaštornyj: *altı er almađıñ* 'Šesteryx mužej s soboj ty ne vzjal!'; Mori: *altı qılmađıñ* 'altı kiři yapmadınız'; Vasil'ev: /LtIKILmDñ/.

All these readings and interpretations are wrong and do not fit the context. The front-vocalic sign /t/ in the first word is only a scribal error for the back-vocalic /T/. Therefore the first three letters should be read *alıtı*. What we have here is the verbal phrase *alıt-ı qıl-* which means 'to let someone be taken (by the enemy)'.

20. F 2. *özl(ü)k (a)t*. Radloff: *özlāk at* 'Reitpferd'; Orkun: *özlük at* 'kendi at' (!); Malov: *özlāk at* 'sobstvennaja (ili žirnaja?) lošad''; Kljaštornyj: *özlük at* 'skakun', i.e. 'thoroughbred, fast-running horse'; Mori: *özlük at* 'hızlı koşabilen at'.

All these interpretations are wrong. As is known, the adjective *özlük* means 'private, personal'. Together with the following *öz:(i)n* the whole phrase means 'the life of (your) private horse'.

21. F 2. *öz:(i)n*, i.e. *öz(i)n*. The word is wrongly separated by a colon; cf. *b(ä)ŋ:güm(i)n* (R3).

This *öz*, too, has been misinterpreted by the previous editors. It means 'life', and not 'self', like the one in the passage *biz uya (a)lp (ä)r özin* ... in line 5 (see note 18 above). It means that the deceased young warrior saved, with his self-sacrificing actions, not only the lives of his three brothers but also the life of his favorite horse in the battle which cost him his life.

22. F 2. *üč (ä)r(i)g*. Finnish Atlas, Radloff: /Üčrg/, Vasil'ev: /ÜčKg/. Radloff, together with the preceding *özin*: and the following *almadıŋ: öčürüg* 'selbst nicht vernichtet'; Orkun: *üç eriğ* 'üç adamı'; Malov: *üč äriğ* 'trex mužej'; 'the three men'; Kljaštornyj: *üç qag* 'trex sosudov'; Mori: *üč äriğ* 'üç savaşıyı'.

The second word is spelt /Kg/ in Korpus. This can be read as *q(a)g* and it could be regarded as the accusative form of *ka* 'kinsmen, relatives', not of *ka* 'vessel, container' as Kljaštornyj thinks, because such a word does not fit the context here.

In my opinion, *üč (ä)r(i)g* is correct. The runic sign identified as /K/ by Vasil'ev looks like a slightly different form of the front-vocalic /r/ in the photograph.

"The three men" whose lives were saved are no doubt the three brothers of the deceased, i.e. slayed warrior.

23. F 2. *(e)z(ä)nčüm ä küz(ä)nčüm ä*. Radloff: /znččÜmA kÜznčÜmA/; Vasil'ev: /znčÜmA:kÜznčÜmA/.

Radloff: *äzänčümä küzänčümä* 'von meinen Gewohnheiten und Wünschen', Orkun: *ezinçüme közüñçüme* 'şöhretimden'; Malov: *äzänčümä közüñčümä* 'ot moix dobryx obyčaev i moix želanij (ili javlenij)'; Kljaštornyj: *ezünčüm (?) a közüñčüm a* 'O, moja dragocenost'! O, moe sokrovisče!' (1976: 261, 263); Mori: *ezünčüm a közüñčüm a* 'Benim hazinem (?), ah! Benim servetim, ah!' (1986: 6).

Radloff's interpretations cannot be accepted, for a word like *äzänč* or *äzänčü* meaning 'Gewohnheit' does not exist and the verb meaning 'to wish, desire' is *küsä-* with /s/ in Old Turkic. Orkun wrongly compares his *ezinçü* with MK's *idınču* meaning 'something which is let loose'. As for the second word read *közüñčüm* by Kljaštornyj and Mori, it is obvious that they wanted to identify this word with *közüñč* 'Schätze',

thus read and interpreted in U I (6: 14). Here it should be remembered that Röhrborn reads the same word as *küzünč* 'Kostbarkeiten' (1979: 113).

These are in fact the most puzzling two words of this inscription, and I must confess having had real difficulties trying to find a solution. Since these two words occur immediately after *yita*, the well-known exclamation of sorrow, pity or regret, the former can be read *(e)z(ä)nčüm ä* or *(e)z(i)nčüm ä* and this can be compared with Kar.H. *ezeñts* 'disgusting, detestable, abominable' and Kar.T. *izinč* 'unpleasant, disgusting'. These two Karaim words go back to an older and original **ezänč* or **ezinč*. The hypothetical **ezänčü* or **ežincü* can then be compared with *erinc* 'wretched, miserable' occurring in the second Altinköl inscription: *(ä)rd(ä)ml[(i)g] bols(a)r bod(u)n is(i)rk(ä)yü (e)rm(ä)di. (e)r(i)nč(i)m, i[s](i)z(i)m ä* 'Since I had manly qualities, the people regretted (so much the loss through my death) and did not despise (me). How miserable it is! Alas (for my life)!' (line 4). But this *(e)r(i)nč* is a noun in *-(X)nç* derived from *(y)er-* 'to loathe, blame' (OTWF: 284). The form **ezänčü* occurring in the inscription, on the other hand, is a noun in *-(X)nçU* derived from an obsolete base **ezä-* which seems to have survived only in Khak. *izä-* 'to destroy, demolish' and in its derivatives *izäg* 'destruction, demolition, wreck', *izäl-* 'to be destroyed', *izälig* 'ruin, devastation'. The second group of letters can in this case be read *küz(ä)nčüm ä* and *küzänčü*, too, can be regarded as a noun in *-(X)nçU* derived from **küzä-*, an obsolete verb which seems to have survived only in Khak. *küzä-* 'to defame, cover with shame'. The passage *ezänčüm ä küzänčüm ä* may then be rendered as 'O, how unpleasant it is! O, how shameful it is!'.

Another possibility is to regard the first word as a derivative of **ezä-*, perhaps an old variant of *yezä-* 'to patrol, go on sentry, watch, guard' (EDPT: 985) without the initial /y/. In Uygur there are examples of the loss of the initial /y/ before /e/, e.g. *yer-* / *er-* 'to loathe, oppose, despise', *yerinč* / *erinč* 'wretched, miserable', *yerinčü* / *erinčü* 'sin, something loathsome, a thing to be condemned', *yelkür-* / *elkür-* 'to soar, flutter', *yelküt-* / *elküt-* 'to excite, activate', etc. The verbal noun **ezänčü* could then be understood as 'advance guard, vanguard', like *yezäk* 'advance guard' in MK (EDPT: 986). The verb *yezä-* seems to have passed from Turkic into Mongolian where it has the form *jese-* and *jise-* 'to mount guard; to go on sentry duty; to watch, guard' (Doerfer 1965-1975, IV: 164).

The second word *küzānčü* or *küzünčü*, too, could in this case be regarded as a noun derived from **küzä-* ‘to watch, guard, protect’. The latter has not yet been attested in the Old and Middle Turkic texts, but I think it lives on, among other languages, in Kirghiz *küzö-* which is synonymous with *küzöt-* ‘to guard, watch, protect’ < OT *küzäd-* < *küzä-d-*. Thus, *küzānčü* or *küzünčü* could be an agent noun meaning ‘guardian, watcher, protector’, like *akınču* ‘raider, raiding party’ in MK. The passage *ezānčüm ä küzānčüm ä* may in this case be translated ‘O, my advance guard! O, my protector!’. Here I would like to point out that Arat read the word *közünčüm* occurring in MK II 8-14 as *küzünčüm* and translated it into Turkish as ‘koruyucum’ (Arat 1965: 17).

For contextual reasons, I am more inclined toward the second solution because of the word *ögürd(i)m* or *ögürd(ü)m* ‘(O,) my joy (in life)!’ occurring immediately after *(e)z(ä)nčüm ä küz(ä)nčüm ä* (see note 25 below).

24. F 2. *s(ä)čl(i)nm[ä]*. FATlas: /sčlnmÜ/; Radloff: /IčlnmÜ/; Vasil’ev: /sčlnmÜ/.

Radloff: *ič älinmü* ‘beim inneren Volke?’; Orkun: *iç elinmü* ‘... mi?’; Malov: *ič älinmü (?)* ‘u vnutrennego naroda (?)’; Kljaštornyj: *sečilenmü* (without translation); Mori: *es ečili inim ö* ‘ağabeyim ve küçük erkek kardeşlerim, beni düşün.’ Vasil’ev: /sčlnmÜ/.

The first letter in this group is the front-vocalic sign /s/, and not /l/. The last sign, on the other hand, is /Ü/ everywhere, including Korpus. But it must be a scribal error for /A/. Thus, what we have here is actually /sčlnmA/, i.e. *s(ä)čl(i)nmä* ‘do not be parted!’, i.e. a form synonymous with the preceding *adrılma* ‘do not be separated!’. The verbal binary *adrıl- säčlin-* is found also in the Kızıl Çıra I inscription (lines 1-2): *oɣl(u)ma yut(u)z(u)ma adr(ı)lt(ı)m s(ä)čl(i)nt(i)m* (Vasil’ev 1983: 30). The binary *adır- säč-* ‘to choose, select, pick out’ is quite common in Kutadgu Bilig, e.g. *adr-a säč-ä yör-*, *adr-a säč-ä tut-*, etc.

25. F 2. *ögürd(i)m*. FATlas, Radloff: /IYÜrdmm/; Vasil’ev: /ÜgÜrdm/.

Radloff: *ıyu ärdämim*; Orkun: *iyü erdemim* ‘erdemim’; Malov: *iyü (?) ärdämim* ‘moja doblest’; Kljaštornyj: *ögürdim[iz]* ‘My (prežde) radovalis’; Mori: *ıyu ärdämim* ‘Ah! (?) Benim (?) cesaretim’.

Regarding the last group of letters of this line there is a great difference between the old copies and the one produced by Vasil’ev. But the photograph in Korpus is clear and supports Vasil’ev’s reading:

/ÜgÜrdm/. As far as I can see, there is no /z/ at the end of it, as Kljaštornyj assumes.

The group of letters /ÜgÜrdm/ could of course be read as *ögürd(ü)m* 'I became happy'. But neither this nor Kljaštornyj's *ögürdimiz* 'We were (formerly) happy' would fit the context.

In my opinion, what we have here is a verbal noun in *-DI* or in *-DU* derived from *ögür-* 'to be joyful, rejoice', i.e. *ögürdi* or *ögürdü* 'joy' (cf. Uyg. *ögdi* 'praise', *alkadı* id., MK *tamdu* 'blaze', *umdu* 'request, begging' etc.). Such a noun would perfectly fit the context and be in harmony especially with the preceding imperative forms *adrılma säčlinmä* 'Do not be parted and separated (from us)!'

26. F 3. *b(a)rs t(ä)g(i)m ä (ä)rd(ä)ml(i)g(i)m ä*. FATlas: /BRrtgmA : rdmlgmA/; Radloff, Vasil'ev: /BRStgmA:rdmlgmA/.

Radloff: *bars tägimä ärdämligimä* 'Bei meinem auf der Erde lebenden Tigergeschlechte, bei meinen mit Trefflichkeit Begabten'; Orkun: *bar ertigime erdemligime* 'yerdeki var olduğuma, erdemliğime'; Malov: /BRr(ili S)tgmA/ *bar ärtigimä* (ili *bars tägimä*) 'moim byt'em i (vsem) moim doblestnym ili moim zemnym rodom "bars"'; Kljaštornyj: *bars etigim a erdemligim a* 'moimi dejanijami i moej doblestvju-ja, Bars,'; Mori: *Bars ätigim a ärdämligim a* 'Yeryüzündeki Bars, benim davranışım, ah! Benim cesurluğum, ah!'

The first group of letters under scrutiny is clearly /BRStgmA/ in Vasil'ev (1983) and in the photograph on p. 103. This can only be read as *bars tägimä* or *bars tägim ä*. *Yerdäki bars tägimä, ärdämligimä bök-mädim* would make a good sentence, but the dative suffix after the 1st p. poss. suffix of the singular is +KA in this and the second Altinköl inscription: *il(i)mkä* (line 1), *oyl(u)mqa, bod(u)n(u)mqa* (Altinköl II, front 2), but *y(a)s(ı)ma* (front 4). In my opinion, what we have here are two phrases, i.e. *bars tägim* 'my tiger-like one' and *ärdämligim* 'my virtuous (one)', both followed by the interjection of address A. It should be kept in mind that the speaker here is not the deceased, but his mother. She employs the same simile in speaking of her deceased son on the right side of the same inscription: *t(a)d čına b(a)rs(ı)m* (line 1).

27. F 3. *bökm[(ä)d(i)m]*. After the first three letters there is a punctuation mark in the shape of /::/, perhaps indicating that the line ends there. But the word is incomplete. Obviously, the scribe simply forgot to add the letters /dm/. Radloff reads this *bükmä(dim)*; Malov, Kljaštornyj and Mori correctly read the word with /ö/ and repaired the text by adding

/dm/, i.e. *-dim*. Orkun wrongly read this as *bükme (?)* and translated it as ‘doyma’.

28. R 1. (a)ltun soḡa y(ı)s. FAtlas, Radloff, Vasil’ev: /LTUNSUḡAYŞ/. Radloff: *altun šunda yaš* ‘Sechzehn sind hier seine Jahre’; Orkun: *altun suña yaş* ‘altın turna’; Malov: *altun suña yas* ‘zolotnyx utok i molodyx gazelej’; Kljaštornyj (together with the following *keyiki*): *altun suḡa yış* ‘O, dič’ zolotoj černi Sunga’, Mori (together with the following *keyiki*): *altun Soḡa (Suña?) yış* ‘Altın Songa (or Sunga?) ormanının yabani hayvanı’.

The second word is written with the back-vocalic sign for /s/. Therefore, all the readings with /š/ are wrong. The third word spelt /Yš/ should be read y(ı)š as Kljaštornyj and Mori did.

Obviously, there are two place names here: *altun yış* ‘the Altay mountains’ and *soḡa yış* ‘the Songa mountains’. The name *soḡa* could be related to or be the same as Chag., Trkm. *sona* ‘a female wild duck’, Kirg. *sono* id., Bashk. *huna*, *huna öyräk* id., Trk. *sunā* id., etc. But the nasal /ḡ/ in *soḡa* precludes such a relationship.

29. R 1. (a)rt (o)ḡ(u)l t(a)š (o)ḡ(u)l. FAtlas: /RTGL:TšGL/; Radloff: /RTGLTUGL/; Vasil’ev: /RTGLTšGL/.

Radloff (together with the following /TDçINA/): *artıylatu aylatdaçına* (*ga?*); Orkun: *artıglat ogul* ‘çoğalt’; Malov. *art ıylat uylat* ‘razyskivaj, zastavljaj plakat’ i rydat’; Kljaštornyj: *artııl toḡııl* ‘... množ’sja! Roždaj (svoe potomstvo)!’; Mori: *artııl toḡııl* ‘çoğalt, doğur!’. It goes without saying that all these readings and interpretations are wrong.

The mother of the deceased, after having described her son with the phrase (a)ltun soḡa y(ı)š *k(e)y(i)ki* ‘the wild animal of the Altay and Songa Mountains’ at the beginning of this line, continues her description with the phrases (a)rt (o)ḡ(u)l ‘(my) last son’ and t(a)š (o)ḡ(u)l ‘(my) son abroad’. The use of the noun *taš* ‘outside, exterior’ attributively is not unusual; cf. *taš y(a)ḡı* ‘the outer enemy’ (Kežeelig-Khovu, 4), MK *taš ton* ‘outer garment’, etc. The only problem here is that the initial vowel of *oḡul* is not written in the two occurrences of the word. But in this inscription and others there are many cases similar to this, e.g. *tört (i)n(i)l(i)gü* (right 2), *(ä)r (ä)r<d>(ä)m (ü)č(ü)n* (right 3), *(ö)l(ü)rm(ä)-dük(ü)m* (Elegest I, line 8), etc. Therefore, we may assume that the scribe simply forgot to write the initial vowel of *oḡul* in its two occurrences.

30. R 1. *t(a)d čina b(a)rs(ı)m*. FAtlas, Radloff, Orkun, Malov: /TDčI:NABRSm/; Vasil'ev: /TDčUNA:BRSm/ in the transliteration (p. 25), but /TDçINABRSm/ in the runic copy (p. 64).

Radloff (together with the preceding /GL/) *aɣlatdačina(ga?) barsım*; Orkun: *tad ečinä barsım* 'yabancı içine'; Malov: *udačina barsım* 'Po ego mogućestvu moj bars'; Kljaštornyj (together with the following *adrilu bardı*): *at ud ačun a barsım adırılı bardı* 'Moj Bars pokinul konej i bykov, (ves' étot) mir, on ušel!'; Mori (together with the following *adrilu bardı*): *tad ičiqä barsım adırılı bardı* 'Başka boy(?)un içine, benim Bars ayrılıp gitti.'

It goes without saying that all these readings and interpretations are wrong.

In my opinion, the first two letters may be read *t(a)d* and this may be compared with *tadun* 'a calf, one or two years old', or *ta:dun* 'a one-year-old calf' occurring in MK. This full form of the word survives in Kach., Koyb. *tazın* 'der Ochs', Khak. *tazın* 'an ox, a gelded bull' (EDPT: 457). It also survives in Kzk. *tayınša* 'a one-year-old calf' < **tadun*+*ča*.

The Middle Turkic *tadun* is in all likelihood a secondary form derived from **tad*, meaning originally 'a young calf or bull', but used here in the sense of 'young' (cf. OT *bod* 'tribe' and *bodun* 'tribes, people').

As for the second word here, I believe that it is *čina* meaning 'little'. The word *čina* has not yet been attested in the Old Turkic texts, but it is found in modern languages as the base or the first element of a group of words meaning 'the little finger', e.g. Kzk. *šinaşak* 'the little finger or little toe' < **čina*+*čak*, Crim.Tat. *činajak* id., Blk. *činačik* id., Nog. *šinatay* id. < *čina tay*, etc.

31. R 2. *tört (i)n(i)l(i)gü. (ä)rt(i)m(i)z*. FAtlas: /tÜrtnlgÜ : rtmz/; Radloff: /tÜrtnlgÜ : rtm/; Vasil'ev: /tÜrtnlgÜ:rtmz/.

Radloff: *tört änligü ärtim* 'die vier (Winkel) habe ich erstrebt'; Orkun: *tört inilgü ertim[iz]* 'dört kardeşli idik'; Malov: *tört änilgü (?) ärtim(iz)* 'Naš bylo četyre mladšix brata'; Kljaštornyj: *tört inelgü ertimiz* 'Naš bylo četvero vysokorodnyx'; Mori: *tört inälgü ärtimiz* 'Dört meşhur aileden idik'.

All these previous readings and interpretations are wrong. What we actually have here is the word *ini* plus the Orkhon Turkic comitative suffix +*IGU* (see Tekin 1990).

32. R 3. *(ä)r (ä)r<d>(ä)m (ü)č(ü)n*. In the inscription we have only the group of letters /rrmčn/. It is obvious that the scribe simply forgot to

inscribe the letter /d/ after the second /r/. He also omitted the initial /Ü/ of the following (ü)č(ü)n.

33. R 3. *uy(u)r(ı)n üč(ü)n*. Radloff, Orkun, Malov, Kljaštornyj: *uyarın*; Mori: *uyurın*. Of these, only Mori's reading is correct, for the word is a present participle in -yUr. The fully vocalized form of this word occurs in Irk Bitig: (ä)dgüsi *uyurı* 'his good and capable (men)' (IB 28).

34. *b(ä)ŋ:güm(i)n*. The word is wrongly separated into two parts. Cf. *öz:(i)n*, i.e. *özin* 'the life of' (front, line 2).

Abbreviations

Bashk.	Bashkir	MK	Maḥmūd al-Kāşārī
Blk.	Balkar	Nog.	Noghay
Chag.	Chaghatay	NUyg.	Modern Uyghur (= New Uyghur)
Crim.Tat.	Crimean Tatar	OT	Old Turkic
Kach.	Kacha / Qača	Sag.	Saghay
Kar.H.	Karaim, Halič dialect	Trk.	Turkish
Kar.T.	Karaim, Troki dialect	Trkm.	Turkmen
Khak.	Khakas	Uyg.	Uyghur
Kirg.	Kirgiz	Uzb.	Uzbek
Koyb.	Koybal	Yak.	Yakut
Kzk.	Kazakh		

References

- Arat, Reşid Rahmeti 1965. *Eski türk şiiri*. (Türk Tarih Kurumu yayınlarından. VII. Seri, No. 45.) Ankara: Türk Tarih Kurumu.
- Doerfer, Gerhard 1965-1975. *Türkische und mongolische Elemente im Neupersischen I-IV*. Wiesbaden: Franz Steiner Verlag.
- DS = *Türkiye halk ağzından Derleme Sözlüğü II*. (Türk Dil Kurumu yayınları 211/2.) Ankara: Türk Dil Kurumu. 1965.
- EDPT = Clauson, Sir Gerard 1972. *An etymological dictionary of pre-thirteenth-century Turkish*. Oxford: Clarendon Press.
- FAtlas = *Inscriptions de l'Énisséi* (recueillies et publiées par la Société Finlandaise d'Archéologie.) Helsingfors. 1889.
- Kljaštornyj, S. G. 1976. *Stely zolotogo ozera (k datirovke enisejskix runičeskix pamjatnikov)*. In: *Turcologica*, Leningrad. 256-267.
- Korpus = Vasil'ev, D. D. 1983. *Korpus tjurkskix runičeskix pamjatnikov bassejna Eniseja*. Leningrad: Akademija Nauk SSSR.

- Malov, S. E. 1952. *Enisejskaja pis'mennost' tjurkov*. (Akademija nauk SSSR. Institut Jazykoznanija.) Moskva, Leningrad: Akademija nauk.
- Mori, Masao 1986. Arutun-Kyuru dai-ichi hibun kōshaku [An interpretation of the first Altın-Köl inscription]. *Tōhō-gaku* 72, 1-17.
- Orkun, Hüseyin Namık 1936-1941. *Eski türk yazıtları*. 1-4. (Türk Dil Kurumu Yayınları.) İstanbul.
- OTWF = Erdal, Marcel 1991. *Old Turkic word formation: A functional approach to the lexicon*, 1-2. (Turcologica 7.) Wiesbaden: Harrassowitz.
- Radloff, W. 1895. *Die alttürkischen Inschriften der Mongolei*. St. Petersburg: Kaiserliche Akademie der Wissenschaften.
- Radloff, W. 1896. *Atlas der Alterthümer der Mongolei*. Dritte Lieferung. St. Petersburg.
- Radloff, Wb I = Radloff, W. 1893. *Versuch eines Wörterbuchs der Türk-Dialecte* I. St. Petersburg.
- Röhrborn, Klaus 1979. *Uigurisches Wörterbuch*. Fascicle 2. Wiesbaden: Franz Steiner Verlag.
- Tekin, Talât 1964. On a misinterpreted word in the Old Turkic inscriptions. *Ural-Altaische Jahrbücher* 35, 134-144.
- Tekin, Talât 1968. *A grammar of Orkhon Turkic*. (Uralic and Altaic Series 69.) Bloomington, The Hague: Indiana University Publications.
- Tekin, Talât 1990. The comitative case in Orkhon Turkic. In: Brendemoen, Bernt (ed.) *Altaica Osloensia. Proceedings from the 32nd Meeting of the Permanent International Altaistic Conference, Oslo, June 12-16, 1989*. Oslo: Universitetsforlaget, 355-359.
- UI = Müller, F. W. K. 1908. *Uigurica* 1. (Abhandlungen der Preussischen Akademie der Wissenschaften 1908: 2.)

The toponym *Takla-makan*

Gunnar Jarring

Jarring, Gunnar 1997. The toponym *Takla-makan*. *Turkic Languages* 1, 227-240.

Takla-makan, the name of the great Central Asian desert is not an old name. It appeared for the first time in 1865. Its present-day official form is Taklimakan. Several interpretations of this toponym, more or less substantiated, have been made in the course of time. The author of this article discusses the historical and toponymic material available together with a critical review of the etymological theories so far presented. His own suggestion is *Takla-makan* < *taqlar makan*, an interpretation expounded in detail in the article.

Gunnar Jarring, Pontus Ols väg 7, S-260 40 Viken, Sweden.

In my *Central Asian Turkic place-names* I state: “*taqla makan*, *Takla-makan* ‘desert’ < *taqlar makan*; *taq* < A. *ṭāq* ‘an arch, ruins with arches’, -*lar* plural suffix; *makan* < A. *makān* ‘place, habitation’, thus, ‘the place of ruins’” (Jarring 1997: 447).

How I have arrived at this conclusion needs an elucidation. *Takla-makan* as a place name is not old. It is mentioned in the literature for the first time by a British surveyor and explorer, W. H. Johnson, who paid a visit to Khotan in 1865. In his *Report on his Journey to Ilchí, the Capital of Khotan, in Chinese Tartary* he mentions the great desert in the following terms:

“At a distance of six miles to the north-east of Ilchí is the great desert of Taklá Makán (Gobi) which, with its shifting sands that move along in vast billows overpowering everything, is said to have buried 360 cities in the space of 24 hours. The edge of this desert has the appearance of a low range of broken hills, and consists of hillocks of moving sand, varying in height from 200 to 400 feet. Tea, of which I have brought away a sample, was dug out of one of these entombed cities while I was at Ilchí, and was believed by the natives to be of great age. Gold coins, weighing 4 lbs., and other articles, are

also reported to have been found in some of them, but the positions of these cities are only known to a few persons, who keep it a secret in order to enrich themselves. The only one that is well known is that in which very large quantities of brick tea are found, and which commands a ready sale in the markets, now that all trade with China is stopped. The site of this buried city is a mile to the north of Urangkásh.” (Johnson 1868: 5)

The fact that no instance of Takla-makan is to be found before 1865 does not imply that the name did not exist earlier. It may have been in use years, or even decades, earlier but probably not much more. Takla-makan is not mentioned by the preceding travellers who content themselves with the designation ‘desert’. Marco Polo for example evidently describes what later on is to be called Takla-makan when he reaches Lop. I quote:

“Lop is a large town at the edge of the Desert, which is called the Desert of Lop, and is situated between east and north-east ... The length of this Desert is so great that it is said it would take a year and more to ride from one end of it to the other. And here, where its breadth is least, it takes a month to cross it.” (Yule 1903, 1: 196)

The *Tarikh-i-Rashidi* relates as follows about this desert:

“To the east and south of Káshghar and Khotan are deserts, which consist of nothing but heaps of shifting sands, impenetrable jungles, waste lands and salt-deserts. In ancient times there were large towns in these (wastes), and the names of two of them have been preserved, namely Lob and Katak; but of the rest no name or trace remains; all are buried under the sand. Hunters, who go there after wild camels, relate that sometimes the foundations of cities are visible, and that they have recognised noble buildings such as castles, minarets, mosques and colleges, but that when they returned a short time afterwards, no trace of these was to be found; for the sand had again overwhelmed them. On such a scale were these cities of which, nowadays, neither name nor vestige remains!” (Elias 1898: 295)

It is evident that the desert of Takla-makan until it received its present name had no name of its own. The desert was quite simply called *gobi*, *čöl*, or *qum*, sometimes in the plural *qumlar*, and sometimes because of

its perils *šajtan qum* (cf. Lansdell 1893, 2, map *Shaitan Kum* 'The devil's desert').

The Western interest in Eastern Turkestan¹ begins with Johnson and is followed up by such explorers as Shaw, Bellew and above all by the two Forsyth expeditions of 1870 and 1873 (cf. Forsyth 1887: 54 sq. and 90 sq.) The exploration of Eastern Turkestan is then especially active during the following decades up to the First World War. It is carried out not only by British researchers but also by other Europeans.

The first comprehensive account of Eastern Turkestan is Robert Shaw's book *Visits to High Tartary, Yârkand and Kâshghar* (Shaw 1871) in which the toponym *Takla-makan* appears as follows:

"There stretch out two great arms of habitable country embracing between them the impracticable Desert of Gobi (or the Takla-Makân of the Toorks)" (1871: 34); "The great Desert of Takla-Makân or Gobi" (1871: 37); "This was also bare sand, communicating, I was told, with the great 'Takla-Makân', the central desert of Asia, which, under the name of Gobi, stretches eastward into China" (1871: 155); "Crossing an arm of the great Takla-Makân desert" (1871: 168); "This desert is connected with wonderful superstitions" (1871: 169); "In the sand of the Takla-Makân in the neighbourhood of Aksoo" (1871: 233).

Bellew, who visited Eastern Turkestan in 1873-1874 and published a book about his experiences there with the title *Kashmir and Kashghar* (Bellew 1875) mentions *Takla-makan* only once, namely in the form "Takla":

"The joint streams are further on lost in the desert of Takla, which is a wide spread of sand in whose loose heaps are buried the ancient cities of Khutan." (Bellew 1875: 228)

There are however some instances evidently referring to this desert or parts of it:

¹ I use the terms Eastern Turkestan and Eastern Turki which were current in those days.

“Beyond it we passed amongst scattered homesteads to Cara Cum [qara qum], or ‘Black Sand’”—no doubt a part of Takla-makan; “The latter is built over the head of Arslan Khan of the Baghra Khan family, whose body lies under the shrine of Ordam Padshah at the Cum Shahidan or ‘Martyr’s Sands’.” (Bellew 1875: 287, 310)

The same Cum Shahidan (*qum šahi:dan* ‘the martyr’s desert’) is mentioned again on page 364 and on page 370: “the sand-dunes at Cum Shahidan”. It evidently refers to the westernmost part of Takla-makan (cf. Bellew 1875: 367-377 and Jarring 1935), and deals with the legend of the introduction of Islam into Eastern Turkestan.

In the important source of our knowledge of Eastern Turkestan at the middle and the end of the nineteenth century *Report of a Mission to Yarkund in 1873* Forsyth often mentions the desert under the different names “Táklamakán”, “Takla Makan”, “Táklá Makán” and “Taklá Mukán” but adds no information beyond that given by Shaw and Bellew. Of special interest is however a passage on page 148 (originally by Bellew): “The wind sometimes blows away this sand, and exposes to view domes and minarets”.

Hayward who was in Eastern Turkestan at the same time as Shaw also reports on the secrets and perils of Takla-makan, adding a few details not found elsewhere (cf. Hayward 1870: 78).

The last testimony to the secrets of Takla-makan comes from Skrine who in 1922 was able to interview at Guma a man who had been lost in the desert of Takla-makan without water to drink:

“Suddenly I saw before me great walls in the sand and a gateway in the midst of them. I passed through the gateway and found myself in the outer court of a huge *yamen*² (any palace or large Chinese house is called a *yamen*). I went through more doors and courtyards and at last I entered a great hall (*aivān*) which was full of treasure, gold and coral and pearls. But there was a huge tiger on guard there; flames issued from his mouth and I knew he was an evil spirit. I fainted from fear, and when I came to my senses I was among the sands and there was no *yamen*. Next day I came upon the tracks of wood-cutters and found my way home.” (Skrine 1926: 190)

² Cf. *The Pinyin Dictionary* (1979: 791): *yámen* ‘government office in feudal China’.

The Russian exploration of Eastern Turkestan was from 1867 onwards mainly devoted to the north-western part of the country. The most interesting statement regarding Takla-makan belongs to M. V. Pevcov who speaks of the great Kashgarian desert also called Takla-makan, a name used by the inhabitants of Western Kashgaria, while those of Southern Kashgaria as well as the Chinese have no name for the desert (cf. Pevcov 1892-1896, 1: 62). "Takla-makan", alternating with "Takla-makan kum" [qum] is used constantly by Pevcov and by other Russian explorers and travellers. Kornilov (1903) repeats Pevcov's statement that in Southern Kashgaria, and with the Chinese there is no comprehensive name for the whole desert, different parts of it having different names, such as *šamal qum*, *ala qum* and *qara qum* (Kornilov 1903: 148-149). Pevcov furthermore has the normal information of towns that have disappeared in the desert (Pevcov 1892-1896, 1: 110, 121).

Sven Hedin is the traveller who has devoted more time to the exploration of the desert Takla-makan than any other explorer. In his report on his first exploration trip to Eastern Turkestan in 1891 Hedin only casually mentions the desert Takla-makan. However, it gets much more space in his elaborate description of his second journey performed in the years 1893-1897, which received world-wide attention. The references to Takla-makan in his *Through Asia* (1898) are numerous. But I have found the following instances of special interest:

"They [the begs] told me, that there once existed a large town called Takla-makan in the desert midway between the Yarkand-daria and the Khotan-daria; but for ages it had been buried in the sand. The whole of the desert was now known by this name, although it was sometimes shortened to Takan. They reported further, that the interior of the desert was under the ban of *telesmat*³ (an Arabic word, meaning 'witchcraft', 'supernatural powers'); and that there were towers and walls and houses, and heaps of gold tacks and silver *jambaus*⁴ (*tack*⁵ and *jambau* being Chinese coins). If a man went there with a caravan

³ *Telesmat* < Arabic *ṭilsim*, *ṭilsam* 'talisman, magical image'.

⁴ For *jambau*, cf. Jarring (1964: 147): "*jambu* (< Ch.) a lump of silver, somewhat in the shape of a shoe and stamped on the top with a Chinese stamp"; cf. further Raximov (1970: 267) *yambu*; *The Pinyin Dictionary* (1979: 852) *yuánbǎo* 'a shoe-shaped gold or silver ingot used as money in feudal China'.

⁵ *Tack* 'ingot'; in the original Swedish edition *guldtacka* 'gold ingot', not coins.

and loaded his camels with gold, he would never get out of the desert again; but he kept there by the spirits. In that case there was only one way by which he could save his life, and that was by throwing away the treasure.” (1898, 1: 450)

“Some of them [the men in Masar-aldi] called the desert Dekkan-dekka,⁶ because a thousand and one towns are said to be buried under its wastes of sand. Moreover vast stores of silver and gold might be found in them. It was possible to reach them with camels; and probably water would be found in the depressions.” (1898, 1: 452)

“An old man of eighty, who heard that we were going to try and cross the Takla-makan Desert came to my house, and told me, that in his youth he had known a man who, whilst going from Khotan to Ak-su, lost his way in the desert, and came to an ancient city, where he found innumerable pairs of Chinese shoes⁷ in the houses; but directly he touched them, they crumbled to dust. Another man started out into the desert from Aksak-maral, and by pure chance stumbled upon a town, amid the ruins of which he unearthed a quantity of gold and silver *jambaus* (Chinese coins). He filled his pockets with them, as well as a sack he had with him. As he was going off with his booty, a pack of wild cats rushed out upon him and frightened him so much that he threw everything away, and took to flight; when, some time afterwards, he plucked up courage to venture his luck a second time, he was unable to find the place again. The mysterious town was completely swallowed up in the sand.

A mollah from Khotan was more successful. He had fallen into debt, and went into the desert to die. But instead of dying, he discovered a treasure of gold and silver, and was now an exceedingly rich man. The number of those who had gone into the desert with the same design, and never returned, was legion. The old man solemnly assured me, that the evil spirits must be exorcised, before the hidden treasure could be sought for with any likelihood of success. The spirits bewitch the unhappy beings who venture thither, so that they become confused and bewildered and without knowing what they are doing they go round and round in a circle, retracing their own footsteps, and

⁶ *Dekkan-dekka* an enigmatic name; could it be *Terke-terken*, name of a desert? Cf. Jarring (forthcoming: 462): *terke* with comparative material; or, is it the same as *Takkan*, a contracted form of *Takla-makan*?

⁷ Perhaps the Chinese coins were in the shape of shoes, cf. note 4.

go and go until they fall down from sheer exhaustion, and die of thirst.” (1898, 1: 455-456)

“They [the people of Merket] called the desert Takla-makan; and the general consensus of opinion was, that given strong camels we ought to be able to cross right over it to the Khotan-daria.” (1898, 1: 478)

“... the beginning of the desert proper, which is known under the names of Takla-makan, Jallat-kum,⁸ and Adam-öllturgan-kum [adam öltürgen qum], or the Sand that Slayeth Men.” (1898, 2: 732)

“As for the ancient kingdom of Tu-ho-lo, and its buried cities, I will merely state that, according to the Chinese rules of transliteration, Tu-ho-lo is the same word as Tukhari (or Tokhari), and that Tukhari was used to indicate the people who in the year 157 B.C. dwelt at Bulunghir-gol, but subsequently migrated to West Turkestan, where the existing name of Tokharistan perpetuates their memory. Further, the word Tu-ho-lo is the same word as Takla; and there can be hardly a doubt, that the towns which I discovered, and which the indigenous inhabitants call, as indeed they call the entire desert, Takla-makan, were inhabited by this people. Finally, the little village of Tokhla near Khotan, the place in which the inhabitants of the buried cities found refuge from the invading sand, also keeps alive the name, if not the memory, of that once powerful people, a race who, according to Klapproth and Vivien de S. Martin, were of Tibetan origin.” (1898, 2: 784-785)

Chapter LXIII (Hedin 1898, 2: 788-805) entitled “The Buried City of Takla-makan” is in its entirety devoted to the desert and its buried habitations. I only quote the following passages:

“This city of Takla-makan, for that is the name my guides gave to it—we will retain the name, for it is instinct with a wealth of mysterious secrets, of puzzling problems, which it is reserved for future inquiry to solve.” (Hedin 1898, 2: 801)

Hedin’s theory of a connection between Tu-ho-lo and Takla-makan was summarily dismissed by Aurel Stein:

“That the term Taklamakān, by which all desert ground within the central area of the Tārīm Basin is popularly designated, can neither on linguistic nor on

⁸ I.e. *ǧalla:d* [< Arabic *jallād*] *qum* ‘the executioner’s desert’.

historical grounds be derived from the name *Tu-huo-lu* scarcely needs to be demonstrated to critical students” writes Stein for well-founded reasons (Stein 1907, 1: 435).

The main point of all these descriptions of Takla-makan here referred to, is that this desert is the abode of mysterious vanished cities together with a wealth of gold and precious stones, all of them possessed by supernatural powers and therefore dangerous for all visitors. Only ruins are left, consisting of walls, gateways, domes and the like.

In modern Uighur Takla-makan is rendered with Taklimakan⁹ but one as often meets the form Təklīmakan. It is a form never heard by the explorers of Eastern Turkestan who always keep to Takla-makan. There is no reason to believe that the early explorers such as Johnson, Shaw, Bellew, Hedin or Le Coq would have neglected to register an *i* if they had heard the name pronounced as Taklimakan. As modern normalized Uighur is based on the dialect spoken in northern Sinkiang, this way of writing does not represent the factual pronunciation in southern Sinkiang, at least not as it was a hundred years or more ago. The *-i-* would never appear in the southern dialects although it may have been accepted now after decades of linguistic indoctrination. I have however noted one exception to this rule in a geography textbook,¹⁰ published by the Swedish Mission in Kashghar in 1927 where it is written تكلې مكان i.e. *tekli meka:n*. On the other hand in an Eastern Turki manuscript dating back to the beginning of this century¹¹ the name is written تكله مكان. It is tempting to understand *tekle* as *tekler* but there is no *tek* with a meaning that would suit the desert name. More as a curiosity I refer to a form *Tex̣ti-makan* in a collection of folk tales published in Almota which is said to be the same as Taklimakan (cf. Qadiri 1958: 93).

The main question is: Do we have a palatal *k* or a velar *q* in Takla? Is it *takla* ~ *tekla* or is it *taqla*?

Johnson in his report writes the name Taklá-makán and states that “The first vowel sound in *mama* will be indicated by *a*, the second by *á*.” (Johnson 1868: 47). That would mean *tekla: meka:n*. The long *a:* in *tekla:* probably indicates the rising tone in the second syllable of Eastern

⁹ Cf. *Atlas of the People's Republic of China*, map No. 29.

¹⁰ *Ilmi ǧoyra:fija*, p. 15; cf. further Jarring (1991: 73).

¹¹ In the Lund University Library, Manuscript Division, Prov. 207:II:9.

Turki words which he has understood as a prolonged *a*. But as will be seen from other examples of the pronunciation of Eastern Turki words given by Johnson on page 47, his linguistic credibility is poor. Shaw has another way of writing the name, viz. *Takla-Makân* or *Tâkla-Makân*, i.e. *takla maka:n* or *ta:kla maka:n*, where long vowels are clearly indicated. Menges (1968:14) for unexplained reasons uses the form *Taqlamaqan*. The second part of the word is of course Arabic *maqām* which in Eastern Turki sometimes is confused with Arabic *makān* and receives a final *-n* instead of *-m*. Both forms are semantically very close which explains the confusion.

There are a few instances where *tak* (*taq*) or *takla* (*taqla*) appear as place names. Bellew as quoted before gives *Takla* as the name of the desert but it is probably only a shortened form of *Takla-makan*. Johnson (1868: 3) mentions a town in Khotan called *Ták*. Hedin (cf. Jarring, 1997: 446) has *Tak* [*taq*] as the name of a village, but it is a quote from a Russian source, not his own observation. Deasy (1901: 209-210) noted *Takla* as the name of a village in the mountains south of Yarkand but it is hardly likely that it is related to the desert name *Takla-makan*. I suspect that the little village *Tokhla* near Khotan, registered by Hedin (1898, 2: 785) may be the same as *Takla*. The non-existence in the present literature of place names such as *Tak* [*taq*] or *Takla* [*taqla*] does not exclude that such names exist. A careful inventory of the place names of Eastern Turkestan may well reveal surprises.

Over the years there have been several attempts at an interpretation of the toponym *Takla-makan*. I list them in chronological order.

1. Hoernle in his report to the Government of India on a collection of Central Asian antiquities says that "The *Takla Makan* desert appears to have received its name from the large quantities of broken pottery, which are found strewn about in many places" (Hoernle 1899: XXIV). It is a statement based on information he had received from the Swedish missionary Magnus Bäcklund¹² who is quoted by Hoernle (1899: XXIV, n. 16) as follows: "*Takla Makan* is a peculiar word which the natives apply to places covered with pottery. Such places are very numerous. Also many skeletons can be found in those places." That Bäcklund has understood *takla* as a plural of a word *tak* is evident from a letter published in

¹² Regarding Bäcklund see my paper "Silent Helpers" (Jarring 1983). Bäcklund was proficient in the Eastern Turki spoken in those days.

Stockholm in 1897 where he gives the name as Taklar Makan (Bäcklund 1897: 228).

My feeling is that Bäcklund is on the right track, but his “places covered with pottery” implies that his interlocutors have paid more attention to the pottery than to the ruins surrounding them, including for example *taq*, in the meaning ‘arches’.

2. Hedin (1900: 366): “*Takla-makan* Wüste, *taklamak* sich bewegen, vorrücken, z.B. *kum taklap juradi* [*qum taqlap jüredür*] der Sand bewegt sich vorwärts; auch schütten, z.B. wenn Mais in einem Sieb gereinigt wird; *makan* = Haus, Station”.

Hedin’s interpretation is not acceptable. The verb *taqla-* does not mean ‘to move forwards’ but ‘to thin’ (Jarring 1964: 295). The movement aspect is to be found in the second verb *jür-* ‘to go, to move’. Thus the meaning would be “the sand moves thinning out”, an observation which the practically-minded Eastern Turkestanians scarcely would have paid attention to.

Hedin’s theory of Takla-makan having a Tokharian origin has been dealt with above and is not to be taken seriously.

3. Albert von Le Coq, who spent a long time in Eastern Turkestan doing archaeological and ethnological fieldwork, understood the name as *taqla makān*, putting a question mark on its Arabic or Turkic origin (Le Coq 1922: 116). This name of the desert was according to him entirely unknown to the people of Turfan who only called it *čöl* or *gobi*. On the other hand it was current among the inhabitants of Kucha, Yarkand and Khotan. Le Coq leaves the question open whether it has to be translated with ‘pottery fragments’. He adds that it might be of Jungar origin, a rather ambiguous supposition which I feel can be left aside. The most important thing is that Le Coq has understood *tak* as *taq*.

4. In 1961 Tenišev presented a new interpretation of Takla-makan in its modern Uighur form *täklimakan*, assuming that the first part of the word was Arabic *tark*, in its Uighur form *terk* ~ *tek* (loss of *r* often produces a preceding long vowel) meaning ‘abandoning’; *-li* a shortened form of the Turkic *-lik* suffix. The meaning would then be ‘the abandoned place or, land’ (cf. English ‘bad land’) (Tenišev 1961: 90). It is quite a plausible theory, especially with reference to the meaning of the place name. But there is a snag in it: The *k* in *-lik* would disappear only if the next word begins with a *k* or *q*; the next word is *mekan*, which does not belong to the required phonetic juxtaposition.

5. Murzaev (1966: 350-351) dealing with the place name *Takla-makan* considers that the name is neither of Uighur, or Mongolian nor of Chinese origin. He does not present a theory of his own but refers to Tenishev's theory without objection to it. I understand it to mean that he leaves the question open.

I now come to my own interpretation of *Takla-makan*. For *taq* there are two possibilities:

(a) < "P. ... *tāq*, A tree, a fire of which will burn for seven days." (Steingass 1957: 276); "tak ... der Name eines Baumes (Holzes), der langsam brennt" (Radloff 1893-1911, III: 778). It is no doubt the same tree as (Aitchison 1890: 202) تاغ or تاك the white tamarisk *Haloxylon Ammodendron* (further explained in Aitchison 1890: 203); cf. Doerfer (1963-1975, entry 858): "... *tāg* 'Saksaul, *Haloxylon*'" This word pertains to the Persian-speaking area, its Central Asian Turkic equivalent is *julyun*. The existence of saksauls or tamarisks in the *Takla-makan* desert is never mentioned in the material which I have examined. The main interest concentrates on the existence of ruined towns in the desert. This *taq* is therefore out of the question.

(b) < "A. ... *tāq*, An arch; an arched building, cupola or any kind of vaulted work ... " (Steingass 1957: 806). The ruins in the desert often must have looked to the visitors like arches or entrances to ruined houses.¹³

My theory is, as said at the beginning of this paper, that *Takla-makan* has to be understood as *taqlar makan* < Arabic *tāq* (the loss of vowel length in Arabic and Persian loanwords is a common occurrence) + the Turkic plural suffix *-lar* (with the usual loss of final *r* > *-la*); *makan* < Arabic *makān* (also with loss of the long vowel in loanwords); *taqlar* > *taqla* in *taqla makān* has an adjectival function. The same phenomenon can be observed in for example *muztay ata* where *muztay* stands as an adjective forming an attribute to *ata* 'the ice-mountain father', in its modern form *Muz-tagh-ata*.

The fact that the components of the toponym *Takla-makan* are of Arabic origin points to an intellectual source for its genesis. It is not the language of illiterate peasants, hunters or caravan people who for one reason or another lived near the desert or somehow got acquainted with all its perils and dangers.

¹³ Cf. the rich pictorial evidence in, for example, Stein (1907).

The appearance of Takla-makan as a place name coincides with the rule of Yakub Bek over Eastern Turkestan (1864-1877), when the Andijani¹⁴ influence became pronounced.¹⁵ The toponym Takla-makan may well have been created or invented by some of Yakub Bek's Western Turkestan followers and advisers who no doubt were more literate than the inhabitants of Eastern Turkestan of those days. But this is only a conjecture.

References

- Aitchison, J. E. T. 1890. *Notes on the products of western Afghanistan and of north-eastern Persia*. (Transactions of the Botanical Society 18.) Edinburgh.
- Atlas of the People's Republic of China* [Zhonghua renmin gongheguo fen sheng ditu]. 1977. [Romanized Chinese edition. With geographical name index in Chinese characters and romanization.] Beijing.
- Bäcklund, M. 1897. *Letter from Yarkand*. (Missionsförbundet 15.) Stockholm.
- Bellew, H. W. 1875. *Kashmir and Kashghar. A narrative of the journey of the embassy to Kashghar in 1873-74*. London. [Reprint New Delhi 1989].
- Bruce, C. D. 1907. *In the footsteps of Marco Polo. A journey overland from Simla to Peking*. Edinburgh, London.
- Deasy, H. H. P. 1901. *In Tibet and Chinese Turkestan. Being the record of three years' exploration*. London.
- Doerfer, G. 1963-1975. *Türkische und mongolische Elemente im Neupersischen. Band I-IV*. (Akademie der Wissenschaften und der Literatur. Veröffentlichungen der Orientalischen Kommission XVI, XIX-XXI.) Wiesbaden.
- Elias, N. (ed.) 1898. *A history of the Moghuls of Central Asia being the Tarikh-i- Rashidi of Mirza Muhammad Haidar, Dughlát*. London.
- Forsyth, D. 1875. *Report of a Mission to Yarkund in 1873 under Command of Sir T. D. Forsyth with Historical and Geographical*

¹⁴ Andijan, a town in present-day Uzbekistan.

¹⁵ Cf. Shaw (1871: 163, 468): "The administration is at present chiefly carried on by Andijanees". The Andijani influence continued also after the fall of Yakub Bek (cf. Bruce 1907: 99).

- information Regarding the Possessions of the Ameer of Yarkund.* Calcutta.
- Forsyth, D. 1887. *Autobiography and reminiscences of Sir Douglas Forsyth.* Edited by his daughter. London.
- Hayward, G. W. 1870. Journey from Leh to Yarkand and Kashgar, and exploration of the sources of the Yarkand river. *Journal of the Royal Geographical Society* 40, 33-166.
- Hedin, S. 1898. *Through Asia* 1-2. London.
- Hedin, S. 1900. *Die geographisch-wissenschaftlichen Ergebnisse meiner Reisen in Zentralasien, 1894-1897.* (Dr. A. Petermanns Mitteilungen aus Justus Perthes Geographischer Anstalt, Ergänzungsheft 131.) Gotha.
- Hedin, S. 1916-1922. *Southern Tibet. Discoveries in former times compared with my own researches in 1906-1908* 1-9, Maps I-II. Stockholm, Leipzig.
- Hoernle, A. R. R. 1899. A collection of antiquities from Central Asia 1. *Journal of the Asiatic Society of Bengal* 1. Extra-number 1.
- Jarring, G. 1935. *The Ordam-padishah-system of Eastern Turkistan shrines.* (Hyllningsskrift tillägnad Sven Hedin på hans 70-årsdag.) (Geografiska annaler utgivna av Svenska Sällskapet för antropologi och geografi 17.) Stockholm.
- Jarring, G. 1964. *An Eastern Turki-English dialect dictionary.* (Lunds Universitets årsskrift, N. F. Avd. 1, 56/4.) Lund.
- Jarring, G. 1983. "Silent helpers" in the exploration along the southern Silk Road of Sinkiang. *Materialia Turcica* 7/8, 210-219.
- Jarring, G. 1991. *Prints from Kashghar. The printing-office of the Swedish mission in Eastern Turkestan. History and production with an attempt at a bibliography.* (Swedish Research Institute in Istanbul. Transactions 3.) Stockholm.
- Jarring, G. 1997. *Central Asian Turkic place-names. Lop-nor and Tarim Area. An attempt at classification and explanation based on Sven Hedin's diaries and published works.* (Reports from the scientific expedition to the north-western provinces of China under the leadership of Dr. Sven Hedin – The Sino-Swedish expedition – Publication 56. VIII. Ethnography. 11.) Stockholm.
- Johnson, W. H. 1868. Report on his journey to Ilchí, the capital of Khotan, in Chinese Tartary. *Journal of the Royal Geographical Society* 37, 1-47.

- Kornilov, L. G. 1903. *Kašgarija ili Vostočnyj Turkestan. Opyt voenno-statističeskago opisanija*. Taškent.
- Lansdell, H. 1893. *Chinese Central Asia. A ride to Little Tibet* 1-2. London.
- Le Coq, A. von 1922. Osttürkische Namenliste mit Erklärungsversuch. In: Hedin, S., *Southern Tibet*, 9: 2.
- Menges, K. H. 1968. *The Turkic languages and peoples. An introduction to Turkic Studies*. (Ural-altaische Bibliothek 15.) Wiesbaden.
- Murzaev, E. M. 1966. Kratkij toponimičeskij slovar'. In: *Priroda Sin'-czjana i formirovanie pustyn' Central'noj Azii*. Moskva. 336-357.
- Pevcov, M. V. 1892-1896. *Trudy Tibetskoj ekspedicii 1889-1890 gg. pod načalstvom M. V. Pevcova* 1-3. Sankt Peterburg.
- Qadiri (Bahar), Rozi (ed.) 1958. *Uighur čöčekliri*. Almuta.
- Radloff, W. 1893-1911. *Versuch eines Wörterbuches der Türk-Dialecte* 1-4. St. Petersburg.
- Raximov, T. R. 1970. *Kitajskie elementy v sovremennom ujugurskom jazyke. Slovar'*. Moskva.
- Shaw, R. 1871. *Visits to High Tartary, Yârkand, and Kâshgar (formerly Chinese Tartary) and return journey over the Karakoram Pass*. London. [Reprint with an introduction by Peter Hopkirk, Hong Kong 1984.]
- Skrine, C. P. 1926. *Chinese Central Asia*. With an introduction by Sir Francis Younghusband. London. [Reprints 1971, 1986.]
- Stein, M. A. 1907. *Ancient Khotan. Detailed report of archaeological explorations in Chinese Turkestan ...* 1-2. Oxford.
- Steingass, F. 1957. *A comprehensive Persian-English dictionary*. London.
- Tenišev, È. 1961. O značenii slova Taklamakan. *Akademija Nauk Kazaxskoj SSR. Trudy Instituta Jazyka i Literaturny* 2, 38-40. Almuta.
- Tenišev, È. 1984. *Ujugurskie teksty*. (Akademija Nauk SSSR, Institut Jazykoznanija.) Moskva.
- The Pinyin Chinese-English dictionary*. Edited by Wu Jingrong. Beijing, 1979.
- Yule, H. (ed.) 1903. *The book of Ser Marco Polo the Venetian concerning the kingdoms and marvels of the east* 1-2. London.

Two Turkic-based hybrid languages in northwestern China

Stephen A. Wurm

Wurm, Stephen A. 1997. Two Turkic-based hybrid languages in northwestern China. *Turkic Languages* 1, 241-253.

Centuries ago, several hybrid languages developed in northwestern China, largely as a result of the trading activities and the intensive contacts of various nationalities and their intermixing on the Silk Road. One such language has a Turkic-Uyghur grammar and phonology, and a very large Persian vocabulary. The vocabulary of another is very largely a debased Chinese, and its structure Turkic, with Turkic grammatical elements and Chinese-looking ones which function according to Turkic principles. The relevant characteristics of these two languages, Eynu and Hezhou, are briefly described here and illustrated with analyzed examples. With Hezhou, the similar typological features of Hezhou and Uyghur are contrasted with the very different ones of Chinese.

Stephen A. Wurm, Department of Linguistics, Research School of Pacific and Asian Studies, The Australian National University, Canberra, A.C.T. 0200, Australia.

In the course of the work resulting in the compilation of the large three-volume *Atlas of languages of intercultural communication in the Pacific, Asia, and the Americas* (Wurm, Mühlhäusler & Tryon 1996), attention was drawn to the existence of several creolized hybrid languages in the Xinjiang, Gansu and Qinghai provinces of China. Little information on them has so far been published in China where, with one exception, they are regarded as debased and corrupted forms of Chinese. While there is a strong, usually formally and semantically extensively changed and distorted, Chinese-derived lexical element in most of them, their structures are largely non-Chinese. Even less has been published on them outside China, and some of this has also been written in Chinese (for instance Chen 1986).

Four such languages have been identified in the area mentioned above and recognized for what they are in the course of this work. At the same time, it seems that there are more such hybrid languages in Gansu not far from where two of these identified hybrid languages are located, and which are probably closely related to, or variants of, one of these two, i.e. the language called Tangwang. It appears that three of the four hybrid languages identified may owe their origin at least in part to the trading activities on the continental Silk Road which ceased to function during the seventeenth century. In two of these, which are called Eynu and Hezhou, a Turkic element plays a major role. The third, called Tangwang, is structurally largely based on the Mongolic Santa (or Dongxiang) language, but its speakers may have originally been Chinese who kept much of their phonologically, tonally and semantically dialectal Mandarin vocabulary largely intact. At the same time they adopted much Santa, Arabic and Persian vocabulary (which were given tones), and a largely Santa structure in which the grammatical elements are toneless and some Chinese elements appear which have lost their tones and meanings and fulfil Santa Mongolian grammatical functions, not Chinese ones. A few Turkic grammatical elements are found, for instance the plural suffix of nouns, obviously loaned from the neighbouring Turkic Salar (Lee-Smith 1996a). There are no Turkic elements in the fourth language which is called Wutun and is a highly complex toneless creolized hybrid language based structurally on Bao'an Mongolian, Tibetan, with some Chinese functional elements, and a vocabulary based on Bao'an, Tibetan and distorted toneless Chinese. Its origin cannot be directly attributed to the Silk Road trading activities (Lee-Smith & Wurm 1996).

Eynu, one of the two hybridized creole languages with a strong Turkic element as mentioned above, is spoken in the western part of Xinjiang province by several thousand widely scattered speakers in the area extending from Kashgar to Yarkand and Khotan, and eastwards to beyond Aksu. The existence of the language and its speakers has been superficially known for a century. Grenard (1898) suggested that they were descendants of Persian Shiites who came to Turkestan in the eighth century. He also mentioned that their language was originally Persian giving way grammatically to Turkic Uyghur. Tietze & Ladstätter (1994) mention additional theories about the origin of the Eynu speakers and their language, without offering firm conclusions. Zhao & Haxim (1982) give a short description of the language, but regard it

as an Uyghur dialect. Grenard's view on the origin of the Eynu speaking people (called Abdal by their neighbours) is correct, but is only part of the picture. Of the many Persian traders who had dominated the continental Silk Road trade for centuries, many left at the cessation of that trade, but some remained, married Uyghur women and joined the Abdal people. While the Chinese administration regards them as being of Uyghur nationality, they themselves strongly disagree with this and insist that they are of Persian origin and ethnicity. Their Eynu language is grammatically and phonologically Uyghur, including the typical Uyghur vowel changes $a > e$, $a > i$ etc., but its vocabulary is very largely derived from Persian, though with an Uyghur phonology, except for the presence of voiced final b , d , g in some Eynu words, e.g. Eynu *ab* (from Persian *a:b*) 'water' (Lee-Smith 1996b). Many of the Persian-derived Eynu words differ somewhat in their form from their original Persian equivalents, e.g. Eynu *kes* 'person', *uſtur* 'stomach', *hep* 'seven', *kox* 'mountain' correspond to Persian *kas*, *futur*, *haft*, and *kuh*. (The transcription used in the Eynu section of this article is IPA, but with \ddot{o} and \ddot{u} used instead of ϕ and y for the front rounded vowels, and e to represent the ε -sound.) In words of Persian origin which have a Persian formative suffix, that suffix has been replaced by an Uyghur suffix, but the Persian stem has been maintained, e.g. Persian *a:b-kef* 'seller of water', Eynu *ab-tſi*, Persian *haft-um* 'seventh', Eynu *hep-inĉi*. The Eynu personal, indefinite, reflexive and interrogative pronouns are all Uyghur, e.g. *men* 'I', *biz* 'we', *siler* 'you (pl.)', *her* 'each one', *özimiz* 'ourselves', *kim* 'who', *qajsi* 'which', etc. Eynu verbs are usually a Persian stem + *-la* or *-le*, e.g. Persian *xor-* 'eat', Eynu *xor-la*; Persian *nigar* 'look', Eynu *niga-la*, etc. In Persian, verbs have two stems, essentially for present and past tense, but this feature has been lost in Eynu, except for some very few petrified forms. It has to be remembered that Uyghur has a large number of Persian (and Arabic) loanwords, many of which also occur in Eynu. In addition to its Persian-derived vocabulary, Eynu has a very small number of Turkic Uyghur loanwords, e.g. *tüt-le* from Uyghur *tüt-* 'to take'; also loanwords from Mongolian and Sibe-Manchu, words of unknown origin, as well as some metaphorical expressions resulting from a change in the meaning of Persian, Arabic and other words.

To illustrate what has been outlined above, a number of inflected forms and short sentences are given in Eynu and Uyghur, with explanations and translations. The Eynu Persian-derived words and ele-

ments, and other words and elements in Eynu which are not Uyghur-derived, are underlined:

Eynu		Uyghur
<u>engür</u> -ler	'a lot of grapes'	üzüm-ler
<u>mike</u> -m	'my goat'	öfke-m
<u>miki</u> -si	'his goat'	öfki-si
<u>hatta</u> -da	'at the market'	bazar-da
(hatta is from Sibe-Manchu where it means 'goods')		
<u>hatta</u> -din	'from the market'	bazar-din
<u>xurd</u> -raq	'smaller'	kitfik-rek
<u>niga</u> -la-f-t-i	'they saw each other'	körü-f-t-i
<u>niga</u> -li-d-im	'I have seen'	kör-d-üm
Pedir-im <u>dongyuj</u> xor-la-ŋ.	'My father, eat water-melon!'	Baba-m tawuz je-ŋ.
my father watermelon eat		my father watermelon
(dongyuj is Chinese-derived)		
<u>Batfi</u> -lir-i <u>gijaŋ</u> -li-d-i.	'Their children, they cried'	Bali-lir-i jiyli-d-i.

Eynu *Pedir-im hatta-din jek saŋ atef we jek mike ün-d- i.*
 my father, from the market, one stonefire and one goat he came
 'My father brought one flintstone and one goat from the market.'
 (ün- is an Eynu verb of uncertain origin which indicates leaving something and heading in any direction)

Uyghur *Ata-m (or data-m) bazar-din bir tfaqmaq tef-i we bir öfke epkep-t-u.*
 (tfaqmaq tef-i 'flintstone'; epkep is from elip kelip 'taking-coming')
 'My father brought one flintstone and one goat from the market.'

Eynu *feb-de bad qis-li-d- i.*
 night-in wind it did
 'During the night, a wind blew up.'
 (qis- is an Eynu verb derived from Mongolian xi- or qi- 'to do')

- Uyghur *Ketf-te famal tfiq-t-i.*
 night-in wind it-came-out
 ‘During the night, a wind blew up.’
- Eynu *Xani-da mike hes-mu, nist-mu.*
 in the house, goat exist-question, not exist question
 ‘Is there any goat in the house or not?’
- Uyghur *Öj-de öfke bar-mu joq-mu?*
 in the house goat exist-question, not exist-question
 ‘Is there any goat in the house or not?’

From these examples it should be evident that Eynu and Uyghur are not mutually intelligible.

It seems that the explanation for the origin and continued existence of the Eynu language is as follows: The offspring of the originally Persian-speaking Abdal people (probably of the first immigrants and the Persian traders who stayed in the area after the cessation of the Silk Road trade) from their intermarriage with Uyghur women learned Uyghur from their mothers. At the same time, the Persian fathers who had preserved a very strong feeling of Persian ethnic identity which the Abdal people have traditionally maintained to the present day, taught the children Persian words to use as a symbol of their ethnic identity. The children learned these words with the phonological base of their Uyghur mother tongue, with the language handed down in this hybrid form to subsequent generations as a creole. It may be mentioned that all Eynu speakers, including the children of Eynu-Eynu marriages, are entirely bilingual in Uyghur as their ‘outside’ language, and in Eynu as their ‘inside’ language used within the Eynu community, and as a secret language, and there is no sign of Eynu being in danger of being replaced by Uyghur with Eynu speakers.

Hezhou is the other of the two above-mentioned hybridized creole languages with a strong underlying Turkic element in it. Its name is the old name of Linxia city (which is located south of the Yellow River at the mouth of the Daxia River in Gansu province). The language has been regarded by Chinese scholars, e.g. Ma Shujun (1984) as basically Chinese heavily influenced by local Turkic and Mongolian languages. Dwyer (1992) essentially adheres to that view, but leaves open the possibility that the syntactic pressure from Altaic languages on Chi-

nese syntax may have resulted from imperfect learning of Chinese on the part of non-native Chinese speakers (Dwyer 1992: 173). Lee-Smith (1996c) takes the view that a form of Hezhou is the result of this. It appears that there is a form of it in Linxia city which has three stable tones, and one outside it, spoken as a trade lingua franca in an area south and west of the Linxia Autonomous Region in Sansu province, and in adjacent parts of Qinghai province. There Hezhou speakers constitute a relatively small population while the majority of the population are speakers of the Mongolian languages Santa (or Dongxiang) and Bao'an, the Turkic language Salar, and Tibetan. In that form of Hezhou, the three tones appear to be unstable, and there are tone sandhi which show changes that are not typically those of a Sino-Tibetan language. Rather than the tones becoming unstable and non-semantic, which would be looking at them from the Chinese angle, it seems more likely that the language started off as a non-tonal Turkic language, and is in the process of acquiring tones, which in Linxia city, with its large Chinese population, has progressed much further. The basic structural characteristics of Hezhou are Turkic (Salar and / or Uyghur), but its vocabulary is very largely Chinese-derived. The word order and the rudimentary mirroring in the language of Turkic verbal and suffixal grammar show the thought patterns underlying the syntactic and structural features of Hezhou to be Turkic. Five of the six cases with nouns and pronouns are marked by suffixes derived from Turkic, Tibetan and Chinese forms. A particularly interesting feature of the language is the making up for the absence of verbal suffixation in Chinese by the appearance of Chinese-looking elements used as suffixes added to verbs, but totally divorced from their Chinese meanings and functions. They indicate typical Turkic grammatical functions, e.g. converbs, verbal nouns, intention, tense, necessity, etc.

What has been said above will now be illustrated by some examples. The transcription used for Hezhou is fairly broad. IPA symbols are used, except that, as elsewhere in this article, *ü* is used for IPA *y*. In the Hezhou examples, *s*, *ts* and *z* are almost always retroflexed, *ç* is an alveopalatal or palatal voiceless fricative. No tones are marked. The transcription used for the Uyghur versions is the same as in the section on Eynu. In the examples, Hezhou is given first with explanations, followed by an Uyghur version to demonstrate the Turkic features of Hezhou structure by comparison, and finally by the Mandarin Chinese version in the usual pinyin transcription, without tone marks, to show

the essentially Chinese nature of the Hezhou vocabulary by comparison.

‘I bought this thing for you.’

Hezhou	<i>Tsi tūŋçi fi ŋo ni-xa me-lío.</i> this thing is I you-for buy-past
Uyghur	<i>Bu nersī-ni men saŋa al-d-im.</i> this thing I you-for buy-past-I
Mandarin	<i>Zhe dongxi shi wo gei ni mai de.</i> This thing is I give you buy of (= the bought one)

In this, the *fi* in Hezhou corresponds to Mandarin structure (there is no equivalent word in Uyghur), but *ni-xa* corresponds to the Uyghur *saŋa* ‘for you’. The suffix *-xa* (< Turkic *-ŋa*, *-qa*...) denotes the direct and indirect object.

‘(He) returned from Beijing.’

Hezhou	<i>Betçīŋ-ta xui (-tsə) le-lío.</i> Beijing-from return (-converb) come-past
Uyghur	<i>Bejçīŋ-din qajt-ip kel-d-i.</i> Beijing-from return-converb came-he
Mandarin	<i>Cong, Beijing huilai.</i> from Beijing return-come

In this case, Hezhou corresponds to Uyghur in having the suffix *-ta* ‘from’ after Beijing, and the past tense marker *-lío* added to the final verb. The Hezhou sentence would commonly be *betçīŋ-ta xui-tsə le-lío*, with the Hezhou converb marker *-tsə* (see below) added to *xui* to correspond to the Uyghur *qajt-ip*.

‘He slept until noon.’

Hezhou	<i>Tha sāü-thala sui-lío.</i> he noon-until sleep-past
Uyghur	<i>U tñüt-kiŋe uxli-d-i.</i> he noon-until sleep-past-he
Mandarin	<i>Ta shui dao zhongwu.</i> he sleep until noon

The structural agreement between Hezhou and Uyghur is quite obvious. Hezhou contrasts strongly with Mandarin. The ‘to, until’ marker *-thala* may be Turkic *-da(n)* ‘from’ + Tibetan *-la* ‘to’.

‘My father’

Hezhou	<i>ŋo-ti ata</i> I-of father
Uyghur	<i>Meniŋ ata-m</i> I-of father-my
Mandarin	<i>Wo baba</i> I father

The Hezhou possessive (genitive) marker *-ti* is a Chinese particle.

‘What do you intend to serve them with?’

Hezhou	<i>Ni tham-xa fim̃-a-la khuet-e-li.</i> you they-to (or them) what-with wait upon intention
Uyghur	<i>Sen ular-ni nime-bilen küt-mektŋi-sen.</i> you them what-with wait upon-intention-you
Mandarin	<i>Ni yong sheme zhaodai tamen.</i> you use what serve they

The structural agreement between Hezhou and Uyghur is again striking. The instrumental marker *-la* is the Salar instrumental marker.

‘Say it in Mandarin!’

Hezhou	<i>Ni phuthūxua-la suo.</i> you Mandarin-with say
Uyghur	<i>Sen putuxua-bilen sözle.</i> you Mandarin-with say
Mandarin	<i>Ni yong putonghua shuo.</i> you use Mandarin say

Again, there is agreement between the Hezhou and Uyghur structure in contrast with Mandarin.

The Hezhou marker *-tsə* is a surrogate converb marker added to verbs to mirror the Turkic converbs. It is probably derived from the Chinese particle *zhe* which is tense-oriented and marks action in progress. The Hezhou *-tsə* does not indicate those functions. Examples:

‘(When) the movie finished, I returned.’

Hezhou	<i>Tiejǐng vē-lio-tsə ǰo xui-tsə le-lio.</i> movie finish-past-converb I return-converb come-past (i.e. return-converb = returning came)
Uyghur	<i>kino tüge-p men qaj-t-ip kel-d-im</i> movie finish-converb I return-converb come-past-I
Mandarin	<i>dianying wan le wo jiu huilai le</i> movie finished, I then return-come past

The Hezhou *-tsə* can be added to the past tense marker *-lio*.

‘(I) went to the market, purchased things and returned
(i.e. returning came).’

Hezhou	<i>Kesā-xa ṭchi-lio-tsə tūŋçi mesā-lio-tsə xui-tsə le-lio.</i> market-to go-to-past-converb thing purchase-past-converb return- converb come-past
Uyghur	<i>Bazar-ya ṭfiq-ip nerse-ler-ni elip qajt-ip kel-d-im.</i> market to go-out-converb things (Obj) buy-converb return-con- verb come-past-I
Mandarin	<i>Shangjie qu mai le dongxi jiu huilai le.</i> market go out buy past thing then return-come past

The equivalent use of converbs in Hezhou and Uyghur contrasts with the Mandarin structure. Uyghur: *setiwelip* = *setip elip* ‘sell-converb-take-converb’.

The Hezhou verb *si* ‘to say’ in the Hezhou converb form *si-tsə* mirrors the Uyghur *dep* (= intention, ‘in order to’). Examples with *si*:

‘He says (or said) he has (or had) no time today.’

Hezhou	<i>Tha j̣iŋkətsi mə ḳiŋfu si(-lio).</i> he today not time say
Uyghur	<i>U ḅügün waxt-im joq dej-d-u.</i> he today time-my there-is-not say-past-he
Mandarin	<i>Ta shuo ta jintian meiyou gongfu.</i> he say he today there-is-not time

The identical Hezhou and Uyghur sentence structures and word orders contrast with the Mandarin.

‘He went out (intending) to buy a book.’

Hezhou	<i>Su me-li si-tsə ṭshu-ṭchi-lio.</i> book buy-intention say-converb go-out-go-to-past
--------	---

Uyghur	<i>Kitap al-i-men dep tñiq-t-i.</i> book buy-intention-say-converb go-out-past-he
Mandarin	<i>Ta chuqu mai shu le.</i> he go-out-go-to buy book past

The structure and word order are equivalent in Hezhou and Uyghur and contrast with Mandarin. Uyghur *al-i-men* indicates a close future sense and intention. Hezhou go-out-go-to mirrors the Chinese lexical usage.

Uyghur verbal nouns in *-if* are mirrored in Hezhou by the verbal noun marker *-ti* which is probably derived from Mandarin *de*. The Uyghur necessity marker *kerek* in its various functions (positive, negative, interrogative) has mirroring Hezhou equivalents, i.e. Uyghur positive: verbal noun *-if* + *kerek* = Hezhou verbal noun *-ti* + *joli*. The Uyghur negative: verbal noun *-if* + *kerek emes* = Hezhou verbal noun *-ti* + *pujo*, and the Uyghur interrogative: verbal noun *-if* + *kerek-mu* = Hezhou verbal noun *-ti* + *jola*. Examples:

‘Is it necessary to go to hospital?’

Hezhou	<i>Jiüē-li tñhi-ti jola?</i> hospital-inside go-to-verbal noun necessity-interrogative
Uyghur	<i>Doxturxani-ya ber-if kerek-mu?</i> hospital-to go-verbal noun necessity-interrogative
Mandarin	<i>Yaobuyao shang yiyuan?</i> need-not-need go hospital

The structure and word order are equivalent in Hezhou and Uyghur and contrast with Mandarin.

‘It is necessary to go to hospital.’

Hezhou	<i>Jiüē tñhi-ti joli.</i> hospital-inside go-to-verbal noun necessity-positive
--------	---

- | | |
|----------|---|
| Uyghur | <i>Doxturxani-xa ber-if kerek.</i>
hospital-to go-verbal noun necessity-positive |
| Mandarin | <i>Yao shang yiyuan.</i>
necessary go hospital |

Structural and word order contrast as above.

'It is not necessary to go to hospital.'

- | | |
|----------|--|
| Hezhou | <i>Jiüē-li tçhi-ti pujo.</i>
hospital-inside go-to-verbal noun necessity-negative |
| Uyghur | <i>Doxturxane-xa ber-if kerek emes.</i>
hospital-to go-verbal noun necessity-negative |
| Mandarin | <i>Bu yong shang yiyuan.</i>
not need go hospital |

The structural and word order contrast is again as above.

The examples given above show clearly that the Hezhou language, which superficially appears to be Chinese, has distinctly Altaic, especially Turkic, general grammatical and structural characteristics, though there is little formal agreement. Had Hezhou been originally Chinese, it would probably have Turkic loanwords, tonally and otherwise correct Chinese words and more formal Turkic grammatical forms as loans. A situation like this happened with the Tangwang creole in which the originally Chinese speakers kept much of their tonally and semantically correct Chinese vocabulary but adopted much of the Mongolian Santa (or Dongxiang) grammar, though much of it incorrectly. In Hezhou the situation is different: Tonally and, in cases, semantically and otherwise incorrect Chinese vocabulary has been superimposed upon a correct Altaic Turkic structure and word order, with formally Chinese elements divested of their tones and functions and re-employed to mirror elements of Turkic structure in a rudimentary way. It seems therefore plausible to assume that ancestral Hezhou developed as a simplified trade and intercommunication language between speakers of an originally Turkic language (Salar and / or Uyghur) and Chinese speakers at the western end of the involvement of

the latter in the Silk Road trade. The Turkic speakers appear to have attempted to acquire as much Chinese vocabulary as possible, while preserving much of the Turkic grammatical features, expressing them through Turkic syntactic principles, and in part through Turkic elements, and partly through Chinese-looking syllables used as suffixes and elements with Turkic functions.

References

- Chen, Nai-Xiong 1986. An outline of Wutun linguistic structure (text in Chinese). *Journal of Asian and African Studies* 31, 33-52. Institute for the Study of Languages and Cultures of Asia and Africa (ILCAA). Tokyo: Tokyo Gaikokugo Daigaku.
- Dwyer, Arianne M. 1992. Altaic elements in the Linxia dialect: Contact-induced change on the Yellow River Plateau. *Journal of Chinese Linguistics* 20 (1), 160-179.
- Grenard, F. 1898. Le Turkestan et le Tibet. Second part of: Dutreuil de Rhins, J. L. 1898. *Mission Scientifique dans La Haute Asie 1890-1895*. Appendice: Races Particulières de Turkestan Chinois. Paris. 303-315.
- Lee-Smith, Mei W. & Wurm, Stephen A. 1996. The Wutun language. In: Wurm & Mühlhäusler & Tryon (eds.) 1996, 883-897.
- Lee-Smith, Mei W. 1996a. The Tangwang language. In: Wurm & Mühlhäusler & Tryon (eds.) 1996, 875-882.
- Lee-Smith, Mei W. 1996b. The Eynu language. In: Wurm & Mühlhäusler & Tryon (eds.) 1996, 851-863.
- Lee-Smith, Mei W. 1996c. The Hezhou language. In: Wurm & Mühlhäusler & Tryon (eds.) 1996, 865-873.
- Ma, Shujun 1984. Hanyu Hezhouhua yu A'ertai yuyan (Comparisons between the Hezhou dialect of Chinese and Altaic languages). *Minzu Yuwen* 2, 50-56.
- Tietze, Andreas & Ladstätter, Otto 1994. *Die Abdal (Äynu) in Xinjiang*. (Österreichische Akademie der Wissenschaften. Philosophisch-historische Klasse. Sitzungsberichte, 604. Band.) Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Wurm, Stephen A. & Mühlhäusler, Peter & Tryon, Darrell T. (eds.) 1996. *Atlas of languages of intercultural communication in the Pacific, Asia, and the Americas*. 3 vols. Berlin: Mouton de Gruyter.
- Zhao, Xiang Ru & Haxim 1982. Xinjiang Ainuren de yuyan (The language of the Xinjiang Eynu people). *Yuyan Yanjiu* 1, 259-279.

The Central Asian Languages Corpora project (CALC). I: Modern Uzbek

Marc Vandamme & Hansje Braam

Vandamme, Marc & Braam, Hansje 1997. The Central Asian Languages Corpora project (CALC). I: Modern Uzbek. *Turkic Languages* 1, 254-261.

The CALC project aims to produce medium-sized (500,000-1,000,000 tokens) computer-readable text corpora of modern Central Asian Turkic languages. The texts are meant to be used as linguistic data as well as language teaching material. Care is taken to collect texts from a broad range of text types and usage domains. The Uzbek part has been finished recently and the Kazakh, Kyrgyz and other components are *in statu nascendi*.

The Uzbek corpus contains some 1,100,000 words, distributed over 243 corpus texts, covering about 45 text types. The larger part of the texts dates from after 1990. The corpus is available for analysis at Utrecht. In the future it will be possible to consult the data collection over the Internet.

Marc Vandamme & Hansje Braam, Department of Oriental Studies, Drift 15, NL-3512 BR Utrecht, The Netherlands.

1. Context

The development of descriptive and comparative linguistics of the modern Central Asian Turkic languages has been hampered as a result of the political constellations of the Cold War era on the one hand, and the relatively late introduction of digital information processing to this field on the other. These factors explain why computer readable language corpora of Turkic languages are almost nonexistent, and why we have so few adequate bilingual dictionaries of Turkic apart from Russian. With the disappearance of many if not most of the political obstructions since 1991, initiatives contributing to a solid empirical foundation of descriptive Turkic linguistics would seem apposite. Such initiatives are also necessary in consequence of the growing demand for educational materials and research tools such as adequate dictionaries.

In 1994 a plan was drawn up to meet this goal. This resulted in the CALC (Central Asian Languages Corpora) project, an international co-operation of institutes and scholars from Europe and Central Asia. The Utrecht group is responsible for the coordination and concrete realization of the project activities. At the Department of Oriental Languages and Cultures of Utrecht University the project functions within a broader research frame concerning the philology of the Central Asian Turkic languages, from the older stages (Chagatay) to the modern varieties. CALC focuses on the modern Turkic languages, albeit with a view to link the digital materials to older stages of the relevant languages in a later phase. Of course this only makes sense after lemmatization of the textual materials has taken place. In this process we work from the older towards the more recent stages (Vandamme, Boeschoten & Braam 1989). The lexicon of the text of Rabghūzī's *Stories of the Prophets*, completed in 1310 AD (Boeschoten, Vandamme & Tezcan 1995), serves as point of departure for the lemmatization work.

2. CALC goals

In the first place, the CALC project strives to construct high quality computer-readable medium-sized (0.5 to 1 million tokens) text corpora of modern Central Asian Turkic languages (Uzbek, Kazakh, Kyrgyz, Turkmen, Uyghur and others) and make them accessible to all interested scholars.

It is secondly our intention to function as a clearing house for other electronic language corpora and derived sets of data of the Central Asian languages concerned, yet not originally created by the Utrecht CALC group.

CALC also wishes to provide for the publication of research materials (data, bibliographical information etc.) and results (frequency lists, analytical studies of the data collections etc.), especially in electronic form.

3. Corpus aspects

3.1. Text selection

The data collection must support amongst others lexicographical, area-linguistic, textlinguistic and general comparative studies. The data bases should be flexible in use and of a sufficient size to enable the realization of related projects such as the compilation of dictionaries, language

manuals tailored to various professions and activities (banking, law, agriculture, education etc.) and language courses. For reasons of efficiency we have excluded non-printed sources in this phase of CALC. We have been able to collect some manuscript materials for modern Uzbek, although they are not included in the corpus. Future inclusion of oral materials (from tape, existing transcriptions etc.) is necessary, especially since many dialects are in danger of extinction.

Texts are selected for inclusion in the corpora according to a predefined division into text types and domains. At the first CALC meeting in December 1994 the issue of selecting text types and domains was discussed. This resulted in a preliminary list of proposed text sorts of a more or less ad hoc character. Finally a list was established based on a more systematic approach. This list covered many of the pre-theoretic text sorts mentioned in Gülich & Raible (1975) and Heinemann & Viehweger (1991). Some extra text sorts proposed by participants of the meeting were also included. The problem of systematic classification still remains, as a single generally accepted text-linguistic framework which can be used to derive a clear-cut text type list does not seem to exist.

We had to settle for a trade-off: On the one hand, we wanted linguistically relevant variation (in text construction), on the other, we also had to choose according to lexical demands. We used the cognitive communication theory of Heinemann and Viehweger to produce, by parametric variance, a smallest set of text sorts and language usages. This ensured that texts taken from these categories would (most probably) contain the linguistic forms which are of interest to us. We tried to meet the lexical demands by selecting the texts from specific domains.

Heinemann & Viehweger (1991: 147-149) state that the systematic classification of texts should take into account four different levels: Function, situation, procedure and structuration. The *function* of the text is what is realized by using the text in interaction: To express oneself, to (re)present oneself, to contact another, to inform someone, to guide someone or to act aesthetically. These functions generally occur in combination. The *situation* of the text is not easily classified, the social structure of the communicative action is described using a sociolinguistic model of interaction. Text *procedures* are goal-oriented cognitive procedures applied in the process of text production and interpretation. Three main classes can be distinguished: Text unfolding procedures (to explain using an example, to make an issue more specific by giving extra information, to give a reason for something etc.), strategic procedures (choos-

ing a narrative, a descriptive or an argumentative set-up of the text) and tactical procedures (additional specification or strengthening of the main procedure, for example, emotional strengthening). Text *structures* concern the way in which the different text parts are combined in order to produce a certain text (for example a request has as structure: Letter-head; Letter-nucleus: K, because of G; End-of-letter. G can have a complex argumentative substructure).

This means that text sorts can be distinguished using a four-dimensional structure, with values which can be combined (as a text can have several functions), and not necessarily holding only binary feature values. The sets of possible values for the categories of text situation, procedure and structure are large sets. Although we can draw up a very large systematic matrix of text types, the CALC project cannot cover all these values at the same time. This is due to a lack of money, time, workpower and available texts. For instance, it became clear that it was impossible to find instructive texts on housekeeping products, such as washing instructions and the like. And of course utterances like curses, obscenities etc. are also hard to find in printed sources. Consequently, we had to limit the extent of text selection.

In the end we came up with the following criteria for drawing up the text selection list, in decreasing order of importance:

1. Gather the same collection of text types (in the same domain, if possible with the same subject) for every language concerned.
2. Choose text types to cover the text functions mentioned above.
3. Choose texts which allow maximalizing the expected number of different linguistic phenomena. This means, of course, that one must vary between communicative situation, procedure and structure.
4. Choose texts according to a priority list of domains. This list reflects the main lexical domains we are interested in. For the lists of text types and domains which were actually selected for Uzbek, see section 4 below.

3.2. Representativity

The resulting data collections are intended to be exemplary, not representative (Bungarten 1979: 42). This means that the frequencies of linguistic phenomena occurring in the corpus for language A are not intended to be statistically good approximations of such frequencies in all A language utterances. As mentioned above, we have tried to increase

the odds of finding special phenomena by previous selection (not random choice) of text type and domain.

Care has been taken that the corpus texts, the basic items of the corpus collection, generally contain about 5,000 tokens each. This ensures that the linguistic utterances can be studied in their larger context. However, in the case of texts of types which are generally smaller in size (for example, poetry or advertisements) a corpus text is made up of several of such smaller texts. The corpus texts are each homogeneous in text type and usage domain, although a few corpus texts contain material of several text types, for instance, periodical articles on football with some inserted laudatory verses.

3.3. Representation standards (transliteration, text headers)

The CALC project will eventually use in its electronic data collections the conventions of text encoding of TEI / CES (cf. Dunlop 1995 and Ide 1996) and for further planned explicit language description the conventions as published by the *Eurotyp* group (Bakker et al. 1993) as far as possible. The data will be made available in two formats: A platform independent representation that supports online data retrieval using Internet / WWW connections and also in the current Macintosh form. We are currently investigating the possibilities of a full SGML version of the corpus texts, using TEI / CES. This would ensure 100% transparency also with respect to future developments in hardware and software. At present only the Macintosh version is available for pattern searching at Utrecht.

As complex taggings are for the time being not part of the project, the only problem concerns the alphabets used to write the relevant languages. In order to save space and to be able to work easily with the texts we use a simple transliteration scheme (CATL: Central Asian TransLiteration), which follows in almost all cases a straightforward method. This scheme is strictly one-to-one inside one language or orthography. A few interlingual many-to-one and one-to-many symbol pairs could not be circumvented. In choosing the symbols we have striven to comply with the proposed national standards as far as possible, although no opinion concerning the current discussion on Romanizing the Cyrillic Central Asian writing systems is to be inferred. Example: The poetry line *Ҳорғиним, ипакдай қош-кўзи чангларим*, becomes in CATL *Horğinim, ipakday qoş-közi çanglarım*.

Textual structures have not yet been covered in full detail, due to the enormous amount of work which must be invested in such a process. For example, in a collection of short articles from a newspaper the bibliographical information is marked (using a simple SGML compliant markup coding system), however the internal structure of the text itself is not (headlines etc.). Information about the texts (contents, date, length, domain, text type etc.) is available through a database and through the TEI-textheaders of the corpus texts.

4. The Uzbek corpus

The Uzbek corpus was located, selected, and converted between early 1995 and autumn 1996. Some material was selected from the library collections at Utrecht, Mainz and Frankfurt, but the major part was collected in Uzbekistan by H. Ykema. The bulk of the material was keyboarded in Tashkent on PC's, a smaller amount was scanned using OCR methods. Because of the generally bad quality of print and paper of the Uzbek books and periodicals, the available OCR techniques did not always produce high quality scans; in some cases 99% correctness was obtained, but in most cases the quality was considerably lower.

The Uzbek corpus comprises 1,164,851 tokens, in 243 corpus texts, taken from 388 sources. Poetry and proverb collections are counted as one source each, the sources of slogans are not taken into account here. All corpus texts, except one, date from after 1963. We have 200 corpus texts dating later than 1980, of which 140 corpus texts date later than 1992.

The following text types are covered:

Academic text book	92,976	Newspaper report	24,278
Advertisement	4,789	Novel	52,299
Announcement / TVguide-text	10,197	Official application form	4,606
Autobiography	21,000	Periodical article	73,465
Congratulations	730	Plan	10,426
Cooking recipe	20,086	Poetry	40,262
Decree, permission	28,670	Proverb	9,925
Drama	52,513	Report	20,472
Fairy tale	54,899	Rules of conduct	4,302
Guide	22,128	School book	3,189
Holy text	30,211	School text	57,716
Instruction manual	7,096	Scientific book	55,980

Joke	21,698	Short story	55,903
Jurisprudence	35,198	Slogan	2,383
Law text	113,346	Speech	6,867
Letter	5,045	Story	34,791
Manual	47,390	Survey	21,323
Newspaper article	79,252	Tales & Proverbs	3,730
Newspaper article, interview	26,272	Texts on soap boxes etc.	738
Grand total of tokens	1,156,889		

The following domains are covered:

Arts & Culture	130,753	Politics&Economics	91,643
Daily life	76,872	Public administration	4,606
Education	110,856	Religion	71,062
Health	70,216	Society	28,401
Law	177,214	Sports	45,402
Literature	296,007	Technology	57,583

Grand total of tokens 1,164,851

The grand total by text type is slightly smaller than the grand total by domain because a few corpus texts contain parts belonging to different text types. These corpus texts were not taken into account in the first grand total.

5. Accessibility

As mentioned above, the Uzbek corpus is for the time being only accessible on location at Utrecht University. However, we shall try to answer requests for data extracts within reasonable limits. It is our intention to make the Uzbek corpus available worldwide via the Internet.

6. Status of other languages concerned

Corpora of other languages are still being put together: Kazakh (selected 80%, converted 20%), Kyrgyz (selected 70%, converted 5%), Uyghur (selected 30%). In addition to these planned corpora, CALC offers to function as an archive or clearing house for language materials from other Central Asian languages. CALC was already so fortunate to re-

ceive into its care samples of Yellow Uyghur, Tuvinian and some others.

Acknowledgement

Our special thanks to T. Atabaki, R. Dor, E. Gürsoy-Naskali, L. Johanson, M. Kirchner, A. Nauta, M. Ploeger, M. Roos, C. Schöning, U. Shamiloglu, H. Ykema and others for cooperation, advice and supplying textual materials.

References

- Bakker, D. & Dahl, Ö. & Haspelmath, M. & Koptjevskaja-Tamm, M. & Lehmann, C. & Siewierska, A. 1993. *Eurotyp guidelines*. (Eurotyp Working Papers.) Berlin, Strasbourg: European Science Foundation.
- Boeschoten H. E. & Vandamme, M. & Tezcan, S. 1995 (eds.) *Al-Rabghūzī. The stories of the prophets: Qışaş al-Anbiyā', an Eastern Turkish version*. Volume I. Leiden, etc.: Brill.
- Bungarten, T. 1979. Das Korpus als empirische Grundlage in der Linguistik und Literaturwissenschaft. In: Bergenholtz, H. & Schaefer, B. (eds.) *Empirische Textwissenschaft. Aufbau und Auswertung von Text-Corpora*. Königstein / Ts.: Scriptor.
- Clear, J. H. 1993. The British national corpus. In: Landow, G. P. & Delany, P. (eds.) *The digital word: Text-based computing in the humanities* 1. Cambridge, Mass., etc.: MIT press. 163-187.
- Dunlop, D. 1995. Practical considerations in the use of TEI headers in a large corpus. *Computers and the humanities* 29, 85-98.
- Gülich, E. & Raible, W. (eds.) 1975. *Textsorten. Differenzierungskriterien aus linguistischer Sicht*. Frankfurt a. M.: Athenaion.
- Heinemann, W. & Viehweger, D. 1991. *Textlinguistik*. Tübingen: Niemeyer.
- Ide, N. 1996. *Corpus encoding standard*. <http://www.lpl.univ-aix.fr/projects/multext/CES/CES1.html>.
- Vandamme, M. & Boeschoten, H. & Braam, H. 1989. Editing and linguistic analysis of a medieval Turkic text with the aid of computer facilities. In: Sagaster, K. (ed.) *Religious and lay symbolism in the Altaic world and other papers*. Wiesbaden: Harrassowitz. 77-99.

A new attempt to classify the Turkic languages (2)

Claus Schönig

Schönig, Claus 1997. A new attempt to classify the Turkic languages (2). *Turkic Languages* 1, 262-277.

4. Interactive areas between Central Turkic and non-Central Turkic units

Some features reveal the strong areal interaction between Central Turkic and non-Central Turkic units. Thus the forms of the agent noun of the type (verbal noun) +čI show a distribution more or less identical to that of the genetic *tağliğ-feature (see 3.1.). Among the genetically connected Central Turkic branches (see 3.1.) we find in Oghuz -IjI, in Kipchak (including Uzbek) *-UwčI and in South East Turkic New Uigur -GUči; in the areal group North East Turkic we have *-A(:)ččI, see 3.2.2.¹ As for non-Norm Turkic, Khalaj has the same suffix -GUči as New Uigur, see also 4.1.3. The Chuvash form -ĂvśĂ resembles the Kipchak form (see also 4.2.1.).

The picture of internal connections can be refined by means of another feature, the development of first syllables consisting of a palatal vowel and a weak consonant, of which at least one element is labial, see Schönig (1992a). Again we obtain an inner segmentation of Kipchak not greatly different but more differentiated than that based on *tağliğ, see 3.1. Oghuz, mainly Western Oghuz, behaves differently from Kipchak, which itself resembles South East Turkic, see 4.1. Besides the special position of Uzbek between Kipchak and South East Turkic, see 3.1., this feature also stresses the transitory position of Turkmen between Oghuz and Western Central Asian Turkic (WCA, see 4.1.1.2.). At the same

¹ See Schönig (1991). In some South Siberian Turkic units we find additional forms like -(I)GčI in Tuvan (TuvGr 313) or a suffix -čI (< *-(I)GčI?) in Khakas.

time we mainly see Western Oghuz and the non-Norm Turkic units designated by a higher degree of archaicism. Again, neither South Siberian Turkic nor North East Turkic appear as relevant groups. In Kipchak, except for Kirghiz-Kipchak, these sound groups have converged in *üy*, in Volga Kipchak in *öy*; in Mishar Tatar and Karaim the “Far West” (see 6.) has variants with (sometimes unetymological) final *-w*, see Berta (1989) and Schönig (1992a). In Kirghiz-Kipchak we find *üy* besides contracted forms with long labial vowels. Here, New Uigur and Uzbek behave like Kipchak languages of the non-Kirghiz type, but Uzbek has not consistently changed *äv* into *üy*, *öy*. In the case of *äv* Altay Turkic uses, besides *öy*, forms in which labiality still has not crossed over to the vowel. Here, Yenisey Turkic and Fu-yü with *ib* are different from modern Kipchak (including Kirghiz-Kipchak) and Sayan Turkic and more closely resemble the afore-mentioned languages in the west and south. New Uigur sometimes has kept *ŋ* or produces an unetymological *g*. South Siberian Turkic, excluding Altay Turkic (see 3.2.4.), preserves *-g* and *-ŋ* in word-final position. While Turkmen shows its affinity to Western Central Asian Turkic (see also 4. and 4.1.1.2.) by having developed many *y*-variants, Western (and Khorezmian) Oghuz (and some New Uigur dialects) together with the non-Norm Turkic units Lena Turkic and Chuvash demonstrate a clear tendency to preserve the feature of nasality, e.g. **söŋgök* ‘bone’ > Azeri *sümük*, Turkmen *süŋk*; Chuvash *šämă*; Yakut *uŋuox*; Uigur (dialectal) *söŋäk*. Khalaj is also conservative regarding *äv* and nasality.

4.1. Oghuz and non-Oghuz Central Turkic

Some features can be used to tie Kipchak and South East Turkic closer together while at the same time separating them from Oghuz; of South Siberian Turkic at least its Kipchakoid constituents demonstrate connections to Kipchak-South East Turkic. Most of the non-Oghuz Central Turkic units express (im-)possibility of performing an action by use of the verb connection *-A al(-ma)-*; besides, in many units forms like *-p bol-* or *-sA bol-* exist, see Schönig (1987a). The form *-A al(-ma)-* is attested in most of the Kipchak and South East Turkic languages, in Salar, Yellow Uigur, Tuvan and perhaps in Chuvash.² South Siberian Turkic

² For Chuvash *-ay-* < **-A al-*, see Benzing (1959c: 721) and Levickaja (1976: 54-55).

Altay, Yenisey and Chulym Turkic use *-A/p al-*. All these *-A/p al-* and *-A/p- bol-* forms seem to be absent in Oghuz, see also 4.1.2. Lena Turkic has the enigmatic suffix *-(A:)yA-*.

For the verb ‘to cry’ we find in Kipchak and South East Turkic **yǵla-* forms (< **hiǵla-*, see Doerfer 1995) like Tatar *yıla-*, Karačay-Balkar *ǵıla-*, Kazakh *žıla-*, Uzbek *yǵla-*, New Uigur *ǵǵla-*. Kirghiz-Kipchak (and Bashkir) has *ıyla-*, which as **iǵla-* matches the South Siberian Turkic forms like Yenisey Turkic Shor, Khakas *ilǵa-*, Sayan Turkic, Tuvan, Karagas *iǵla-* etc. *Fu-yü jǵlǵi- ~ yǵlǵi-* is a (non-Kirghiz-) Kipchak South East Turkic form with Yenisey Turkic metathesis. Yellow Uigur (*yǵla-* ~ *iǵla-*) has intermediary forms between Kipchak-South East Turkic and South Siberian Turkic. Again non-Norm Turkic Chuvash *yēr-* and Lena Turkic Yakut *ıta-* are deviant forms; Khalaj *hiǵla-* shows preservation of *h-*, see 2.2.1. The Oghuz forms do not follow a common pattern. In Western Oghuz Qaşqa’i of Firuzabad (Doerfer 1990a: 114) and the Turkmen literary language we find *aǵla-*, in Afshar (*h*)*aǵla-* (Doerfer 1989b: 399), in Sonqori and Khorasan Turkic *yǵla-*. We may assume that the Oghuz *aǵla-* forms are secondary or go back to a stem different from **hiǵla-*.

4.1.1. Oghuz

Here I only list Oghuz features not already mentioned in previous chapters. Common to all modern Oghuz languages is preservation of the verb **bāŋzā-* ‘to resemble’. The word **čojuq* for ‘child’ mainly exists today in the Oghuz area.³ Among its modern features, Oghuz has produced suffixes with initial vowel, which require binding consonants (see Schönig (1995a) and 4.1.3.2.). In all Oghuz units *n* is used as a binding consonant for the genitive suffix; the other binding consonants for dative and accusative are different at least between Turkish, Azeri and Turkmen.

³ Severtjan (IV: 28) gives *jujuq* for neighboring Uzbek. Róna-Tas reads *čojuq* in one of the Nagyszentmiklós inscriptions (1990: 21), which would mean it is also attested in Bolgar Turkic.

4.1.1.1. The internal segmentation of Oghuz

Oghuz can be divided into Western (and Southern) and Eastern Oghuz (mainly Turkmen).⁴ Western Oghuz has a binding consonant *y* in the post-vocalic forms e.g. of the dative in *-(y)A*, the gerund in *-(y)Ip*, the verbal noun in *-(y)Iš* or the participle in *-(y)An*, whereas Eastern Oghuz Turkmen has long vowels due to contraction; but the Western Oghuz future suffix *-(y)AĵAK* is *-ĴAK* in Turkmen.⁵ To express impossibility Turkish has preserved the Old Turkic *-A u-* construction (from **u-* 'to be able to') in *-(y)AmA-*, whereas it uses *-A bil-* to express possibility. In Azeri (im-)possibility is expressed by the biverbal construction *-A bil(-me-)*, the same as in Kipchak, South East Turkic and Chuvash. Replacement of the vocalic gerund by the gerund in *-B* is found in Turkmen, Baraba and North East Turkic (see 4.1.). In most branches of modern Turkic except Oghuz, case-marked **qay-*stems can be used to ask for places, directions, aims, sources and the like. In Oghuz we find interrogative pronouns like Turkish *nerede* or Turkmen *nirede* (derived from *ne* 'what'), while Azeri employs the same derivational element *+rA* in **hara+* on a **qa(n)-*stem, see Schönig (1995c).

4.1.1.2. Turkmen and Western Central Asian Turkic

Turkmen has a whole set of features separating it from Western Oghuz and tying it somehow closer to other Central Turkic and Border Turkic units. Thus, it has, like Kipchak, South East Turkic and South Siberian Turkic, replaced the perfect participle in **-mIš* by **-GAn* (> *-An*). Of these **-GAn*-Turkic groups only modern South East Turkic has not preserved the Old Turkic 3rd ps.poss. accusative form **+(s)In* consistently. Like in Kipchak and South East Turkic (Literary Uzbek and New Uigur), there are no formally analogized negative *-mA-*forms of the gerund in *-B* and the vocalic gerund, instead **-mAyIn*-forms appear, whereas South Siberian Turkic often has *-BAy*-forms. We may assume that these features were preserved in or passed over to Turkmen by areal interaction in the Western Central Asian Turkic area (see 4.). This is supported by the fact that we indeed find common Turkmen and WCA Kip-

⁴ For a more detailed internal segmentation of Oghuz, see Doerfer (1990b).

⁵ The *-AĵAk*-forms in Crimean Tatar, Caucasus Kipchak, Tatar and Uzbek must be exported Ottoman forms; perhaps Nogay *-AyAK* belongs here, too.

chak-South East Turkic features not found in Kipchak and South East Turkic units outside this area, e.g. the marking strategies in relative clauses, the headword of which is not referentially identical with the subject of the relative clause. We can observe the Western Central Asian Turkic units Uzbek, Kazakh, Turkmen (and even Salar, which perhaps originated there or at least has some genetical ties to this area) together with Kirghiz-Kipchak and Lena Turkic regularly using constructions in which a possessive suffix on the headword refers to the subject of the relative clause. The neighboring areas of the “Far West” (Chuvash and Volga-Ural-Caucasus Kipchak) and the “Far East” (New Uigur, Yenisey Turkic and Karagas) do not employ such constructions regularly, if they use them at all. They employ types with no possessive subject marker—neither on the headword of the relative clause nor on the participle, which serves as a nonfinite predicate of the relative clause. Only Western Oghuz and Khalaj have regular possessive subject marking on the *-DIK*-participle, a form almost exclusively used for such types of relative clauses. Sayan Turkic is divided into two groups: The first group is represented by the Tuvan literary language, which has all three types. The second group is represented by Karagas following the Chuvash-Kipchak model, see Schönig (1992b and 1993a).

Many of these WCA Turkmen features can be described as made up of Oghuz material along Western Central Asian Turkic patterns. Although Turkmen has reflexes of Old Turkic *ń* (see 3.2.4.1.), it has a *y*-form *haysi* for the attributively used Old Turkic pronoun *qa:ńo* ‘which’ like the other units in the Western Central Asian Turkic area and has no *n*-forms like Western Oghuz-Khalaj (see 4.1.3.3. and Schönig 1995c). Additionally, the marking system of perfective versus cursive participles reveals the same isoglosses. Chuvash, Lena Turkic and Western Oghuz show no formal connections between the opposing forms. In Turkmen the cursive participle *-yAn* is marked against the perfect participle *-An* by the same marker *y*, which is used to mark the renewed present tense form *-yAr* against the aorist form *-(A)r*. This means Turkmen uses the sign of cursivity common to Oghuz and going back to the verbal connection **-A yori-* similar to the way Kipchak(oid) and South East Turkic languages use the marker **-A tur-* (or at least contain traces of such a use): They render the oppositions expressed as **-GAn : *-A turgan* with

the participles, as $-(V)r : *-A \text{ turur}$ with the present tense forms.⁶ Also, the Old Turkic verb $\ddot{i}d-$ 'to send', which has survived in North East Turkic and Chuvash, and its "long forms" going back to $\ddot{i}du \text{ ber-}$ 'id.' in Kipchak and South East Turkic are both missing in Western Oghuz, but Turkmen has—like the Western Central Asian Turkic units—a (non-auxiliary) verb $ibär-$ 'to send'; as an auxiliary in comparable function it uses $goyber-$, which may go back to $*qoyu \text{ ber-}$, which is perhaps the original Oghuz counterpart to $\ddot{i}du \text{ ber}$. Incidentally, one can see that Turkmen seems to have adopted the system of biverbal constructions expressing actionality from Western Central Asian Turkic.

Sometimes the transitory position of Turkmen between Western Oghuz and Central Turkic leads to a whole variety of forms bearing one function, as in the case of biverbal compositions with the verb $başla-$ 'to start, to begin'. The preceding verb can appear with gerunds in $-A$ and $-B$ as well as with the dative of the verbal noun in $-mAK$; the literary language has $-(y)Ip \text{ başla-}$, for $-A \text{ başla-}$ and $-mAGA \text{ başla-}$, see Benzing (1939) and 4.1.3. In the case of the sound group $*äv$, Turkmen, different from Western Oghuz (like Kipchak and some South East Turkic units), tends to shift the feature $[+labial]$ onto the vowel, so that we often

⁶ If we also take into account the Lena-Sayan Turkic data (see 3.2.4.1.) we may assume that North East Turkic became an interactive area after the development of special strategies of cursivity marking in Kipchak(oid), South East Turkic and the different Oghuz branches (see Johanson 1976b). Sayan Turkic stopped on a very archaic level, Lena Turkic never made an attempt of recursivation, perhaps because it was isolated from the other Turkic units while the recursivation of the present tense forms was going on. We find cursivity in Western Oghuz marked with the same element $y < *-A \text{ yorī-}$ as in Eastern Oghuz. The Kipchak(oid) and South East Turkic units employ the element $*-A \text{ tur-}$. Especially in the east of this area, a second wave of renewal must have taken place, in which the connections $*-A/p \text{ yata turur}$ played a particularly important role. Only after these renewals had taken place, could the before-mentioned constellation have been formed. Chuvash shows a renewed cursive present tense form of the Kipchak type $*-A \text{ tur-}$. With respect to the many Tatar influences on Chuvash (see 4.2.1.), it seems very likely that the renewed present tense form is also due to such an influence. The marking system of participles in Chuvash does not show a corresponding element, on the contrary: At least today the perfect and cursive participles $-n\ddot{A}$ and $-Akan$ do not formally establish a privative opposition, but see fn. 11.

find results like *öy*. Different from Western Oghuz, in Turkmen we sometimes find preservation of **G* after the end of the first syllable, perhaps under Western Central Asian Turkic influence, see Doerfer (1990b: 32).

Turkmen has not adopted the paradigms of politeness, which are typical of non-Oghuz Western Central Asian Turkic, see 4.1. and 6.

4.1.1.3. Western Oghuz

Western Oghuz often shows traces of its relatively isolated position in the extreme southwest. Azeri has had intensive interaction with Persian, Turkish has had additional interaction with Greek and other Indo-European languages of the Balkans, Gagauz has been in contact with Slavic languages. As results of such interaction we find e.g. elaborate systems of complex conditional forms absent in other branches of Turkic or subordinated clauses with finite predication (the latter mainly in New Turkic Azeri and Gagauz, see Schönig (1993b)).

Within Western Oghuz the transition from Turkish (representing Western Anatolian Turkic) to Azeri (Eastern Anatolian Turkic) can be demonstrated by features such as preservation of deep vowels, e.g. Old Turkic *bädük* > (Azeri) *böyük* > (Turkish) *büyük* 'big', nasalization of word-initial **b*- in Azeri if a nasal consonant follows (see fn. 16) and the penetration of the personal marker of the 1st ps.pl. *-*K* into other paradigms than those of the *di*-preterite and the conditional in Azeri. For these and other features see Doerfer (1990b: 14).

Another set of features appears in Azeri and Turkmen, but is missing in Turkish. Quite a number of these are Central Turkic, like the reflexive pronoun *öz* or the postvocalic accusative in *-nI*.⁷ A radical form of analogization is the negation of the aorist *-mAr*, which has penetrated the paradigms of both languages, but can sporadically be found in other units, too, e.g. in Western Siberian Tatar, see Axatov (1963).

One of the individual features of Turkish e.g. in the area of lexicology is the replacement of the old word *söğök* 'bone' by *kämik*.

⁷ In transitory Azeri dialects +*yI* appears. Khorasan Turkic (according to Tulu 1989) behaves, with postvocalic dative and accusative forms +*yA* and +*nI*, like Azeri.

4.1.2. Archaic features in Oghuz and Border Turkic

As can be seen from the paragraphs above, Oghuz, especially Western Oghuz, has a special position within Central Turkic by having preserved many more Old Turkic features than other Central Turkic units. In this respect Oghuz often behaves like Border Turkic (see above 3.2.4.1. and 4.). Here, as with Oghuz / non-Norm Turkic features (see 4.1.3.), in most of the cases the question, whether or how an Old Turkic feature is preserved, can be used to separate Western Oghuz from Turkmen. Thus, Old Turkic vowel length is preserved in long vowels or diphthongs in Turkmen as well as in Lena Turkic and Khalaj (see Doerfer 1971) and long **ö*: as (*ä*)*va* in Chuvash. Western Oghuz, Sayan Turkic, Salar and Yellow Uigur have short vowels but contain reflexes of vowel length in the consonants following them, see Janhunen (1980), Johanson (1986b), Schönig (1991).

Many of the archaisms of Turkmen are common to Eastern Border Turkic: The Turkmen form of the 1st ps.sg. points to *-*AyIn*, the Western Oghuz ones to *-*AyIm*, see Schönig (1987b). Like Lena-Sayan Turkic and Chuvash it has no *-*K*-marker for the 1st ps.pl. imp., see 3.2.5. Turkmen and Lena Turkic are the only units which today use -*ŋ* exclusively to designate the imperative of the 2nd ps.pl.

In the case of the word for 'lip', Oghuz (*dudaq*, *dodaq*) together with Khalaj (*dudaq*) and Chuvash (*tuta*), but also far-eastern Salar (*dodax*) point to a form **tutaq*, while most of the other Turkic languages have **ärin*,⁸ again Lena Turkic with *uos* (< *ağiz* [+ okanie]?) stands apart.⁹ Western Oghuz has preserved a short form *äl* of Old Turkic *älig* 'hand' like Khalaj and Salar (for the long forms see 3.2.). Normally it is replaced by **qol* meaning both 'hand' and 'arm', see also 4.1.3.

4.1.3. Oghuz and non-Norm Turkic

Another set of features which is even absent in Border Turkic is attested in non-Norm Turkic and Oghuz (in most cases in differing shapes in Western Oghuz and Turkmen). Like Lena Turkic, Khalaj and Western Oghuz, Turkmen still uses Old Turkic -*mAdOK* as the negation of the

⁸ For these two words, see Doerfer (1988: 59, 104, 174-176, 237). Maybe the words for 'fishing pole' in Turkish (*olta*) and Chuvash (*välta*) belong here too.

⁹ See Doerfer (1988: 101-102). For *okanie* in Yakut, see Ivanov (1980).

perfect participle, although it has given up the *-mİš*-participle in this function, see 4.1.1.2. Western Oghuz has kept the *-mİš*-form, but shows, besides *-mAdOK*, completely analogized negative forms—in the case of the participles as well as the gerunds. Of the latter, only forms like Turkish *-mAdAn* remind us of older stages with common negations. Regarding the sound group *äv* Western Oghuz is more conservative than Turkmen (see 4.) and has—like non-Norm Turkic—preserved the labial consonant, e.g.: *äv* ‘house, home’: Ottoman *äv*, Gagauz *yev*, Azeri *ev*, *öv*, Turkmen *öy*, Khalaj *häv*. Perhaps the use of **qizil* for ‘gold’ in Azeri, Turkmen dialects and Lena Turkic belongs to these features, too. Different from Turkmen, the Azeri literary language does not have the word *altın*, see Doerfer (1965: 85), Cincius & Bugaeva (1979), Schönig (1990).

Especially Western non-Norm Turkic and Oghuz employ biverbal forms for ‘to begin to x’ by using the verbal noun in *-mA(K)* of the verb meaning ‘x’. In Oghuz we find *-mAGA başla-*, while in Khalaj *-mAKKA häna: sa- / ba’sla-* is rare, see Doerfer (1988: 136). Chuvash has *-mA pušla-* and an alternative form **-A başla-*, which connects it with many Kipchak units. Dative-marked forms of other verbal nouns can be found e.g. in Tatar *-(V)rGA kěřěš- / totin-* (aorist) or in New Uigur which uses dative forms of the verbal noun in *-(I)š* and *başla-*. Turkmen has (like in the case of the (im-)possibility form, see 4.1.1.1.) a form with the gerund in *-B* (which is also in use on the border of the Kipchak area and in Yenisey Turkic) sometimes alternating with *-A*-forms like in Kumyk. Khalaj and New Uigur have preserved the archaic use of *-GAll* or *-GIII* in such constructions. Replacement of *başla-* by other verbs can be found in Khalaj and Lena-Sayan Turkic.

Another common Oghuz-Chuvash-Khalaj feature is the necessitative suffix Oghuz *-mAll*, Chuvash *-mAllĀ*, Khalaj *-mAlU(G)* (competing with the form *-GUlUk*, see Doerfer (1988: 145), which is also used in New Uigur). The sporadic appearance of *-mAll*-forms in Tatar or in Uzbek may be due to later Ottoman language export.

4.1.3.1. Western Oghuz and non-Norm-Turkic / Border Turkic

Some Old Turkic features have only survived in Western Oghuz and non-Norm-Turkic. Thus, only Western Oghuz, Khalaj, Lena Turkic

have preserved the participle in **-DOK*,¹⁰ *ürüŋ* ‘white’ has only survived in Lena Turkic (Yakut *ürüŋ*), Khalaj (*hürüŋ*) and Anatolian dialects (see *ürün* (I + II) in DS XI: 4071, and Schönig (1987a)). The **-mlš*-participle is still used only in Western Oghuz, Lena Turkic and Salar. Additionally, the participle has survived functionally narrowed in some Western Central Asian Turkic languages (mainly in Eastern Oghuz Turkmen and South East Turkic; for Uzbek see Kononov (1960)). Moreover Western Oghuz Gagauz and Lena Turkic still employ non-analogized forms of positive and negative participles like the Old Turkic **-mlš*, **-DOK* : **-mADOK*-system. Chuvash has a common negative form *-mAn* for both the cursive and the perfect participle.¹¹ Another common feature of Western Oghuz (and perhaps under its influence Crimean Tatar) and Border Turkic units like South Siberian Altay Turkic and Yenisey Turkic Khakas (especially Sagay) or non-Norm Turkic Lena Turkic is the existence of an instrumental case suffix of the type *+(I)nAn* ~ *+nAn* or the like, see Schönig (1997).

4.1.3.2. The Oghuz-Chuvash connection

We can isolate a set of features somehow connecting Oghuz units and Chuvash closer together. The word for ‘navel’ in these two branches goes back to something like **gö:bäk*, while all the other languages point to something like **kindük*. Most of the South Siberian Turkic units in addition have a “short form” *kin*, Lena Turkic and Karagas only have *ki(:)n*. In Radloff’s materials we also find for Yenisey Turkic Khakas and Sagay only *kin* ‘Nabel des Moschustieres, Moschusbeutel’ (II 1344), but no *kindik*, whereas an alternative “short form” *kin* is only attested for the Altay Turkic dialects Altay, Teleut and Lebedin (Quu). However, these data need careful interpretation and should not be taken

¹⁰ **-DOK* > Yakut *-TAx* (very limited in use, see JakGr 237), Khalaj *-DUK* (rare, see Doerfer 1988: 129), Oghuz *-DIK* (very common).

¹¹ If we assume as the underlying proto-Chuvash system (perfect(-neutral) participle : cursive participle) : negative participle = (**-An* : **-AGAn*) : **-mAn*, we obtain a formal marking system, which resembles on the positive side of the complex opposition (in brackets) the marking system of cursivity against perfectivity / neutrality in Kipchak, South East Turkic and Turkmen, see 4.1.1.2. As a whole, the complex opposition parallels the marking system of *-B-* and the vocalic gerunds in Kipchak, South East Turkic and South Siberian Turkic.

as a hint for internal differences between South Siberian Turkic idioms. Especially Western Oghuz is tied closer to Chuvash by the existence of the verb *barın-* 'to take shelter, to lodge', which seems to be paralleled only by Chuvash *purăn-* 'to live'. Besides lexical features we find an astonishing parallel in suffix structure, mainly in that of case suffixes: Chuvash and Oghuz have kept the Old Turkic type of genitive suffix $^{*}+(n)I\eta$ with the structure $+(C)VC$; most of the modern units have produced by analogy genitive suffixes of the frequent type $+CVC$.¹² In Chuvash and Oghuz the $+(C)VC$ -type has spread onto the dative and the accusative suffixes too. In other Turkic units it can be found only sporadically. So we find an accusative suffix of this Chuvash-Oghuz type $...V.nI$ in Lena Turkic. Salar has a dative suffix $+(G)A$ (see Dwyer 1997), which perhaps has emerged independently by internal analogy. For Khalaj, see Doerfer (1988: 87-88).

I assume the Oghuz-Bolgar connection to date back before the collapse of the Khazar empire (or even earlier).

4.1.3.3. The Oghuz-Khalaj connection

The beginnings of the Oghuz-Khalaj connection may date back to the period before Kāšgārī. Doerfer assumes that the modern Khalaj language is connected with the language of Kāšgārī's Arghu tribe (see Doerfer 1987). There are no real facts known about the history of the Khalaj. But what we can see is that areal interaction between Oghuz and Khalaj has been remarkable. I think we may assume generally that in all cases of the Khalaj-Oghuz (at least during the last few centuries: Khalaj-Azeri) exchange of features, the small group of Khalaj speakers has been on the receiving end.

Both branches have preserved the Old Turkic negative aorist in $-mAz$ instead of today's more common $-mAs$,¹³ present tense forms going back to $^{*}-A\ yor\dot{i}-$, future forms in $-(y)(A)J\dot{A}K$ or gerunds in $-(y)(A)-r\dot{A}K$. The Khalaj personal endings of the 2nd ps.pl. $-(rs/y)A\eta I_z$ clearly resemble Oghuz forms like $^{*}+sI\eta I_z$.

¹² Today the suffix is $+nI\eta$ in most of the units. Khalaj has $^{*}+(U)\eta$, Salar $+niyi$ (see Dwyer 1997). The genitive is missing in Lena Turkic, see Schönig (1990).

¹³ Chuvash has preserved this form only in the negative present tense copula *mar* < $^{*}\dot{a}rm\dot{a}z$. The Lena Turkic $-BA\dot{t}$ -forms could also have developed after the sound change $-z > -s$.

We find numerous hints especially for Western Oghuz-Khalaj interaction. Only in this area do we encounter a systematic development **b-* > *v-* or *Ø-* in the four words *var-* ‘to arrive’, *var* ‘exists’, *ver-* ‘to give’ and *ol-* ‘to be, to become’.¹⁴ From the Old Turkic pronoun *qa:ño* ‘which’ (see Doerfer 1988: 108) attributively used **qan*-forms have only survived in Khalaj (*qa:ni(si)*) and Western Oghuz, e.g. Turkish *hangi*, Gagauz *angî* (< **qanu* + *GI* ?) or Azeri *hansî* (< **qanu* + *si*); for Qumanda *qanji* see Schönig (1995c). Perhaps another archaic element preserved in the Khalaj-Oghuz area is the 2nd ps.sg. imperative suffix *-Gil*, which is attested only in Khalaj (in the language of children’s games, see Doerfer (1972: 300)), in Chulym Turkic (Pritsak 1959a: 627) and—enlarged by a suffix *-An/än*—in Azeri dialects (Agazade 1967: 88). It seems that constructions consisting of a finite verbal form and the conditional copula **ärsä* are limited to the Western Oghuz-Khalaj area and to Sayan Turkic (Tuvan and Karagas, not Toja, see Rassadin (1978: 229–230)). Another feature common to Western Oghuz and Khalaj is the regular use of possessive marked *+DIK*-participles as verbal cores of relative clauses, see also 4.1.1.2.¹⁵

Some archaic features are preserved only in Khalaj and Turkish and a few other units, e.g. the Old Turkic 1st. ps.pl. imperative suffix **-Allm* (which also exists in the Yenisey Turkic Khakas Beltir unit) or preservation of the verb **bul-* ‘to find’ in Khalaj, Turkish and Lena Turkic, which in the other units is replaced by **tap-* which exists in Yakut, too, and means ‘to hit (the target)’.

On the whole, the before-mentioned segmentations are quite well reflected in the distribution of different postpositionally or enclitically used elements meaning ‘with’. In Western Oghuz the form *ilä* predominates in the literary languages, while e.g. some Anatolian dialects also use *bilä* and *birlän*. At the same time, enclitic forms of the type *+(y)lä(n)* (sometimes already showing sound-harmonic forms) can be found in all types

¹⁴ In the case of *var-* we also find a characteristic shift in meaning because the Western Oghuz-Khalaj forms correspond to *bar-*, which in other and in older Turkic units means ‘to go (to a point)’. In Yellow Uigur and Salar we sporadically find *ol*-forms of **bol-*, but only as one of a few alternatives like *bol-* or *vol-*, see Tenišev (1976a and 1976b).

¹⁵ In Tuvan sometimes possessive marking on the headword appears together with other participles, see Schönig (1992b).

of substandards and dialects. The same holds true for neighboring Khalaj. The transitory state of Turkmen again is apparent in the area of postpositions, with Turkmen having Western Oghuz *ilä* besides (mainly) *bilä(n)* and (rarely) *birlä(n)*. The only other modern Turkic unit which has preserved a *birlä*-form to some extent is non-Norm Turkic Chuvash with *përle*. The more successful form was *bilä(n)*, which exists in various forms with final *n* in Standard Tatar, Misher Tatar, Uzbek, New Uigur, Yellow Uigur and—as the only South Siberian Turkic unit—Chulym Turkic Küärik (+*BllAŋ*). According to its transitory state between Volga-Ural-Caucasus Kipchak and Western Oghuz, Crimean Tatar has *bilän* besides *ilän*. We find oscillation between (postpositional and enclitic) **b*-forms with and without final *n* in Kumyk, Balkar, Baraba and Salar; at least in Chuvash dialects an enclitic form +*pala(n)* can be found besides the before mentioned postposition. While South Siberian Sayan Turkic prefers postpositional *bilä*-forms, the remaining Kipchakoid South Siberian Turkic shows a strong tendency to use **bilä* enclitically. In this respect Chulym Turkic Küärik (see above) resembles more closely the Kipchak units Crimean Tatar, Kumyk, Misher and Orenburg Tatar, Baraba and Kirghiz-Kipchak Altay Turkic with their **(+)mlnAn*-forms. In some of these units as well as in Karaim and Chuvash we have short forms of the type **+mA(n) ~ *+BA(n)*. WCA Kipchak (Kazakh, Karakalpak, Kirghiz) is especially characterized by **menen*-forms showing a tendency to become enclitic, too. Fu-yü gains a special position by having shortened postpositional forms *bil*, *bul*. Lena Turkic and Yenisey Turkic Khakas (probably the Sagay dialect) stand apart by not having a postpositional unit going back to *bi(r)lä(n)* but using at least one grammatical suffix somehow connected with the Old Turkic suffix of the instrumental case; this again may point to a (temporary?) closer connection between Lena Turkic and components of Yenisey Turkic Khakas at least in earlier times, see Schöning (1997) and 3.2.4.

4.1.3.4. The Turkish-Lena Turkic connection

Turkish in the extreme southwest and Lena Turkic in the extreme northeast of the Turkic area are connected by some exclusive features like preservation of the sequence *b-...-n* in Turkish and some Yakut dia-

lects.¹⁶ Mainly in the same area the verb **tön(ün)-* is used to express ‘to turn around, return, go home’ and cognates of the Old Turkic reflexive pronoun *käntü* have survived.

The latter two features can be used to classify all of Turcia in a way more or less according to patterns already described. Turkish, Lena Turkic and sometimes Tatar (see TTAS III, 288) employ forms of a verb **tön(ün)-*,¹⁷ Chuvash uses *tavrän-* < Old Turkic *tägzin-*, Kipchak, South East Turkic and non-Turkish Oghuz (= non-Turkish Central Turkic) mainly have forms going back to Old Turkic *qadit-*. South Siberian Turkic is individually characterized by forms of Old Turkic *yan-*. The forms of the reflexive pronouns produce a parallel isogloss. The *qadit*-languages have *öz*, the *yan*-languages forms of Old Turkic **bod*. If Chuvash *xă+* cannot be connected with Turkish *kendi* < Old Turkic *käntü*, Turkish is the only living Turkic language still using the Old Turkic form as a reflexive pronoun. The form *kini* in Lena Turkic may go back to *käntü*, but it is used as the 3rd ps.sg. personal pronoun, so that the **tön(ün)*-languages except Tatar are at least connected by preservation of the word **käntü*.

4.2. Kipchak and non-Norm Turkic

Kipchak shares some features with the non-Norm Turkic units Chuvash and Lena Turkic. A lot of Chuvash-Kipchak features can be explained by relatively young but intensive contacts in the Volga area. Lena-Kipchak features may be coincidental or point back to early contacts. Some

¹⁶ Very common is the change of initial *b-* to *m-*, if a nasal (or—especially in South Siberian Turkic, but sometimes also in Kipchak—a guttural) follows at the end of the first syllable. Another type is perhaps caused by a Samoyedic substratum in South Siberian Turkic (see Menges 1955-56): Regressive nasalization of word-initial **y-* > *n-*, *ń-* by a nasal or a guttural fricative at the first syllable border in some South Siberian Turkic units, e.g. Tuba *ńa:s*, Karagas *ńes* < **yığač* ‘tree’, Qumanda, Quu *ńan-*, Shor *nan-* < **yan-* ‘to return (home)’. See Schönicg (1993c). But cases like Salar *neme*, *nemä* < *yemä* ‘food’ demonstrate that this type of assimilation exists outside the Turkic-Samoyedic area, too.

¹⁷ According to Kāšgārī *tön-* was Oghuz (see Dankoff & Kelly III: 197) and can be found only sporadically in other Turkic languages, see also Clauson (1972: 515a) and Tenišev (1961: 241).

of these features even belong to a greater Northern Turkic interactive area, see 5.

4.2.1. The Kipchak-Chuvash connection

Chuvash has one of the classical features in common with Kipchak, the development of *tag* > **taw* (Chuvash *tu / tāv*+). At the same time, the development of this feature helps to isolate Lena Turkic (which has contracted forms), whereas the remaining Turkic units have preserved *-g*, *-x* or at least non-labial reflexes of it. Another Chuvash-Kipchak feature is the 2nd ps.sg. imperative particle noted as *čU* or *šU* in Kāšgarī (today sometimes reduced to *-č* or *-š*). Besides in Chuvash (*-čă*), it is attested in Kirghiz, Nogay, Caucasian Turkic, Bashkir, Tatar dialects and Uzbek, see Schönig (1987b: 206).

Volga-Ural-Caucasus Kipchak (VUC Kipchak), i.e. Tatar, Bashkir, Karachay-Balkar and Kumyk, has another feature in common with Chuvash. In both groups we find **+IGIz* instead of **+IŋIz* for the possessive suffix of the 2nd ps.sg., see also 5. Other common features are the use of the same types of relative clauses (see 4.1.1.2.) or use of **-A turur* as a renewed present tense form (see 4.1.1.1.). The fact that the system of participle marking in Chuvash did not follow the Kipchak model makes it impossible to say whether its ancestor joined a Volga-Ural-Caucasus Kipchak area perhaps even before the renewal of cursive finite forms had begun merely resisting a renewal of the participle system (see 4.1.1.2.). One must also bear in mind that the system of phase-specifying biverbal constructions with gerunds seems somehow copied from a Volga-Ural-Caucasus Kipchak model.

4.2.2. The Kipchak-Lena Turkic connection

There seems to be an old connection between Lena Turkic and Kipchak. Thus, especially Lena Turkic and the Kipchak languages weaken *p* and *K* while they keep *t* in intervocalic position. In addition, Lena Turkic shows some structural similarities with the Kipchak languages in the area of phonotactical rule sets (see 5.1.). It seems possible that the Lena Turkic-Kipchak connection is more precisely a Lena Turkic-Kirghiz-Kipchak connection, see 5.1.1.

4.3. South East Turkic and Khalaj

A common New Uigur-Khalaj set of features seems to consist of the use of the agent noun **-GUČI*, the necessitative suffix **-GUIUK*, *-GAI* *-GIII*-forms as the connecting deverbal unit in the case of constructions with *bašla-* to express 'to begin to x', and generally doubled intervocalic consonants in numerals, see 4., 4.1.3. and 6.1. It is still unclear whether these common features indicate a closer connection in earlier times between the ancestors of South East Turkic and Khalaj.

(To be continued.)

Corrigenda to part 1:

p. 121, line 5: ... (see Erdal 1993) use of a gerundial unit *-sA*...
should be: ... (see Erdal 1993), and the use of a gerundial unit *-sA* ...

p. 124, footnote 4: ... Mrass-dialect of Shor...
should be: ... Kondoma-dialect of Shor...

p. 125, line 7: ... loss of the plural marker *+lAr* ...
should be: ... use of the plural marker *+lAr* ...