Werk

Titel: Articles

Ort: Wiesbaden

Jahr: 2006

PURL: https://resolver.sub.uni-goettingen.de/purl?666048797_0010|LOG_0010

Kontakt/Contact

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

info@digizeitschriften.de

Reciprocal constructions of Turkic languages in the typological perspective

Vladimir P. Nedjalkov

Nedjalkov, Vladimir P. 2006. Reciprocal constructions of Turkic Languages in the typological perspective. *Turkic Languages* 10, 3-46.

In this paper, on the basis of the previously published papers on reciprocals in Karachay-Balkar, Yakut and Kirghiz (*Turkic Languages* 6, 19-80; 7, 30-104; 7, 181-234), and also other publications on reciprocals, Turkic languages are compared according to the range of polysemy of reciprocal markers as well as the productivity of their different meanings. Typological parallels of these sets of polysemy or the absence of such parallels are established. The specific nature of this polysemy and its unique marking as part of a five-member voice system is brought out cross-linguistically. For discussion, eight issues are selected. Naturally, the characteristics cited below are not exhaustive, and the paper is to a certain degree fragmentary.

Vladimir P. Nedjalkov, Institute of Linguistic Research, Russian Academy of Sciences, Tuchkov per. 9, St. Petersburg, 199053 Russia. E-mail: nedjalkov@typology.spb.su

1. Two types of expressing the reciprocal meaning in Turkic languages

1.1. Reciprocal forms: verbal and pronominal reciprocals

As is known, in all the Turkic languages, there are two main reciprocal markers, suffixal (see (1) and (8)) and pronominal (see (2), (3), and (5)) and thus two types of reciprocals, verbal, marked with suffixes, and pronominal.[1] They enter into two main types of reciprocals distinguished cross-linguistically, where verbal reciprocals can also be marked with other types of affixes (prefixes, infixes, circumfixes), clitics, root reduplication, grammaticalized doubling of a clause, etc.

Suffixal markers (*-ïš* or *-ïs*) are polysemous. Positional variants for the consonant are *-s/-h-* in Yakut, *-š/-z/-č* in Khakas, *-š/-ž/-č* in Tuvan, and *-š/-š'-* in Chuvash. In Yakut, sometimes a reduplicated marker *-sïs* with the same meaning(s) is used. The

[1] This project was partly financed by a grant of the President of the Russian Federation for the support of leading research schools (Grant No. NSh-2325.2003.6). Certain aspects of this paper were contained in a paper I read at the Max-Planck-Institute for Evolutionary Anthropology, Department of Linguistics (June 2002), at whose invitation I spent three productive months. Parts of the paper are also influenced by the ideas of this Institute's project entitled WALS (World Atlas of Linguistic Structures). An earlier version of this paper was presented at the International Symposium LENCA-2 in Kazan', May 2004.

vowel preceding the consonant is subject to vowel harmony or it can be absent under certain conditions. The suffixal marker, though varying phonemically, is genetically the same in all the Turkic languages, cf. Karachay-Balkar (1a) and Yakut (1b) respectively:

(1) a. *tanï-* 'to know' → *tanï-š-* 'to get acquainted with each other'
b. *taptaa-* 'to love'→ *tapta-s-* 'to love each other'.

Pronominal reciprocal markers are monosemous. They are marked for case and person (the latter is indicated by the reflexive-possessive marker, e.g. *-leri-* in (2a), *-ï* in (2b), etc.). In contrast to suffixal markers, they are of at least two varieties. In this paper, a limited number of Turkic languages are taken into account. As there are about 40 living Turkic languages (see, for instance, Tenišev 1997), there may be other types of reciprocal pronouns.

Type one. This type of reciprocal pronouns is derived from reflexive pronouns by means of root reduplication and is typical of North-Eastern Turkic languages. This variety has two subtypes, illustrated below by Yakut and Tuvan 3rd person reflexive and reciprocal pronouns (*-leri-n* = REFL.POSS.3PL-ACC; *tar-ï-n* = PL-REFL.POSS.3-ACC; all the accusative forms for the three persons of both reciprocal pronouns are given in (3)):

(2) a. Yakut *beye-leri-n* '(they) themselves' → *beye-beye-leri-n* '(they) each other'.
b. Tuvan *bot-tar-ï-n* '(they) themselves' → *bot-bot-tar-ï-n* '(they) each other.'

The other languages of this area display reciprocal pronouns either of the first subtype (see (2c, d)), or the second subtype (see (2e, f)), the variation being possible even among the dialects of one language (Shor-1 is a dialect spoken along the River Kondoma and Shor-2 on the River Mrassu; Tenišev 2002: 540); cf. (*-lar-, -tar-* = PL; *-ï-* = REFL.POSS.3; *-n* = ACC):

c. Altai *boy-ï-n* '(they) themselves' → *boy-ï-boy-ï-n* '(they) each other'
d. Shor-1 *poy-lar-ï-n* '(they) themselves' → *poy-poy-lar-ï-n* '(they) each other'
e. Shor-2 *pos-tar-ï-n* '(they) themselves' → *pos-pos-tar-ï-n* '(they) each other'
f. Khakas *pos-tar-ï-n* '(they) themselves' → *pos-pos-tar-ï-n* '(they) each other.'

In Khakas, alongside the pronoun (2f) the adverb *udur-tödir* <entgegenzurück> 'each [opposite] other' is also used as a reciprocalizer, without the support of other reciprocal markers (Letučij, forthcoming). In this respect it seems to be the only adverb in the Turkic languages that functions as a reciprocal marker on its own (however, it requires additional study). For instance, the respective adverb *udur-dedir* in Tuvan cannot function as a reciprocalizer on its own (Kuular, forthcoming).

In the Altai language, a marker of the second type like *biri-n biri* in (5), i.e. the type of marker uncharacteristic of this area, is also attested; cf. (*-si* = REFL.POSS.3; *-n* = ACC):

h. *pir-si pir-si-n* 'each other', lit. 'one-their one-their-ACC' (Čankov 1961: 202).

Here are the forms of all three persons of reciprocal pronouns in the accusative case for two languages (*-biti-,-ïvïs-*=REFL.POSS.1.PL;*-ɣiti-,ïŋar-*=REFL. POSS. 2.PL; *-n, -tï, -nï* =ACC):

(3)	Yakut		Tuvan
	a. 1PL	*beye-beye-biti-n*	*bot-bot-tar-ïvïs-tï* '(we) each other'
	b. 2PL	*beye-beye-ɣiti-n*	*bot-bot-tar-ïŋar-nï* '(you) each other.'
	c. 3PL	*beye-beye-leri-n*	*bot-bot-tar-ï-n* '(they) each other.'

The Yakut reflexive and reciprocal pronouns are Mongolic borrowings resulting from long areal contact (and borrowed together with the noun *beye* 'man', 'body' from which they are derived). As regards Tuvan, it uses the Turkic root *bot* 'body', and copies the Mongolic pattern of reciprocal formation; cf. Buryat (*Ø-ee* = ACC-REFL.POSS):

(4) *bey-Ø-ee* 'themselves' → *beye-bey-Ø-ee* 'each other.'

Type two. In this type reflexive and reciprocal pronouns have different roots. In other than North-Eastern Turkic languages, the reciprocal pronouns are a reduplication of the numeral *bir* 'one'. As for the reflexive pronouns, they have at least four roots across these languages.

This type of reciprocal pronouns was registered as early as in the 11th century, in the works of Mahmûd al-Kâšgarî.

Examples of Karachay-Balkar, Kirghiz, Turkish and Chuvash reflexive and reciprocal pronouns respectively (*-leri-, -dör-, -lerin-, -se-* = REFL.POSS.PL.3; *-n, -ün, -i, -ne* = ACC):

(5)	a. *kes-leri-n* '(they) themselves'	→ *biri-n biri* '(they) each other'
	b. *öz-dör-ün* '(they) themselves'	→ *biri-n biri* '(they) each other'
	c. *kendi-lerin-i* '(they) themselves'	→ *bir-birin-i* '(they) each other'
	d. *hăy-se-ne* '(they) themselves'	→ *pĕr-pĕr-ne* '(they) each other.'

If we consider both types of reflexive and reciprocal pronouns, we see that across the Turkic languages there are at least six reflexive pronouns derived from different roots and at least three reciprocal pronouns with different roots as well. The causes of such variety are unknown.

(My thanks to I. A. Nevskaja and A. M. Ščerbak for the consultations on the data cited in this section).

Both the first and second types of the reciprocal pronouns described here are widely represented in the world's languages. Examples of the first pattern from Lezghian (Haspelmath 1993: 55) and Twi (Boadi 1975: 55) respectively:

(6) a. *čeb* '(one)self' → *čpi-čeb* 'each other'
b. *hō* '(one)self' → *hō hō* 'each other.'

The second type of reciprocal pronouns, i.e. the type not derived from reflexive pronouns, is represented by two-member combinations of words meaning 'one', 'person', '(an)other', etc. Compare Malayalam (Jayaseelan 2000: 119), English, Russian and Latvian examples respectively:

(7) a. *taŋŋaL* 'themselves' — *oraaL ... oraaL* lit. 'one person ... one person'
b. *oneself, myself ... themselves* — *each other*
c. *sebja* 'oneself, myself ...' — *drug druga* lit. 'another another(ACC)'
sebja 'oneself ....' — *odin odnogo* lit. 'one one(ACC)'
d. *sevi* 'oneself ....' — *cits citu* 'another another(ACC)', etc.

All the reciprocal pronouns cited above have a two-component structure, which iconically reflects the reciprocal situation with two participants (or two groups of participants) and their respective actions.

1.2. Relationship between verbal and pronominal reciprocals

With respect to the means of expressing reciprocity, the following exposition shows the place of the Turkic languages among the world languages. As mentioned, cross-linguistically, languages can employ pronominal and/or verbal reciprocal markers. With respect to the use of these types of devices, the following types of languages can be distinguished:

a) languages employing reciprocal pronouns only, as is the case in Basque, English, Finnish, Lezghian, Georgian, etc.;

b) languages employing only verbal reciprocal markers, e.g. Yukaghir, Quechua, Ainu, Mundari, Amele, etc.;

c) languages employing both types of devices, but in this case the latter can be used

i) only separately (German, Lithuanian, Polish, etc.) or

ii) either separately or simultaneously with reciprocal pronouns (Chukchi, Japanese, Mongolic, etc.).

The Turkic languages are of type (c.ii). Respective Yakut illustrations: (8) where reciprocity is expressed by the suffix alone (*-s* > *-h* in intervocalic position), (9) with both markers used simultaneously and (10) with reciprocity expressed by the reciprocal pronoun alone:

(8) *Ïhïkmï-h-an* *kebis-ti-ler* (Kïïs Debeliye 1993: 222)
let.go-REC-CONV AUX-PAST-3.PL
'(They) let each other go.'

(9) *Beye-beye-leri-n* *kör-s-ön* *tur-but-tar-a ...*
each-other-3.POSS-ACC look-REC-CONV AUX-PAST-PL-CONV
'(They) having looked at each other ...' (ibid., p. 132).

(10) *Beye-beye-leri-n* *xolun-nar-al-lar* (Xaritonov 1963: 35)
each-other-3.POSS-ACC run.down-PL-PRES-3.PL
'They are running each other down.'

Across Turkic languages, suffixal and pronominal reciprocal markers differ in their usage.

For instance, in Yakut and Tuvan, the main reciprocal marker is the suffix, the reciprocal pronoun being used much less frequently and, often, pleonastically with a suffixed reciprocal. Thus, for instance, (9) is the only case in the 6,000 lines of the epic poem "Kïïs Debeliye", and, characteristically, the verb also carries the reciprocal suffix. In another Yakut epic poem entitled "Modun Er Soγotox" (The Powerful Er Sogotox"; over 6,400 lines), the reciprocal pronoun does not occur at all. It is also indicative that in the Russian translation of this same epic on p. 217 there are eight pronominal reciprocals with *drug druga* 'each other' corresponding to the suffixed reciprocals of the original. A similar state of things is observed in Tuvan. Thus, in two Tuvan heroic epics ("Hunan-Kara") and "Boktus-Kiriš, Bora-Šeelej" (about 10,500 lines) the reciprocal pronoun occurs only once. Here is this sentence:

(11) *Olar-bile* *ol* *bügü-nü* *töögü-ž-üp,* *čugaala-ž-ïp,*
they-with s/he all-ACC tell-REC-CONV talk-REC-CONV
bot-bot-tar-ï-n *kaya,* *kanča-p,*
each-other-PL-3.POSS-ACC where/when what.to.do-CONV
tïp-č-ïp *alïr-ïn* *dugurž-up...*
find-REC-CONV AUX-CONV agree-CONV
'She discussed everything, talked it over with them. Having come to an agreement where and when they would meet.'

A similar situation is observed in the Kirghiz heroic epics. Thus, for instance, in "Manas" (book 4, 1995, 10-366) in more than 14,500 lines there are only three occurrences of the reciprocal pronoun:

(12) a. *bir* *biri-nen* *sura-š-ïp ...*
one one-ABL ask.questions-REC-CONV
'asking each other questions' (ibid., p. 170)
b. *biri-n* *biri* *kara-š-ïp...*
one-ACC one look.at-REC-CONV
lit. 'exchanging glances between themselves' (ibid., p. 175)
c. *biri-n* *biri* *kör-ö* *al-ba-y...*
one-ACC one see-CONV AUX-NEG-CONV
'(people) could not see each other [in the smoke]' (ibid., p. 244).

Analogous combinations of pronominal and suffixed reciprocals are mentioned in the works of Mahmûd al-Kâšgarî (cited from Nigmatov 1973: 51):

d. *olar ikki bir bir-ig suv-qa batr-uš-dï* (II. 236)
they two each other-ACC water-DAT dip-REC-PAST.3
'They both dipped each other into the water.'

Alongside the differences in the frequency of pronominal and verbal reciprocals, there is a problem of choice between the two markers, and we ought to bear in mind the possibility of both markers being used pleonastically (see (9) above and also *bot-bot-tar-ï-n ... tïp-č-ïp* 'to meet', lit. 'to find each other each other' in (11)), and also those cases where one of the markers is possible or preferable (see (a) and (b)) or their joint usage changes the meaning (see (c)). The following illustrates the main cases (this issue requires special research).

(a) A suffixed reciprocal is preferable or is used exclusively. Thus, for instance, in Kirghiz the meaning 'to kiss (each other)' is expressed, as a rule, by a suffixed reciprocal (13a). A pronominal reciprocal (13b) is clear in meaning, but it is not used.

(13) a. *öb-üš-* 'to kiss (each other)'
b. *biri-n biri öp-* 'to kiss each other.'

(b) A pronominal reciprocal is preferable or it alone is possible. For instance, in the meaning 'to deceive each other' the pronominal reciprocal (14a) is preferable, and in the meaning 'to hate each other' the pronominal reciprocal (14b) is possible only.

(14) a. *biri-n biri alda-* 'to deceive each other'
b. *biri-n biri žek kör-* 'to hate each other' (*žek kör-* 'to hate', lit. hatred look)

(c) When used simultaneously, the markers express different meanings; this happens when the reciprocal pronoun is attached to a verbal derivative with the assistive meaning; cf.:

(15) a. *žïyna-* 'to gather something'
→ b. *žïyna-š-* 'to help someone to gather something'
→ c. *biri biri-ne*(DAT) *žïyna-š-* 'to help each other gather something' (T. Abdiev, p.c.)

It should be noted that the differences in the frequency of usage between pronominal and verbal reciprocals are also observed in many languages unrelated to Turkic. For instance, in colloquial German, verbal reciprocals of the type *sich lieben* 'to love each other' are prevalent, while pronominal reciprocals like *einander lieben* are more typical of written German and sound bookish, lofty and solemn in colloquial speech (Berger et al. 1972: 544).

Reciprocal pronouns are often preferable or exclusively possible if the base construction contains a locative constituent marked by a locative preposition, postposition and/or a locative case (as, for instance, *kini-tten* 'from him' in (16a)) and the underlying verb does not specify unambiguously the spatial characteristic of the action. Verbal reciprocals, on the contrary, are most frequently derived from bases with the inherent spatial meaning of joining something to something (see section 8 below). This distribution is typical both of the Turkic languages and German (see Nedjalkov 2000: 102-117) and many other languages. Consider a Yakut example (N. M. Artem'ev, p.c.):

(16) a. *Aγa-m* *kini-tten* *kuot-ta-* Ø
father-my s/he-ABL run.away-PAST-3SG
'My father ran away from him.'
b. *Kini-ler* *beye-beye-leri-tten* *kuot-ta-lar*
s/he-PL each.other-REFL.POSS-ABL run.away-PAST-3PL
'They ran away from each other.'

Opposite to Yakut in this respect are Karachay-Balkar and Turkish.[2] Thus, for instance, in Karachay-Balkar there is no counterpart for the Yakut reciprocal in (8): the reciprocal suffix in Karachay-Balkar, which covers about 60 items, is unproductive, no new suffixed reciprocals being derived (though new derivatives are comprehensible to native speakers). There are reciprocals meaning 'to hit each other', 'to bite each other', but there are no reciprocals meaning 'to praise each other', 'to love each other', 'to kiss (each other)', etc. In this language, in contrast to Yakut, sentences with reciprocal pronouns are much more common. Here is one of the three reciprocal pronouns encountered on one page only of a folklore text of the Balkar variety of Karachay-Balkar (Boziev 1962: 124):

(17) *Alay* *anï* *bla* *erkin* *oyna-rγa* *bir-biri-n*
but s/he.GEN with freely play-INF each-other-ACC
žiber-me-y *e-di-le.*
let.go-NEG-CONV AUX-PAST-3.PL
'But (they) did not let each other play with him freely.'

1.3. The weakening of the nominal properties of reciprocal pronouns

Constructions (8), (9), (10) raise the problem of nominal properties of the reciprocal pronoun, such as the ability to occupy the same syntactic positions as nouns proper, excepting the subject position. (8) is an intransitive construction with a suffixed reciprocal where a direct object expressed by a noun cannot be added. (10) is a transi-

[2] As Kornfilt (1997: 159) asserts, "<...> there are only a certain number of <...> reciprocal verbs which are related to the corresponding simple verbs in a transparent fashion".

tive construction where the direct object position is occupied by the reciprocal pronoun, the valency structure of the sentence being preserved. Hence, the problem of the status of construction (9), which contains an intransitive reciprocal verb, as does (8), and also the reciprocal pronoun in the accusative, like (10). Thus, we may tentatively assume that the nominal properties of the reciprocal pronoun in cases like (9) are weakened: it functions in (9) as a kind of adverbial although it has acquired fewer adverbial properties, as is the case in some other languages, e.g. Bulgarian *edin drug* 'each other' and French *l'un l'autre* 'each other', which cannot be used as reciprocal markers alone with non-reciprocal verbs, i.e. verbs without a reflexive clitic (with the exception of verbs taking a prepositional object).

Thus, the French sentences (18a) and (18b) are grammatical and (18c) is not:

(18) a. *Les parents* *s'aiment* 'The parents love each other/themselves'
b. *Les parents* *s'aiment l'un l'autre* 'The parents love each other'
c. **Les parents* *aiment l'un l'autre* (same intended meaning).

In this connection it may be relevant to note that according to Kuular (forthcoming), Tuvan suffixed reciprocals are used mostly without the reciprocal pronoun *bot-bot-tar-ï-n*, while the latter is not used alone, as a rule, but with suffixed reciprocals.

Note that suffixed reciprocals derived from three-place transitives usually retain the direct object, which means they are not intransitivized by the reciprocal suffix (as it eliminates a non-direct object), as shown in the following Yakut example:

(19) *Ikki* *inibii* *kur-dar-ï-n* *bïld'a-s-pït-tar* *ühü.*
two brother belt-PL-POSS-ACC take.away-REC-PAST-3.PL they.say
'They say two brothers are taking belts from each other' (Pekarskij 1959: 616).

It should be added that derived transitive verbs with the non-reciprocal meanings of the suffix *-š* (*-s*), such as the sociative, comitative and assistive, retain the direct object of the base verb (see (20) and (23)).

2. Four main meanings of the Turkic suffix *-š* (*-s*)

The four meanings discussed in this section are regarded as the *main* ones, all the other meanings entering the polysemy of reciprocal markers are termed here *secondary*, whatever their productivity.

It may be suitable to open this section with the following quotation from Sevortjan (1962: 528), which, though it concerns Azerbaijani, is also applicable to many other Turkic languages:

> "In respect of the variety of its meanings, both productive and unproductive, the reciprocal-sociative voice [vzaimno-sovmestnyj zalog] is the richest, exceeding in this the reflexive voice. At the same time, this voice may be regarded as the least grammati-

> calized because its productive meanings in Azerbaijani are not derived mechanically but are dependent on certain conditions." (my translation - V.N.)

The four main meanings of the suffix *-š* (*-s*), which may be expressed in many, albeit not all, Turkic languages, can be illustrated by a Tuvan derivative in (20). These meanings have been repeatedly distinguished in the literature (see, among others, Böhtlingk 1989 (1851): 393-394; Xaritonov 1963: 16-50; Zinnatullina 1969: 182-197). These meanings were already distinguished in the works of Mahmûd al-Kâšgarî (Fazylov 1965: 78-96; Nigmatov 1973: 51).

Note that this suffix can involve all the possible valency changes, including zero change, as in the sociative meaning (with simultaneous increase of the number of participants in this case, like in those with valency increase, i.e. comitative and assistive). Thus:

a) The subject of reciprocals and sociatives is plural (at least semantically), as a rule (see, however, (21)):

(20) *üpte-š-*	i. 'to rob each other'	reciprocal (valency decreases)
	ii. 'to rob someone together'	sociative (valency is retained)

b) The subject of comitatives and assistives can be either singular or plural; the valency increases (at least semantically).

iii. 'to rob someone together with someone'	comitative
iv. 'to help someone rob someone'	assistive

The context usually disambiguates any reciprocal form. It goes without saying that one or another meaning of the reciprocal suffix is determined by the lexical meaning and valency of the base verb and pragmatic situation; it is also related to the transitivity or intransitivity of the derivative. Thus, for instance, the Yakut derivative *tapta-s-* in (1) can acquire only the reciprocal meaning 'to love each other', like the reciprocals *ubura-s-* 'to kiss each other', *öydö-s-* 'to understand each other' (Xaritonov 1963: 33), while Tatar *šaula-š-* 'to rustle' (about leaves) in (24) can be sociative only (in this example, as in many other analogous examples, the sociative sense is difficult to render in the English translation). Sometimes, a derivative is conventionally fixed in one of the possible meanings; cf. Yakut *bar-ïs-* used in the comitative meaning 'to go with sb' exclusively and never, for unclear reasons, in the sociative meaning 'to go together' (N. M. Artem'ev, p.c.). But the latter seems to be untypical: usually, derivatives which may have the comitative meaning can also express the sociative meaning, in the same way that derivatives with the assistive meaning can be used in the comitative and sociative meanings (the opposite is not necessarily true). As regards sociatives and reciprocals, their lexical range overlaps.

Cross-linguistically, reciprocal constructions can be classified according to the possibility of expressing one of the reciprocants by a non-subject constituent. In all

the Turkic languages, reciprocal constructions with a non-subject constituent seem to be possible, and thus they may involve not only valency decrease but also retain the number of the arguments. However, in this latter case demotion of the second argument takes place. This kind of constructions can be called *secondary* reciprocal constructions. The predicate agrees with the subject participant only, as in the following Tatar example (taken from Isxakova 1974: 278):

(21) a. *Fərid belən Mostafa kočakla-š-a-lar.*
F. and M. embrace-REC-PRES-3.PL
'Farid and Mustafa embrace each other.'
b. *Fərid Mostafa belən kočakla-š-a-Ø.*
F. M. with embrace-REC-PRES-3.SG
(same) lit. 'Farid embraces each other with Mustafa.'[3]

This property of Turkic reciprocals noticeably distinguishes them from the reciprocal constructions of Romance languages and also of German, where, with very few exceptions, type (22b) constructions are ungrammatical (these exceptions are related to lexicalization; cf. German *Peter schlug sich mit Hans* 'Peter fought with Hans'); cf. the following German example:

(22) a. *Peter und Hans umarmten sich.* 'Peter and Hans embraced each other.'
b. **Peter umarmte sich mit Hans* lit. 'Peter embraced each other with Hans.'

As a rule, derivatives with the reciprocal, assistive and comitative meanings are more or less easy to translate into English. The same applies to transitive sociatives. In these cases the derived meanings are distinctly different from the base meanings. As (23c) and (23d) show, the comitative and the assistive meanings involve different marking of the object: the postposition *kïtta* with the accusative or dative case respectively.[4] Cf. the following Yakut examples (N. M. Artem'ev, p.c.):

(23) a. *Aha-m ikki min ot tiey-di-bit.*
father-my and I hay carry-PAST-1PL
'Father and I carried hay.'

[3] Such constructions were already registered by Mahmûd al-Kâšgarî (cited from Nigmatov 1973: 51): *ol meniŋ birlä quč-uš-dï* (II, 98) lit. 'he embraced with me'.

[4] As it happens, the dependence of the comitative and assistive meanings on the explicit expression of the second participant was already noted by Mahmûd Kâšgarî (reported by Sevortjan 1962: 539). Xaritonov (1963: 23) terms all the three meanings as *jointness* (= Russian term sovmestnost') and distinguishes among them (lit.) *total jointness* (= sovokupnaja sovmestnost'), i.e. sociative in the terminology used in this paper, and (lit.) *adjoining jointness* (= primykajuščaya sovmestnost') which covers the comitative and assistive meanings, in the terminology used here.

b. *Aha-m* *ikki* *min* *ot* *tiey-is-ti-bit.* sociative
father-my and I hay carry-REC-PAST-1PL
'We carried hay together.'
c. *Aha-m* *miig-in* *kïtta* *ot* *tiey-s-t-e.* comitative
father-my I-ACC with hay carry-REC-PAST-3.SG
'Father carried hay with me.'
d. *Aha-m* *mie-xe* *ot* *tiey-s-t-e.* assistive[5]
father-my I-DAT hay carry-REC-PAST-3.SG
'Father helped me carry hay.'

The verb 'to cart', 'to carry', etc.', like other two-place transitive verbs with an inanimate direct object, often requires lexical indication of the reciprocal meaning; e.g.:

e. *Aha-m* *ikki* *min* *xardarïta* *ot* *tiey-is-ti-bit.*
father-my and I by.turns/mutually hay cart-REC-PAST-1PL
'Father and I carted hay to each other.'

The sociative meaning of intransitive derivatives, on the contrary, can be (not infrequently) elusive and therefore may be difficult to translate, because translations with words like 'together' may sound clumsy and overemphasized, especially if the subject is inanimate, e.g. in sentences with meanings like 'the stars are twinkling', 'the firewood is burning', 'apples were hanging down from the branches', etc. In derivatives of intransitives the sociative meaning seems to evolve in the direction of a kind of emotive or intensifying colouring. Compare the following Tatar (Zinnatullina 1968: 185) and Karachay-Balkar (A. A. Xasanov, p.c.) examples:

(24) *yafrak-lar* *šaula-š-tï.*
leaf-PL rustle-REC-PAST
'The leaves started rustling.'

[5] Cross-linguistically, the development of the assistive meaning (sharing semantic and syntactic (viz. valency increase) properties with the comitative and causative) on the reciprocal marker in the Turkic languages is not accidental. It is in fact semantically motivated. It may be noted in passing that the assistive meaning can be expressed not only by a special marker, like the Aymara suffix *-jaya/-jaa* (Middendorf 1891: 145; note that this suffix seems to contain the causative suffix *-ya/-aa*; cf. ibid., p.148), but also by markers with such meanings as comitative (cf. the suffix *-ysi* in Bolivian Quechua; van de Kerke (1996: 28-29); to be precise, its meaning 'to accompany sb' is not a pure comitative meaning, but it implies the sense 'for the purpose of giving assistance'; see Bills , Vallejo, & Troike 1969: 306-307), or the causative (cf. the circumfix *a-...-in-* in Georgian (Gecadze, Nedjalkov & Xolodovič 1969: 149-150) and the suffix *-idz-* in Shona (Aksenova 1990: 172)). Cf. also the assistive meaning of the Buryat suffix *-lsa* in (26d).

(25) *ala dьbьrtla-š-dь-la.*
they gallop-REC-PAST-3.PL
'They broke into a noisy run (in a hurry, in a disorderly way, kind of overrunning each other').

In connection with (24) and (25), it may be useful to cite the following comments on Yakut sociatives by a linguist who is also a native speaker of Yakut: "a joint action is understood as one common action in which two or more acting persons participate in equal measure"; "verbs of joint action manifest in their meaning the presence of some inner bond between the participants"; "the linking bond is probably the unity of the cause and the unity of the situation (and of psychological conditions) of the action performed" (Xaritonov 1963: 24-25). Compare the opinion on Tatar sociatives concerning their difference from the plain plural number: "*Čьrьldašabьz, köleš□bez, šaulašabьz* – they do not mean simply '(we) squeal', '(we) laugh', '(we) make noise', though there is no other way of expressing their meaning in Russian. Due to the suffix, they present the actions as a sum of single acts of a great number of people and not as a single process <...> The actions are expressed as dimensional, dynamic: the action of each person has its own peculiarities, it is performed at different intervals." (Zinnatullina 1969: 194-195). The very last statement about different intervals should be understood, I think, as a series of swallow-tailing acts, close to the distributive meaning, but this is only one of the types of situations denoted by sociatives. On the relations between the sociative and the distributive meanings and markers see Kemmer (1997: 231-249).

With respect to the four main meanings of Turkic reciprocal markers, judging by the specialist literature that was availabale to me (over 200 titles), this set of meanings is not attested in reciprocal markers of other world languages so far, excepting the areally close Mongolic languages, e.g. in Buryat, where these meanings can be expressed by the suffixes *-lda* (in Buryat and Khalkha) and *-lsa* (in Buryat), *-lca* (in Khalkha). Compare the following examples taken from Sanžeev (1963: 240); Cydendambaev (1979: 109) and Kuz'menkov (1984: 75) respectively:

(26) a. *Tani-lsa-aar tata-lda-xa, xara-lsa-aar*
recognize-REC-CONV pull-REC-PART look-REC-CONV
xaza-lda-xa bolo-xo-mnaj gü, übgen?
bite-REC-PART AUX-PART-our Q old.man
'Why is it, old boy, as soon as we meet we start fighting, as soon as we see each other we start biting each other?'

If the subject is plural the construction allows two readings: (a) sociative, if all the participants are thought to be named by the subject (see (i) below); (b) comitative, if the subject is interpreted as one collective participant and the second participant is considered as omitted (see (ii) below). This ambiguity is observed in Turkic languages as well.

b. *Xuragša-d-Ø ... xüdelmeriše-d-öör žel büri übhe-Ø xur'aa-lsa-dag.*
pupil-PL-NOM worker-PL-INST year every hay-NOM make-REC-PART
i. 'Every year all the pupils make hay together as workers.'
ii. 'Every year the pupils take part in making hay as workers.'

The following sentence is comitative in meaning, but the translation '...after him' instead of '...with him' is a more precise description of the situation when two persons go through the door together'.

c. *Parxae ... xoino-hoo-n' gara-lsa-ba.*
P. behind-ABL-his go.out-REC-PAST
'Parxaj went out after him.'

The following Khalkha sentence is assistive in meaning, but the participant who receives help is denoted by a genitive attribute of the direct object rather than by a dative object; therefore, the syntactic valency is unchanged while the semantic valency has increased, as in the previous instance.

d. *Bat Dorži-in ger-iig bari-lsa-na.*
B. D.-GEN yurt-ACC rig.up-SOC-PRES
'Bat will help to rig up Dorji's yurt.'

It has been noted in Turkology that a number of Turkic languages lack the assistive meaning. The polysemy covering all the four meanings is characteristic of Yakut, Tuvan, Khakas, Tofalar, Tatar, Turkmenian, Uzbek, Karakalpak, Kazakh, and Salar, but not of Azerbaijani, Turkish, Karachay-Balkar, Gagauz and Karaim, where the assistive meaning is lacking (see also Sevortjan 1962: 532). It is unclear whether the second group of languages has lost this meaning or failed to develop it in due time. It is possible that there are different explanations for the different languages of this group.

In conclusion, note that the four meanings at issue differ in the degree of proximity between them.

In the sequence of meanings *reciprocal – sociative – comitative – assistive*, the ones closest to each other are the *sociative* and *comitative*, and those most removed from one another are the *reciprocal* and *assistive*. The *reciprocal* and *sociative* meanings are close enough because of the plural subject, and the *comitative* and *assistive* are very close because acting with someone often involves assisting. If we look for the semantic proximity of any three of the four meanings, i.e. for features they share, which bring them close together, we may establish the following triplets of meanings: *reciprocal – sociative – comitative* and *sociative – comitative – assistive*. Expressing each of these two triplets of meanings by the same marker seems to be common enough cross-linguistically (cf. (27) and (28)), while the common marker

for all four meanings is registered, so far, in the areally close Turkic and Mongolic languages.

Here are examples from Rwanda (Aksenova 1994: 160, 177; *-an* = REC; *n'-* (before vowels; < *na*) = preposition 'with') and Ancient Greek (Dvoreckij 1958: 1542; the prefix *sym-* more or less corresponds to the German detachable prefix *mit-*; the translation of the 1SG by means of the infinitive is conventional) respectively:

(27) a. *tu-reb-an-a*	'they look at each other'	reciprocal
b. *tu-kor-an-a*	'they work together'	sociative
c. *n-kor-an-a n'-ûmwâna*	'I work with the girl.'	comitative
(28) *féro* 'to carry' → *sym-fero*	i. 'to carry sth together'	sociative
	ii. 'together with sb carry sth'	comitative
	iii. 'to help sb carry sth.'	assistive

3. The Turkic suffix *-š (-s)* and three main types of polysemy of the reciprocal markers in the world languages

Note in the first place that there are numerous languages with monosemous reciprocal markers, e.g. Chukchi, Even, Yukaghir, Itelmen, Ainu, Nivkh, etc., which do not concern us in this section.

As regards the meanings most closely related to the reciprocal, three are the most important: reflexive (shared feature – anaphoric relations), sociative (shared feature – multiplicity of participants) and iterativity (shared feature – multiplicity of actions). Accordingly, the following three main types of polysemy can be postulated for reciprocal markers. They also happen to be widespread across languages.

3.1. Reflexive-reciprocal polysemy

The reflexive-reciprocal polysemy is a result of the polysemy of reflexive markers rather than reciprocal markers, because the main meaning of this type of markers is reflexive, the reciprocal being a later development (cf. Geniušienė 1983: 140-141, 1987: 344-348; Kemmer 1993: 151-200). This type of polysemy is typical of many languages of various families, e.g. French, German, Slavic, Mansi, Maasai, Mizo, Limbu, Amharic, Shoshone, etc. The following examples are Bulgarian (29a), German (29b), Sumbwa ((30); Capus 1898: 64; *-i-* = REFL) and Russian (31) respectively:

(29) a. *Te me gledat*	'They watch me.'	
→ *Te se gledat*	i. 'They watch themselves (in the mirror, etc.).'	
	ii. 'They watch each other.'	
b. *Sie lieben mich*	'They love me.'	
→ *Sie lieben sich*	i. 'They love themselves.'	ii. 'They love each other.'

(30) *-shim-a* 'to love' → *-i-shim-a* i. 'to love oneself', ii. 'to love each other'
-gay-a 'to hate' → *-i-gay-a* i. 'to hate oneself', ii. 'to hate each other.'

In Russian, the reflexive and reciprocal meanings of the postfix *-sja/-s'* are in complementary distribution relative to the sets of verbs, with one or two exceptions.

(31) a. *Oni brejut ix* 'They shave them.'
→ *Oni brejutsja* 'They shave [themselves].'
b. *Oni obnimajut ix* 'They embrace them.'
→ *Oni obnimajutsja* 'They embrace each other.'

As is known, the Turkic reflexive marker *-ïn-* does not have the reciprocal meaning.

3.2. Reciprocal-sociative polysemy

This type of polysemy is also typical of numerous languages, e.g. Turkic, Bantu, Tagalog, Halkomelem, Mongolic, Palau, etc. Here are examples from Buryat (Cheremisov 1973: 63, 361) and Karanga (Marconnès 1931: 195):

(32) a. *asuu-* 'to ask someone' → *asuu-lda- / asuu-lsa-* 'to ask each other'
b. *oro-* 'to enter' → *oro-lda- / oro-lsa-* 'to enter together.'

(33) *Ti no chek-an-a [nyama]*
we NO cut-REC-IND meat
i. 'We shall cut each other' reciprocal without the object
ii.'We shall cut the meat together' sociative with the object included

I wish to stress that the Turkic languages represent this type of reciprocal polysemy; cf. the Yakut example (cf. also Xaritonov 1963: 31):

(34) *Kiniler [is taŋah-ï] mïïlala-s-t-ïlar.*
they interior clothes-ACC soap-REC-PAST-3.PL
i. 'They soaped each other' reciprocal without the object
ii. 'They soaped the underwear together' sociative with the object included

On this type of polysemy see also Lichtenberk (1985: 19-41).

3.3. Iterative-reciprocal polysemy

This type of polysemy has been recently described in the Chinese data as having developed from the iterative; similar polysemy is also attested in some Oceanic languages. The examples are from Chinese (Liu 1999: 124, 126; *dă* 'to hit', *V-lái-V-qù* = iterative-recipocal where *lái* = 'come' and *qù* = 'go'), Samoan (Mosel & Hovdhaugen 1992: 180-183), Sobei (Sterner 1987: 53; *re-* = 'they'; *-re-/-ro-* = REC infix).

(35) *dă-lái-dă-qù* i. 'to fight several times' ii. 'to beat each other.'

(36) a. *a'a* 'to kick' → *fe-a'a* 'to kick again and again'
b. *sogi* 'to kiss' → *fe-sogi* 'to kiss each other.'

(37) a. *re-fedfadnar* 'they jump' → *re-f-re-dfadnar* 'they jump repeatedly'
b. *re-soro* 'they help' → *re-s-ro-ro* 'they help each other.'

A variety of this type is probably the distributive-reciprocal polysemy of a reciprocal marker.

Besides these three types, other types of polysemy of reciprocal markers are also possible which seem to be less widespread. Such types can be accounted for by the reason that one of the three concomitant meanings has gone out of use, as is the case in present-day Kirghiz, where the sociative meaning is lost.

3.4. An instance of a secondary meaning of polysemous markers

Each of the three types of polysemous markers in question can possess some other meanings, e.g. competitive, anticausative and that of plurality, and, though differing in productivity, they are also attested in some other genetically related and non-related languages. They will be discussed further on (see sections 7.3, 7.2 and 8.4 respectively). Here I will mention only one of such meanings, namely, the meaning known as "antipassive", "depatientive", "absolutive"; see, for instance, Dixon (1980: 434, 440, 445-448ff.), Lichtenberk (1991: 171-183), Geniušienė (1987: 87, 249, 314), etc. (In Russian grammatical tradition, this meaning is sometimes termed "active-objectless" (aktivno-bezob"ektnoe); cf. Janko-Trinickaja 1962: 198-202.) In this case the predicate, while generally retaining its meaning, loses its object (is intransitivized) if in a given situation the object is obvious from the situation or context or is irrelevant (= "absolutive"); it may further evolve into the meaning of a habitual property of the subject-referent, which also makes the object irrelevant (= "habitual"). I propose the general term "absolutive" for this meaning, with the term "habitual" for a submeaning within.

In most Turkic languages there occur only occasional derivatives (or none at all) of this type. Thus, for instance, in Yakut there are only a few such derivatives (among them, *bultaa-* 'to hunt' → *bulta-s-* 'to be engaged in hunting', *et-* 'to scold' → *et-is-* 'to be scolding' (vi); *irdee-* 'to recover (debts)' → *irde-s-* 'to be engaged in recovering debts'). This meaning is not attested in Khakas (Letučiy, forthcoming) and Karachay-Balkar.

On the other hand, in Tatar, judging by the information in Zinnatullina (1969: 192-193), absolutives are rather productive, being represented by several dozens of items. It is likely that verbs of this type have similar productivity in the closely related Bashkir language, judging by the list of Juldašev (1981: 253). As a rule, they denote habitual activity which may have become a distinctive feature of the subject referent, though this may not be easy to show in the translation, as the implication of habituality may be rather slight (ibid., p. 193). Compare:

(38) *alda-* 'to deceive' → *alda-š-* 'to deceive, be in the habit of deceiving'
ələklə- 'to inform, sneak (on)' → *ələklə-š-* 'to be a habitual sneaker'
yaz- 'to write' → *yaz-ïš-* 'to be engaged in writing'
könlï- 'to be jealous' → *könlï-š-* 'to be jealous, envious.'

Needless to say, derivatives of this kind may be polysemous and also have other meanings typical of the reciprocal suffix; e.g. (Tatar-Russian dictionary 1966: 591):

(39) *urla-*	'to steal'		
urla-š-	i. 'to be engaged in stealing, to be a thief'		absolutive
	ii. 'to steal from each other'		reciprocal
	iii. 'to help sb in stealing'		assistive
	iv. 'to take part in stealing'		comitative

The meaning in question is attested in languages other than Turkic as well, and it may be expressed by markers with reciprocal-sociative (see Swazi (40); Ziervogel 1952: 73; *-an* = REC/SOC) and reflexive-reciprocal polysemy (see Russian (41); *-sja* = REFL/REC). Verbs with negative colouring are prevalent here, as in Tatar.

(40) *-tondz-a* 'to hate somebody' → *-tondz-an-a* 'to hate each other'			reciprocal
-lum-a 'to bite somebody' → *-lum-an-a*		i. 'to bite severely'	intensive
		ii. 'to be fierce'	absolutive
-hlaɓ-a 'to gore' → *-hlaɓ-an-a*		i. 'to gore severely'	intensive
		ii. 'to have the habit to gore'	absolutive

(41) *Byki bodajut-sja*	'The bulls are butting [each other]'	reciprocal
Smotri, byk bodaetsja!	'Look, the bull is butting!'	absolutive
Etot byk bodaet-sja	'This bull butts.'	habitual

Katanov (1894: 31) asserts that in most of its uses the suffix *-š* indicates jointness of identical actions of several agents or repeated performance of the same action presented regardless of the agent(s). Zinnatullina's (1969: 193) explanation of the origin of the meaning I call absolutive is based on this assertion; she also assumes that it further develops into the meaning of its habitual permanence as a characteristic of the subject referent.

3.5. Note on etymology

In conclusion of this section, a few words about the current theories of the origin of the reciprocal meaning in the three types of polysemy are due. In the cases of reflexive-reciprocal and also iterative-reciprocal polysemy, the primary meaning is most frequently reflexive or iterative respectively. Evolutions of this kind have been registered in the observable historical period. As regards the reciprocal-sociative polysemy, the most likely primary meaning is reciprocal, the sociative being a result of its extension. It is likely that first sociatives from intransitive verbs appeared,

which was facilitated by the predominance of intransitives among reciprocals (transitive reciprocals illustrated in (19) are rather rare cross-linguistically); compare Yakut: *kiniler bil-is-ti-ler* 'they knew each other' — *kiniler kül-üs-tü-ler* 'they (all of them) started laughing'.

The evolution of the sociative meaning from the reciprocal is supported by the existence of sociative markers formed by means of reciprocal markers in a number of languages. (The opposite way of derivation is not known to me.) Besides, a sociative form is a reciprocal derived from a comitative.[6] Compare the following Haya (Dammann 1954: 168), Ainu (Alpatov, Bugaeva & Nedjalkov forthcoming) and Kabardian (Apazhev et al. 1957: 167, 59, 102) examples (in (42b), (43b) and (44b) the applicative affixes are used in the comitative meaning):

(42) a.	*-nyw-a*	'to drink'	
b.	*-nyw-el-a*	'to drink with somebody' (vt)	comitative
c.	*-nyw-el-an-a*	'to drink together' (vi)	sociative
(43) a.	*rewsi*	'to stay somewhere overnight'	
b.	*ko-rewsi*	'to stay the night with somebody' (vt)	comitative
c.	*u-ko-rewsi*	'to stay the night together' (lit. 'with each other') (vi)	sociative
(44) a.	*kIuɜn*	'to go'	
b.	*dɜ-kIuɜn*	'to go with somebody'	comitative
c.	*zɜ- dɜ-kIuɜn*	'to go together'	sociative

Above, I said that in the case of reciprocal-sociative polysemy the sociative meaning is most likely a secondary development from the reciprocal. This leads to the question of the primary, initial meaning of the suffix *-š(-s)*. In the preceding paragraph, it is stated that in the case of reflexive-reciprocal polysemy the most likely initial

[6] If we accept the path of derivation as a reflection of the semantic structure of the sociative, we have to admit that the sociative meaning is semantically more complex than the reciprocal. It can be added that the comitative meaning in itself is essentially reciprocal; cf.: *I am going with you = You are going with me = We are going together*. In the last sentence, *together* can be semantically explained as 'with each other' (stylistically, this phrase is unacceptable), i.e. the sociative meaning can be explained via the reciprocal. The opposite, i.e. explaining the reciprocal meaning by means of the sociative, is hardly possible. In this connection it may be appropriate to mention that in Nivkh, the concept 'together' is denoted by the reciprocal form of the verb with the meaning 'to go with sb, accompany sb', which is a kind of lexical comitative: *i-γrɜ-t* 'accompanying him' (*i-* = 3.SG, *-t* = CONV) → *u-γrɜ-t* 'together', lit. 'accompanying each other' (*u-* = REC). Other instances of the morphological derivation of the sociative marker are attested in some Bantu languages, cf. Dabida: *-kund-a* 'to love' → *-kundan-a* 'to love each other' and *-damb-a* 'to travel' → *-damb-any-a* 'to travel together' (Rjabova 1989: 111).

meaning is reflexive, and in the case of iterative-reciprocal polysemy it is the iterative meaning (these meanings are more elementary than the reciprocal). It is interesting to note that among the proposed hypotheses on the etymology of Turkic *-š* there are two that assume that the initial meaning of this suffix was either reflexive or iterative. The first hypothesis is advocated by Xaritonov (1982: 274), who suggested that before the suffix *-n* came to be used as the principal reflexive marker, the suffix *-s* had been used in the reflexive function, which Xaritonov supports by the argument that in numerous ancient fossilized derivatives with the reflexive marker *-n-* the latter is preceded by the reciprocal marker *-h (< -s)* (cf. *oŋor-* 'to do' → *oŋo-h-un-* 'to do for oneself' (here, the fossilized causative suffix *–r* is replaced by the suffix *-h*); *sup-* 'to close, cover' → *sab-ïh-ïn-* 'to overhang something', 'to close/ cover something for oneself'; Xaritonov 1963: 47). But this assumption does not explain the origin of the presumable reflexive meaning of the suffix *-s*. Another hypothesis is supported by Serebrennikov and Gadžieva (1979: 208), and Sevortjan (1962: 138), who argue that one of the most ancient meanings of this suffix was plurality of actions, i.e. iterativity (one of the arguments in its favour is the use of the suffix *-š* in the complex suffix *-š-tïr* denoting iterativity in many Turkic languages; cf. Tatar *uyna-š-tïr-* 'to play from time to time'; see also footnote 8). (For a detailed discussion of this question see Ščerbak 1981: 113-115.)

The issue becomes even more complicated if we take into account the assertion that "A verbal reciprocal suffix **-ĺ(č)-* is widely represented in Turkic", being a reflex of the Proto-Altaic reciprocal suffix **-ĺ-* (Starostin et al. 2003: 200). The possibility of *-š* originating from *-l* was pointed out by Sevortjan (1962: 137).

4. Turkic voices and their correspondences in other languages of Eurasia

While spread over vast territories, mostly in Asia, the Turkic languages cover part of a larger area where a great many other languages are spoken. Of course, the Turkic languages are not in immediate contact with all of them. In the Turkic languages, traditionally, a five-member voice system is distinguished (see (45) for the derived voices). The functions of each of the derived voices include not only the one after which they are named but also a number of other functions. Thus these markers are polysemous and in some of their meanings, e.g. anticausative, they are synonymous, see (81).

(45)

Turkic voices	Suffixes
reciprocal	*-ïš-*
reflexive	*-ïn-*
passive	*-ïl-*
causative	*-dïr-, -t-*

Further on, I will briefly consider the existence and expression of the meanings of three of the derived voices in the Eurasian area. In the first place, as it seems, none of non-Turkic languages of the areas in question possesses an analogous five-member

system of voices, each with its own marker.[7] However, the system of the Kolyma Yukaghir languages has a significant similarity to it, differing in that the counterpart of the Turkic passive is the object-oriented resultative, a category which is very close to the passive.

(46)	prefixes	reciprocal	*n'i-*
		reflexive	*met-*
	suffixes	resultative	*-o:(l)*
		causative	*-š-*

Generally, three major language domains adjacent to the Turkic area can be distinguished. These three areals, and the languages within each area, differ from one another in the affixal devices used for the reciprocal, causative and reflexive meanings (alongside affixal devices, clitic expression, or, more generally, devices descending from reflexive pronouns are considered here as functionally equal to affixal devices).

Note that reciprocal pronouns like *each other* are not considered here at all. The category of passive is not included in the discussion either.

4.1. The eastern area

The following unrelated languages and language groups are assigned to this territory: Mongolic, Tungusic, Chukchi, Itelmen, Eskimo, Yukaghir, Nivkh (Gilyak), Ainu, Ket, Aleut, Japanese. In fact, this area is the original territory of the Turkic languages. Being genetically unrelated and occupying a vast territory, these languages exhibit similarity in the morphological expression of the reciprocal, causative and reflexive meanings.

1) Most languages of this area, as well as the Turkic languages, possess reciprocal affixes: suffixes (Buryat *-lda, -lsa*; Evenki *-maat, -ldə*, Chukchi *-wəly, -čit*; Eskimo *-uta*; Japanese *-a/-aw*) or prefixes (Yukaghir *n'i-*; Itelmen *lu-;* Nivkh *u-*; Ainu *u-*). An exception are Ket and Aleut, which have no reciprocal affixes.

Among the languages of this area, three groups of languages possess highly polysemous reciprocal markers, which testifies to their ancient character. These are the Turkic, Mongolic and Eskimo languages. In the other languages the reciprocal marker is either monosemous or has a weak polysemy.

2) Like Turkic, all the languages of the area except Ket have causative affixes.

3) Besides in the Turkic languages, reflexive affixes are attested in three languages of the area only: Yukaghir (*met-*), Ainu (*yay-*) and Nivkh (*p'(i-*)).

[7] Outside Eurasia, the analogous system is registered in Swahili where there are "five voice categories: active, passive, reflexive, reciprocal and causative", each with its own marker: zero for the active, passive *-w*, reflexive *-ji-* (placed in the slot of the object agreement marker), reciprocal *-an* and causative *-ish* (Vitale 1981: 177).

4.2. The western area

This area comprises Indo-European languages. As is known, in many languages of this family the reciprocal meaning is expressed by a reflexive clitic, reflexive pronouns (or, more generally, devices descending from reflexive pronouns) like German *sich*, Bulgarian *se*, etc. (see examples (29)-(31)). In contrast to the languages of the eastern area, where most of the reciprocal markers are monosemous, in the western area there are no monosemous reciprocal markers. In contrast to the languages of the eastern area, in which the Turkic languages are also included, the Indo-European languages lack morphological causatives (exceptions: the Baltic, Armenian and Indo-Iranian languages).

In some languages of this area a reflexive-reciprocal clitic [or a preposition] developed into an affix:

a) a postfix in East Slavic languages (see (31)) and Latvian, and a postfix (in unprefixed verbs) or interfix (in prefixed verbs) in Lithuanian (see, for instance, Geniušienė 1987: 19); cf. Lithuanian:

(47) a. *stumdyti* 'to push' → *stumdyti-s* 'to push each other' (imperfective)
b. *pa-stumdyti* 'to push' → *pa-si-stumdyti* 'to push each other' (perfective)

b) a prefix, also descending from a reflexive pronoun, in a language of the Rheto-Romance group, viz. Surselvan (Stimm 1977: 70, 84); cf.:

(48) *jeu selavel* (vi) 'I wash [myself].'

c) In Celtic languages, a preposition genetically related to the German *um*, Old English *ymbe*, Latin *ambi* 'around', became a verb prefix (see among others Lewis & Pedersen 1961: 264). Though non-reflexive by origin, this prefix displays a polysemy typical of the Indo-European reflexive clitics and including the reflexive and reciprocal meanings. This prefix is practically lost in Irish, but it is still preserved in a number of derivatives in Welsh (*ym-* < *ambi-*) and Breton. Here are examples from Welsh (Evans & Thomas 1981: 203-206):

(49) a. *ymolch* 'to wash oneself'
b. *ymweld* 'to see/visit each other.'

4.3. The "intermediate" area

This is the territory occupied by Uralic (Finno-Ugric and Samoyedic) languages. This area may be regarded as a kind of intermediate because, like the languages of the eastern areal, the Uralic languages possess causative affixes, on the one hand, and on the other, they have reflexive-reciprocal markers, like the Indo-European languages (though suffixes, not clitics). Examples from Mansi (Rombandeeva 1973: 148-149), Mari (Galkin 1966: 333, 817) and Udmurt (Tepliashina 1966: 273) respectively are cited below:

(50) *lowt-xat-* 'to wash [oneself]'
titt-xat- 'to feed each other'
say-xat- i. 'to plait one's hair' ii. 'to plait each other's hair'

(51) *mušk-ïlt-* 'to wash [oneself]'
šupšal-alt- 'to kiss [each other].'

(52) *düs'a-s'kï-* 'to dress [oneself]'
vera-s'kï- 'to talk, converse'

4.4. The continuum from causative and reciprocal towards reflexive-reciprocal

A simplified overall picture can be shown by the following schema, in which the brackets signify weak use or absence of the respective verbal derivatives in the languages of the area:

Eastern area	"Intermediate" area	Western area
CAUSATIVE	CAUSATIVE	(CAUSATIVE)
RECIPROCAL	(REFLEXIVE-RECIPROCAL)	REFLEXIVE-RECIPROCAL
(REFLEXIVE)	(REFLEXIVE-RECIPROCAL)	REFLEXIVE-RECIPROCAL

As mentioned, the causative and reciprocal markers (suffixes and prefixes) are more prominent in the eastern area (where the Turkic languages belong) and the reflexive-reciprocal markers (clitics, etc.) in the western area. The "intermediate" area manifests markers (suffixes) which link the other two areas.

The loss of morphological causatives in the majority of Indo-European languages has caused the development of other devices for expressing semantic causative oppositions, including development of labile verbs (like English *to break* (vi/vt)), of anti-causative forms (cf. Russian *slomat'* 'to break' (vt) → *slomat'sja* 'to break' (vi)), and also the use of auxiliary causative verbs (cf. German *kommen* 'to come' → *kommen lassen* 'to cause to/let come') (see, among others, Terasawa 1985: 133-143; Zubizarreta 1985: 247-289; Cannings & Moody 1978: 331-362; Nedjalkov 1976).

A few words about the Caucasian area, which is also adjacent to the Turkic area (Karachay-Balkar, Kumyk, Noghai, Azerbajani, Turkish) are due. In this area languages of five families neighbour the Turkic languages. As for morphological causatives, reflexives and reciprocals, they are absent in Ossete, an Indo-European language. Morphological causatives are attested in Tati, another Indo-European language, and also in Georgian, Chechen-Ingush and some Daghestan languages, morphological reflexives and reciprocals being absent. And only in one family, Abkhaz-Adyghe, do all three meanings find morphological expression, the markers of reflexives and reciprocals being genetically related.

5. Productivity of the reciprocal, sociative and assistive meanings of the suffix *-š-(-s)* in some Turkic languages

Between the Turkic languages, there are of course differences in both the sets of meanings of the suffix *-š(-s)* and their productivity. Among the most prominent differences, the following should be named in the first place (the first two facts that have been mentioned above are repeated here in the general context of the differences).

(a) The reciprocal meaning is productive in Yakut (Uyghur, or North-Eastern group) and unproductive in Turkish (Oguz group) and Karachay-Balkar (Kipchak group).

(b) The assistive meaning is productive in the languages of the Uyghur group and many languages of the Kipchak group (Tatar, Kazakh, Kirghiz), and also in Turkmenian (Oguz group), and it is absent in Karachay-Balkar and Karaim (Kipchak group) and Turkish and Azerbaijani (Oguz group).

(c) The sociative meaning is highly productive in Yakut; it is lost in present-day Kirghiz (which may be due to the employment of this suffix as the 3PL marker; see 7.1). In Karachay-Balkar and Tatar it occurs mostly on intransitive bases (Nedjalkov 2002: 54-62; Isxakova 1974: 283-284).

As we see, the differences do not necessarily correspond to the classsification of these languages: sometimes, there is an overlap between them with respect to the features mentioned (true, this also pertains to some other properties of Turkic languages). It would be interesting to find out whether the unproductivity of the reciprocal and the assistive meanings was initially characteristic of the marker or developed later (probably in connection with the advancement of the Turkic tribes westward and interaction with the local languages).

As an instance of different productivity of the meanings under discussion in various Turkic languages, Tatar might be cited: according to Zinnatullina (1969: 190, 187, 193), the assistive meaning is much more common than the reciprocal and sociative. In her corpus of about 660 derivatives in *-š* these three meanings are distributed as follows: 335 assistives, 110 reciprocals and 130 sociatives.

6. "Geography" of the productivity of the meanings of the reciprocal markers: reciprocal-sociative (Turkic languages) and reflexive-reciprocal (Indo-European and Uralic languages)

By somewhat simplifying the overall picture, one may propose the following observations.

1) Among the Turkic and Finno-Ugric languages, the productivity of the reciprocal use of the markers in question diminishes and even becomes lost in the western part of the area (cf. Yakut and Karachay-Balkar and, on the other hand, Mansi and Finnish).

2) In the Indo-European languages, on the contrary, the productivity diminishes or becomes lost in the opposite direction, i.e. in the eastern part of the areal (Baltic,

East Slavic languages, also Armenian (suffix *-v*) and Modern Greek (middle (passive) conjugation), Indo-Iranian, (and probably Rumanian and Albanian), while in French, Spanish, Italian, and also German and in West and South Slavic languages this meaning is highly productive.

3) It might as well be added that some Indo-European (Celtic, English, Scandinavian, Dutch) and Uralic (Finnish, Estonian, Samoyedic) languages, i.e. languages of Northern Europe and a part of north-western Asia, display the absence or unproductivity of reflexive-reciprocal markers.

4) If we compare these tendencies in the first two areas, we conclude that the productivity of the reciprocal function of reciprocal markers (i.e. markers one of whose meanings is reciprocal) diminishes in the direction towards a certain point in Eastern Europe.

7. On secondary meanings of the Turkic reciprocal suffix

7.1. Introductory remarks

Alongside the four main meanings which may cover large lexical groups of verbs though not in all Turkic languages, the reciprocal suffix may have a number of secondary meanings (one of these is considered above in 3.4).

7.1.1. Unproductive secondary meanings

These meanings are represented by limited sets of derivatives numbering two or more items in one or several Turkic languages only; cf., for instance, the converse meaning in Yakut (Xaritonov 1963: 45):

(53) a. *atïïlaa-* 'to sell something to somebody'
→ *atïïla-s-* 'to buy something from somebody'
b. *tüülee-* 'to lease (meadowland)'
→ *tüüle-s-* 'to take meadowland on lease' (arch.)
c. *naymïïlaa-* 'to hire'
→ *naymïïla-s-* 'to apply for work'

or the imitative meaning in Tuvan (Kuular 1986: 76):

(54) a. *sadïgla-* 'to work as a salesperson'
→ *sadïgla-š-* 'to play shopping'
b. *baškïla-* 'to be a teacher, to teach'
→ *baškïla-š-* 'to play teachers'
c. *emčile-* 'to work as a doctor'
→ *emčile-š-* 'to play doctors',

or the pseudo-reciprocal meaning in Tuvan (Kuular, forthcoming):

(55) a. *emzir-* 'to breast-feed (a baby)'
b. *Ava-škï-lar* *emzir-ž-ip* *olur-gan-nar*
mother-COLL-PL suckle-REC-CONV AUX-PAST-PL
'Mother was breast-feeding her child.'
lit. ≈ 'Mother and her children breast-fed each other.'
[more precisely, *ava-škï-lar* means 'mother and her child(ren) ...']

It goes without saying that the child could not breast-feed her mother, but this referent is active in the situation described, and this makes it possible to describe this situation in a simplified way without specifying the roles of the referents.

Such meanings are of typological interest as they manifest clear-cut semantic oppositions and they happen to be attested in other unrelated languages; the parallel to (53) is Ancient Greek (56a) (Dvoreckij 1958: 1035, 342) and to (55) it is Ancient Chinese (57) (Jaxontov, forthcoming):

(56) a. *daneidzo* 'to lend someone money' → *daneidzomai* 'to borrow money from someone'. (the reciprocal meaning is also sometimes expressed by the middle inflection:
b. *loidoréō* 'to scold somebody' → *loidorémai* 'to scold each other.')

The middle inflection in Ancient Greek expresses reciprocity and the competitive meaning in a very limited number of cases.

(57) *Yàn qué ... zǐ mǔ xiāng bǔ yě* (*xiāng* = REC; cf. *xiāng suí* 'to follow each other')
lit. 'Swallows and sparrows ... children and mothers feed each other.'

It is but natural that it is only mothers that feed their younglings.

As to the imitative meaning (see (54)), I fail to find it on a reciprocal marker in other languages; nevertheless, this meaning can be expected due to a certain associative link with the reciprocal meaning, as it denotes a situation with two pseudo-symmetrical participants, one who is being imitated and the other who imitates.

The pseudo-reciprocal "meaning" occupies a special place: the point is that this meaning is probably unrelated to the development of polysemy of reciprocal markers as it is registered in Chinese, where the reciprocal pronoun *xiāng* is monosemous. Therefore it is not quite correct to regard "pseudo-reciprocal" as a special meaning like the imitative or the converse or the sociative. Most likely, this is simply a special extended use of the reciprocal function. This use is somewhat similar to the "pseudo-inclusive" use of the 1st p. plural pronoun in expressions such as *And now we shall put on this pretty dress* (mother to daughter).

7.1.2. Productive secondary meanings

Secondary meanings may be productive and represented by numerous enough derivatives, as, for instance, the absolutive meaning in Tatar, numbering a few dozens

of derivatives (see section 3.4 above), or the competitive meaning in Karachay-Balkar, numbering at least 40 derivatives. They may even achieve grammaticalization, having become inflectional markers on an unlimited number of verbs, as is the case with the 3PL meaning in Kirghiz.

Needless to say, a Turkic language may have forms in *-š (-s)* with unproductive and productive secondary meanings.

Most likely, the meanings of this kind may have developed or become productive during the development of the individual Turkic languages. The following meanings deserve special mention.

7.2. The 3PL meaning of the suffix -š in Kirghiz

The suffix *-š* functions in this case as an agreement marker, i.e. as an inflection (though an optional one). This usage is due to the coincidence of the 3SG and 3PL verb forms which have zero marking in Kirghiz. The other personal forms (1SG and 1PL, 2SG and 2PL) have sharply distinct special person-number markers each. The suffix *-š* occupies its usual place in the verb form, both in its reciprocal and 3PL functions: it precedes the tense marker, while other agreement markers follow the tense markers. Compare (I owe the examples below to T. Abdiev, p.c.):

(58)	Singular	Plural
1p	*at-tï-m*	*at-tï-k*
2p	*at-tï-ŋ*	*at-tï-ŋar*
3p	*at-tï-Ø*	*at-tï-Ø (at-ïš-tï)*

Due to the absence of the sociative in modern Kirghiz, the suffix *-š*, when used in the 1PL and 2PL forms, can be interpreted as reciprocal only (see (59a)). Needless to say, I have in mind cases when the lexical meaning of the base and/or construction allows the reciprocal interpretation; otherwise the form is ungrammatical (see (59b), where the direct object prevents reciprocal interpretation and the plural interpretation is ruled out by the 1st person).

(59) a. *Biz* *at-ïš-tï-k.*
we shoot-REC-PAST-1.PL
'We fired at each other.' (-ïš = REC)

b. **Biz* *dušman-dï* *at-ïš-tï-k.*
we enemy-ACC shoot-REC-PAST-1PL (cf. (59d))

In the case of 3PL forms, this suffix on two-place base verbs allows both interpretations, as a reciprocal and a 3PL marker (see (59c)). However, if the syntactic structure of the transitive base construction is retained, the reciprocal interpretation is ruled out and the suffix is plural in meaning; cf. (59d):

c. *Alar at-ïš-tï-Ø.*
they shoot-REC/3PL-PAST-3
i. 'They fired at each other.' (*-ïš* = REC; *-Ø* = 3PL)
ii. 'They fired at somebody.' (*-ïš* = 3PL)
d. *Alar dušman-dï at-ïš-tï-Ø.*
they enemy-ACC shoot-3PL-PAST-3
i. *'They fired at each other.' (*-ïš* = REC; *-Ø* = 3PL)
ii. 'They fired at the enemy.' (*-ïš* = 3PL)

Possible, though somewhat unusual, is the simultaneous use of both markers in succession (see (59e)). In causative constructions derived from reciprocals the plural marker *-ïš* follows the derivational marker (see (59f).

e. *Alar at-ïš-ïš-tï.*
they fire-REC-3PL-PAST
'They fired at each other.' (*-ïš-* = REC; *-ïš-* = 3PL)
f. *Alar ... at-ïš-tïr-ïš-tï.*
they fire-REC-CAUS-3PL-PAST
'They made somebody fire at each other.' (*-ïš-* = REC; *-ïš-* = 3PL)

Also possible is the ambiguity between the assistive and 3PL meanings:

(60) a. *Alar čop tamï-š-tï.*
they hay cart-REC/3PL-PAST
i. 'They helped someone cart hay' ii. 'They carted hay.'
b. *Alar čop tamï-š-ïš-tï.*
they hay cart-REC-3PL-PAST
'They helped someone cart hay.'

This marker may also denote the 3PL meaning on converbs. This happens only in those cases when a converb is related to a 3PL subject:

(61) *čunaŋda-š-ïp kel-iš-e-t biri-n-biri teb-iš-e-t.*
press.ears-3PL-CONV come-3PL-PRES-3 each-other kick-REC-PRES-3
'They (= horses) approach each other pressing their ears, kicking each other.'

For more details see Nedjalkov (2003b: 205-207).

Among other Turkic languages, this 3PL usage of the reciprocal marker is also characteristic, though to a lesser degree, of the areally adjacent Uzbek and Kazakh languages. It has been noted in the literature that this usage of the suffix *-š* is also attested in Ancient Turkic (see, among others, Sevortjan 1962: 355; Kondrat'ev 1970: 25; and especially Blagova 1976: 46-59). However, judging by the cited Ancient Turkic examples, the suffix *-š* may also be interpreted as sociative.

The reciprocal marker in the function of the PL marker also occurs, for instance, in Samoan (Churchward 1951: 77, 78); cf.:

(62) *gagana* 'to speak (of one person)'
→ *fe-gagana-a'i* 'to speak (of two or more persons)'
(cf. reciprocal use:)
ilo 'to see' → *fe-ilo-a'i* 'to see each other.'

7.3. The competitive meaning in Karachay-Balkar

7.3.1. Introductory remarks

The competitive meaning can be defined as follows: "to try to surpass each other in some activity". As we see, it contains the reciprocal component of meaning. In those Turkic languages where it is registered, the competitive meaning may have a broad and/or varied derivational base: competitives may derive from intransitives and transitives; cf. Kirghiz:

(63) a. *taskakta-* 'to trot fast (of horses)'
→ *taskakta-š-* 'to compete in horse trotting races'
b. *tart-* 'to pull, to drag'
→ (*ulak*) *tart-ïš-* 'to compete in goat-pulling (of riders).'

The semantic change may be individual, i.e. with an individual semantic relation with the base; cf. Tatar and Tuvan respectively:

(64) a. *yar-* 'to chop, saw' → *yar-ïš-* 'to compete'
b. *či-* 'to eat something' → *či-š-* 'to compete.'

And sometimes, derivation of a reciprocal occurs via lexicalization; cf. Karachay-Balkar and Kirghiz respectively:

(65) a. *ayt-* 'to speak' → *ayt-ïš-* i. 'to speak to each other'
ii. 'to try to surpass each other in talking' iii. 'to compete in wit'
b. *ayt-* 'to speak' → *ayt-ïš-* i. 'to speak to each other'
ii. 'to compete in improvization (about folktale narrators).'

There are two main types of derivatives which denote the following:

1) competition *per se* ("non-specified" competition; cf. Kirghiz *žeŋ-* 'to win' → *žeŋ-iš-* 'to compete, try to win'), in which case the activity the participants compete in is expressed by a dependent complement (cf. Tuvan *tïvïzïk-ka či-š-* <riddle-DAT eat-REC> 'to compete in asking riddles'), or

2) both the competition and the activity in which the participants compete ("specified competitive"; cf. Kirghiz *at-* 'to shoot' → *at-ïš-* 'to compete in shoot-

ing'). Languages may differ in the productivity of "specified" competitives, "non-specified" being usually very limited in number (generally, three at the most).

2a) If the competitive meaning is productive, the marker can form a respective derivative from practically any base verb (or at least many bases) for which one can imagine the situation of contest (including *ad hoc* situations; cf. Karachay-Balkar *sava-* 'to milk' → *sava-š-* 'to compete in milking').

2b) If the competitive meaning is non-productive, the derivatives, being far from numerous (generally 4-10 items), usually denote conventional contests practiced by the native speakers. In this case derivation of competitives is lexically restricted.

In both types of competitives, there are derivatives from bases denoting actions that naturally involve competition to a greater or lesser degree (see (63a), (66a), (69a), (68a, c)) or an action which is a competition itself (see (70) and (71)).

Among the Turkic languages, the competitive meaning is lacking, for instance, in Azerbaijani, Turkish and Gagauz (Sevortjan 1962: 533). Karachay-Balkar is prominent among Turkic languages as one whose "specified" competitives are highly productive. For comparison, languages with unproductive "specified" competitives will be mentioned in 7.3.3.

7.3.2. Productive competitives: evidence from Karachay-Balkar

In Karachay-Balkar, derivatives with the "specified" competitive meaning can be formed from a variety of both transitive and intransitive verbs. Note that the reciprocal meaning of the Karachay-Balkar suffix *-ïš-* is unproductive. (66a) are common competitive events, while (66b) denote *ad hoc* competitive events.

(66) a.	*čab-* 'to run'	→	*čab-ïš-*	'to compete in running'
	mara- 'to shoot'	→	*mara-š-*	'to compete in shooting.'
b.	*artï-* 'to peel'	→	*artï-š-*	'to compete in potato peeling'
	tig- 'to sew'	→	*tig-iš-*	'to compete in sewing'
	tükür- 'to spit'	→	*tükür-üš*	i.'to spit at each other' ii. 'to try to surpass each other in spitting farther.'

Here is an example from Balkar folklore (Boziev 1962: 60):

(67) *Zü, qart kiši üfgür-üš-ey-ik.*
come.on old man blow-REC-IMP-1.PL
'Come on, old man, let's compete in which will blow the other off.'

For more details see Nedjalkov (2002:61-64).

7.3.3. Unproductive competitives: evidence from other languages

Three main cases deserve to be mentioned, all of them distinguished with a degree of simplification.

1) Competitives are represented by a few (from five to a dozen) derivatives formed from verbs denoting common actions, like 'to run', 'to jump', 'to shoot', etc., which are often practiced as competitive events by the native speakers; their number is limited pragmatically, and they may be specific to only one or a few languages; cf. Yakut (68a), Kirghiz (68b) and Tuvan (68c) respectively. As (68b) and (68c) show, a derivative may have a parallel reciprocal meaning.

(68) a. *kïlïy-* 'to jump on one foot' → *kïlï-s-* 'to compete in jumping on one foot'
b. *at-* 'to shoot' → *at-ïš* i.'to compete in shooting'
ii.'to exchange shots'
eŋ- 'to touch ground (of riders)'→ *eŋ-ïš-* 'to compete in dismounting the rival riders'
c. *tïvïrïkta-* 'to ask sb riddles' → *tïvïrïkta-š-* i. 'to compete in asking riddles'
ii. 'to ask each other riddles.'

This kind of competitive derivatives are also registered in languages other than Turkic (the marker is not necessarily reciprocal in its main meaning). Here are examples from Tagalog (*mag-...-an-* = REC), Japanese (*-a/-aw/-at* = REC; in some districts of Japan competitions in laughing are traditional; T. Tsunoda, p.c.) and Ancient Greek (Dvoreckij 1958: 164; in (69c) the prefix *anti-* and in (69d) middle inflection is used):

(69) a. *mag-panah-an*
REC-shoot.arrows-REC
'to shoot arrows in contest' (cf. reciprocal *mag-ibig-an* 'to love each other').
b. *Gakusei wa warai-at-ta.*
student TOP laugh-REC-PAST
i. ?'The students laughed together.' sociative
ii. 'The students competed in laughing.' competitive
(cf. reciprocal *daki-a-u* 'to embrace each other').
c. *theō* 'to run' → *anti-theō* 'to compete in running'
(cf. reciprocal *ant-eraō* 'to love each other').
d. *diatokseúō* 'to shoot arrows' → *diatokseúomai* 'to compete in arrow-shooting'
(cf. reciprocal (56b) above).

2) There occur derivatives from verbs meaning 'to overcome', 'to win'. These derivatives denote competition only, without indicating the action in which the contestants are competing. Compare Kirghiz:

(70) *žeŋ-* 'to win' → *žeŋ-iš-* 'to compete, try to win.'

Similar derivatives with the general competitive meaning are also registered in languages other than Turkic as well, the main meaning of the marker not necessarily

being reciprocal. Cf. the following examples from Buryat (Čeremisov 1973: 512), Khmer (Dictionnaire 1962) and Chukchi respectively:

(71) a. *ürdi-* 'to overtake' → *ürdi-lde-* 'to compete'
b. *cɛ:ŋ* 'to overcome' → *prɒ-cɛ:ŋ* 'to compete'
c. *yala-* 'to overtake sb' → *yala-čet-* 'to compete.'

3) There are derivatives from the bases meaning 'to overcome sb in *some action/activity*', and their acquired meaning may be either

(a) 'to compete in the activity denoted by the stem' (cf. the meaning (i) in (72a, b, c)) or

(b) 'to compete' in the generalized unspecified sense only (cf. the meaning (ii) in (72a, b, c)).

Note that the meaning of the base verb generally implies a similar action of the second human participant: a sentence like *A is outrunning B* implies that B is running, too.

Here are examples from Yakut, Tuvan and Tatar respectively:

(72) a. *kuot-* 'to outrun sb' → *kuot-us-* i. 'to compete in running' ii. 'to compete'
b. *kag-* 'to overcome sb in wrestling' → *ka-aš-* i. 'to overcome each other by turns' ii. 'to compete'
c. *uz-* 'to outrun' → *uz-ïš-* i. to compete in running' ii. 'to rival, try to overstrip sb in sth.'

A typologically interesting semantic parallel is observed in Bulgarian where the competitive meaning is expressed by the reflexive-reciprocal marker (cf. (29)) on verbs that denote overcoming (cf. the meaning of the base verbs in (73)), e.g. 'to win', 'to overtake', which is expressed by means of the prefix *nad-* 'over'. Thus, the meaning of competition is determined by the prefix in the first place. Cf. (Ivanova 1973: 171-179):

(73) a. *A nadpiva B* + *B nadpiva A* = *A i B se nadpivat*
'A outdrinks B' + 'B outdrinks A' = 'A and B compete in drinking' (lit. ...outdrink each other).

A few more examples:

b. *Te me nadbjagvat* 'They outrun me'
→ *Te se nadbjagvat* 'They compete in running'
c. *Te me nadpluvat* 'They overtake me in swimming'
→ *Te se nadpluvat* 'They compete in swimming'
d. *Te me nadxitrjat* 'They outwit me'
→ *Te se nadxitrjat* 'They compete in outwitting each other'

e. *Te me nadžatvat* 'They overtake me in reaping'
→ *Te se nadžatvat* 'They compete in reaping.'

There are at least 25 such derivatives, while other Indo-European languages which use reflexive-reciprocal markers lack such derivatives. An exception is Serbian, another South-Slavic language, where at least ten such derivatives, also with a prefix *nad-*, are registered (Tolstoy 1970: 274-276); cf.:

(74) a. *nadgovariti* 'to win in an argument'
→ *nadgovariti se* 'to compete in arguing'
b. *nadskakivati* 'to win in jumping'
→ *nadskakivati se* 'to compete in jumping'
c. *nadlagivati* 'to overcome in lying (slandering)'
→ *nadlagivati se* 'to compete in lying.'

Two such synonymous derivatives with the prefix *před-* in the analogous meaning are registered in Czech (Melnikov et al. 1968: 586, 587):

(75) a. *předháněti* 'to outrun'
→ *předháněti se* i. 'to race (with) one another' ii. 'to compete'
b. *předbíhati* 'to outrun'
→ *předbíhati se* 'to race (with) one another (about children)'

In conclusion of this section, I will note that although the Turkic languages have preserved the reciprocal suffix, they display significant variation not only with respect to the main meanings but also in the domain of secondary meanings. (I am grateful to S. Say and A. Letučij for their critical remarks on the early version of the section on competitives.)

8. Spatial reciprocals. The meaning of joining

Spatial transitive reciprocals, like Yakut *baay-* 'to tie two things together' in (76d), crucially differ from proper reciprocals semantically (cf. the pairs of derivatives in (78a) and (78b)), as their reciprocal arguments are objects denoting patients manipulated by the subject referent(s). Despite obvious differences between them, the fact that in many languages both types of reciprocals share the same markers, affixal or pronominal, shows their semantic affinity: both types of reciprocals imply symmetrical arguments.

Prototypical transitive spatial reciprocals are derived from three-place transitives as a result of co-reference of both objects (usually a direct and nondirect object). The non-direct object, though being a spatial argument, is not an adverbial constituent proper, because it is implied by the lexical meaning of the predicate. Generally, it denotes the goal or destination of the direct object referent.

Cross-linguistically, there are at least four types of marking for spatial reciprocals: 1) affixes with locative meanings (see 8.2); 2) reciprocal affixes (8.3); 3) a causative affix attached to anticausatives (8.4); 4) reanalyzed combination of a reciprocal and a causative affix (8.5). The latter two cases are of primary interest because they are represented in Turkic languages. Below, all four types of marking are illustrated. Their discussion is preceded by that of unmarked (= lexical) spatial reciprocals with which a reciprocal pronoun may be optional (8.1). As is shown below, a language may employ several types of spatial reciprocals.

Marked transitive spatial reciprocals are subject to lexicalization: if unmarked three-place transitives of the type *baay-* 'to tie two things together' in (76a) can be used as two-place spatial reciprocals (see (76d)), then, on the other hand, marked two-place spatial transitives also begin to be used as three-place transitives.

8.1. Unmarked spatial reciprocals

In this case one and the same verb functions both as a three-place with the meaning of adding or joining one object to another (which may be different or of the same class; cf. *to stick a sheet of paper on the wall* and *to stick a sheet of paper to another*) and as a two-place with the meaning of joining two objects of the same class (cf. *to stick two sheets of paper together*). Compare Yakut *baay-* i. 'to tie something to something', ii. 'to tie two things together'. Here are examples (N. M. Artem'ev, p.c.)

(76) a. *Kini maŋan kuru qara kur-ga baay-da-Ø.*
s/he white belt black belt-DAT tie-PAST-3SG
'He tied the white belt to the black belt.'

= b. *Kini kara kuru maŋan kur-ga baay-da-Ø.*
s/he black belt white belt-DAT tie-PAST-3SG
'He tied the black belt to the white belt.'

= c. *Kini maŋan kuru ikki qara kuru baay-da-Ø.*
s/he white belt and black belt tie-PAST-3SG
'He tied the white belt to the black belt.'

= d. *Kini ikki kuru [beye-beye-leri-ger] baay-da-Ø.*
s/he two belt each.other-POSS.3PL-DAT tie-PAST-3SG
'He tied two belts [to each other] together.'

In the translation of (76d) the word *together* is used. It should be kept in mind that this English adverb, like its counterparts in many languages (for instance, Yakut *biirge* and Tatar *berg□*; cf. *berg□ kuš-* lit. 'to join something together'), has at least two meanings, viz. sociative of *joint action* and spatial-reciprocal, that of (making) *spatial contact* (cf. Lasersohn 1990: 179-206) or, figuratively, mental contact (see (84c)).

8.2. Affixes with locative meanings

Such affixes are attested in many languages, among others, in languages which do not possess affixal reciprocal markers. Generally, they belong to the sets of numerous (at least 15-20) affixes with various locative meanings. Here is an example from Russian:

(77) a. *On pri-kleil A k B* — 'He pasted A to B'(three-place)
b. *On s-kleil A i B* 'He pasted A and B together' (two-place).

8.3. Reciprocal affixes

As just mentioned, in numerous languages, spatial reciprocals can be formed with the same means as reciprocals proper, as is shown in the Kabardian (78a) (Apazhev et al. 1957: 99, 106) and Swahili (78b) (Ovir 1896: 258) examples below:

(78) a. *guəun* 'to shout at sb' → *zə-guəun* 'to shout at each other'
kIərədən 'to sew sth onto sth' → *zə- kIərədən* 'to sew two pieces together'
b. *-nen-a* 'to speak' → *-nen-an-a* 'to speak with each other'
-fung-a 'to join' → *-fung-an-a* 'to join something together.'

This device does occur in the Turkic languages, but it is extremely rare. In contrast to Kabardian, where the number of reciprocals of type (78b) reaches 150, in Kirghiz only two such derivatives are registered, one of them functioning both as a spatial transitive reciprocal and as an intransitive anticausative (taken from Judaxin 1, 1965; 326, 312-313; 443):

(79) a.	*kak-*	'to knock (once)'	
→ b.	*kag-ïš-*	i. 'to knock two things against each other'	spatial reciprocal
		ii. 'to collide'	anticausative
(80) a.	*ege-*	'to saw', 'to grind'	
b.	*ege-š-*	'to rub one thing against the other.'	

8.4. A causative affix attached to anticausatives

The term anticausative is applied to the member of the semantic causative opposition that is non-causative in meaning and formally marked: this is the member denoting a process the subject referent undergoes without any exterior force (see Nedjalkov & Silnicky 1969 (1973): 20). Anticausatives can be derived not only by means of reciprocal markers but also (and even more productively) by reflexive and passive markers producing, not infrequently, synonymous derivatives. Here is such an example from Yakut (Nedjalkov 2003a, 7: 85, 99-101):

(81) a.	*silimnee-*	'to paste sth and sth together'
→ b.	*silimne-s-*	'to get pasted together'
c.	*silimne-n-*	'to get pasted together.'

Anticausatives derived by means of reciprocal markers are rather widespread among languages; they are attested, among others, in Zulu (Dammann 1954: 164), Muna (van der Berg 1989: 206, 314) and Mbay (Keegan 1997: 66). Here are examples from Mbay (ibid., p. 66) with the reciprocal-sociative marker and from Lithuanian with the reflexive-reciprocal marker respectively:

(82) a.	*tōl-n nàā*	'they killed each other'	reciprocal
	gō-n nàā	'they laughed together'	sociative
	ɓ̄ndə nàā	'become entangled'	anticausative
b.	*jie ap-si-rengė*	'they dressed [themselves]'	reflexive
	jiedu ap-si-kabino	'they (two) embraced each other'	reciprocal
	durys at-si-darė	'the door opened'	anticausative

Across languages, reflexive-reciprocal markers are more productive as anticausative markers than are reciprocal-sociative markers. For instance, among Turkic languages the number of anticausatives with reciprocal-sociative markers does not exceed 20 or 30, while in Lithuanian the number of anticausatives with the reflexive-reciprocal marker is about 800 (Geniušienė 1987: 97).

The reciprocal-causative derivatives can be more or less close in meaning to the base transitives; cf. Kirghiz (Judaxin 1965, 2: 2, 304) and Tatar (Tatar-Russian dictionary 1966: 543):

(83) a.	*ula-*	'to join sth to sth', 'to join the ends of sth and sth'	
→ b.	*ula-š-*	'to join sth/sb'	anticausative
→ c.	*ula-š-tïr-*	'to tie sth and sth together'	causative of anticausative
(84) a.	*tiŋlï-ü*	i. 'to make equal to sth', ii. 'to compare'	
→ b.	*tiŋlï-š-ü*	'to become equal'	anticausative
→ c.	*tiŋlï-š-tïr-ü*	'to compare'	causative of anticausative

Some such pairs, like (a) and (c) in these examples, are sometimes used as synonyms in Russian-Turkic dictionaries to translate one and the same verb; for instance, (84a) and (84c) are used as equivalents of the Russian verb *sravnit'* 'to compare' in Ganiev (1997: 588).

Combinations of a reciprocal and a causative suffix tend to turn into markers of joining together of two objects. In this respect it is significant that although the meaning 'to knock one thing against another' can be expressed by means of the suffix *-š* alone (see translation (i) in (85b)=(79b)), Kirghiz also derives from the anticausative with the meaning 'to collide' (see translation (ii) in (85b)=(79b)) one more spatial reciprocal by means of the causative suffix, viz. (85c):

(85) a.	*kak-*	'to knock (once)'	
→ b.	*kag-ïš-*	i. 'to knock one thing against another'	spatial reciprocal
		ii. 'to collide'	anticausative
→ c.	*kag-ïš-tïr-*	'to knock one thing against another.'	

In the case of verbs denoting joining, the first and the third members of the derivational chains are close in meaning; therefore, verbs of type (c) may oust the base verb, or they may undergo lexicalization easily, thus breaking the synonymy of the members named.

Besides chains like (84), there are triplets in Turkic languages, for example in Tatar, in which the standard semantic relations between the pairs are shifted to a greater or lesser degree (though the semantic relation between (a) and (b) is discernible). This concerns, in the first place, the relations in pairs (a)-(b). Compare (*Tatar-Russian dictionary* 1996: 241, 242; 698, 699; 548):

(86) a. *kat-* i. 'to twist (threads)', ii. 'to flavour (e.g. with double cream)'
b. *kat-ïš-* 'to get mixed (up)'
c. *kat-ïš-tïr-* 'to mix something up.'
(87) a. *yab-* 'to cover'
b. *yab-ïš-* 'to stick to something'
c. *yab-ïš-tïr-* 'to stick, glue something to something.'
(88) a. *tot-* 'to hold, grasp'
b. *tot-aš-* 'to join to something (vi)'
c. *tot-aš-tïr-* 'to join something to something.'

In the following cases there are only forms (b) and (c), base (a) being non-existent (the component *-š* is identified as a reciprocal marker by the meaning of verbs (b)). Compare (ibid., pp. 240, 38):

(89) a. no transitive base verb
b. *katn-aš-* 'to get mixed (up)'
c. *katn-aš-tïr-* 'to mix something up'
(90) a. no transitive base verb
b. *arala-š-* 'to get mixed (up)'
c. *arala-š-tïr-* 'to mix something up.'

The data of this kind cited above may lead us to interpret the *-š-tïr-* complex of suffixes as a marker of derived verbs of joining.

8.5. Reanalyzed combination of a reciprocal and a causative affix

This combination may function as a single suffix, which in the prototypical case changes (potentially) three-place transitives into two-place transitives, i.e. in this case, like in proper reciprocals, valency decrease takes place (see, however, (93d)). It can be attached immediately to the transitive base: an "intermediate" form with the reciprocal suffix is either absent (mostly due to the meaning of the base) or expresses some other, e.g. sociative or assistive, rather than the anticausative meaning. And these meanings of the intermediate form do not correspond semantically to that of the derivative with the reanalyzed reflexive-causative suffix. Here are Kirghiz examples (from Judaxin (1965, 1: 96-97) and T. Abdiev (p.c.) respectively):

(91) a.	*bayla-*	'to tie something to something'	three-place
b.	no anticausative		
→ c.	*bayla-š-tïr-*	'to tie (e.g. horses) together'	two-place
(cf. d.	*bayla-š-*	'to tie something together (= with sb)')	

(92) a.	*kuy-*	'to pour something into something'	three-place
b.	no anticausative		
→ c.	*kuy-uš-tur-*	'to pour from several vessels into one'	two-place
(cf. d.	*kuy-uš-*	'to help to pour').	

Here are a few more analogous examples (T. Abdiev, p.c.) with derivative (d) omitted (identical translations in some pairs below do not rule out slight differences in meaning and the range of meanings of the forms):

(93) a.	*kotor-*	'to move horses from one pasture to another'
	kotor-uš-tur-	'to move many horses from the whole pasture to one place'
b.	*sal-*	'to put something into something'
	sal-ïš-tïr-	'to put several things one into another'
c.	*tüy-*	'to tie some things into a bundle'
	tüy-üš-tür	'to tie (e.g. several bundles) together'
d.	*žïyna-*	'to gather something'
	žïyna-š-tïr-	'to gather something.'[8]

[8] It is interesting to note that in Turkic languages there is a homonymous suffix in which the causative component also lacks the causative meaning; the meaning of this complex suffix is iterative; cf. Kirghiz *sura-* 'to ask' → *sura-š-tïr-* 'to question again and again' (Judaxin 1965, 2: 166); Tatar *yama-* 'to patch' → *yama-š-tïr-* 'to patch many times'. The suffix *-š-tïr-* is synonymous to the iterative suffix *-gula-/-kala-/...* and they frequently co-occur in either order; cf.: *at-* 'to shoot' → *at-kala-* 'to shoot from time to time', *at-kala-š-tïr-* (same); *boraula-* 'to bore, drill' → *boraula-š-tïr-gala-* 'to bore, drill many times' (*Tatar-Russian dictionary* 1966: 709, 43, 78; see also Il'minskij 1863: 15-18 and Sevortjan 1962: 356-358). In this suffix, and also in the homonymous suffix in (91c), (93), etc., the common component of multiplicity may be discerned: multiplicity of actions in the former and multiplicity of objects in the latter. It may be relevant to mention that in Tuvan the suffix *-š-tïr* in combination with the suffix *-gula- /-kula-/...* materially identical with the iterative suffix is used in the meaning of comparison (as is known, comparison is a reciprocal concept); e.g.: *sogun* 'arrow' → *sogun-gula-š-tür* 'like an arrow', *oor* 'thief' → *oor-kula-š-tïr* 'like a thief, in a stealthy manner'. In the Todža dialect of Tuvan, the component *-š-* appears in one more complex suffix which also contains *-š-tïr*, viz. *-šïla-š-tïr*; cf.: *balïk* 'fish' *balïk-šïla-š-tïr* 'like a fish, in a fishlike manner'. The first component of this complex suffix also occurs as an independent suffix with the attached component *-y*: *balïk-šïlay* 'like a fish' (Čadamba 1974: 95). It is not clear if the component *-š-* in this complex suffix is related to the reciprocal suffix.

Analogous Tatar examples (*Tatar-Russian Dictionary* 1966: 85, 86, 630):

(94) a. *buta-* 'to mix up'
buta-š-tïr- 'to mix up'
b. *čal-ïp bəylə-* 'to tie something criss-cross' (*-ïp* = CONV, *bəylə-* 'to tie')
čal-ïš-tïr-ïp bəylə- 'to tie something criss-cross.'

Forms superficially and semantically analogous to Kirghiz *bayla-štïr-* are attested in Japanese, e.g. (Hasselberg 1996: 46, 47):

(95) a. *har-u* 'to paste something to something'
→ b. *hari-aw-ase-ru* 'to paste sth and sth together.'
(96) a. *nu-u* 'to sew sth'
→ b. *nui-aw-ase-ru* 'to sew two things together.'

Here, *-aw* (allomorph *-a*) is a reciprocal suffix and *-ase-* is a causative suffix; cf. (97a) and (97b). In (95b) and (96b) the complex *-aw-ase-* functions as a single morpheme, because in Japanese, causatives cannot derive from reciprocals (and, besides, there are no reciprocal forms of the verbs *haru* (95a) and *nuu* (96a), i.e. the forms **hari-a-u* and **nui-a-u* are non-existent), i.e. formations like (97c) are incomprehensible to native speakers:

(97) *damas-u* 'to deceive somebody'
→ a. *damasi-a-u* 'to deceive each other'
→ b. *damas-ase-ru* 'to make somebody deceive somebody'
c. **damasi-aw-ase-ru* (intended meaning) 'to make someone deceive each other.'

This raises the question: where do forms (95b) and (96b) come from? As a matter of fact, these forms are compounds of two verbs, the base verb followed by the verb *aw-ase-ru* i. 'to join', ii. 'to coordinate', iii. 'to compare' (there are about 80 such compounds in Japanese; see Himeno 1982: 17-52; Hasselberg 1996: 46-51). The latter verb *aw-ase-ru* is the causative form of the verb *a-u* 'to meet', 'to come up', i.e. the verb that was the source of the reciprocal suffix *-a/-aw*. Incidentally, some linguists who are themselves native speakers consider reciprocal derivatives as compounds with the verb *a-u* as well (Nishigaushi (1992: 157) calls it "the reciprocal verb *-aw*"). But in other publications the component *-a/-aw* is regarded as a suffix (derived from the verb *au* 'to meet'; Iwasaki 2002: 144).

The Japanese examples in (95) and (96) as a precise formal (though not morphological) and semantic counterpart of the Turkic material in (91)-(93) reveal the same tendency to derive two-place spatial reciprocals from three-place bases. (My thanks to M. Shibatani for his advice on the Japanese data.)

9. Conclusions

In summary, the following issues discussed above should be stressed.

The Turkic reciprocal suffix can decrease, preserve and increase the valency of the verb depending on the meaning of the derivative: reciprocal, sociative, comitative or assistive (see section 2). A specific feature of the polysemy of this marker is the assistive meaning. The Turkic languages (though not all of them) seem to be the only family of languages, with the exception of the areally adjacent Mongolic languages (see (26)), whose reciprocal marker displays this particular pattern of polysemy covering these four meanings.

With respect to the two main types of reciprocals, with pronominal and affixal markers, the Turkic languages have analogues among other languages and thus differ from others having either only pronominal or only affixal reciprocals (see 1.2). These two types of reciprocals may have their preferential or obligatory semantic domains of usage, alongside their possible pleonastic use. There is also an areal distribution of these two types of reciprocals: in the western part of the Turkic area pronominal reciprocals are prevalent and verbal reciprocals have lost their productivity. The loss of the productivity of verbal reciprocals is observed in some languages of both the Oguz and the Kipchak groups (see section 5). In the eastern Turkic area, on the contrary, the reciprocal meaning of the suffix *-š* has preserved productivity, as has the assistive meaning, which is lacking in the Turkic languages with unproductive reciprocal meaning.

Interestingly enough, the unproductivity of the reciprocal meaning in the *western area* of the Turkic languages is paralleled by unproductivity of this meaning in the neighbouring *eastern area* of Indo-European languages which have a reflexive-reciprocal marker (see section 6).

Differences between the Turkic languages are observed not only in the set and/or productivity of the four main meanings of the reciprocal suffix, but also in the set of secondary meanings. Some languages display idiosyncrasies, such as the unexpected productivity of the competitive meaning in Karachay-Balkar (see 7.3). Another idiosyncrasy is the 3PL agreement function of the reciprocal marker in Kirghiz (see 7.2). The differences also concern such less prominent meanings as, for instance, the anticausative, converse, imitative and pseudo-reciprocal use of the reciprocal marker (section 7.1.1).

The analysis of the polysemy of the reciprocal markers, though fragmentary, reveals a rather varied picture in the Turkic languages selected here. A comparison with non-Turkic languages shows some features of the Turkic reciprocal markers in a broader perspective. Certain meanings of the reciprocal marker attested in some Turkic languages and lacking in others, find parallels in genetically unrelated languages with a different type of basic polysemy. On the whole, the Turkic data may contribute to the investigation of the typology of the polysemy of reciprocal markers,

i.e. to establishing the limits of variation in the domain of polysemy cross-linguistically.[9]

It is clear from the above that we cannot speak of *one type* of reciprocal markers and the limits of variation of its polysemy, three basic types being distinguished, the reflexive-reciprocal, reciprocal-sociative and iterative-reciprocal (see section 3), and the polysemy of each type of markers being determined to a significant degree by their origin. Therefore, these three types should be investigated separately. On the other hand, all three types can share some meanings. Their source is mostly the intransitivizing function of the markers (e.g. the absolutive meaning represented in Tatar and Bashkir and absent in Karachay-Balkar and Khakas (see 3.4) is productive not only in the areally adjacent Russian, but also in some Bantu languages (see example (40)). Another example may be the complex reciprocal-causative marker *-štir* which derives two-place spacial transitive reciprocals from three-place verbs (section 8.5): even this special device has a typological parallel at least in one language, viz. Japanese.

To repeat, a researcher of the typology of reciprocal markers and their polysemy can find extensive material in the Turkic languages.

Acknowledgments

I am grateful to Sergey Say for his valuable comments on the paper that have helped me avoid some pitfalls. Very special thanks are due to my wife Emma Geniušienė, my rigorous critic, for her unstinting help and support.

References

Aksenova, I. S. 1995. Glagol. In: Aksenova, I. S. & Toporova, I. N. (eds.) 1995. *Vvedenie v bantuistiku*. Moskva: Nauka. 127-240.

Alpatov, V. A. & Bugaeva, A. & Nedjalkov V. P. forthcoming. Reciprocals and sociatives in Ainu. In: Nedjalkov, V. P. (ed.) forthcoming. *Typology of reciprocal constructions*.

Apažev, M. L. et al. 1957. *Kabardinsko-russkij slovar'*. Moskva: Gosudarstvennoe izdatel'stvo inostrannyx i nacional'nyx slovarej.

Berger, D. & Drosdovski, G. & Grebe, P. & Müller, W. (eds.) 1972[2]. *Duden 9. Zweifelsfälle der deutschen Sprache*. Mannheim: Bibliographisches Institut.

Bills, G. D. & Vallejo, B. & Troike, R. C. 1969. *An introduction to spoken Bolivian Quechua*. Austin, London: The University of Texas Press.

Blagova, G. F. 1976. Zametki o vzaimno-sovmestnom zaloge. *Sovetskaja tjurkologija* 6, 46-59.

[9] Compare: "We have reason to assume that at least for some grammatical categories in different languages there exists a certain limit [...] of possible polysemy. [...] According to the range of various meanings expressed by comparable forms in them, individual languages differ from one another and can be subjected to classification" (Nedjalkov 1964: 301-302).

Boadi, L. A. 1975. Reciprocal verbs and symmetrical predicates. *The Journal of West African Languages* 10, 1. (West African language monograph series.) Institute of Modern English Language Studies. University of Leeds in association with The West African Linguistic Society. 55-77.

Böhtlingk, O. 1989 (1851). *O jazyke jakutov* [translation from the German "Über die Sprache der Jakuten". St. Petersburg, 1851]. Novosibirsk: Nauka.

Boktug-Kiriš, Bora-Šeelej. 1997. In: *Tuvinskie geroičeskie skazanija*. Novosibirsk: Nauka. 298-527.

Boziev, A. Ju. (ed.) 1962. *Materialy i issledovanija po balkarskoj dialektologii, leksike i fol'kloru*. Nal'čik: Kabardino-Balkarskoe knižnoe izdatel'stvo.

Čadamba, Z. B. 1974. *Todžinskij dialekt tuvinskogo jazyka*. Kyzyl: Tuvinskoe knižnoe izdatel'stvo.

Čankov, D. I. (ed.) 1961. *Russko-xakasskij slovar'*. Moskva: Gosudarstvennoe izdatel'stvo inostrannyx i nacional'nyx slovarej.

Cannings, P. & Moody, M. D. 1978. A semantic approach to causation in French. *Linguisticae Investigationes* 2, 331-362.

Capus, P. A. 1898. Grammaire de Shisumbws ou Lusumbwa. *Zeitschrift für afrikanische und oceanische Sprachen* 4, 1-123.

Čeremisov, K. M. 1973 *Burjatsko-russkij slovar'*. Moskva: Izdatel'stvo Sovetskaja ėnciklopedija.

Churchward, S. 1951. *A Samoan grammar*. Melbourne: Spectator.

Cydendambaev, C. B. 1979. *Grammatičeskie kategorii burjatskogo jazyka*. Moskva: Nauka.

Dammann, E. 1954. Reziprok und Assoziative in Bantusprachen. *Zeitschrift der Deutschen Morgenländischen Gesellschaft*. 104, 163-174.

Dictionnaire 1962. *Dictionnaire Cambodgien* 1. Pnompen: L'Institut Bouddhique.

Dixon, R. M. W. 1980. *The languages of Australia*. Cambridge: Cambridge University Press.

Dvoreckij, I. X. 1958. *Drevnegrečesko-russkij slovar'*. Moskva: Russkij jazyk.

Evans, H. M. & Thomas, W. O. 1981. *Geiriadur Newydd*. Llandybie: Hughes and Son.

Fazylov, E. 1965. K istorii vzaimno-sovmestnogo zaloga v uzbekskom jazyke. In: Ibragimov, S. (ed.) *Issledovanija po grammatike i leksike tjurkskix jazykov*. Taškent: Nauka. 78-96.

Galkin, I. S. (ed.) 1966. *Russko-marijskij slovar'*. Moskva: Izdatel'stvo Sovetskaja ėnciklopedija.

Ganiev, F. A. 1997. *Russko-tatarskij slovar'*. Moskva: Insan.

Gecadze, I. O. & Nedjalkov, V. P. & Xolodovič, A. A. 1969. Morfologičeskij kauzativ v gruzinskom jazyke. In: Xolodovič, A. A. (ed.) 1969. *Tipologija kauzativnyx konstrukcij. Morfologičeskij kauzativ*. Leningrad: Nauka. 131-152.

Geniušienė, E. 1983. *Refleksivnye glagoly v baltijskix jazykax i tipologija refleksivov*. Vilnius: Vilnius University.

Geniušienė, E. 1987. *The typology of reflexives*. Berlin, New York, Amsterdam: Mouton de Gruyter.

Haspelmath, M. 1993. *A grammar of Lezghian*. Berlin: Mouton de Gruyter.

Hasselberg, I. 1996. *Lexikon japanischer Verbalkomposita*. Hamburg: Helmut Buske.

Himeno, M. 1982. Auxiliary verbs denoting symmetry relations (about *au* and *awaseru*). *Nihon-gakoo ronshuu* 9, 17-52 (in Japanese).

Hunan-Kara 1997. In: *Tuvinskie geroičeskie skazanija*. Novosibirsk: Nauka. 54-297.

Il'minskij, N. I. 1863. Učaščatel'naja forma tatarskogo glagola. In: *Učenye zapiski Kazanskogo Universiteta* 1, 15-18.

Isxakova, X. F. 1974. Diatezy i zalogi v tatarskom jazyke. In: Xolodovič, A. A. (ed.) 1974. *Tipologija passivnyx konstrukcij. Diatezy i zalogi.* Leningrad: Nauka. 274-288.

Ivanova, K. 1987. Nyakoi slučai na vzaimodejstvie meždu kategorialno različni javlenija v oblasta glagola. *Slavističen sbornik.* Sofia, 171-179.

Iwasaki, Sh. 2002. *Japanese.* (London Oriental and African Language Library 5.) Amsterdam, Philadelphia: John Benjamins.

Janko-Trinickaja, N. A. 1962. *Vozvratnye glagoly v sovremennom russkom jazyke.* Moskva: Izdatel'stvo Akademii Nauk SSSR.

Jaxontov, S. J. fortcoming. Reciprocal constructions in Ancient Chinese. In: Nedjalkov, V. P. (ed.) forthcoming. *Typology of reciprocal constructions.*

Jayaseelan, K. A. 2000. Lexical anaphors and pronouns in Malayalam. In: Lust, Barbara C. & Wali, Kashi & Gair, James W. & Subbarao, K. V. (eds.) 2000. *Lexical anaphors and pronouns in selected South Asian languages.* (Empirical Approaches to Language Typology 22.) Berlin, New York: Mouton de Gruyter. 113-168.

Judaxin, K. K. 1965. *Kirgizsko-russkij slovar'.* Moskva: Sovetskaja ėnciklopedija.

Juldašev, A. A. 1981. Kategorija zaloga. In: Juldašev, A. A. (ed.) 1981. *Grammatika sovremennogo baškirskogo literaturnogo jazyka.* Moskva: Nauka. 242-264.

Katanov, N. F. 1894. Iz ėtimologii tatarskogo jazyka. In: Voskresenskij, A. 1894. *Russko-tatarskij slovar'.* Kazan': Tipolitografija Universiteta. 15-105.

Keegan, J. M. 1997. *A reference grammar of Mbay.* (Lincom Studies in African Linguistics 14.) München, Newcastle: Lincom Europa.

Kemmer, S. 1993. *The middle voice.* Amsterdam, Philadelphia: John Benjamins.

Kemmer, S. 1997. Collective and distributive marking, or: where unity meets multiplicity. In: Melby, A. K. (ed.) *The twenty-third LACUS forum 1996.* Chapel Hill, NC: Linguistics Association of Canada and the United States. 231-249.

Kīīs Debeliye: Jakutskij geroičeskij ėpos. 1993. Novosibirsk: Nauka.

Kondrat'ev, V. G. 1970. *Očerk grammatiki drevnetjurkskogo jazyka.* Leningrad: Izdatel'stvo universiteta.

Kornfilt, J. 1997. *Turkish.* London, New York: Routledge.

Kuular, K. forthcoming. Reciprocals, comitatives and assistives in Tuvan. In: Nedjalkov, V. P. (ed.) forthcoming. *Typology of reciprocal constructions.*

Kuular, K. B. 1986. Sovmestno-vzaimnyj zalog v tuvinskom jazyke. *Sovetskaja tjurkologija* 2, 73-82.

Kuz'menkov, E. A. 1984. *Glagol v mongol'skom jazyke.* Leningrad: Izdatel'stvo Leningradskogo universiteta.

Lasersohn, P. 1990. Group action and spatio-temporal proximity. *Language and Philosophy* 13, 179-206.

Letučij, A. forthcoming. Zalog i aktantnaja derivacija v xakasskom jazyke.

Lichtenberk, F. 1985. Multiple uses of reciprocal constructions. *Australian Journal of Linguistics.* 5, 19-41.

Lichtenberk, F. 1991. Reciprocals and depatientives in To'aba'ita. In: Blust, R. (ed.) 1991. *Currents in Pacific linguistics: Papers on Austronesian languages and ethnolinguistics. In honour of George W. Grace.* (Pacific Linguistics C-117.) 171-183.

Liu, M. 1999. Reciprocal marking with deictic verbs *come* and *go* in Mandarin. In: Frajzyngier, Z. & Curl, T. S. (eds.) 1999. *Reciprocals: Forms and functions.* Amsterdam, Philadelphia: John Benjamins.

Manas 1995. Volume 4. (Kirghiz heroic epic poem). Moskva: Nasledie.

Marconnès, F. S. J. 1931. *A grammar of Central Karanga. The language of Old Monomotapa. As at present spoken in Central Mashonaland, Southern Rhodesia.* Johannesburg: Witwatersrand University Press.

Melnikov, E. et al. 1968. *Česko-ruský slovník.* Praha: Státní pedagogické nakladatelství.

Middendorf, E. W. 1891. *Die Aimara-Sprache.* Leipzig: Brockhaus.

Modun Er Soyotox: Jakutskij geroičeskij ėpos. 1996. Novosibirsk: Nauka.

Mosel, U. & Hovdhaugen, E. 1992. *Samoan reference grammar.* Oslo: Scandinavian University Press.

Nedjalkov, V. P. 1964. O svjazi kauzativnosti i passivnosti. In: Skrebnev, J. M. (ed.) 1964. *Voprosy obščego i romano-germanskogo jazykoznanija.* Ufa: Baškirskoe knižnoe izdatel'stvo. 301-310.

Nedjalkov, V. P. 1976. *Kausativkonstruktionen.* (Studien zur deutschen Grammatik 4.) Tübingen: Gunter Narr.

Nedjalkov, V. P. 2000. *Zametki o nemeckix reciprokal'nyx konstrukcijax.* In: *Jazyk i rečevaja dejatel'nost'* 3, 102-117.

Nedjalkov, V. P. 2002. Karachay-Balkar reciprocals. *Turkic Languages* 6, 19-80.

Nedjalkov, V. P. 2003a. Yakut reciprocals. *Turkic Languages* 7, 1-77.

Nedjalkov, V. P. 2003b. Kirghiz reciprocals, assistives and plural. *Turkic Languages* 7, 181-236.

Nedjalkov, V. P. & Silnicky, G. G. 1969. Tipologija morfologičeskogo i leksičeskogo kauzativov. In: Xolodovič, A. A. (ed.) 1969. *Tipologija kauzativnyx konstrukcij. Morfologičeskij kauzativ.* Leningrad: Nauka. 20-50. [English version: The typology of morphological and lexical causatives. In: Kiefer, F. (ed.) 1973. *Trends in Soviet theoretical linguistics.* Dordrecht, Boston: D. Reidel Publishing Company, 1-32.

Nigmatov, X. G. 1973. Zalogi glagola v vostočno-tjurkskom jazyke XI-XII vekov. *Sovetskaja Tjurkologija* 1, 46-61.

Nishigaushi, T. 1992. Syntax of reciprocals in Japanese. *Journal of East Asian Linguistics* 1, 157-196.

Ovir, E. 1896. Die abgeleiteten Verba in Kiswahili. *Zeitschrift für afrikanische und oceanische Sprachen* 2, 249-266.

Pekarskij, Ė. K. 1959. *Slovar' jakutskogo yazyka* 1-3. Moskva: Akademia Nauk SSSR [1st edition 1907-1930].

Rjabova, I. S. 1989. Fragmenty grammatiki jazyka dabida. In: Ismailova, R. N. (ed.) 1989. *Afrika: kul'tura i obščestvo* 2. *Jazykoznanie, literaturovedenie.* Moskva: Nauka. 101-113.

Rombandeeva, E. I. 1973. *Mansijskij (vogul'skij) jazyk.* Moskva: Nauka.

Sanžeev, G. D. 1963. *Sravnitel'naja grammatika mongol'skix jazykov.* Moskva: Izdatel'stvo vostočnoj literatury.

Sanžeev, G. D. 1960. *Sovremennyj mongol'skij jazyk.* Moskva: Izdatel'stvo vostočnoj literatury.

Ščerbak, A. M. 1981. *Očerki po sravnitel'noj morfologii tjurkskix jazykov: Glagol.* Leningrad: Nauka.

Serebrennikov, B. A. & Gadžieva, N. Z. 1979. *Sravnitel'no-istoričeskaja grammatika tjurkskix jazykov.* Baku: Maarif.

Sevortjan, E. 1962. *Affiksy glagoloobrazovanija v azerbajdžanskom jazyke.* Moskva: Izdatel'stvo vostočnoj literatury.

Shkarban, L. I. & Rachkov, G. E. forthcoming. Reciprocal, sociative and comitative constructions in Tagalog. In: Nedjalkov, V. P. (ed.) forthcoming. *Typology of reciprocal constructions*.

Starostin, S. & Dybo, A. & Mudrak, O. 2003. *Etymological dictionary of the Altaic languages*. (Handbook of Oriental Studies. Section eight, Central Asia 8,1). Leiden, Boston: Brill.

Sterner, J. K. 1987. Sobei verb morphology reanalyzed to reflect POC studies. *Oceanic Linguistics*. 26, 1-2, 30-54.

Tatar-Russian dictionary 1966. = *Tatarsko-russkij slovar'* [editor not indicated]. Moskva: Sovetskaja ėnciklopedija.

Tenišev, Ė. R. (ed.) 1997. *Jazyki mira. Tjurkskie jazyki*. Moskva: Indrik.

Tenišev, Ė. R. 2002. Kyrgyzskaja gruppa. In: Tenišev, Ė. R. (ed.) 2002. *Sravnitel'no-istoričeskaja grammatika tjurkskix jazykov. Regional'nye konstrukcii*. Moskva: Nauka. 476-599.

Tepljašina, T. 1966. Udmurtskij jazyk. In: Lytkin, V. I. et al. (eds.) 1966. *Jazyki narodov SSSR* 3. *Finno-ugorskie i samodijskie jazyki*. Moskva: Nauka. 261-281.

Terasawa, J. 1985. The historical development of the causative use of the verb *make* with an infinitive. *Studia Neophilologica* 57: 133-143.

Tolstoj, I. I. 1970. *Serbskoxorvatsko-russkij slovar'*. Moskva: Sovetskaja ėnciklopedija.

van de Kerke, S. 1996. *Affix order and interpretation in Bolivian Quechua*. Meppel: Krips Repro.

van den Berg, R. 1989. *A grammar of the Muna language*. Dordrecht: Foris Publication.

Vitale, A. J. 1981. *Swahili syntax*. Dordrecht, Cinnamirson: Foris Publications.

Xaritonov, L. N. 1963. *Zalogovye formy glagola v jakutskom jazyke*. Moskva, Leningrad: Akademii Nauk SSSR.

Xaritonov, L. N. 1982. Sovmestno-vzaimnyj zalog. In: Korkina, E. I. et al. (eds). *Grammatika sovremennogo jakutskogo literaturnogo jazyka*. Moskva: Nauka. 268-274.

Ziervogel, D. 1952. *A grammar of Swazi*. Johannesburg: Witwatersrand University Press.

Zinnatullina, K. Z. 1969. *Zalogi glagola v sovremennom tatarskom literaturnom jazyke*. Kazan': Tatarskoe knižnoe izdatel'stvo.

Zubizarreta, M. L. 1985. The relation between morphophonology and morphosyntax: The case of Romance causatives. *Linguistic Inquiry* 16, 247-289.

Ös tilɪ: Towards a comprehensive documentation of Middle and Upper Chulym dialects

Gregory D. S. Anderson & K. David Harrison

Anderson, Gregory D. S. & Harrison, David K. 2006. *Ös tilɪ*: Towards a comprehensive documentation of Middle and Upper Chulym dialects. *Turkic Languages* 10, 47-72.

This paper offers data that represent progress in our comprehensive (text/audio/video) documentation of *Ös tilɪ*, a moribund Turkic language spoken along the Middle and Upper Chulym river basin. We present data on newly documented genres of texts (including spontaneous speech) and report on the progress of the development of the incipient indigenous Ös orthography. We present an analysis of data collected during a field expedition in June and July 2005 as they reflect idiosyncratic innovative and archaic features of Turkic, a loss of structures attested in the previous literature on the language, and the effects of contact with the socially dominant Russian.

Gregory D. S. Anderson and K. David Harrison, Department of Linguistics, Swarthmore College, 500 College Ave. Swarthmore, PA 19081, USA. Email: Info@livingtongues.org

0. Current status

This paper presents results from a July 2005 field expedition to document speakers of the Ös language of the Middle and Upper Chulym river basin. Native ethnonyms include [*bistɪŋ kiʃɪler*] ('our people') and [*øs kiʃɪler*] 'Ös people'. Exonyms include 1. [*xakas*], 2. [*jasatʃnije*] (Russian 'tribute-payers'), 3. [*tadar*] 'Tatar'. The native designation for the language is almost always [*øs til(ɪ)*] ~ [*øs til(ɪ)*], and less frequently [*tadar til(ɪ)*].[1] The former, while indeed partially coinciding with the Turkic stem [*øs*] 'self' (albeit in this language always appearing in a possessed form such as [*øzym*], [*øzyŋ*], or [*øzy*]) is probably primarily derived from the native designation for the Chulym River, viz. [*øs*] (and further upriver, [*ys*], whence the Russian designation *Ijus* for its southeastern tributary in northern Khakasia). We refer to the language herein as Ös or Ös tilɪ, in deference to established practice and stated preference within the native-speaker community. The Ethnologue (Gordon 2005) code is [CHU]. The language is critically endangered and moribund with probably fewer than 30 fluent speakers and an additional 20 to 30 semi-speakers remaining. The Middle Chulym dialect now has fewer than twenty speakers; the youngest was 54

[1] Transcription conventions follow IPA usage. We transcribe nine Ös vowel phonemes: front [i y ø e æ] and back [ɯ u o a] and one allophone [ɪ].

years old in 2005. Pending future field expeditions, we have no solid estimate for Upper Chulym dialect speakers. We found no evidence that any speakers remain of the Lower Chulym variety (which is in fact rather different from the Middle and Upper Chulym varieties, showing more Siberian Tatar-like structures).

Ös speakers reside primarily in villages in Tomsk oblast' (Tegul'det, Novošumylovo, Belij Jar, Ozjornoje) and neighboring Krasnojarsk kraj (Pasečnoje, Čindat, and Tjuxtet). Results of our 2003 expedition, including speaker demographics, were reported in Harrison and Anderson (2003). Prior to a linguistic Turkicization of the region, the area was formerly inhabited by various peoples speaking the Arin and Pumpokol languages of the Yeniseic family, a fact that is amply attested by the predominance of hydronyms in the region with clear Yeniseic origins, viz. Latat, Tegul'det, Kul', etc. It is at least possible that some of the characteristic features of Ös tilı might reflect this linguistic substratum, although this remains to be adequately demonstrated by future research.

1. Previous research

Our work builds upon a long tradition of research into the Ös people and their language. According to Dul'zon (1966: 446) the first Chulym forms to be mentioned come from a few toponyms in Russian documents of the 17th century. The Swedish explorer J. P. Falck visited the Chulym Turks during his Siberian expedition of 1768-1773 and reported on their demographics, material culture, subsistence lifeways, and traditions such as the lunar calendar. In his posthumously published account, Falck (1786: 554-557) wrote:

"Die Jurten sind kleine tatarische, meistens kubische Hütten, aus einer Stube, mit der Tatarischen breiten Banke, Kamin, eingemauerten Grapen, und Fensterlöchern mit Quappenhäuten überspannt, oder auch mit eingesetzen Eise. Solche Eisfenster nennen sie Ulugi Siderno."

"Die Tschulymer schlafen auf Birkenrinde und decken sich mit ihren Kleidern zu. Kinder packen sie in Rindekörbe und umschütten sie mit zerriebenen olmigen Holz..."

"Ihr erstes Gewerbe ist die Fischerey, die sie in den Flüssen und Seen das ganze Jahr treiben. Hiernächst beschäftigen sie sich des Winters sehr mit der Jagd. Viele haben auch einen kleinen Ackerbau und nur wenige besitzen einige Kühe und Schaafe."[2]

[2] "The yurts are small Tatar-style, mostly cubical huts, with one room, with the Tatar-style broad bench...and window-holes covered with burbot skin, or made with inset ice. These ice-windows are called Ulugi Siderno."
"...The Chulym sleep on birch bark and cover themselves with their clothes. They pack their children into bark baskets and pack ground rotted wood around them."
"Their first industry is fishery, which they practice in the rivers and lakes the year round. In the winters they do hunting. Some have little gardens and only a few have cows or sheep."

The first significant collection of lexical materials dates to Messerschmidt's journal from the early 18th century, a significant portion of which were published on pages 224-226 of J. Klaproth's (1823) *Asia Polyglotta*. Lexical materials also may be found in the *Linguarum totis orbis vocabularia comparativa* (Comparative vocabulary of the languages of the whole world) commissioned by Catherine the Great and appearing in 1789 under the editorship of P. Pallas. Some 150 words and 60 expressions appeared in the anonymously authored *Jazyk čulymskix inorodtsev* (Language of Chulym Aborigines) from the annals of the Tomsk *Gubernija* of 1858. V. V. Radloff visited the Chulym in 1863 and published an excerpt from an epic tale in the Lower Chulym variety "Taska Mattyr" in the second volume of his text samples of (Siberian) Turkic languages (1868: 689-705). He added some brief phonological and lexical materials in his *Opyt slovarja tjurkskix narečij* (An attempt at a dictionary of Turkic dialects) (1882-1899) and *Phonetik der nördlichen Türksprachen* from (1882).

A modest amount of Ös data appears in N. F. Katanov's 1903 study of Tuvan and in S. E. Malov's 1909 field report. The scholar A. P. Dul'zon renewed the study of Ös in the 1940s and 1950s, undertaking field expeditions to the Chulym region, and producing a range of short works (cf. Dul'zon 1952a, 1952b, 1954, 1956, 1957, 1966, 1973). His student R. M. Birjukovič produced some studies in the 1970s and 1980s (e.g., 1972, 1973, 1975a, 1975b, 1979a, 1979b, 1980a, 1980b, 1981a, 1981b, 1984, 1997, Serebrennikov and Birjukovič 1984); these are based both on her own field notes as well as on Dul'zon's field notes. A couple of brief texts were published by Abdraxmanov (1970) with Russian translations but no interlinear glossing or grammatical analysis in the obscure Tomsk-based journal *Jazyki i Toponimija Sibiri*. Except for one very brief article in German (Pritsak 1959) and the extremely limited amount of Ös data in Radloff (1882), no language data were previously accessible to non-Russian speakers prior to Harrison and Anderson (2003) and Anderson and Harrison (2004a). Ethnological and historical research has been conducted by E. L. Lvova (1972, 1978), V. Novokšonov (1995), V. P. Krivonogov (1998) and others.

In addition to the publications mentioned above, the Siberian Languages Laboratory in Tomsk has ca. 700 pages of unpublished field notes taken by R. A. Pečjorskaja, R. A. Boni, R. M. Birjukovič and A. P. Dul'zon. These are found in four and a half field notebooks and several thousand index cards dating back 60 years, containing individual lexical items and phrases of the Middle, and Upper Chulym dialects of Ös (although the two dialects are not differentiated in the notebooks, a careful inspection of the materials reveals this to be the case). These have been kept locked away in the lab for over thirty years. There were originally seven such notebooks; two of Dul'zon's original field notebooks have vanished. One of the notebooks (mainly dedicated to Nganasan, a Northern Samoyedic language spoken in the Tajmyr district of Krasnojarsk Kraj) contains four Ös texts of varying lengths, two traditional tales (one long and one short) and two brief instructive tales. A further two of the notebooks are dedicated to lexical materials, elicited from Russian, which contain in them a considerable number of Russian loan words (e.g., for many cultural

items of Russian provenience). These were gathered in Pasečnoje, Krasnojarsk Kraj and seem to reflect an Upper Chulym variety. The two remaining volumes have approximately forty additional pages of texts, approximately 900 elicited sentences, two hundred forty pages of additional lexical materials organized according to the alphabetical order of the Russian originals and approximately two hundred fifty pages of specialized lexical lists, e.g. plant names, animal names, trees, etc. collected in Tegul'det and Pasečnoje. The total number of speakers consulted in producing these materials is fewer than 10, and the materials mainly reflect the speech of four informants, all now deceased, representing both the Middle and Upper Chulym varieties of Ös. These valuable legacy materials have been digitized and are currently being analyzed under the auspices of our ongoing documentation project of Ös (for more on this see section 5 below).

The Russian ethnographer V. P. Krivonogov may have some recordings of Ös language materials collected during his ethnographic expedition of the early 1990s, and there is mention of audio recordings done by Tomsk scholars, but their current whereabouts are unknown to researchers in Tomsk and abroad. The list of field notebooks, file cards with lexical materials, etc., unpublished but housed in Tomsk, is listed in Il'jašenko et al. (1998).

2. Necessary and sufficient documentation

Despite the aforementioned body of work, Ös (Chulym Turkic) remains only minimally documented in terms of contemporary standards for documentation (Himmelman 1998, Woodbury 2003), and also minimally described. As A. Filtchenko (2001) points out, not only is Chulym Turkic one of the most endangered Turkic languages (possibly second only to Karaim), but it is also underdocumented:

"In terms of documentation, Chulym Turkic is also one of the least studied Turkic languages; apart from casual collections of scattered data which have been included in some Turkological reference works, the serious study of Chulym Turkic in its own right began only after the World War II. However, a dictionary of any kind is still nonexistent for Chulym Turkic which greatly impedes the study of this dwindling and, in view of its demonstrable early contacts with now extinct members of the Yenissejan language family also a really important language. The number of published texts remains also very small."

Our current documentation builds upon all previous work, whether published or unpublished, professional or amateur. A good deal of unpublished data from the 1970s is archived in Tomsk (Il'jašenko et al. 1998) and provides a valuable comparative dimension to the work. In light of the accelerated changes that moribund and obsolescent languages may undergo—for example changes observed in Tofa—(Harrison and Anderson forthcoming-a, -b, Anderson and Harrison forthcoming), it has been fruitful to compare the previous studies with the current state of the language. But we cannot rest upon legacy data. Virtually no spontaneous Ös speech has been recorded prior to our recent field expeditions, and the set of morphological and syn-

tactic features that can be gleaned from either the elicited sentences in the field books or the two traditional tales[3] and other brief textual materials is quite limited. A full, multi-media documentation must be undertaken while remaining speakers are available. New paradigms are emerging in documentary linguistics (Himmelman 1998, Woodbury 2003) that call for the fullest possible collection and long-term archiving of multi-media annotated materials and metadata, not simply written notes or scholarly articles.

3. New findings from expedition to the Chulym, July 2005

We report here a sample of the range and quality of data we were able to collect on an expedition in June-July 2005 to the villages of Belij Jar, Novošumylovo, Tegul'det, and Ozjornoje. We made video and audio recordings of five fluent speakers and two semi-speakers. We interviewed an additional two fluent speakers. We made a total of 8 hours of video and 10 hours of audio recordings and 150 pages of handwritten transcriptions and field notes, all with associated metadata for archiving. We also collected, translated and annotated 15 texts written by a native speaker, as described in section 4. Our findings give an emerging picture of dialect differences between the so-called Middle and Upper Chulym varieties of Ös, both phonological and morpho-lexical differences (3.1), features of these two speech varieties reflecting possible obsolescence effects due to the advanced moribund state of the language (3.2), a very brief discussion of the presence of micro-variation even on the idiolectal level of a type that is not uncommonly found in severely endangered languages (3.3), and finally some brief comments on the influence of the socially dominant Russian language on the structure of the contemporary Ös language (3.4).

3.1 Dialect differences

A range of phonological, morphological and lexical differences may be readily observed between speakers of the Middle Chulym dialect of Ös, and those hailing from the Upper Chulym dialect regions. During our field expeditions in 2003 and 2005, we recorded six speakers of the Middle Chulym dialect of Ös and two speakers of the Upper Chulym dialect, some now deceased. Due to significant inter- and intra-speaker variation, and the fact that the two Upper dialect speakers have lived for many years in the Middle dialect speech community, it is not possible to conclusively claim that the attested differences are solely indicative of a dialect isogloss. Nonetheless, we present a sample of the differences here.

3.1.1 Phonological differences

A range of phonological differences has emerged in our field corpus between the so-called Middle Chulym and Upper Chulym varieties of Ös. These emerged during an

[3] Albeit these texts are clearly an invaluable source as this genre of traditional tales now appears to be extinct.

elicitation of a Swadesh-type word list with two speakers, each representing a different dialect area, but both currently living in the tiny village of Ozjornoje (total population under forty). A selection of these is offered in (1)-(5) below.

(1)	Middle Chulym	Upper Chulym	
	[ʃ]	[tʃ]	
	[ʃoʃqa]	[tʃoʃqa]	'pig'
	[iʃ]	[itʃ]	'drink!'
	[ʃarɯn]	[tʃarɯn]	'back'
	[qeʃ]	[qætʃ]	'child'
	[paʃlama]	[patʃtɯlama]	'dirty'

(2)	Middle Chulym	Upper Chulym	
	[ʃ]	[s]	
	[ʃij]	[sij]	'write!'
	[tɯʃpa-zɯ]	[tɯspa-zɯ]	'knee'-3

(3)	Middle Chulym	Upper Chulym	
	[b]	[m]	
	[bojnum]	[mojnum]	'neck'-1
	[bunzun]	[munzun]	'that one'-3.ACC

(4)	Middle Chulym	Upper Chulym	
	[ŋ]	[tʃ] ~ [ʃ]	
	[ŋaa]	[tʃaa]	'new'
	[ŋaaɣɯ]	[tʃaaɣɯ]	'jaw'-3
	[ŋøme]	[ʃøme]	'thing'

(5)	Middle Chulym	Upper Chulym	
	[m]	[p] ~ [b]	
	[møːre]	[pøːre] ~ [bøːre]	'wolf'

With respect to the *b : m* and *m : p / b* and *ŋ : ʃ / tʃ* correspondences, it appears that there is a certain degree of regularity, although of a type that is not necessarily what might be 'expected'. The Middle Chulym variety of Ös shows a distant nasal assimilation with the palatal but not with the labial series to a following nasal, while the Upper Chulym variety shows just the opposite; contrast the forms in (3) and (4). The labial-initial forms have the nasal sound triggering the distant assimilation still present, but the first two palatal-initial forms have lost the nasal sound that triggered the nasal assimilation process: etymological (but not synchronic) **-ŋ-* is regularly lost in intervocalic position, leaving a long vowel (the nasal that triggers the nasal assimilation in the palatal forms does remain in the final example in (4)). On the other hand, the word for 'wolf' has an unexpected nasal-initial in Middle Chulym that is lacking in the Upper Chulym variety. Regarding the *ʃ : s* correspondence, the Upper Chulym variety shows a more Khakas-like structure; note that the so-called

Xyzyl (Kyzyl) variety of Khakas is spoken in the region just to the south of the Upper Chulym region (currently separated by Russian speaking populations).

Certain phonological processes appear to be different, though there is also variation within dialects and speakers. These include velar elision,

(6)	Middle Chulym	Upper Chulym	
	[*tʃa-a*]	[*tʃaɣ-ɯ*]	'fat'-3

final vowel epenthesis,[4]

(7)	Middle Chulym	Upper Chulym	
	[*tur-ubulu*]	[*tur-ubul*]	'stand'-PRES
	[*ajd-ɯbɯlɯ*]	[*ajd-ɯbɯl*]	'say'-PRES
	[*bil-ibili*]	[*bil-ibil*]	'know'-PRES

and diphthongization.

(8)	Middle Chulym	Upper Chulym	
	[*sooq*] ~ [*su^{w}aq*]	[*sooq*]	'cold'

3.1.2 Morphological differences

The following morphological differences have been observed in our field corpus. The Upper Chulym dialect of Ös speakers shows an atypical category of plural *possessor* marking with nouns encoding 2nd plural and 3rd plural possessors. This characteristic feature is lacking in the variety favored by speakers of the Middle Chulym dialect of Ös. Not only is a specifically plural possessor form unusual for Altai-Sayan Turkic languages, the form itself is also of an unusual or marked type. It appears to have a fixed low round vowel segment [o] in an otherwise plural-looking morphemic shape, regardless of the quality of the preceding vowel; this vowel triggers a back and round harmony pattern to spread the general (second or third) person possessive marker that follows.

(9)	Middle Chulym	Upper Chulym	
	[*qaraɣ-ɯ*]	[*qaraq-tor-u*]~[*qaraq-lor-u*] 'eye'-PL:POSS-3	'their eye' (singular)

[4] While from a diachronic perspective it is possible that the longer variants actually reflect the second syllable of the originally bi-syllabic auxiliary, deriving as this present tense formation does in Ös from an auxiliary construction in **-Ip olur*, attested to varying degrees in a range of other Altai-Sayan Turkic languages (Anderson 2004b), from a synchronic perspective it is not possible to know whether this is a quasi-archaic retention or an innovation of a process of final-vowel epenthesis. Given the small pool of speakers of both varieties, it is not clear that this issue can even really be resolved satisfactorily.

Note that for plural forms of possessa, the special plural possessor element does not appear and the forms are thus ambiguous to whether they refer to a plural possessor of a plural possessum, or to a singular possessor of a plural possessum.

(10)	Middle Chulym	Upper Chulym	
	[*qaraq-tar-ɯ*]	[*qaraq-tar-ɯ*]	'their eyes' (plural)
		'eye'-PL-3	or 'his eyes'

Another morphological feature that shows variation between the Middle and Upper Chulym dialects of Ös is the instrumental suffix/enclitic. In fact, one of the variants of the Upper Chulym variety appears to reflect the putative proto-form of the instrumental underlying all the variants of this element found in Altai-Sayan Turkic languages in general, and those of the complex dialect continua of Khakas in particular as reconstructed and discussed in Anderson (2003; 2004a, and forthcoming).

(11)	Middle Chulym	Upper Chulym	
	[*-(b)la*]	[*-bɬæŋ*]	'with', 'by means of'

3.1.3 Lexical differences

Lexical differences (including some that are actually phonological) were observed and also pointed out to us by the speakers themselves, who are sensitive to such dialect differences.

(12)	Middle Chulym	Upper Chulym	
	[*tuɣba-m*]	[*bija-m*]	'(elder) sister'-1
	[*æptʃi*]	[*æpʃi*]	'wife'
	[*moɣalaq*]	[*qatʃqɯ*]	'bear'
	[*tudun*]	[*ʃon*]	'clothing'
	[*kæær i*]	[*ton*]	'old' (of a person)
	[*ʃaŋ*], [*ŋaa*]	[*tʃaa*]	'new'

With regards to the element meaning 'bear', the Middle Chulym form is of currently unknown origin, while the Upper Chulym [*qatʃqɯ*] literally means 'devil, evil spirit', a euphemism word of a type commonly found designating bears across many languages of Central Siberia.

3.2 Obsolesence effects

As we collected our corpus of texts and observed (and participated in) the spontaneous conversations of our consultants, it became clear to us that certain features that had been reported in the previous literature on Ös were not appearing. Thus, it seems that certain 'expected' features of the language have either been lost or are extremely rare. These may in fact reflect the general moribundity of the language and the concomitant loss or reduction of certain features that this particular sociolinguistic state

commonly entails. Such features include the use of the common Turkic -DI past and the unaccomplished aspectual formation.

In the case of the -DI past, one elderly female speaker of the Upper Chulym variety of Ös produced one token of a verb form with this suffix, viz. [*di-di-m*] 'I said' ('say'-REC.PST-1). Other speakers in elicitation and consultation sessions explicitly reject such forms even in the context of the two features that seem to favor its use in other Turkic languages of the region (cf. Anderson 1998, Anderson and Harrison 1999, Anderson and Harrison in preparation, Anderson 2005 for Khakas, Tuvan, Tofa, and Abakan Khakas respectively), that is, recent past action with a first singular subject (this is in fact the exact context in which the above-mentioned speaker spontaneously produced the one token attested so far in our field corpus). Given the textual frequency of the -GAn past in comparison to the -DI past in many Altai-Sayan Turkic languages, it is perhaps not overly surprising that in the speech of the final generation speakers of the moribund Ös language, the latter element has been replaced in its usual functional contexts by the former.

With respect to the unaccomplished aspect affix, this too was acknowledged and produced under elicitation by exactly one male speaker of the Middle Chulym variety of Ös. When prompted by a Russian token, he produced the following forms:

(13)	[*mæn*	*uz-alaq-ɯm*]	[*sæn*	*jeʃo*	*uzu-ba-a-ŋ*]
	I	sleep-UNACMPL-1	you	still/yet	sleep-NEG-PST-2
	'I haven't slept yet'		'you haven't slept yet'		
	(GDSA field notes 24, 118)		(GDSA field notes 94)		

It is interesting to note that this unaccomplished element seems to have been already nearly lost when the systematic study of the language began in the mid-20th century; almost exactly the same two variants occurred in Dul'zon's (1960) description of the verbal system of Middle Chulym.

(14)	*men*	*uz-a:lak-im*	*men*	*iʃo*	*uzu-ba-a-m*
	I	sleep-UNACMPL-1	I	yet	sleep-NEG-PST-1
	'I haven't slept yet'		'I haven't slept yet'		
	(Dul'zon 1960: 121)		(Dul'zon 1960: 121)		

The loss of this relatively marked grammatical feature occurs in other threatened or high-contact speech varieties of Altai-Sayan Turkic, *viz.* the variety of Khakas spoken in Russian-dominated Abakan as described in Anderson (2005, chapter 4).

(15)	a.	Abakan Khakas				
		[*olɣan-nɨɣ*	*jeʃo*	*ojna-ba-an*	*pol-ɣan-nɨɣ*	*futbol*]
		boy-PL	yet < Russ.	play-NEG-PST	AUX-PST-PL	Soccer
		'The boys haven't played soccer yet.'				

b. Abakan Khakas

[*min*	*jeʃo*	*toos-pa-a-m*	*poz-ɨm-nɨŋ*	*toɣɨz-ɨm-nɨ*]
i	still/yet	finish-NEG-PST-1	self-1-GEN	work-1-ACC

'I haven't finished my work yet.'

Another possible obsolescence feature attested in our field corpus on Ös is the apparent overgeneralization of the third singular possessive form in combination with overt possessors of other person/number combinations. Such examples as the following occur relatively frequently in our corpus, with *ton-u* instead of the expected **ton-uŋ*.

(16) [*ljovuʃka, seŋ ton-u purzaj køj-ybs-qan*]
Lyovushka you:GEN coat-3 completely burn-PRF-PST
'Lyovushka, your coat has burned up completely.'[5] (GDSA 58)

3.3 Micro-variation

In the speech of a single speaker, examples of velar deletion or non-deletion, variants of case markers, distant nasal assimilation, *b* ~ *m*, etc. may be attested. This kind of micro-variation, even on the idiolectal level, is commonly attested in moribund languages like Ös.

(17) a. [*bojnum*] ~ [*mojnum*] 'neck'-1
b. [*siler-ge*] ~ [*siler-ga*] 'you.PL'-DAT

3.4 Code mixing and Russian loans

Russian influence is abundant in Ös, as evidenced in semantic and syntactic calques, loanwords, code-mixing and code-switching. The following are examples of loanwords or code-mixing. Words of Russian origin and their translations are underlined. As is commonly the case in languages, adverbs, discourse particles, and elements of various functions deriving from Russian that occur at the clause periphery freely pepper the spontaneous speech of Ös speakers. (In the data, the notations GDSA and KDH followed by page numbers refer to our unpublished field notebooks).

[5] Note the switch of harmony pattern from front to back in *køj-ybs-kan*, possibly triggered by the now lost vowel of the perfective suffix, which etymologically is [*ɯ] (Anderson 2004b) or alternatively, by the emergence of back vocalism in suffixes as a possible new default.

(18)	[*paluuq tʃiige na primer*]	'To eat fish, for example.'	(KDH 53)
	[*sæŋ je ʃo uzubaɹŋ*]	'You haven't slept yet?'	(KDH 54)
	[*a mæn amda tʃo̱*]	'What (will) I (do) now?'	(KDH 64)
	[*gabaf sæ ajduup ili qajduɹɣ*]	'Is your name Gabov, or what?'	(KDH 64)
	[*qartooʃqa kroʃɪt foʃqalarga*]	'Crush potatoes for pigs.'	(KDH 64)
	[*ol ajduubuul mæn snaju*]	'He said, I know.'	(KDH 16)
	[*nuutʃɪktermm*]	'My grandchildren.'	(KDH 30)
	[*nuutʃɪɣm*]	'My grandchild.'	(KDH 30)

The syntax of Ös also frequently displays canonically Russian word order:

(19) [*mæn bil-ɪbɪl-ɪm ineeɣ-ɪ-n sa-arga*]
I know-PRES-1 cow-3-ACC[6] milk-INF
'I know how to milk a cow.' (KDH 21)

(20) [*mæn itpek kees-tɪr-ɪm tʃidɪɣ piʒek=læŋ*]
I bread cut-PRES-1 sharp knife=INS
'I cut bread with a sharp knife.' (KDH 28)

(21) [*pis-ter kel-ge-bɪs sler-ge*]
we-PL come-PST-1PL you.PL-DAT
'We came to you.' (GDSA 25)

4. An emergent Ös orthography

We have previously reported (Anderson & Harrison 2004b, Harrison & Anderson 2003, 2005) on the emergence of a native Ös orthography, including details of how it was later abandoned under pressure from local Russians. A selection of the text of this story, which has accompanying video and metadata to be archived in the Endangered Languages Archive (www.hrelp.org/archive/), is offered below.

(22) The invention and abandonment of Ös writing
(as told by V. M. Gabov, July 2003)

[*men tajɣa-da tʃejit bol-ɣa-m*]
I taiga-LOC young be-PST-1
'When I was young out in the taiga.'

[6] Note the use of the third possessive accusative here in a non-possessed, non-definite context. Such overgeneralization of third possessive accusative forms is a characteristic feature of contemporary use. An understanding of the rules for the use of this element requires further research.

[*tajɣa-da* *siree* *tsør-ej-m*]
taiga-LOC always go-ASP-1
'I always used to go about in the taiga.'

[*no* *anda* *meeŋ* *qaɣas* *bol-ɣan*]
and (< Russ.) there 1:GEN paper be-PST
'And there I had some paper.'

[*men* *ærtæn* *jesli* *bar-za-m* *aaliʃ-ti-ige*]
I morning if (< Russ.) go-COND-1 moose-VSF-INF
'In the morning if I would go moose hunting,'

[*no* *nøøn[e]* *bol-za* *tav-arɣa* *iir-ge*
and (< Russ.) thing be-COND find-INF evening-DAT

kæl-ze-m]
come-COND-1
'And if anything happened, when I returned in the evening.'

[*tygæde* *ʃija-a-m* *po* *qaɣas-qa* *ʃija-t-ɯm*
always write-PST-1 that paper-DAT write-HAB-1

nø-ny *æt-ke-m*]
thing-ACC do-PST-1
'I would always write, I would write what I did on that paper.'

[*ærtæn øs* *saam* *andɯn* *men*
morning self (< Russ.) from.there I

pajdaq *anda* *ʃi-j* *bol-ɣa-m*]
much there write-CV AUX-PST-1
'In the morning I myself had written a lot there'

[*a* *po* *kyd-ys-ke-m* *qaɣeez-in* *qazaq* *kiʒ-ee*]
and that show-PRF-PST-1 paper-3:ACC Russian person-DAT
'and I showed it, the paper, to a Russian person'

[*a* *ol* *ajd-ɯbɯl* *nø-ny*
and s/he say-PRES thing-ACC

mɯnaar *ʃij-ip* *sal-ɣa-ŋ*]
to.here write-CV AUX-PST-2
'and he said, what's that thing you have written here?'

[*a*	*men*	*anɯ*	*no*
and	I	it-ACC	DISC (< Russ.)

aŋdɯn	*køre-p*	*sal-ɣa-m*]
3:ABL	throw-CV	AUX-PST-1

'So I threw it away.'

[*a*	*ʃij-ba-a-m*	*anzon-da*]
and (< Russ.)	write-NEG-PST-1	afterwards-EMPH

'And since then I haven't written (anything).'

[*men*	*sybyr*	*no*	*tɯŋna-p*	*pa-ɣa-m*]
I	bad	DISC (< Russ.)	hear-CV	AUX-PST-1

'After I heard that bad thing:'

[*andɯn*	*sen*	*ʃij-ip-tɯr-zɯŋ*	*sen*]
from.there	you	write-CV-EVID-2	you

"What have you written there?"

[*tadar*	*til-le*	*qaja*	*ʃij-ik-tɯr*]
Chulym	language-INS	why	write-FUT-ASSRTV

'Why would (you) write in Chulym?'

[*men*	*andɯn*	*ʃij-ba-a-m*]
I	3:ABL	write-NEG-PST-1

'After that I didn't write (any more)'

[*anzon*	*køre-p*	*sal-ga-m*	*po*	*qaɣas-tɯ*]
afterwards	throw-CV	AUX-PST-1	that	paper-ACC

'and threw away that paper.'

[*men*	*kyt-ys-ke*	*moʒet*	*siler-ge*	*pol-ɣa-m*]
I	show-PRF-INF	may be (< Russ.)	you.PL-DAT	AUX-PST-1

'Maybe I would have shown it to you.'

[*azɯ*	*tʃoɣul*	*qajda*	*tʃat*	*qɯl-ɣan*
or	NEG:COP	where	lie	AUX-PST

anda	*køre-p*	*sal-ɣa-m*]
there	throw-CV	AUX-PST-1

'But it doesn't exist anymore; it's still there where I threw it away.'

[*æmde*	*ʃij-ar*	*tʃoɣ-um*]
now	write-P/F	NEG-1

'Now I don't write.'

[*men* *siler-ga* *ajd-ɯrgam* *pol-ga-* *anɯ*]
I you-DAT tell-INF AUX-PST-1 it:ACC
'I wanted to tell you this.'

[*men* *tygeedi* *øs* *til-ni* *qɯnaa-dɯ-m*]
I always Ös language-ACC love-HAB-1
'I have always loved the Ös language.'

[*ʧurtta bɯl-mɯn* *øs* *til-ni* *saad-ɯbɯl-mɯn*]
live-PRES-1 Ös language-ACC speak-PRES-1
'I live, and speak the Ös language.'

[*øs* *til-ni* *ʧaqʃɯ* *saad-ɯbɯl-mɯn*]
Ös language-ACC beautiful speak-PRES-1
'I speak the Ös language very well.

[*ii* *idʒaa-m* *bol-gan* *toʒe* *ajt-qan-nar*]
and (< Russ.) mother-1 be-PST also (< Russ.) say-PST-PL
'And my mother also used to say'

[*ʃto* *øs-tɯŋ* *til-ni* *tadar* *til-ni*
that (< Russ.) Ös-GEN language-ACC Ös language-ACC

saat-arga *kerek*]
speak-INF NEC
'that it's necessary to speak the language of the Ös, the Tatar (Ös) language.'

[*qazaq* *qazaq-tar-ga* *tadar* *tadar-lar-ga*]
Russian Russian-PL-DAT Ös Ös-PL-DAT
'Russian is for the Russians and Ös is for the Ös.'[7]

We regret to say that some negative pressures against Ös orthography persist. These have taken the form of comments made by Russian academics and by Indigenous Siberian community representatives from Tomsk oblast' to the effect that since no

[7] In the preceding text, many of the issues we discussed in the previous sections are made manifest but only briefly commented on here. For example, the speaker shows variation in forms, the dative form of the second plural pronoun may be [*siler-ge*] or [*siler-ga*], the past tense form either shows spirantization of the velar to [ɣ] after [l] or not, there is a Khakas-like (Anderson 1998) zero allomorph of the *-p* converb [*ʧal kɯl-*], the negative present form (here with a first singular subject) shows the typical Ös (but somewhat unusual) structure */-Ar ʧoɣ-um/*, and the form of the word 'always' may appear with a final [-e] or a final -i], use of Russian discourse elements, complementizers, etc. With regard to variant pronunciations even in the speech of a single person, as mentioned above, hyper-micro-variation even on the idiolectal level is a characteristic of Ös, as it is of other moribund languages.

scholar has bestowed a writing system upon the Chulym, they cannot be thought of as having a real writing system or *pis'mennost'*. We categorically reject this absurd and colonialist view on the legitimacy of indigenously created orthographic traditions, as do our Ös consultants.

Despite these pressures from within and without the community (and his statements to the contrary in the text above which was recorded in 2003), the inventor of the Ös orthography, V. M. Gabov, continues to use it enthusiastically and prolifically to record his own stories and those told by others. Since 2003, the author has made minor emendations to his writing system. Specifically, he decided to introduce two new symbols, {н} for the velar nasal [ŋ] and {ғ} for the voiced velar fricative [ɣ]. It should be noted that this was Gabov's own decision and creation. He does not have familiarity with any other orthographic system of Turkic languages of the region (or other regions for that matter), nor did we suggest any such emendations.

(23) пис ичабла чуртабыс чаң элда
we mother-INS live-PST-1PL new village-LOC
'I lived with (my) mother in new village.' (KDH 45)

(24) пистя андағ алуу полгабыс
we-EMPH very stupid be-PST-1PL
'We were very stupid.' (KDH 45)

Our research goal as linguists is to observe the process of orthography emergence and to avoid influencing it. We believe that the organic emergence of this writing system—without interference from outsiders—will greatly increase the prestige of the language and enhance its viability. It will also provide a laboratory in which to examine the psychological reality of phenomena such as word boundaries, enclitics, vowel harmony, etc. We encouraged Gabov to write as much as possible. He has now produced more than a dozen texts of various genres, and we continue to observe the process of stabilization and standardization in his writings. We also field-tested the new orthography by asking speakers who had never seen it before to try reading it. The reception was positive, and the readers made suggestions directly to the author on how to improve it. We believe we have already witnessed the effect of the new orthography (and texts produced in it) on the relative prestige of the language.

4. 1 New texts and genres

Our field expedition resulted in a large quantity of newly written texts, some of them representing never before documented genres. We collected songs (*častuškas*), biography, humorous stories, hunting tales, etc. These were primarily written down by hand by V. M. Gabov's using his own orthography, as shown in the following sample.

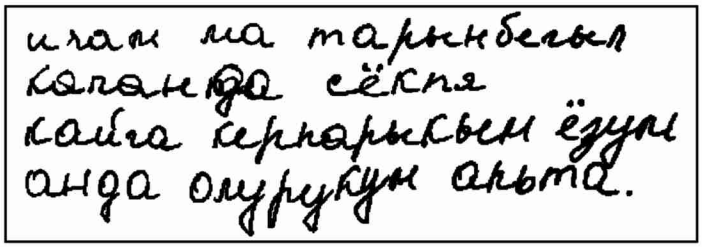


Fig. 1 Častuška "Mother don't scold" in V. M. Gabov's handwriting (July 2005)

The full text of the above is as follows:

(25) ичам ма тарынбегыл
качанда сёкпя
кайга керпарыкым ёзум
анда олурукум апьта.

We provide a close phonetic transcription here, adding morpheme boundaries, interlinear glossing and translation.

(26) [*idʒa-m* *ma* *tarɯm-be-ɣɪl*]
mother-1 I:DAT be.angry-NEG-IMP
'Mother, don't ever be angry with me.'

[*qadʒan-da* *søk-pæ*]
ever-EMPH scold-NEG
'Don't ever scold (me).'

[*qajɣa* *ker-par-ɯq-ɯm* *øz-ym*]
when grow.old-AUX-FUT-1 self-1
'When I grow old myself.'

[*anda* *olur-uq-um* *æp-te*]
there sit-FUT-1 house-LOC
'I will stay there at home.'

Another genre we collected was aphorisms, like this one written by V. M. Gabov (July 2005).

(27) [*sæn* *tʃer-dɯn* *ʃɯq-qa-ŋ*]
you earth-ABL come.out-PST-2
'You came from the earth'

[*tʃɐr-ge* *bar-ɯq-sɯŋ*]
earth-DAT go-FUT-2
'and you will return to it.'

[*saa* *køpespe* *ʃij-ɨk* *sal-gan*]
you:DAT as.many write-FUT AUX-PST
'As much as it would be written (for you).'

[*aan-ɯŋ* *artɯq* *tʃurta-b-uq-suŋ*]
that-GEN extra live-NEG-FUT-2
'You won't live any more than that.'

4.2 Some linguistic commentary on the *častuška* and aphorism texts

Before turning to some of the sociolinguistic consequences of our current work among the last speakers of the moribund Ös language, let us briefly comment on the structure of the language of these two short texts. Line one of the song text (the častuška) has a negative imperative verb, as does the second line. However, the verb in the first line has the archaic imperative marker in *-GIl* that harks all the way back to Old Turkic (Gabain 1974: 110). Thus, Ös, although divergent in many respects, shows extremely archaic features as well. That this element is optional is shown by its lack in the second sentence.

Other noteworthy features of the first two sentences include the disharmonic form for 'mother' [*idʒam*], which appears in an obligatorily possessed form, even in its vocative function as here (cf. the body part and kin term forms cited in 3.1 above). Furthermore, the dative case of the first singular pronoun shows back vocalism and loss of the final nasal (and probable vowel lengthening), e.g. [*maː*], possibly reflecting a lost **-ŋ-* (or perhaps **-ɣ-*). The final two sentences of the quatrain show Russian-type syntax with post-verbal elements. In addition, the characteristic and peculiar Ös future formation in *-(C)IK* is found in lines three and four, where C may show full assimilation (and consequent degemination, that is, its ultimate loss) or may reflect a voicing or nasal assimilation to a preceding stem (although both of the examples here are of the former type). Lastly the word 'when' in this text has the unexpected form [*qajɣa*].

The second text, the aphorism, also exhibits a range of noteworthy and characteristic contemporary Ös grammatical features. The ablative case form in Ös has a high vowel (*-DIn*) as is found in certain Old Turkic texts (Gabain 1974: 88-89) and various Middle Turkic languages as well (also in modern and Old Uighur), not a low vowel as in most other modern Altai-Sayan (and other) Turkic languages (except in lexicalized formations, cf. Tofa [*munun*] 'from here'). The second singular pronoun in line three of this text shows a similar development to that seen with the first person pronoun, namely a shift to back vocalism and an apparent loss of an original intervocalic velar, realized as [*saː*]. Also, both of the conjugationally determined

variants of the second person subject marker are seen in the verbs in the first two sentences, viz. *-ŋ* and *-s/zIŋ*, respectively.

Another verb with the uniquely Ös future form is found in line two of the aphorism. In line four its corresponding negative form is attested with a fixed high round segment (*y/u*), which itself starts a new round harmony pattern, preceded by a consonantal allomorph of the common negative morpheme *-B[A]-*. Lastly, note the future form of the verb that is attested with a following auxiliary rather than a converb form–a feature also found in the moribund Tofa language (Harrison and Anderson forthcoming-a/b, Anderson and Harrison forthcoming), where it appears to represent an innovation (in Tofa with the emergent default auxiliary [*ber-*] 'give') in the speech of terminal generation speakers.

4.3 Floodgating, memory and dormancy

Spontaneous conversation in Ös is vanishingly rare, considering there is only one household where husband and wife both speak it natively (although the man has recently become deaf, rendering conversation somewhat difficult). There is one additional household where intergenerational speakers (father and daughter) live, but truth be told, even they rarely communicate together in their ancestral tongue. Speakers who are neighbors in the same village reported that they use it only infrequently for conversation. Part of our research method includes bringing speakers together to talk. We also trained one speaker in basic interviewing techniques, and agreed with him upon a list of topics of cultural relevance and ecological validity to the Ös. Topics included social history, ethnic identity and ethnonymy, oppression of the language, discrimination against the Ös in education and other spheres, language loss, hunting, medicinal praxis, religious praxis, shamanism, indigenous technologies (fishing nets, trapping, hunting and resource management).

By arranging and recording conversations on these topics, interviews conducted in this manner yielded a rich set of registers, genres, and grammatical forms, and considerably more enthusiasm from the speakers than elicitation sessions ever could. We provide a very brief sample of one such conversation here. The following segment of conversation is led by V. M. Gabov (VG), his interlocutors are A. Badeyeva (AB) and I. Skoblin (IS).

(28) An Ös conversation about the traditional calendar (July 2005)

a. (VG) [*a* *sæn* *vot* *mæn* *sur-ga-am*]
and you DEIC (< Russ.) I ask-PAST-1
'And you, well I asked,

b. (VG) [*bis-tɪŋ* *dil-læ* *qajdɯɣ* *vot*]
we-GEN language-INS what DEIC (< Russ.)
'How is it in our language?'

c. (VG) [*qazaq* *dil-læ* *aj* *aj*]
Russian language-INS month month
''Month' in the Russian language?'

d. (IS) [*bis-tɪŋ* *aj* *qajdɯɣ* *bo-or*
we-GEN month how be-P\F
dylgy *aj-ɯ* *pol-gan*]
fox month-3 be-PST
'How is our month? There was fox month.'

c. (IS) [*qajdɯɣ* *øs-yŋ* *bil-ɪp-sɪŋ* *qajdɯɣ*
which self-2 know-PRES:EVID-2 Which
aj *bistɪŋ* *dil-læ* *aj*]
month we-GEN language-INS Month
'Which you yourself should know what months, month in our language.'

e. (AB) [*aj* *eh*]
month DISC
'Oh, month.'

f. (IS) [*qajdɯɣ* *dylgy* *aj-ɯ*]
how fox month-3
'how, fox month,'

g. (AB) [*dylgy* *aj-ɯ* *kyzyɣen* *aj-ɯ*]
fox month-3 chipmunk month-3
'fox month, chipmunk month,'

h. (AB) [*kyzyɣen* *aj-ɯ* *qarga* *aj-ɯ*]
chipmunk month-3 raven month-3
'chipmunk month, raven month'

i. (IS) [*køk* *køk* *aj-ɯ*]
green green month-3
'green, green month'

j. (AB) [*køk* *aj-ɯ* *eta* *maj*
green month-3 DEIC (< Russ.) May
mæn *an-ɯ-n* *di-dir-mæn*]
I it-3-ACC say-EVID.PST-1
'that's May, I said that, it seems.'

k. (AB) [*køk* *aj* *maj* *køk* *aj-ɯ*
green month May green month-3

æmdæ *qajnar*]
now where.to
'May is green month, green month. What else?'

l. (AB) [*partʃa* *aj-ɯ-n* *bil-ge-m*
some month-3-ACC know-PST-1

dædæ-m *ʒe* bes]
father-1 EMPH (< Russ.) all (< Russ.)
'I knew some of the months, my father, (he knew them) all.'

m. (IS) [*bil-ge-m* *men* *toʒ*]
know-PST-1 I also (< Russ.)
'Me too, I used to know.'

n. (AB) [*bil-ge-m* *ej* *mæn* *und-up* *pa-ɣa-m*]
know-PST-1 DISC I forget-CV AUX-PST-1
'I knew them. Oh, I've forgotten them.'

o. (VG) [*und-up* *pa-ɣa-ŋar* *bajdaq* *qanʲeʃna*]
forget-CV AUX-PST-2PL much of.course (< Russ.)
'You've forgotten a lot...of course.'

Dialogues such as this offer a unique glimpse into Siberian prehistory, including remnants of traditional knowledge about a now all-but-vanished belief system and worldview. They illustrate both the scope of what has already been forgotten and what may yet be documented. It is interesting to compare the now only partially remembered Ös calendar system with the fuller one reported by Falck (1786: 557). The following is a facsimile of Falck's original account, in German.

Der erſte Schnee-
fall iſt ihr Neujahr. Von demſelben zehlen ſie 12 Monathe (Ai), Ka-
rakal Ai iſt meiſtens unſer Septemb., Gariſch Ai Octobr., Kitſcha Ai
(kurzer Monath) Novemb., Ulu Ai (großer Monath) Decembr., Jel
Serta (halber Winter) Januar, Tulg Ai (Fuchsmonat) Febr. Kutſchu-
gen Ai (Adler Monath) März, Karga Ai (Krähen Monath) April,
Koi Ai (Kukuks-Monath) May, Kitſchi Schilgai (kleiner Sommer)
Jun., Ulu Schelgai Ai (großer Sommer) Jul. und Urgai Ai (lan-
ger Monath) Auguſt, weil er bis zum Schnee dauert.

As was typical of Siberian lunar calendars, the Ös had twelve months including one longer, flexible month that was synchronized with the first snowfall to reset the calendar:

"The first snow-fall is their New Year. From there they count 12 months *(ai)*, *Karakal ai* is essentially our September, *Garisch ai* October, *Kitscha ai* (short month)

November, *Ulu ai* (big month) December, *Jel Serta* (half winter) January, *Tulg ai* (fox month) February, *Kutschu-gen ai* (eagle month) March, *Karga ai* (raven month) April, *Koi ai* (cuckoo month) May, *Kitschi* Schilgai (small summer) June, *Ulu Schelgai ai* (big summer) July and *Urgai ai* (long month) August, because it lasts until the snow falls."

A few brief comments can be made on this first-ever spontaneous conversation to be published in the moribund Ös language. First, on the content of the conversation: only four months are enumerated, but of these only two coincide in more or less the same form as the Falck text, viz. [*dylgy aj(ɯ)*] and [*karɣa aj(ɯ)*]. Rather than [*koj aj*] 'cuckoo month' as listed in Falck's text, speaker AB (a user of an Upper Chulym mixed with a Middle Chulym dialect) has [*køk aj*] 'green month'–a designation known from other traditional Siberian calendrical systems. Furthermore, she used the term [*kyzygen*] 'chipmunk' not *Kutschugen* 'eagle' as listed in Falck. In this case, it could be an actual dialectal phonological variant that has triggered this reinterpretation with [z] for expected [(d)ʒ], (cf. the variants listed in (2) above in voiced variants).

Other characteristic linguistic features of contemporary Ös are attested in this short text. For example, a shortened form of the evidential is seen in line five, realized as what appears to be a (historical) converb alone. No possessive form is found after the genitive of the first plural pronoun in lines two, four, and five. Finally, a range of Russian elements is found throughout the text, often, as mentioned previously, at clause periphery, as in lines one, two, eleven, thirteen, fourteen, and sixteen.

5. Ethics and mobilization of resources

Discussions among documentary linguists increasingly focus on the ethical dimension of the process, recognizing that linguists have responsibilities beyond data collection and publication of results. Within the scope of our ELDP-funded project to document Ös, we have earmarked resources for community language revitalization goals, dissemination of information in/about Ös to a broader audience, and the creation of a local, community-accessible archiving.

Community ownership of Chulym intellectual property is a primary consideration in all our work. Digital recordings housed at Tomsk remain under the auspices of the Chulym community itself, which will retain full ownership and copyright, and will grant permission (both individually and collectively) for any scholarly use or dissemination. Because the Chulym community itself is not yet connected to the Internet, we also produce and disseminate materials in alternative media (e.g., print, audio tape, VHS video tape) so that community members who wish to see and hear the language spoken may do so readily.

While we regard a full revitalization (Hinton 2001, Wurm 1998) unlikely, we have already taken measures to support projects that will enhance awareness of the language within the community and beyond. First, we arranged meetings at which two or more native speakers who do not normally have the opportunity to converse

could do so. Separated by geography, nearly all remaining speakers of Ös lack regular opportunities to speak it. By arranging such meetings, we assisted in a process of floodgating, allowing speakers to call up from memory long dormant words and conversational skills.

Second, we have agreed at the request of members of the Ös council (and schoolteachers in the ethnically Ös-dominant village of Novoshumylovo) to produce an ABC primer and storybook using original stories collected from consultants, drawings by local children, and V. M. Gabov's orthography.

Third, to ensure long-term archiving, we have digitized and begun to transcribe and systematize legacy materials collected in the 1970s and written by hand in field notebooks. To provide a local repository readily available to the Ös community and to scholars, we purchased a state-of-the-art computer and peripherals (printer, scanner, external drive) for the Tomsk Laboratory of Languages of Siberian Peoples. This equipment, funded by the ELDP grant, reflects the consensus among funding agencies that research projects should fund not just data collection, but also mobilization of resources for indigenous communities. The Ös community, located just a few hours by car from Tomsk, has been made aware of this resource and invited to access it.

Fourth, we will archive our original field materials with the Hans Rausing Endangered Languages Project (www.hrelp.org). Any interested party may now find Ös data (for example, a Swadesh word list of the Middle and Upper Chulym dialects, with associated soundfiles) on the Internet, without waiting for future publication of our results. Much of course remains to be done in documenting Ös. What we have outlined above merely represents the first steps in this direction.

Acknowledgements

Fieldwork, research, archiving and support for Ös language revitalization are funded by a grant from the Endangered Languages Documentation Programme of The Hans Rausing Endangered Languages Project (HRELP), under the direction of Professor Peter Austin. The authors gratefully acknowledge assistance from many members of the Ös community and from scholars at the Faculty of Languages of Siberian Peoples, Tomsk State Pedagogical University as well as Professor Lars Johanson for encouraging us in this pursuit. All linguistic and cultural material presented herein should be regarded as the intellectual property of the Ös people.

References

Abdraxmanov, M. A. 1970. Teksty čulymsko-tjurkskogo jazyka. Srednečulymskij dialekt. *Jazyki i toponimija Sibiri* 2, 58-70.

Anderson, Gregory D. S. 1998. *Xakas.* (Languages of the world materials 251.) München: Lincom Europa.

Anderson, Gregory D. S. 2003. Yeniseic languages from a Siberian areal perspective. In: Vajda, Edward & Anderson, Gregory D. S. (eds.) 2003. *Studia yeniseica: Typological studies on Yeniseic.* Special issue of *Sprachtypologie und Universalienforschung*, 12-39.

Anderson, Gregory D. S. 2004a. *Auxiliary verb constructions in Altai-Sayan Turkic*. Wiesbaden: Harrassowitz.

Anderson, Gregory D. S. 2004b. The languages of Central Siberia: Introduction and overview. In: Vajda, Edward (ed.) 2004. *Languages and prehistory of Central Siberia*. Amsterdam: Benjamins. 1-122.

Anderson, Gregory D. S. 2005. *Language contact in South Central Siberia.* (Turcologica 54.) Wiesbaden: Harrassowitz.

Anderson, Gregory D. S. forthcoming. Towards a typology of the Siberian linguistic area. In: Matras, Yaron (ed.) *Linguistic Areas*.

Anderson, Gregory D. S. and K. David Harrison. 1999. *Tyvan.* München: Lincom Europa.

Anderson, Gregory D. S. and K. David Harrison. in preparation. *A Grammar of Tofa*.

Anderson, Gregory D. S. 2004a. Shaman and bear: Siberian prehistory in two Middle Chulym texts. In Vajda, Edward (ed.) 2004. *Languages and prehistory of Central Siberia.* Amsterdam: Benjamins. 179-197.

Anderson, Gregory D. S. 2004b. Na(t)ive orthographies: Recent findings from Central Siberia. Paper presented at Pacific American Modern Language Association, Portland, Oregon.

Anonymous. 1858. Jazyk čulymskix inorodtsev. *Izvestija tomskoj gubernii 1858 goda.* Tomsk.

Birjukovič, R. M. 1972. O zvuke *č* v srednečulymskom dialekte. In: *Fonetika i morfologija jazykov narodov Sibiri.* Novosibirsk: Akademija Nauk SSSR Sibirskoe Otdelenie. 40-43.

Birjukovič, R. M. 1973. O refleksax pervičnyx dolgix glasnyx v odnosložnyx slovax čulymsko-tjurkskogo jazyka. In *Proisxoždenie aborigenov Sibiri i ix jazykov.* Novosibirsk: Akademija Nauk SSSR Sibirskoe Otdelenie. 108-12.

Birjukovič, R. M. 1975a. Palatalizatsii soglasnyx v srednečulymskom dialekte čulymsko-tjurkskogo jazyka. *Jazyki i Toponimija Sibiri* 7, 18-21.

Birjukovič, R. M. 1975b. O pervichnyx dolgix glasnyx v čulymsko-tjurkskom jazyke. In: *Tjurkologicheskie issledovanii v SSSR.* Baku In *Sovetskaja tjurkologija* 6: 55-67.

Birjukovič, R. M. 1979a. *Morfologija čulymsko-tjurkskogo jazyka I.* Saratov: Saratov State University.

Birjukovič, R. M. 1979b. *Zvukovoj stroj čhulymsko-tjurkskogo jazyka.* Moskva: Akademija Nauk SSSR.

Birjukovič, R. M. 1980a. Struktura i semantika form nastojaščego vremeni indikativa čulymsko-tjurkskogo jazyka. In: *Jazyki Narodov SSSR.* Novosibirsk: Akademija Nauk SSSR Sibirskoe Otdelenie 31-43.

Birjukovič, R. M. 1980b. Semantika i forma vyraženija sposobov glagol'nogo dejstvija v čulymsko-tjurkskix jazyke. *Sovetskaja Tjurkologija* 1, 68-77.

Birjukovič, R. M. 1981a. Tipologičeskie xarakteristiki čulymsko-tiurkskoj aktsentuatsii. In: *Fonetika jazykov Sibiri i sopredel'nyx regionov.* Novosibirsk Akademija Nauk SSSR Sibirskoe Otdelenie. 110-14.

Birjukovič, R. M. 1981b. *Morfologija čulymsko-tjurkskogo jazyka* II. Saratov: Saratov State University.

Birjukovič, R. M. 1984. *Leksika čulymsko-tjurkskogo jazyka.* Saratov: Saratov State University.

Birjukovič, R. M. 1997. Čulymsko-tjurkskij jazyk. In: In Ė. R. Tenishev et al. (eds.) *Jazyki Mira : Tjurkskie Jazyki.* Moskva: Indrik. 491-497.

Dul'zon, A. P. 1952a. Očerk fonetiki čulymsko-tjurkskogo jazyka. *Učenye zapiski Tomskogo pedagogičeskogo instituta* IX: 34-63.

Dul'zon, A. P. 1952b. Čulymskie tatary i ix jazyk. *Učenye zapiski Tomskogo pedagogičeskogo instituta* IX, 76-211.

Dul'zon, A. P. 1954. Terminy rodstva i svojstva v jazykax Narymskogo kraja i Pričulym'ja. *Učenye zapiski Tomskogo pedagogičeskogo instituta* II. 108-119.

Dul'zon, A. P. 1956. O proisxoždenii čulymsko-tjurkskogo jazyka. In: *Doklady i soobščenija Instituta Jazykoznanija Akademija Nauk SSSR* 9: 99-103.

Dul'zon, A. P. 1957. Tjurkskie jazyki i dialekty zapadnoj sibiri. *Učenye zapiski Tomskogo pedagogičeskogo instituta* 17: 110-145.

Dul'zon, A. P. 1960. Lično-vremennye formy čulymsjo-tjurkskogo glagola. *Uchenye zapiski XakNIIJaLI* VIII. 101-145.

Dul'zon, A. P. 1966. Čulymsko-tjurkskii jazyk. In: N. A. Baskakov et al. (eds.) *Jazyki Narodov SSSR: Tjurkskie Jazyki.* Moskva: Nauka. 446-66.

Dul'zon, A. P. 1973. Dialekty i govory tjurkov čulyma. *Sovetskaja Tjurkologija* 2, 16-29.

Falck, Johan Peter 1786. *Beyträge zur topographischen Kenntniss des russichen Reichs.* Dritter Band. Edited by Johann Gottlieb Georgi. Sankt-Peterburg: Kayserl. Akademie der Wissenschaften. Digitized version is available from Göttingen State and University Library. Call number SUB Göttingen <4 H NAT III, 1450:1-3>, accessed July 2005 at <http://www-gdz.sub.uni-goettingen.de/cgi-bin/digbib.cgi? PPN335894348>.

Filtchenko, Andrei 2001. http://www.owlnet.rice.edu/~andreif/project/Chulym%20Turkic%20language.htm. Accessed July 2005.

Gabain, Annemarie von 1974[3]. *Alttürkische Grammatik.* Wiesbaden: Harrassowitz.

Gordon, Raymond G., Jr. (ed.) 2005[15]. *Ethnologue: Languages of the world.* Dallas, Tex.: SIL International. Online version accessed July 2005. <http://www.ethnologue.com>.

Harrison, K. David & Anderson, Gregory D. S. 2003. Middle Chulym: Theoretical aspects, recent fieldwork and current state. *Turkic Languages* 7, 245-56.

Harrison, K. David & Anderson, Gregory D. S. 2005. Na(t)ive orthographies and language endangerment. Paper presented at Linguistic Society of America meeting, Oakland, California.

Harrison, K. David & Anderson, Gregory D. S. forthcoming-a. Tofa language change among last-generation speakers. To appear in Filtchenko, A. Yu. et al. (eds.). *XXIV Dul'zonovskie Čtenija.* Tomsk: TGPU.

Harrison, K. David & Anderson, Gregory D. S. forthcoming-b. Language change and terminal generation speakers: The case of Tofa. To appear in Dwyer, Arienne & Harrison, K. David & Rood, David (eds.) *A world of many voices: Lessons from documented endangered languages.*

Himmelman, Nicholas P. 1998. Documentary and descriptive linguistics. *Linguistics* 36, 161-195.

Hinton, Leanne 2001. Language revitalization: An overview. In: Hinton, Leanne & Hale, Ken (eds.) *The Green Book of Language Revitalization in Practice.* New York: Academic Press. 3-18.

Il'jašenko, Irina A. & Kim, Aleksandra A. et al. (compilers) 1998. *Katalog polevyx zapisej jazykov narodov Sibiri (fondy Laboratorii jazykov narodov Severa Tomskogo gosudarstvennogo pedagogičeskogo universiteta).* Tomsk: Tomskij Gosudarstvennyj Pedagogicheskij Universitet.

Katanov, N. F. 1903. *Opyt issledovanija urjangxajskogo jazyka s ukazaniem glavnejšix rodstvennyx otnošenij ego k drugim jazykam tjurkskogo kornja.* Kazan': Kazan' University.

Klaproth, Julius Heinrich von 1823. *Asia Polyglotta.* Paris: A. Schubart.

Krivonogov, V. P. 1998. *Etničeskie protsessy u maločislennyx narodov sredney sibiri.* Krasnojarsk: Krasnojarsk State Pedagogical University Press.

Lvova, E. L. 1972. *Materialy po izučeniu šamanizma u korennogo naselenia srednogo čulyma: Nekotorie voprosy istorii Sibirii.* Tomsk: Tomskij Gosudarstvennyj Universitet.

Lvova, E. L. 1978. On Shamanism of the Chulym Turks. Diószegi, Vilmos & Hoppál, Mihály (eds.) *Shamanism in Siberia* 237-244. Budapest: Akadémiai Kiadó.

Malov, S. E. 1909. Otčet o komandirovke studenta Vostočnogo Fakul'teta S. E. Malova. *Izvestija russkogo komiteta dlja izučenija Srednej i Vostočnoj Azii*, Aprel' 1909: 9.

Messerschmidt, D. G. 1723. *Reisejournal.* Bd. II Arkhiv Akademii Nauk 98/I, 2.

Novokšonov, V. 1995. *Čulymtsy—aborigeny raiona.* Tegul'det: Kraevedečeskij muzei.

Pallas, Peter S. 1787-9. *Linguarum totis orbis vocabularia comparativa.* St. Petersburg.

Pritsak, Omeljan 1959. Das Abakan-Tatarische, das Schorische und das Chulymtürkische. In: Deny, Jean & Grønbech, Kaare & Scheel, Helmuth & Togan, Zeki Velidi (eds.) 1959. *Philologicae turcicae fundamenta* 1. Aquis Mattiacis: Steiner. 622-629.

Radlov, Vasilij V. 1868. *Obraztsy narodnoj literatury tjurkskix plemen.* II. St. Peterburg.

Radlov, Vasilij V. 1882 *Phonetik der nördlichen Türksprachen.* Leipzig.

Radlov, Vasilij V. 1882-1899. *Opyt slovarja tjurkskix narečij.* I/II. St. Petersburg.

Serebrennikov, B. A. and R. M. Birjukovič 1984. Iz istorii nekotoryx glagolnyx vremen čulymsko-tjurkskogo jazyka. *Sovetskaja Tjurkologija* 6, 3-14.

Various 1972. Middle Chulym language materials. Unpublished notebooks.

Woodbury, Tony 2003. Defining documentary linguistics. In: Austin, Peter (ed.) 2003. *Language Documentation and Description 1*: 35-51. London: SOAS.

Wurm, Stephen A. 1998. Methods of language maintenance and revival, with selected cases of language endangerment in the world. In: Matsumara, Kazuto (ed.) 1998. *Studies in endangered languages.* Tokyo: Hituzi Syobo. 191-211.

Abbreviations

1	First person	HAB	Habitual
2	Second person	IMP	Imperative
3	Third person	INF	Infinitive
ABL	Ablative	INS	Instrumental
ACC	Accusative	LOC	Locative
ASP	Aspectual	NEC	Necessitive mood
ASSRTV	Assertive	NEG	Negative
AUX	Auxiliary	P\F	Present/Future
COND	Conditional	PL	Plural
COP	Copula	PL:POSS	Plural Possessor
CV	Converb	PRES	Present
DAT	Dative	PRF	Perfective
DEIC	Deictic	PST	Past
DISC	Discourse marker	RUSS	Russian (loan)word
EMPH	Emphatic	REC	Recent
EVID	Evidential	UNACMPL	Unaccomplished
FUT	Future	VSF	Verb stem formant
GEN	Genitive		

Indirective sentence types

Lars Johanson

Johanson, Lars 2006. Indirective sentence types. *Turkic Languages* 10, 73-89.

The paper deals, mainly on the basis of Turkish data, with levels of grammatical analysis of indirective (evidential) sentences, the relationship between forms and functions, and between sentence types, speech-act types and registers. It is claimed that indirective markers of the broad Turkish type permit speakers to be vague about the source of evidence for the propositional content.

Lars Johanson, Seminar für Orientkunde, Johannes Gutenberg-Universität Mainz, D-55099 Mainz, Germany. E-mail: johanson@uni-mainz.de

Introduction

The present paper will briefly discuss the interaction of various sentence types with markers of indirectivity, a specific type of evidentiality. It will deal with levels of grammatical analysis of evidential sentences, and the relationship between forms and functions, between sentence types and speech-act types. The discussion will mainly be illustrated with examples from Turkish.

Evidential modalities and categories

The evidential sentences dealt with here represent a certain kind of modal sentences, and are, as such, opposed to non-modal or 'indicative' sentences. Evidentials may be taken to belong to the attitudinal modalities, which express attitudes towards the content of the proposition. It is, however, important to note that they do not primarily express evaluation or assessment of the truth value of the utterance. Evidential categories, as defined here, state the existence of a source of evidence for a propositional content (cf. Aikhenvald 2003: 1).

The term 'evidential category' will be reserved for cases in which evidentiality finds unique expressions in a language. Evidential categories have the indication of evidence as their primary meaning, not only as a pragmatic inference. Languages lacking them may use various evidentiality strategies based on optional expressions. The linguist has to decide whether and by what means a given language codes evidentiality. Evidential modalities may be grammatically expressed by markers such as verbal inflectional affixes or particles. The realization of evidential sentences comprises two components, one proposition operator expressing the evidential meaning, and another component representing the propositional content. Evidential operators do not contribute to the description of the narrated event, but just add an attitudinal

specification. They represent meanings grounded in a conceptualizer's subjective perspective and awareness with respect to the content of the proposition.

Though evidentials do convey epistemic notions and may express the speaker's cognitive attitude towards the propositional content, they are not attitudinal or 'subjective' modalities in the evaluative or volitional sense. As already mentioned, they do not primarily concern the assessment of the truth of the propositional content, i.e. the personal opinion that the content is more or less certain, probable or possible. They do not express volition, i.e. the wish or hope that the content is or will be realized. Evidentiality will thus not be used here in the sense of marking the speaker's 'attitude towards his/her knowledge of reality' (Chafe 1986: 271). Expressions of epistemic stance in the sense of dubitatives, presumptives, assertives, etc., may include comments on the source of information. But this does not conversely imply that evidential sentences, which primarily state the existence of a source of evidence, also express the reliability of the information in terms of certainty and doubt or the strength of commitment to the content.

Various kinds of evidential meanings may be expressed grammatically. The languages of the world display a broad variety of simpler and more complex types of evidentiality systems. Traditional treatments are mostly based on distinctions between information based on first- or second-hand evidence, witnessed or unwitnessed evidence, visual or auditory evidence, inference, etc. It is often difficult to judge on the relationship of these notions to each other and to use them for cross-linguistic descriptions. Many of them can probably be ordered on a few dimensions, if sufficiently abstract definitions are applied. It is clear, however, that the traditional notions do not cover the whole range of possible evidential meanings.

Indirectivity

The type of evidentiality dealt with in the present paper will be referred to as indirectivity (see, e.g., Johanson 1998a, 2000b, 2003). In languages possessing indirective categories, the propositional content is presented in an indirect way. A content marked for indirectivity is characterized by reference to its reception by a conscious subject. The result is two-layered information: 'it is stated that the narrated event is acknowledged by a recipient'. The recipient, who is or becomes aware of X, may be the speaker as a participant of the speech event, or a participant of the narrated event, e.g. a protagonist in a narrative. The core meaning may be periphrased in a stereotype way as 'it appears to the recipient that X is the case'.

Specification of the source of information—the way in which the event is acknowledged by the recipient—is not criterial for indirectivity as such. The reception may be realized through hearsay, inference or perception. The perceptive uses cannot, of course, be subsumed under 'non-first-hand knowledge'. Indirectives thus do not fit into evidential schemes distinguishing between the speaker's non-first-hand and first-hand information. Their primary task is not to express the foreign origin of the speaker's knowledge. Indirective specifications are possible in all persons. A

definition of indirectivity as the expression of the speaker's non-first-hand information would exclude the first person except in cases such as lack of awareness, consciousness or control due to inattention, sleep, drunkenness, coma, etc. However, with a definition based on the presentation of the event by reference to its reception by a conscious subject, it is by no means contradictory to use indirectives with first-person referents who are aware of the event.

Indirectivity is a characteristic feature of Turkic, though not unique to this family. The crucial element of indirectivity, the presentation of an event by reference to its reception by a conscious subject, may be basic to a number of other evidentiality systems, and even qualify as a crosslinguistic definition of evidentials (Comrie 2000: 1).

Some languages with which Turkic has been in close contact display conspicuously similar traits in their evidentiality systems (see Johanson 1998, 2002a: 99-100, 144-145). Features of Turkic evidential systems have proven highly attractive in language contact situations and have been copied into non-Turkic languages of Southwestern and Central Asia, Southeastern and Northeastern Europe, etc. (On the concept of code-copying, see Johanson 1999b, 2000a: 8-19, 2002b.) Indirective categories similar to the Turkic ones typically appear in contact areas such as the Balkans, Anatolia, Caucasus, Central Asia and the Volga region, e.g. in Bulgarian, Macedonian, Albanian, Kurdish, Armenian, Georgian, Tajik and Eastern Finno-Ugric. For example, Northern Tajik has developed a comprehensive evidential system on the Southeastern Turkic model as represented by Uzbek. Indirective functions have been copied onto postterminals of the 'perfect' type, and onto related copulas (see below). The Balkan Romance languages Aromanian, Megleno-Romanian and Daco-Romanian also display items signaling indirectivity. Megleno-Romanian utilizes an inverted perfect that contains an auxiliary verb meaning 'have' occurring after the past participle. This construction is formally similar to the Albanian so-called admirative. In the Albanian evidentality opposition, the admirative is a marked indirective with reportive, inferential and perceptive readings.

Sentence types

The use of indirectives depends on properties of certain basic sentence types that can be distinguished across languages. To cite some Turkish examples, declarative sentences are primarily used for representative speech acts such as stating, asserting and claiming, e.g. Turkish *Ali geldi* <A. come-PAST> 'Ali came / has come', *Ali geliyor* <A. come-INTRA> 'Ali comes / is coming'. Interrogative sentences are typically used for eliciting information. Polar interrogative sentences involve the same proposition as the corresponding declarative sentences, e.g. *Ali geldi mi?* <A. come-PAST Q> 'Did Ali come? / Has Ali come?'. There are alternative questions such as *Ali geldi mi gitti mi?* <A. come-PAST Q leave-PAST Q> 'Did Ali come or leave? / Has Ali come or left?', and there are constituent interrogative sentences which elicit answers that provide the information specified by an interrogative word they contain, e.g. *Kim*

geldi <who came-PAST> 'Who came / has come?'. Imperative sentences will be left out, since they do not interact with the kind of indirective operators dealt with here.

Each sentence type has a wide range of uses. The connections between them and their illocutionary functions are often complex. Declarative sentences do not only have descriptive functions, but may also be used for speech acts of accusing, promising, complaining, etc., e.g. Turkish *söz veriyorum* <word.give-INTRA-1.P.SG> 'I promise'. Interrogative sentences are, as is well known, not only used for questions, but also, for example, for more or less polite requests, e.g. *Can you close the window?* Speech act distinctions result from interaction of the sentence types with other structural properties and contextual conditions.

Turkish markers

Some examples from Turkish will be cited in order to illustrate the issue of sources. In Turkish, indirectivity is expressed by the inflectional verbal suffix *-miş* and the indirective copula particle *imiş,* originally a form of a verb 'to be' (see Johanson 2000b, 2003, Csató 2000a). The copula particle is mostly used in its suffixed form *-(y)miş*.

The simple inflectional marker *-miş* is suffixed to verbal stems and capable of carrying high pitch. It is a rather stable marker of indirectivity, mostly with past time reference, e.g. *gülmüş* <laugh-MIŞ> 'has / had evidently laughed'. It is homonymous with the postterminal participle suffix *-miş*, which does not, however, express indirectivity.

The copula particle *imiş* cannot be added to verbal stems, but may follow nominals, including nominal stems of the verb. It is not capable of carrying high pitch. It is a stable marker of indirectivity with reportive, inferential and perceptive uses. Since *imiş* normally exhibits suffixed allomorphs in the shape of *-(y)miş*, some of its written realizations coincide with those of the inflectional marker *-miş,* e.g. *gülmüş* [rose-IMIŞ] 'it is / was evidently a rose'. This copula particle may, for example, form indirective intraterminals, i.e. presents and imperfects, e.g. *geliyormuş* <come-INTRA-IMIŞ> 'is / was evidently arriving', prospectives such as *gelecekmiş* <come-PROSP-IMIŞ> 'will / would evidently arrive' and unequivocally indirective postterminals such as *gelmişmiş* <come-POST-IMIŞ> 'has / had reportedly arrived'. Indirective copula particles are incompatible with the simple past marker *-di* and the copula form *idi* 'was', e.g. Turkish **-diymiş*, **-(y)diymiş*.

The deceptive similarity of certain allomorphs of the two markers *-mIş* and *imiş* has led linguists to confuse them, referring to both as 'the suffix *-miş*', allegedly attachable to both verbal and nominal stems, or to speak of an 'evidential perfect' that is taken to include both markers.

Oppositions

Turkic languages display paradigmatic contrasts between marked indirectives and their unmarked counterparts. The indirective terms of the Turkish oppositions, *gelmiş*

<come-MIŞ> 'has obviously come / obviously came' and *geliyormuş* <come-INTRA-IMIŞ> 'is / was obviously coming, obviously comes' etc. are the marked ones, systematically opposed to the corresponding unmarked terms *geldi* <come-PAST> 'has come / came', *geliyor* <come-INTRA> 'is coming / comes', etc. The marked members of the oppositions represent the positive value of indirectivity. The unmarked members represent non-indirectivity, implying negation of, or neutrality towards, the positive notion on a common basis of comparison.

Thus the Turkish inflectional suffix *-miş,* which mainly displays indirective uses, is opposed to an unmarked simple direct past in *-di,* which negates the notion of indirectivity but also displays neutral uses. Since Turkish lacks a competing pure postterminal (see below), this item covers both 'perfect' and 'preterite' functions, e.g. *geldi* <come-PAST> 'has come / came'.

Sources

As mentioned, specification of the way the event is acknowledged by the recipient is not criterial for indirectivity as such. The reception may be realized through hearsay, inference, or perception. All these readings can be translated by *evidently*, *obviously*, etc. The interpretations may also be intertwined or overlapping.

A quotative or reportive interpretation means that the source is hearsay: the recipient reports what (s)he has heard from someone else. The narrated event or its effect is reported to the recipient. The basis of knowledge is a foreign source. We may translate: 'the recipient is told / has been told that X is the case'. English translation equivalents include *reportedly, allegedly, as they say / said*, etc. Example: Turkish *Bakan hastaymış* <minister sick-IMIŞ> 'The minister is / was reportedly sick', typically referring to a recipient who is informed about the sickness. Note that the recipient's source is not necessarily the original source; the source quoted may have the information from another source.

An inferential or experiential interpretation means that the source is pure reflection, logical conclusion based on indirect evidence or previous personal experience. The narrated event or its effect is inferred by the recipient. We may translate: 'the recipient concludes / has concluded that X is the case'. English translation equivalents include *as far as R understands / understood*, etc. Example: Turkish *Uyumuşum* <sleep-MIŞ-1.P.SG> 'I have / had obviously slept', typically referring to a recipient just waking up.

The source may, however, also be first-hand knowledge based on direct sensory perception: the narrated event or its effect is perceived by the recipient. The source may be what the recipient sees, hears, tastes, smells or feels. We may translate: 'the recipient perceives / has perceived that X is the case'. English translation equivalents include *it appears / appeared that, it turns / turned out that, as R can / could see, hear*, etc. Examples: Turkish *Burnun kanamış* <nose-POSS.2.P.SG bleed-MIŞ> 'Oh, your nose has bled [as I see]', *Elbisem leke olmuş* <dress-POSS.1.P.SG stain become-MIŞ> 'It appears that my dress is stained', *Yanağına domates bulaşmış* <cheek-

POSS.2.P.SG-DAT tomato smear-MIŞ> 'You've got tomato on your cheek', *Ali baktı, yıldızlar silinmiş gitmiş* <A. look-PAST star-PLUR expire-MIŞ go-MIŞ> 'Ali looked and saw: the stars had faded out completely', *Ali iyi çalıyormuş* <A. good play-INTRA-IMIŞ> 'Ali is / was [as I hear / heard] playing well', typically referring to a recipient listening to a musical performance. This last type can also be used when the recipient of second-hand knowledge agrees with the source: 'I observe that the source I got the information from was right'.

As stated above, the perceptive uses cannot possibly be subsumed under 'non-first-hand knowledge'. They express that the event or its effect turns out to be the case, becomes manifest, visible, or apprehended through one of the senses and thus open to the recipient's mind. These uses cannot be derived from reportive or inferential meanings.

Some Turkic languages with more elaborated evidentiality systems, e.g. Turkmen and Uyghur, distinguish 'reportive' and 'non-reportive' (inferential / perceptive) evidentiality. None of them distinguishes other types of sources in a systematic way, e.g. visual information versus other kinds of sensory information.

Aspect and tense

When analyzing the aspecto-temporal systems of indirective and non-indirective sentence types, we will use the viewpoint notions of intraterminality and postterminality. These two aspectual ways of envisaging events with respect to their limits are grammaticalized in Turkic as well as in many other languages (Johanson 1971, 1994, 1996a, 2000a, 2001).

The intraterminal perspective, +INTRA, envisages, at a given aspectual vantage point, an event within its limits, *intra terminos*, i.e. after its beginning and before its end. Non-intraterminality, −INTRA, disregards this view. Intraterminality is an introspective manner of presentation allowing to perceive an event from inside, and not in its totality. It is typical of progressives, present tenses and imperfects. Intraterminals are marked imperfectives expressing 'the state of doing'. Combinations with tenses yield expressions such as −PAST (+INTRA) 'intraterminal-in-present', e.g. English *writes*, *is writing*, and +PAST (+INTRA) 'intraterminal-in-past', e.g. *was writing*.

The postterminal perspective, +POST, envisages, at a given aspectual vantage point, an event after the transgression of its decisive limit, *post terminum*, i.e. after its beginning or its end. Non-postterminality, −POST, disregards this view. Postterminality is typical of perfects and resultatives. Postterminals express 'the state of having done' and can thus refer to preexisting events in an indirect way. The event is totally or partly absent from the view, but it is still relevant at the vantage point, possibly through observable results or traces. Combinations with tenses yield expressions such as −PAST (+POST) 'postterminal-in-present', e.g. English *has written*, and +PAST (+POST) 'postterminal-in-past', e.g. *had written*.

Indirective markers often go back to postterminals. The type of indirectives found in Turkic languages are closely connected with postterminality, the view of the event

after the transgression of its relevant limit. Postterminals can refer in an indirect way to an event that has already, entirely or partly, disappeared from the range of vision, but is still relevant in one way or another. It may extend right up to the orientation point, have left observable traces or effects or other forms of present knowledge of it at this point. This indirect perspective creates an element of distance. It may suggest that information becomes available *post factum*, that a preexisting state is discovered, etc.

Postterminals may exhibit both aspectual and more or less vague evidential uses. In many languages, perfects, expressing past events with present relevance, tend towards indirective readings (Johanson 1971: 280-292, 2000a: 121-123), without being stable evidentials. They may acquire additional secondary evidential meanings through pragmatic inference and use them for evidential strategies. For example, Persian *kærdæ-æst* 'has done' and Eastern Armenian *gnac-el ē* 'has gone' are ambiguous with respect to indirectivity. Genuine indirective categories may develop through further grammaticalization, as a semantic extension in the sense of conventionalized implicatures.

In many cases it may be difficult to decide whether a given perfect-like item is a pure postterminal or an indirective postterminal. The Balkan Slavic oppositions with respect to direct vs. indirect experience are claimed to obtain between confirmative vs. nonconfirmative terms. The basic evidential oppositions of Bulgarian and Macedonian are described as relying on marked 'confirmative' items indicating unequivocal and direct assertion, whereas the corresponding unmarked items convey indirective meanings in particular contexts. According to Friedman, who admits that the label 'evidential' for these oppositions is infelicitous (2000: 357), the choice between confirmative vs. nonconfirmative terms is determined by the speaker's attitude toward the information, confirmativity expressing the speaker's vouching for the truth of the information. The Macedonian past tense formed with the *l*-participle is claimed to be an unmarked past rather than one marked for nonconfirmation. Friedman takes the Albanian 'admirative' to be "a marked nonconfirmative, expressing the three basic types of nonconfirmative meaning described for the Balkan Slavic *l*-form: reportedness, dubitativity, and admirativity" (2000: 343).

The Bulgarian copulaless *l*-periphrasis has been judged to be an unmarked past that has nonconfirmativity as its main contextual variant meaning. Levin-Steinmann suggests that it is still an item whose invariant value is based on 'Zustandskonstatierung': "eine auf das Subjekt bezogene, zu einem vergangenen Moment bereits eingesetzte Handlung als eine ihm bis zu einem bestimmten Zeitpunkt anhaftende Eigenschaft". This is clearly a viewpoint aspectual notion in the sense of high-focal postterminality (resultativity). Though the item is said to have contextual readings such as renarrative, conclusive, admirative, etc., it is not, if the analysis is correct, a fullfledged evidential.

Regarding the question whether the Balkan Slavic systems based on marked confirmatives have emerged through areal contact with Turkic systems based on marked indirectives, Comrie considers the possibility that the semantic distinction can be

reduced to a single prototype with markedness inversion: "one of the systems, almost certainly the Balkan one, has undergone a shift whereby an old indirective was reinterpreted as unmarked, with the originally unmarked non-indirective then becoming a marked confirmative" (2000: 8).

Turkic systems under foreign influence

Despite certain system differences, almost all known older and recent stages of Turkic possess grammatical means of expressing indirectivity. Indirectives play a central part in almost all Turkic languages. However, due to strong influence from Indo-European languages such as Persian, Greek and Slavic, a few languages and dialects lack them. Karaim, spoken in Lithuania, has lost them under Slavic and Lithuanian impact (Csató 2000b). In the Turkish dialects of the Trabzon province on the east Black Sea coast, the inflectional suffix *-miş* exhibits an atypical behaviour under the impact of Greek (Brendemoen 1997).

Some dialects only exhibit evidentiality strategies. The status of the Azerbaijanian inflectional marker *-miš*, which forms a common mixed paradigm with *-(i)b*, is different from that of Turkish *-miş*. Thus *-miš / -(i)b* represents an ambivalent type: a postterminal past ('perfect') with secondary indirective readings, e.g. *gelmišem* <come-POST-1.P.SG> 'I have arrived', *yazïbsïn* <write-POST-2.P.SG> 'you have written'. The unmarked term *-di* tends towards preterite rather than perfect functions, e.g. *geldi* <come-PAST> 'came' versus *gelib* <come-POST> 'has come'. This behaviour is most probably due to Persian influence, e.g. *yapïb* 'has done', cf. Persian *karde(-ast)*. Also in the non-Oghuz language Khalaj, spoken in central Iran, *-miš* does not display any indirective meaning (Kıral 2000). In some East Anatolian dialects that are close to Azeri, the finite inflectional marker *-miş,* which conveys indirectivity in Standard Turkish, tends to express pure postterminality in the sense of a perfect, e.g. *uyumuşum* <sleep-MIŞ-1.P.SG> 'I have slept'. In Cypriot dialects, the simple finite -MIŞ is absent as an indirective suffix and rather used as a postterminal (perfect) marker, like in the Anatolian dialects mentioned, e.g. *Şimdi gelmişem* <now come-MIŞ-1.P.SG> 'I have arrived now'. However, the particle *miş* (< *imiş*), which is placed in front of the predicate core, expresses indirectivity, e.g. *Ali miş gelecek* <A. MIŞ come-PROSP> 'Ali will obviously come'; cf. Standard Turkish *Ali gelecekmiş* <A. come-PROSP-IMIŞ>.

Reduced aspect-tense inventories

An interesting point is that indirective sentences are characterized by a reduced inventory of aspectotemporal markers. In the simple past we find the following opposition in Turkish: *Ali geldi* <A. come-PAST> 'Ali came / has come' vs. *Ali gelmiş* <A. come-MIŞ> 'It appears that Ali came / has come / had come'. The Turkish indirectivity distinction is not, as is sometimes claimed, restricted to the past tense. Other combinations utilize the indirective copula particle *imiş*, which is temporally indifferent, i.e. ambiguous between past and present time reference. Examples: *Ali geliyor* <A.

come-INTRA> 'Ali comes/is coming', *Ali geliyordu* <A. come-INTRA-PAST> 'Ali was coming' vs. *Ali geliyormuş* <A. come-INTRA-IMIŞ> 'It appears that Ali comes / is coming / was coming'.

The Turkish inflectional suffix *-miš* is also temporally neutral between past and non-past. Thus, *Ali gelmiş* <A. come-MIŞ> may also refer to the past tense: 'Ali had apparently come'. The addition of *idi* or *-(y)dI* 'was' would remove the indirective meaning and create a pluperfect in the sense of +PAST (+POST) 'postterminal-in-past': *Ali gelmişti* <A. come-MIŞ-PAST> 'Ali had come'. An example from a novel by Adalet Ağaoğlu: *Kaldırımda, elim hâlâ elindeydi* <pavement-LOC hand-POSS.1.P.SG. still hand-POSS.3.P.SG.>. *Koyvermemişiz* <let.go-NEG-MIŞ-1.P.PLUR> 'On the pavement my hand was still in his hand. [I realized that] we hadn't let go'.

Some other Turkic languages exhibit very similar systems, which will not be dealt with in detail here. Yakut, the northern- and easternmost Turkic language, spoken at the opposite extreme of the Turkic world, has an indirective system that is similar to the Turkish one. The inflectional marker *-bit* conveys reportive, inferential and perceptive nuances, e.g. *kelbit* <come-BIT> 'has obviously arrived', *barbït* <go-BIT> 'has evidently gone'. The temporally indifferent indirective particle *ebit* is similar to Turkish *imiş* and allows similar combinations with intraterminals, postterminals, etc., e.g. *turar ebit* [stand-INTRA-EBIT] 'evidently stands / stood', *kelbit ebit* [come-POST-EBIT] 'has / had evidently arrived' (Buder 1989).

Many Turkic languages utilize the indirective copula particle *eken*, which is also originally a form of a verb 'to be' and equally ambiguous between past and present time reference. Temporal ambiguity is often observed in other languages that possess indirective operators. The Turkish expression of past tense in the opposition *-di* vs. *-miş* corresponds to Bulgarian *čete* 'read' vs. *čel* 'apparently read' or Tajik *kærd* 'did / has done' vs. *kærdæ-æst* 'apparently did/has done'. But Bulgarian intraterminal, i.e. present and imperfect, forms display the same temporal indifference as their Turkish counterparts, e.g. *Ali čete* <A. read-INTRA> 'Ali reads/is reading', *Ali četese* (<A. read-INTRA-PAST> 'Ali was reading' vs. *Ali četjal* <A. read-INTRA-INDIR> 'It appears that Ali is / was reading'.

The Turkish indirective copula particle has close equivalents in Bulgarian *bil* (Johanson 1996b), Tajik *-æ æst*, Armenian *eɣel*, all of which are also forms of 'to be'. These elements show similar temporal neutralizations, e.g. Persian *Æli mekunæd* <A. do-INTRA-NON-PAST> 'Ali does / is doing', *Ali mekærd* <A. do-INTRA.PAST>'Ali was doing' vs. *Æli mekærdæ-æst* <A. do-INTRA> 'It appears that Ali is / was doing'. *Æli kærdæ-æst* <A. do-POST> 'Ali has done', *Æli kærdæ bud* <A. do-POST-PAST> 'Ali had done' vs. *Æli kærdæ budæ-æst* <A. do-POST-INDIR> 'It appears that Ali has / had done'. Bulgarian examples: *Ali čel e* <A. read-post> 'Ali has read', *Ali čel beše* <A. read POST.PAST>'Ali had read' vs. *Ali čel bil* <A. read-POST-INDIR> 'It appears that Ali has / had read'.

Turkish indirective sentence types

Turkic evidentials are limited to main clauses with a stated, contradictable content. Oppositions with respect to indirectivity are not possible in embedded clauses. Turkic languages lack grammatical evidentiality oppositions in embedded clauses: they have no indirective subjunctors, i.e. complementizers, adverbializers and relativizers. It is thus not possible to mark embedded clauses for indirectivity in the sense of 'that X is obvious', 'X being obvious', etc.

Cases such as the following illustrate the purely postterminal value ('having done') of the participle in *-miş*: *Ali ölmüştü* <A. die-POST-PAST-3.P.SG> 'Ali had died' (pluperfect), *Ali kazanmış gibi güldü* <A. win-POST like laugh-PAST> 'Ali laughed as if he had won', *Ali'yi ölmüş sandım* <A.-ACC die-POST believe-PAST-1.P.SG> 'I thought Ali had died'.

One feature of negative sentences is that indirectivity is not within the scope of the negation. It is not the reception by a conscious subject that is negated (*'it does not appear that X is the case'), but the narrated event itself (*'it appears that X is not the case '), e.g. *Ali gelmemiş* <A. come-NEG-MIŞ> 'It appears that Ali did not come / has not come / had not come'.

Indirectives occur in declarative sentences used for representative speech acts which make claims and inform about situations: *Ali gelmiş* <A. come-MIŞ> 'Ali has, as it appears, come'. They also combine with interrogative sentences, i.e. polar interrogative sentences, e.g. *Ali gelmiş mi?* <A. come-MIŞ Q> 'Has Ali, as it appears, come?', alternative questions, e.g. *Ali gelmiş mi gitmiş mi?* <A. come-MIŞ Q leave-MIŞ Q> 'Has Ali, as it appears, come or left?', and constituent interrogative sentences, e.g. *Kim gelmiş* <who come-MIŞ> 'Who has, as it appears, come?'. The interrogative operator takes the indirective operator in its scope. Speakers typically choose indirective questions if they anticipate indirective responses. The meaning 'it appears that ...' of the corresponding declarative sentence is changed into 'does it appear that ...?'. The answer is expected to express a certain reservation, not stating what is the case, but what appears to be the case, e.g. *Ali delirmiş mi?* <A. go.mad-MIŞ Q> 'Has Ali, as it seems, gone mad?'. Indirectives may also be used in questions asked on behalf of someone else. They do not express the speaker's assumption about the addressee's source of information.

Indirectivity is mostly incompatible with sentences used for non-representative speech acts, e.g. acts of accusing, promising, complaining, etc. The strength of a performative utterance such as 'I (hereby) promise', which only refers to the speaker, would be neutralized by an indirective sentence such as Turkish *söz veriyormuşum* <word give-INTRA-IMIŞ-1.P.SG> '*It appears that I hereby promise'.

Turkic indirective markers do not combine with imperatives such as *Gel!* 'come!', which only express the speaker's own command and not some other person's will. On the other hand, indirective markers may occur with subjective moods such as optative, necessitative and debitive, which evaluate the actuality of an event in terms of moral or similar norms: 'it appears that X should be the case'. An optative

marker such as Turkish *-sin* (often erroneously referred to as an imperative) may combine with indirectivity, e.g. *Ali gelsinmiş* <A. come-OPT.3.P.SG.-IMIŞ> 'Ali is apparently requested to come'. In a sentence such as *Gitmeliymişim* <go-DEBITIVE-IMIŞ-1.P.SG> 'It appears / appeared that I ought to go', the indirective takes the objective deontic mood of the debitive marker *-meli* in its scope. The utterances in question do not refer to the speaker's will, but to some other person's will: 'it appears that (according to somebody's opinion) X should be the case'. Combinations with conditionals are also possible, e.g. *Ali gelseymiş* <A. come-COND-IMIŞ> 'if [as reported] Ali would come'.

Contextual meanings

As stated above, evidential categories have the indication of evidence as their primary meaning, and not only as a pragmatic inference. On the other hand, genuine evidential categories such as indirectives may acquire specific uses in discourse and gain additional meanings and extensions in certain speech-acts. The linguist faces the problem to determine how the coded values are used, on the pragmatic side, by discourse participants. Indirectives may, both in declarative and interrogative sentences, suggest nuances of 'subjectivity' or 'stance' that reflect the speaker's cognitive and emotional affective attitude. Such functions are due to interaction with contextual factors and pragmatic properties in specific speech-acts.

The motives for using Turkic indirectives may vary. They may get various contextual interpretations and display various pragmatic extensions of their central meaning. Their primary task is not to express doubt versus certainty. They are neither dubitatives nor presumptives. Indirective sentences do not necessarily signal that the content ranks lower in reliability. Still they may be used as downtoners in strategies of hedging, suggesting the speaker's distance to the information. One function is to disclaim responsibility for the validity by transferring it to a source ('I am just relating, not asserting').

Indirectives may evoke the impression that the recipient does not or did not witness the event, perceive it, or participate in it consciously; that (s)he is or was not present at the event, not in control of it, not directly involved in it. However, despite the indirect way of presentation, these meanings are not signalled explicitly. The indirectively marked event may indeed be apprehended by the recipient through the senses, consciously taken part in, etc. Lack of participation or control is limited to certain contexts and is not the common core meaning of indirectives. The source of information may be direct evidence, personal, even visually obtained knowledge. A sentence such as Turkish *Ahmet gelmiş* [A. come-MIŞ] 'Ahmed has / had [as I note / noted] arrived' can be felicitously uttered by a speaker who has witnessed the arrival in reality. The indirective statement just expresses the conscious reception.

'Distance' is another possible contextual realisation. It has sometimes even been suggested as the common core meaning of Turkic indirectives. Some kind of distance is likely to be involved if a speaker does not refer directly to the event itself, but

rather to the reception of it. One kind of dissociation from the event may be an ironic relation to it, reservation interpretable as sarcasm, disdain, etc. e.g. *Bunu yapacakmışsın* <this-ACC do-PROSP-IMIŞ-2.P.SG> 'You will obviously [as you think] do this'.

The limitation to an indirective statement may be motivated by caution, modesty, need for a summarizing view, etc., e.g. Turkish *Ben her zaman vazifemi yapmışım* <I always duty-POSS.1.P.SG-ACC do-MIŞ-1.P.SG> 'I have [as it appears] always done my duty'.

Combinations with *güya* 'as if, supposedly' yield dubitative meanings, e.g. *Güya görmüşüm* <supposedly see-MIŞ-1.P.SG> 'It is alleged that I have seen it [but I refuse to accept this]', *Güya almış* <supposedly take-MIŞ> 'Supposedly (s)he has taken it [but I have doubts about it]'. Combinations with *sanki* 'as if' may have the same effect, e.g. *Sanki anlamış* <as.if understand-MIŞ> 'It is as if (s)he has understood [but I don't believe it]'. The last sentence can also have non-dubitative interpretations, if the verbal form is stressed: '(S)he seems to have understood'.

Turkish has other ways of expressing supposition or conjecture, for example by adding *-dir* (< *turur* 'stands') to intraterminals and postterminals, e.g. *Ali okuyordur* <A. read-INTRA-DIR> 'Ali is presumably reading', *Ali gelmiştir* <A. come-POST-DIR>'Ali has presumably arrived'. The relationship between *yapmıştır* and *yapmış* thus does not correspond to the relationship between the Bulgarian confirmative *napravil e* and its nonconfirmative counterpart *napravil*, as has sometimes been claimed.

Emotional nuances

Indirectives may also be used to signal personal emotions, affective responses to the propositional content. In expressive speech acts they may be used in an exclamative way, conveying the recipient's surprise at an unexpected or remarkable situation. Their use may, in particular contexts, be interpreted in terms of admiration, new knowledge, sudden awareness of revealed facts, mental unpreparedness, perception contrary to one's expectations, etc. This includes so-called mirative connotations, which follow naturally from the notion of indirectivity. The conscious reception ('as it turns out / turned out') may be sudden or unexpected; what the recipient turns the mind to may come as a surprise. The fact that Turkish indirectives may convey new information that is not yet part of the speaker's integrated picture of the world (Aksu-Koç & Slobin 1986) is compatible with the central value of indirectivity. Examples: *Bebek sütün hepsini içmiş!* <baby milk-GEN all-POSS.3.P.SG-ACC drink-MIŞ> 'Oh, the baby has drunk all the milk!', *Erken gelmişsiniz!* <early come-MIŞ-2.P.PL> 'Oh, you have come early!', *Ali sınavını geçmiş* (<A. exam-POSS.3.P.SG-ACC pass-MIŞ> '[I am surprised that] Ali has passed his exam!', *Ali şampiyon olmuş* <A. champion become-MIŞ> '[How surprising,] Ali has become a champion! In exclamations such as Turkish *Bu kız ne güzelmiş!* <this girl what beautiful-IMIŞ> 'How beautiful this girl is!', *Büyümüşsün!* <grow-MIŞ-2.P.SG> 'Oh, you have really grown!', the indirective marker just adds the meaning 'as I am / become aware of', which is the central value

of the category. Stress and intonation are important in these cases. The Albanian admirative, which is a marked indirective, can be used in the same way, e.g. *Ti kërcyeke shumë mirë!* 'You dance very well!' (Friedman 2000: 343). These uses do not mean that 'mirativity' is the central meaning from which the other uses of indirectives may be derived (DeLancey 1997). Surprise, novelty and contrariness to the speaker's expectation are not necessary elements of indirectivity. On the contrary, so-called 'hot news' is typically expressed by the direct past marker *-di*.

In some Turkic languages, elements of the type *eken,* superficially corresponding to indirective copula particles, may be used as exclamative and corroborative modal particles, expressing astonishment. This is a result of contamination with an old Turkic particle *e(r)ki(n)*. These elements are not genuine indirective copula particles, but utterance-final particles that do not take personal suffixes, i.e. are added to complete main clauses. They may also cooccur with the simple past item *-di*, e.g. Kazakh *Keldi eken!* <come-PAST MOD> 'It has indeed arrived!'. As mentioned above, genuine indirective copula particles are incompatible with the simple past marker *-di* and the related copula forms *edi* and *idi* 'was'. The *eken* markers of this type also occur in interrogative sentences that do not express real questions. They may form rhetorical questions with readings of wondering and hesitation, in the sense of 'I wonder if ...', e.g. Kazakh *Ne ettim eken?* <what do-PAST-1.P.SG MOD> 'I wonder what I have done'. Here they are used as detensive markers to tone down questions, to give them meditative, skeptical or timid connotations, e.g. Kazakh *Keldi me eken?* <come-PAST Q MOD> 'I wonder if it has arrived'. The Noghay interrogative sentence *Nege kelgenler eken?* <what-DAT come-POST-3.P.PL MOD>, is a rhetorical question meaning 'I wonder why they have come'. By contrast, *Nege kelgen ekenler?* <what-DAT come-POST EKEN-3.P.PL> 'Why do they seem to have come?' ('Why have they, as it seems, come?') is an indirective sentence with the copula particle *eken* taking the personal suffix (see Karakoç 2005: 28).

Markedness

In analyses that take the indirectives to signal the speaker's doubt about the correctness of the content, it is sometimes claimed that the non-indirective terms, *geldi*, *geliyor*, etc., are the marked members, signalling confirmation, etc. The reason is that declarative sentences are typically used to express statements, thus suggesting an attitude of belief in the truth of the propositional content. Most speech acts typical of declarative sentences suggest that the proposition is true or will turn out to be true. Utterances stating 'X is the case' and 'X is not the case' suggest that the speaker is certain about the actuality or non-actuality of the event. Evidentially unmarked terms may suggest that the source of information is direct experience, that the speaker takes / has taken part in the event consciously, is / was in control of it, etc. Though the unmarked members of the indirectivity oppositions may imply certainty, they are not marked for direct experience, and they are even indifferent towards this notion. While the Turkish markers *-miş* and *imiş* signal indirective meanings, the markers

-di, -iyor, etc., do not necessarily imply that the source of information is direct experience. They may also be used for unwitnessed, uncontrolled, reported or inferred events. Indirect experience is grammatically marked, whereas direct experience is the default interpretation of the unmarked members of the oppositions. The widespread opinion that unmarked items such as *geldi* 'has come / came' consistently signal 'direct experience' or 'visual evidence' is incorrect. In historical narratives, direct pasts of the type *-di* are used as the basic items, also for events unwitnessed by the speaker, e.g. *Kemal Paşa, Selânik'te doğdu* [A. S.-LOC be.born-PAST] 'Atatürk was born in Salonika'. They just do not signal that the event is stated in an indirect way, i.e. acknowledged by a recipient by means of report, inference or perception. They just lack the two-layered information typical of indirectives, and may thus be used in a neutral way if the speaker considers the evidential distinction unessential.

Discourse types and registers

Discourse-pragmatic factors in terms of registers and genres are important for the realization of indirectives. Indirectives may play various roles according to different discourse types. Two major text types seem to reflect different tendencies.

Indirectives are typical of subjective registers relating to discourse types that focus on the immediate situation and the personal involvement of the speaker. They prototypically represent direct interaction, oral, immediate face-to-face communication, conversations, speeches and oral narratives. The texts in which they occur mostly exhibit elements of proximity, e.g. first- and second-person personal pronouns, and predominantly paratactic clause-combining techniques yielding numerous short simple sentences. These text types offer excellent context-sensitive options with regard to marking propositional contents for indirectivity.

In many languages, attitudinal particles are mostly used in relatively subjective registers. The use of corresponding devices in Turkic, indirectives and epistemic markers such as dubitatives, presumptives and assertives, is subject to similar restrictions.

Indirectives are also typically used in certain types of traditional narrative discourse describing past events and referring to animate participants, e.g. fairy-tales. In traditional story-telling they play the role of detaching the narrator from the narrated events, e.g. *Bir varmış bir yokmuş* <one existing-IMIŞ one nonexisting-IMIŞ> 'Once upon a time there was...', *Evvel zaman içinde bir padişahın üç kızı varmış* <earlier time in one ruler-GEN three daughter-POSS.3.P.SG existing-IMIŞ> 'Once upon a time there was a king who had three daughters'. (On traditional *-miş*-based narratives in Turkish, see Johanson 1971: 79-80.)

Indirectives are not easily employed in objective registers implying relatively precise text conventions, i.e. in more distanced, detached, descriptive texts with more specific informative contents, a higher degree of formality, less personal involvement of the text producer and thus a low degree of expressivity. Objective registers are typical of more planned written discourse with a certain distance between text pro-

ducer and addressee. More or less fixed stylistic structures may leave the text producer with fewer options. Turkic indirectives are hardly used in these registers. The latter are rich in hypotactical devices, and Turkish lacks, as already mentioned, the possibility of embedding indirective clauses.

The main reason for the avoidance of indirectives in the dominant styles of modern media is, however, that these styles request less ambiguous modes of expression. The undifferentiated indirective meaning 'it appears that X is the case' without specification of the source is open to several interpretations.

This ambiguity can be compensated for by more differentiated lexical means that refer to specific sources. The indirect reception of a propositional content can be optionally signalled by higher clauses with complementation markers, e.g. *it seems that ..., I guess that..., they say that..., I hear that....* or by adverbial expressions such as *reportedly, in my experience, apparently, allegedly*, etc. When choosing one of these optional devices, I commit myself to a specific reading, tracing the information back to specific sources. I cannot leave the interpretation open. With an expression meaning *allegedly*, for instance, I disclaim responsibility for the propositional content, stating that somebody else has conveyed it. With an indirectivity marker, however, the question of source remains open. It allows me to be vague about sources that I do not want to lay open to view, which is a valuable linguistic option.

References

Aikhenvald, A. Y. 2003. Evidentiality in typological perspective. In: Aikhenvald, A. Y. & Dixon, R. M. W. (eds.). *Studies in evidentiality.* (Typological Studies in Language 54.) Amsterdam, Philadelphia: John Benjamins. 1-31.

Aksu-Koç, A. A. & Slobin, D. 1986. A psychological account of the development and use of evidentials in Turkish. In: Chafe, W. & Nichols, J. (eds.) *Evidentiality: The linguistic coding of epistemology.* (Advances in Discourse Processes 20.) Norwood, N.J.: Ablex. 159-167.

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

Buder, A. 1989. *Aspekto-temporale Kategorien im Jakutischen.* (Turcologica 5.) Wiesbaden: Harrassowitz.

Chafe, W. 1986. Evidentiality in English conversation and academic writing. In: Chafe, W. & Nichols, J. (eds.) *Evidentiality: The linguistic coding of epistemology.* (Advances in Discourse Processes 20.) Norwood, N.J.: Ablex. 261-272.

Comrie, B. 2000. Evidentials: semantics and history. In: Johanson, L. & Utas, B. 2000. *Evidentials. Turkic, Iranian and neighbouring languages.* (Empirical Approaches to Language Typology 24.) Berlin and New York: Mouton de Gruyter. 1-12.

Csató, É. Á. 2000a. Turkish MIŞ- and IMIŞ-items. Dimensions of a functional analysis. In: Johanson, L. & Utas, B. (eds.) *Evidentials. Turkic, Iranian and neighbouring languages.* (Empirical Approaches to Language Typology 24.) Berlin & New York: Mouton de Gruyter. 29-43.

Csató, É. Á. 2000b. Some typological features of the viewpoint and tense system in spoken North-Western Karaim. In: Dahl, Ö. (ed.) *Tense and aspect in the languages of Europe.*

(Empirical Approaches to Language Typology, Eurotyp 20-6.) Berlin and New York: Mouton de Gruyter. 723-751.

DeLancey, S. 1997. Mirativity: The grammatical marking of unexpected information. *Linguistic Typology* 1, 33-52.

Friedman, V. A., 2000. Confirmative/nonconfirmative in Balkan Slavic, Balkan Romance, and Albanian with additional observations on Turkish, Romani, Georgian, and Lak. In: Johanson, L. & Utas, B. (eds.) *Evidentials. Turkic, Iranian and neighbouring languages*. Berlin, New York: Mouton de Gruyter. 329-366.

Johanson, L. 1971. *Aspekt im Türkischen. Vorstudien zu einer Beschreibung des türkeitürkischen Aspektsystems*. (Studia Turcica Upsaliensia 1.) Uppsala: Almqvist & Wiksell.

Johanson, L. 1994. Türkeitürkische Aspektotempora. In: Thieroff, R. & Ballweg, J. (eds.) *Tense systems in European languages*. Tübingen: Niemeyer. 247-266.

Johanson, L. 1996a. Terminality operators and their hierarchical status. In: Devriendt, B., Goossens, L. & van der Auwera, J. (eds.) *Complex structures: A functionalist perspective*. (Functional Grammar Series 17.) Berlin, New York: Mouton de Gruyter. 229-258.

Johanson, L. 1996b: On Bulgarian and Turkic indirectives. In: Boretzky, N., Enninger, W. & Stolz, Th. (eds.) *Areale, Kontakte, Dialekte. Sprache und ihre Dynamik in mehrsprachigen Situationen*. (Bochum-Essener Beiträge zur Sprachwandelforschung 24.) Bochum: Brockmeyer. 84-94.

Johanson, L. 1998a. Zum Kontakteinfluß türkischer Indirektive. In: Demir, N. & Taube, E. (eds.) *Turkologie heute – Tradition und Perspektive*. Wiesbaden: Harrassowitz. 141-150.

Johanson, L. 1998b. The structure of Turkic. In: Johanson, L. & Csató, É. Á. (eds.) *The Turkic languages*. London: Routledge. 30-66.

Johanson, L. 1999. The dynamics of code-copying in language encounters. In: Brendemoen, B., Lanza, E. & Ryen, E. (eds.) *Language encounters across time and space*. Oslo: Novus Press. 37-62.

Johanson, L. 2000a. Viewpoint operators in European languages. In: Dahl, Ö. (ed.) *Tense and aspect in the languages of Europe*. Berlin & New York: Mouton de Gruyter. 27-187.

Johanson, L. 2000b. Turkic indirectives. In: Johanson, L. and Utas, B. (eds.) *Evidentials. Turkic, Iranian and neighbouring languages*. Berlin & New York: Mouton de Gruyter. 61-87.

Johanson, L. 2001. On three dimensions of aspectual terminality. In: Bisang, W. (ed.). *Aspects of typology and universals*. (Studia Typologica 1.) Berlin: Akademie-Verlag. 53-62.

Johanson, L. 2002a. *Structural factors in Turkic language contact*s. London: Curzon.

Johanson, L. 2002b. Contact-induced linguistic change in a code-copying framework. In: Jones, M. C. & Esch, E. (eds.) *Language change: The interplay of internal, external and extra-linguistic factors*. (Contributions to the Sociology of Language 86.) Berlin: Mouton de Gruyter. 285-313.

Johanson, L. 2003. Evidentiality in Turkic. In: Aikhenvald, A. Y. & Dixon, R. M. W. (eds.) *Studies in evidentiality*. (Typological Studies in Language 54.) Amsterdam, Philadelphia: John Benjamins. 273-290.

Karakoç, B. 2005. *Das finite Verbalsystem im Nogaischen*. (Turcologica 58.) Wiesbaden: Harrassowitz.

Kıral, F. 2000. Reflections on *-miş* in Khalaj. In: Johanson, L. & Utas, B. (eds.) *Evidentials. Turkic, Iranian and neighbouring languages*. (Empirical Approaches to Language Typology 24.) Berlin, New York: Mouton de Gruyter. 89-101.

Levin-Steinmann, A. 2002. *Bedeutung und Funktionen der kopulalosen l-Periphrase im Bulgarischen*. Habilitation thesis, University of Leipzig.

Turkish glosses in the Pahlavi-Oghuz Turkish glossary *Pahlavī-Āmūz:* A linguistic and textual analysis

Fikret Turan

Turan, Fikret 2006. Turkish glosses in the Pahlavi-Oghuz Turkish glossary *Pahlavī-Āmūz:* A linguistic and textual analysis. *Turkic Languages* 10, 90-110.

Pahlavī-Āmūz is an interlinear Pahlavi to Oghuz Turkish glossary by an anonymous scholar composed in the late fifteenth century. It constitutes one of the last parts of an interlinear Persian-Oghuz Turkish grammar and conversation book that is entitled *Risāle-i Zebān*. It offers new glosses, semantic peculiarities of certain Turkic words and includes very rare and unknown Pahlavi vocabulary. This article examines textual peculiarities of the work and explains certain Oghuz words that demonstrate original semantic characteristics. It also lays out a full list of Oghuz vocabulary with their English meanings and their Pahlavi counterparts provided in the text, and points out the questions concerning certain vocabulary items. A full index of Pahlavi words and the facsimiles are presented in the final section.

Fikret Turan, University of Manchester, School of Languages, Linguistics and Cultures, Middle Eastern Studies, Oxford Road, Manchester M13 9PL, UK.
E-mail: fikret.turan@manchester.ac.uk

Pahlavi glossaries are rare among the linguistic and lexical studies that were produced in the formation period of the Turkish language and literature in and around Anatolia between the thirteenth and sixteenth centuries. These Pahlavi lexicological works that are scattered in the Turkish libraries are composed mostly in a Pahlavi to New Persian format. However, among them, one lexical work of select Pahlavi vocabulary entitled *Pahlavī-Āmūz* (PA) presents an original case of a Pahlavi to Oghuz Turkish glossary. It presents not only certain peculiar Pahlavi words unattested in many sources, but it also yields certain Oghuz Turkish lexemes that are either unattested or appear in different meanings in the sources by providing important information on the lexicology of the early Anatolian Turkish. In this article, I will analyze PA by expounding its textual and linguistic characteristics by demonstrating the peculiarities of its Turkish glosses. In addition, I aim to demonstrate the difficulties and the remaining problems of certain lexical elements that should be reconsidered by further inquiries.

PA is a short Pahlavi-Oghuz Turkish interlinear dictionary by an anonymous author composed in the late fifteenth century. It constitutes one of the last parts of an interlinear Persian-Oghuz Turkish grammar and conversation book entitled *Risāle-i*

Zebān. The work was composed in 1485 CE (890 H), and is currently located in Suleymaniye Library, where it is catalogued as Carullah 2043. Its dimensions measure 24 x 34 centimetres. PA consists of six pages placed between the folios 132b-134a of this bound book form that has 138 folios altogether. The book is in fine condition within a leather binding that was probably made and embellished with traditional motifs in later periods. Since it is placed in a Persian grammar and conversation book, it is clear that the students who studied the Persian language were expected to learn some Pahlavi vocabulary and grammar as part of their education as well. Starting from the fourteenth century, as a means of spreading and infusing classical Persian literature and culture in Anatolia, the school curricula included courses on Persian language and literature, and this work must have been used as a student text-book.

In PA, Pahlavi words are listed on odd-numbered lines and their Turkish counterparts are indicated right under them on even-numbered lines. Each page has fourteen lines. The work is written in the *naskhi* style with a relatively regular and clear vocalization. All words are penned in black ink, whereas the title "Pahlavī-Āmiz" is penned in red ink. Throughout the work, Pahlavi words are penned more thickly and clearly than their Turkish counterparts. Phonetically, like in many other earlier works, the consonants are often voiced word-initially and word-finally in Turkish words. Although this might be considered a result of misspelling, some words are consistently and regularly penned with voiced consonants like *avuǰ, bulamaǰ, busu, burǰaq, giǰi*, etc. And because of this consistency in spelling I take them to be characteristic of this specific dialect of Oghuz, and transcribe and list them as they are.

In the manuscript, the vocabulary is organized in no particular order. However, it seems that all synonymous words are listed successively and linked by the conjunction *ve* 'and'. Sometimes words ending in the same phoneme follow each other, but this is not carried out consistently. Otherwise, words are listed randomly. Sometimes one single Oghuz word represents several synonymous or near synonymous Pahlavi words that are usually shown together one after the other. For example, on line 7-8 (131v) the Oghuz word *qurbaγa* 'frog' is used for the rendition of the Pahlavi words *γūk, vazaγ, bazaγ, ǰere'* all of which are either synonymous or signify different kinds of the animal 'frog'. A similar case is observed on line 9-10 (131v), where the Oghuz *qablu baγa* 'tortoise, turtle' is used for the Pahlavi words *seng-pošt* and *kešef*, and so on.

The Pahlavi words that are explained in *Pahlavī-Āmiz* are mostly nouns or words of nominal classes. Only one verb is given as a word entry, namely the second person imperative form *žāla* 'answer the question'. There is also one single frozen expression of exclamation, *ėy besā* 'Oh how many, how many there are!'. Although the Pahlavi entries are usually simple root or derived words like *sāl* (*yıl*: year), *imsāl* (*bu yıl*: this year), *āteš* (*od*: fire, blaze) or *āteš-dān* (*oǰaq*: hearth), there are a few word groups in izafa constructions and frozen compound word groups such as *sāl-i āyende* (*gelesi yıl*: next year) or *pek ve lek* (*güǰüklenmek*: conceited, ignorant), etc.

About fifteen Pahlavi nouns are translated into Oghuz as *masdar* forms, e.g. *aγlamaq* (*mūye*: crying), *aňmaq* (*yāδ*: remembrence), *edeblemek* (*yārsīn*: training, education), *χoš dutmaq* (*nuvāχt ve nuvāz*: caressing, blandishing), *ınanmaq* (*bāver*: believing), *ič burmaq* (*kuzāyiš*: twisting of the stomach), *qaynamaq* (*tūy?*: boiling), *qurtılmaq* (*rustī*: escaping), *suda yüzmek* (*senāb ve šināh ve šenā ve šenāy*: swimming), etc. And also, because this work is aimed to teach Pahlavi to students, the author often uses nominal and adjectival phrases to explain the words more clearly and more descriptively, e.g. *aγaǰ čivisi* (*biγāz*), *baγarsuq yaγı* (*rūze*), *ǰānver azūsı* (*yešk*), *ǰılq qoyun* (*sitān*), *dikeǰek aγaǰ* (*derχāl*), *er görmiš qız* (*kālim*), *eyer āleti* (*sitāmī*), *giǰi barmaq* (*kālūǰ ve kihīn*), *qatı γavγa* (*χilālūš*), *müšk otı* (*nāk*), *üzüm tegegi* (*tāk*), etc. In this work, approximately 215 Turkish word or phrase entries are given as counterparts or explanations of approximately 250 Pahlavi lexemes, whereas other Pahlavi words are used as loanwords in Turkish and rendered as such. In addition, some of the Turkish words that are used as counterparts to Pahlavi words are originally words borrowed from New Persian or from another dialect of Persian, as is the case with *čašnı* for *numūne* (a sample), *čuvālduz* for *bunduz* (a big packing needle), *fèrište* for *surūš* (an angel), *merǰimek* for *dānze* (a lentil), *moγ-zār* for *māγ* (meadow), etc.

Notes on Turkic glosses

aq-palas (*šāl*: 133a 1-2) In Oghuz dialects the Persian word *šāl* is used widely but mostly with a short /a/ as *šal* meaning 'a shawl or a mantle made of Tibetan goat's wool' or 'a common shawl'. Here *aq-palas* should be the name of a specific kind of shawl that is made of white hair-cloth. Today, in Oghuz dialects the word *palas* or its variant *palaz* usually signifies a kind of kilim, sack, a bed or a coarse cloth with no reference to 'shawl', see Oružov, Redhouse (1890), TS, Tekin et al. (1995). Nakhjivānī (1366) also translates it as 'çul ve küçüçek palas' (sack and small kilim).

aqšama (*mevīzāb*: 134a 9-10) This word is neither vocalized in the text nor is it attested in other sources. It is given as the Oghuz counterpart of *mevīzāb* which means 'an acid and inebriating drink made from dried grapes and water', which is *boza* in Ottoman Turkish, and *aχsuma* or *aχsama* in Persian. The word *aqšama* probably derives from the Persian *aχsuma* or *aχsama*. Although the change s>š in syllable-initial position is rare in Oghuz dialects, few examples of it are seen in Turan (2001) as well. From the orthography of the word *aqšama* we can assume that that the copyist or the translator thought it to be a compound word consisting of *aq* and *šama*, see *Burhan*.

arsuz (*reste*: 131b 9-10) This word is shown as the counterpart of the Pahlavi *reste* meaning 'pure without and within'. It seems that another copyist crossed the *arısuz* and wrote *arsuz*. In Old Turkic we find *ır/ir* or *ırra* meaning 'shame', which Clauson (1972) connects with Arabic *'ār*. Agreeing with him, I also believe the word *arsuz* breaks down as *ar* < *'ār* (Arabic) and the privative +*suz* (*arsuz* < *'ārsuz*) meaning 'pure, anything without shame or fault'. Although its variant *arsuz/arsız* has

negative denotations in Oghuz dialects, where is denotes 'shameless, immodest, carefree, light-hearted', the opposite meaning is semantically plausible as is the case in the New Persian *bī-'ār* meaning 'blameless' or 'faultless', see Steingass (1957).

giši (*kākā*: 133b 7-8) Although this word means broadly 'a man, a human being' in Oghuz dialects, here it signifies specifically 'a mature, grown up male person' as it is used in modern Azeri dialects and is encountered in some Old Anatolian Turkish texts. It is clearly the semantic remnants of the Pahlavi word *kākā* meaning 'big brother, uncle or grown up male relative', see Steingass (1957) and TS. It is also interesting to see the phonetically changed variant of this Pahlavi word in the Iğdır dialect of Azeri as *γağa,* a colloquial word meaning 'a close male friend, buddy' in friendly male banter (my unpublished notes, FT).

güǰüklenmek (*pek ve lek*: 134a 5-6) This appears to be a variant of *gökčeklenmek* meaning 'to be proud, to show himself/herself very beautiful or handsome, to be vainglorious, egotistical', which is widely observed in Old Anatolian Turkish (OAT) texts. In the texts of the same period, it has variants like *göǰeklenmek* or *gökǰeklenmek* as well, see TS. It is clearly derived from *gökček* (beautiful, handsome, nice), which is still commonly used in Azeri and Eastern Anatolian dialects: (<*gökček+le-n-*). Being a variant of it, *göček* is dominant in Iğdır and Nakhçevan Azeri.

iliši yıl (*pīrār*: 132a 11-12) This compound means 'the year before last, two years ago'. This form rarely appears in OAT texts. However, in Azeri dialects of North-Eastern Anatolia, its variant *nišil* (< *lišil* < *iliši il/yıl*) is used extensively, especially in the dialect of Iğdır, and Nakhchevan.

keler (*perχū*: 134a 1-2) This is a variant of *kiler,* which is still used in Anatolian and Balkan dialects of Turkish as a borrowed word of Greek origin, see Eren (1999). However, the variant *keler* is used in the Niğde dialect of Turkish, see DS. Nakhjivānī (1366) translates it as *taχtadan anbār* (a storehouse or granary built of wooden boards).

qıba (*hazār*: 132b 3-4) This must be a vernacular variant of the Arabic *qubbe* (a dome, cupola, vault) in Turkish meaning 'a tomb' or 'dome'.

qıma (*χuvāze*: 132a 1-2) While this word has a very close semantic association with the preceding word *qıba,* it may have a different etymological history. Based on the semantic and phonetic proximity, in my judgement, this word *qıma* is a variant of the Persian *χamba/χumba* (a vine-arbour, trellis, a gallery, a dome), see Steingass (1957).

qoqu (*χʷonǰze*: 133b 13-14) In Old and Middle Turkic dialects *qoq-* means issuing an unpleasant smell of burning meat or food, see Clauson (1972). Thus, it demonstrates the exact meaning of the *χʷonǰze,* which Halimī renders as the smell of burning suet or fat (*čerbeš*). However, this should be only one meaning of the word because the other word *yıδ/yıδıγ* that is encountered in Old and Middle Turkic texts often shows the same meaning 'an unpleasant smell', see Kaşgari. These words plainly must have signified 'any kind of smell' as well as a 'bad smell' as was the case in OAT. In OAT, *qoqu/qoχu* and *yiyi* (<*yıδıγ/yıδ*) are used interchangeably, and

they sometimes appear within the same couplet as in *eser yėl bigidir getürür eyi / gülefden qoχu vü 'alefden yiyi*. (It is like a blowing wind it brings well / The pleasant fragrance of roses and sweet aroma of victuals), see Dilçin (1991). After about the sixteenth century *yiyi* developed into *iy* and became dominant in Azeri with the meaning a good and pleasant smell, while Ottoman and Modern Turkish retained *qoqu* for both pleasant and unpleasant smells.

yardam (*yāver*: 132b 8-9) This is a variant of *yardım* (help, assistance, helper) used in Ottoman, Modern Turkish and Azeri dialects. This variant is, however, common in modern Turkmen, see Tekin et al. 1995. The word has probably derived from Persian *yār* < **yār+da-m* (friend, helper) as Nakhjivānī (1366) translates it with the same Persian word *yār*.

Turkish glosses

(*a.* Arabic, *g.* Greek, *p.* Persian)

a, 'a
ačuq: 1. Open, 2. Naked. (furūje: 133a 3-4)
ādem boqı: Human excrement (kūh: 131b 3-4)
ajla: Hungry, having an empty stomach (nāšitā: 133a 11-12)
ayaj čivisi: Wedge for splitting timber (biγāz: 133b 13-14)
aylamaq: Crying, weeping (mūye: 133a 3-4)
alu: Exalted, eminent, lord, commander (vālā: 132a 5-6)
añmaq: Remembering, remembrance (yāδ: 132b 7-8)
añduruj: Keepsake, anything memorable, a present (yāδigār: 132b 8-9)
aq-palas: Shawl made of white hair-cloth: (šāl: 133a 1-2)
aqšama: *p.* Light alcoholic drink made of dried grapes and water (mevīzāb: 134a 9-10)
arı: Pure, clear, (vīž: 133b 1-2)
arsuz: Pure without and within, freed (reste: 131b 9-10)
'arš: *a.* Sky, heaven (girizmān: 131b 13-14)
aşıl: *a.* Root, origin (nižāz: 134a 7-8)
asma: Vine, grapevine (ābā: 131b 11-12)
ašuq: Bone of the elbow, ankle or the heel, hucklebone (qāb: 132a 1-2)
'avrat: *a.* A woman, married woman (pūt ?: 133a 9-10)
avuj: Palm of the hand (keyfe: 133a 7-8)
avurd: Pouch of the cheek (žem: 133b 9-10)
'ayb: *a.*. Fault, a shameful thing, a shortcoming (sikāl ?: 132b 11-12)

b
balıq duzaγı: Net or fish hook (šast: 134a 1-2)
baγarsuq: Intestine (rūze: 132b 3-4)
baγarsuq yaγı: Intestine fat, suet, lard (čazdere: 132b 3-4)
bāz: *p.* Back, again, anew, afresh (bāz: 132a 13-14)
beñzerü: Alike, resembling (qoš: 131b 13-14)
bezir-χāne: Oil-mill, mill producing linseed oil (γınk: 131b 5-6)
bıldır: Last year, past (pār: 132a 11-12)
bilek: Wrist (γarbūn ?: 134a 5-6)
boq: Dung, excrement (sekāle: 131b 5-6)
boynuz: Horn (šūγ: 134a 7-8)
boyunduruq: A yoke for oxen (yūγ: 134a 7-8)
bögrek: Kidney (kurde: 132b 1-2)
bulamaǰ: A kind of pottage, soup (ardihāla ve kāčī: 132a 9-10)
buqaγu: Fetter, chain (zāvilāne: 132b 8-9)
burǰaq: Vetch (mīšū: 131b 11-12)
burušuq: Wrinkle, a wrinkled thing (kerunǰ: 133a 7-8)
busu: Ambush (kemīn: 133a 5-6)
bu yıl: This year (imsāl: 132a 11-12)

č
čadır: Tent (čere ?: 132a 5-6)
čārsız: Without remedy (nā-guzīr: 133a 11-12)
čašnı: *p.* Sample, example (numūne: 133a 13-14)
čevük: Strong, swift (kelendür: 132a 1-2)
čuvālduz: *p.* Large packing needle (bunduz: 131b 3-4)

d
dar: Narrow, sad (teng: 132a 13-14)
degenek: Club, shepherd's staff (pāhu: 133b 5-6)
delü: Mad, insane, crazy (šeydā ve dīvāne ve āšufte: 133a 1-2)
dırmıq: Rake (esīb: 132a 7-8)
dikeǰek aγaǰ: Young tree, sapling (derχāl: 132a 3-4)
dübdüz: Very plain, equal, even place or thing (yeksān ve hemvāre: 131b 11-12)
dünya: *a.* The world, the earth (gītī: 133a 7-8)

e
edeblemek: To teach politeness (yārsīn: 132b 7-8)
ehl ?: *a.* Family, people, community (balaχ ?: 132a 3-4)
ekinǰi: Farmer, cultivator (dihqān: 132b 7-8)
elek: Sieve (māšū ve pervīzen: 133a 9-10)
emek: Trouble, hardship, grief (renǰ: 134a 5-6)
enegim saγmal: Rainbow (tīrāže ve nūse: 131b 13-14)

er görmiš qız: Widow or divorced woman (kālim: 131b 3-4)
ėš: Friend, partner, confidant (hemdem: 133b 9-10)
ėy niče: Oh how much, how many there are! (ėy besā: 132a 11-12)
eyer āleti: A part or parts of a saddle (sitāmī: 132a 3-4)

f
fėrište: *p.* Angel (surūš: 133b 5-6)
fiske: *p.* Striking the cheeks lightly with the hand, flipping (zīger: 132a 3-4)

g
gelesi yıl: Next year (sāl-i āyende: 132a 11-12)
gėn: Open, wide (firāχ: 132a 13-14)
giǰi barmaq: Little finger (kālūǰ ve kihīn: 133b 7-8)
giǰi biter üzüm: A small kind of grape (feršeng: 132a 3-4)
giši: Big brother, uncle or grown male member of the family (kākā: 133b 7-8)
gizlü: Anything concealed or hidden (nuhuft: 133b 1-2)
gökǰek: Beautiful, elegant, good (naγz: 133a 11-12)
gömeǰ: A kind of bread baked in ashes (sekārvā: 132b 12-13)
gözik: Delicate, beautiful, handsome (geš: 134a 1-2)
güǰeste ?: 1. A musical instrument with six strings, 2. Children's toy similar to a ball (šešχanǰ: 134a 7-8)
güǰli / güǰlü: A strong one (tīve: 131b 11-12, kelendür: 132a 1-2)
güǰlü ve ǰevük: Strong and nimble person (kelendür: 132a 1-2)
güǰüklenmek: Being conceited or ignorant, excessively proud (pek ve lek: 134a 5-6)
gümiš: Silver (māχ: 134a 7-8)
gün dönmesi: Sinking of the sun after noon, sunset (čāvāle ?: 131b 1-2)
günlük: Clear, manifest (mebīn ?: 133b 13-14)
güyegü: Son-in-law (dilād: 132b 5-6)

h, ḥ
ḥalvaǰı: Maker or seller of sweetmeats (šīrīn-kār: 133a 3-4)
heybet: *a.* Fear, awe, respect, dignity (sitūh: 133a 1-2)
ḥuǰre: *a.* Small house, cell or hall (kāšāne ve gurīǰ: 133a 3-4)

ı
ınanmaq: To believe, believing (bāver: 131b 5-6)

i
ič burmaq: Twisting of the stomach, contortion (kuzāyiš: 133a 7-8)
ič yaγı: Suet, lard (pīh: 132b 3-4)
iči boš: Empty, hollow, rotten inside (kavak: 133b 7-8)
iliši yıl: The year before last, two years ago (pīrār: 132a 11-12)
iš-güǰ: Work, labour, occupation (kiyā: 133b 11-12)

iškenǰe: *p.* Torture, torment, pain (efzāh ?: 132a 13-14)
išüñ temāmı: Concluding a work, completing a task (siperī: 134a 11-12)
'ivaż: *a.* Exchange, reward, return for good or evil (keyfer: 131b 9-10)

ǰ
ǰalıq: (Usually a horse) going sideways and skipping (ǰehende: 132b 3-4)
ǰānver azūsı: Tooth or tusk of a wild animal (yešk: 131b 11-12)
ǰevāb vėr su'āla ?: Answer the question (žāla ?: 131b 1-2)
ǰılq qoyun: A weak sheep (sitān: 133b 9-10)
ǰiger: *p.* Liver (ǰiger: 132b 1-2)
ǰubıq yapraγı: Bough of a tree (ferhāñǰ: 131b 5-6)

k
kečiler: Goats (dekihān: 133a 7-8)
kef: Potash, pumice stone, froth (kelāšekere: 133b 11-12)
keler: I. *g.* Buttery (peryχū: 134a 1-2)
keler: II. Lizard, alligator, porpoise (sūsmārā: 131b 9-10)
kemend: *p.* Noose, lasso (χāmī ve uqru'a ?: 132a 5-6)
kešür: Carrot (zerde-zemīn: 132b 8-9)
kilindir: *g.* A kind of wine pitcher mostly made of metal, a firkin (bulbulī: 132a 1-2)
kiri: *p.* Reward given to the bringer of good news (muždegānī: 133a 11-12)
kirpi: Porcupine, hedgehog (χārpošt: 131b 7-8, čīzū: 132b 5-6)
kösegü: Poker, half-burnt wood (āteš-kāv ve āteš-engīz: 132a 9-10)
kükremek: Making tumult, clamour (demdeme: 132b 5-6)
külük ?: Swift, fast, speedy (sefeǰ ?: 134a 3-4)

l
loχusa: *g.* A woman in childbed (zāǰ: 132a 5-6)

m
māla: *p.* Trowel (χišl ve riḥš ?: 131b 3-4)
masχaralıq: Irritation, vexation, derision (efsūs: 132a 7-8)
mekr: *a.* Fraud, imposture (destān: 132b 5-6)
merǰimek: *p.* Lentil (dānze: 132b 7-8)
meydān: *a.* An open field for public gatherings, course (isperiš: 132a 3-4)
morγ-zār: *p.* Green field, or meadow (māγ: 133a 9-10)
mumsuz bāl: Pure honey (šehd: 131b 3-4)
müšk otı: A sort of musky herb (nāk: 133b 7-8)

n
nite ki[1]: As, like (133b 11-12)

o
oǰaq: Fireplace, hearth, furnace (kālāne: 133a 5-6, āteš-dān: 132a 9-10)
od: Fire (āteš: 132a 9-10)
od gözi: Live coal, firebrand (zelek ve aχker: 134a 3-4)
oγul bal: Virgin honey, the first honey of a new swarm of bees (šān: 131b 3-4)
oynaš: Sweetheart, paramour, gallant (mol: 133b 9-10)

ö
öd: Gall-bladder, bile (zehre: 132b 1-2)
ög: Intellect, understanding capacity, judgement (hūš: 133b 3-4)
ögendire: Ox-goad (γāvšeng: 132a 1-2)
ögürmek: To retch, make a noise when vomiting, bellow (bāzbīḥ ?: 132a 5-6)
öksürük: Cough (χafa ve surfa: 132b 1-2)
örnek: Model, specimen (numūdār: 133a 13-14)
öyken: Lungs, lights (šuš: 132b 1-2)
öykünmek: To follow a custom, to be a faithful observer, (verziš: 133b 1-2)

p
palan atı: Pack-horse (māγče: 133b 5-6)

q
qabīle: *a.* Tribe, family (tabār: 132a 3-4)
qablu baγa: Tortoise, turtle (seng-pošt ve kešef: 131b 9-10)
qalem: *a.* Reed, pen (kilk: 133a 5-6)
qanat: Wing, feather (peẕīr: 134a 3-4)
qanā'atlu: Content, satisfied, pleased (χorsend: 133b 11-12)
qanlu qabarǰuq: Bloody swelling, pus (hav: 133b 3-4)
qānūn: *a.* Rite, custom, common law (āyīn: 132a 7-8)
qapaq: Lid of a pot or furnace (nuhunbīn: 133b 1-2)
qaravaš: Female servant, servant (perestār: 133b 9-10)
qarbuz: Watermelon, pumpkin (hinduvāle: 133b 3-4)
qarın: Belly, paunch, stomach (šikem: 132b 3-4)
qarıš: Span, length of a span (vīže: 133b 3-4)
qasurγa: Hurricane, whirlwind (dūle ve hūle: 133b 13-14)
qatı γavγa: Tumult, riot, uproar (χilālūš: 133b 11-12)
qavun: Melon, watermelon (χarbuze: 133b 3-4)

[1] Meaning "like the preceding one."

qayγulu: Sad, unhappy, afflicted (mustemend: 132a 1-2 < arabic)
qayın-ana: Mother-in-law (χasvīte ?: 131b 11-12)
qaynamaq: Boiling, to boil (tūy ?: 134a 5-6)
qıba: Tomb, ossuary, dome (hazzār: 132b 3-4)
qıma: Triumphal arch, tomb, scaffold (χuvāze: 132a 1-2)
qolay: Ease, leisure, rest (pervā: 134a 1-2)
qoltuq iči: Arm-pit (keš: 134a 1-2)
qoqu: Unpleasant burning smell , smell of burnt meat or food, bad odour (χ^{w}onǰīze: 133b 13-14)
qunduzǰuq: Beaver (čafāla: 131b 3-4)
qurbaγa: Frog (γūk ve vazaγ ve bazaγ, ve ǰere': 131b 7-8)
qurd ešeni: Pimple, eruption, rash (pīšterm: 133b 13-14)
qurtılmaq: Escaping, being liberated (rustī: 132b 5-6)

r
raχt: *p.* Burden placed on a pack animal (šilek ?: 134a 3-4)
revnaq: *p.* Beauty, comeliness, elegance, lustre (fer: 131b 5-6)
riše: *p.* Fiber, fibrous fringe (kaχša ?: 133a 5-6)
rufadan yumurda: Poached egg (nīm-birišt: 133a 13-14)

s, ṣ
saqaq: Dewlap or wattles (γabγab: 133a 3-4)
saqsıγan: Magpie (kelāže: 133b 11-12)
ṣarp: Very steep, very hard or difficult (pelek: 134a 3-4)
savaš: War, battle (barχāš: 131b 5-6)
sevinmek: Rejoicing, being joyous or happy (šāδ-kām ve χurrem: 132a 7-8)
sıva: Plaster, mortar (ežend: 132a 7-8)
sidük: Urine (ǰehīn: 132b 3-4)
sikiš: Coition, sexual intercourse (gān: 133a 5-6)
soγulǰan: Earthworm (send: 132b 11-12)
suda yüzmek: Swimming (senāb ve šināh ve šenā ve šenāγ: 132b 11-12)
süǰi: Wine (siyegī ve mul: 132b 12-13, bāδe: 132b 12-13)
sülük: Leech (mālīz ve zālū: 133a 9-10)

š
šāb: *a.* Rock-alum, vitriol (zeme: 132b 8-9)
šalvar: *p.* Breeches, drawers (išim ve rānīn: 131b 13-14)

t, ṭ
ṭalaq: Spleen, milt (supurz: 132b 1-2)
ṭamla: Drop, a drop of water (tek: 133b 5-6)
ṭanušıq: Consultation, deliberation (kengāǰ: 133a 5-6 <t. keñeš ?),
ṭatlu: Sweet, milky, pleasant (šīrīn: 133a 1-2)

ṭatlulıq: Being sweet, milky or pleasant (šīrīnī: 133a 3-4)
ṭavul: Drum, cattle drum (tabīra: 133b 9-10)
taχt: *p.* Seat, bench, stool (nešīmen: 133a 11-12)
ṭayı: Maternal uncle (χālū: 132a 13-14)
temām ?: *a.* Whole, complete, full, perfect (bezdum ?: 132a 13-14)
ters: Inverted, reversed (bāšgūne: 132a 13-14)
ṭılaq: Lingula vulvae, clitoris, throes of childbirth (čeklīz: 132b 5-6)
toñuzlan qurdı[2]: Bombardier beetle, dung-beetle, bracinus crepitans (kuṣṣa: 131b on margin)
ṭuman: Smoke, fog, vapour, mist (nežīm ve hir ve tārmīχ: 133a 11-12)
tü ürpermek[3]: Horripilation, tremor of fever (ferāšā: 133a 7-8)

u
ulu barmaq: The large finger, thumb (sutrug: 133b 7-8)

ü
üzüm tegegi: Tendril of a vine, vine (tāk: 133b 5-6)
[üzum tegegi] bašı: Upper branches and leaves of a vine (mivlāleng: 133b 5-6)

χ
χāne: *p.* House, room, office (χāne: 132b 3-4)
χatun: Princess, lady (bānū: 132a 13-14)
χayran: *a.* Astonished, amazed (āseme: 132a 7-8)
χoš dutmaq: Caressing, blandishing, comforting (nuvāχt ve nuvāz: 133a 13-14, 133b-1)

y
yaba: A farmer's fork for winnowing (hiδ: 133b 9-10)
yara: Wound, pus (hiyam: 132a 5-6)
yašıl ṭaš: Green stone, agate (sevīšem: 131b 5-6)
yapıǰı: Builder (dezār: 132b 7-8)
yardam: 1. Help, assistance 2. Friend, helper (yāver: 132b 8-9)
yėlek: Running, trotting, a runner (pūyā: 134a 5-6)
yeli: Mane of a horse (poš: 134a 1-2)
yengeǰ: Crab (χarǰenk: 131b 7-8)
yer bayarsuyı: Earth-worm (χarāšīn: 134a 1-2)
yıl: Year (sāl: 132a 11-12)
yoyı: Cylinder or waterstone for consolidating clay-roofs (hīzīne: 133b 3-4)

[2] It is shown on the margin.
[3] It is spelled as " tü ürperbek " in the text.

yorγan: Quilt (henguft: 133b 1-2)
yürek: Heart, courage (činīne: 132b 1-2)
yüzgüǰi: Swimmer (senāber ve šenāver: 132b 12-13)

z
zėvle: Wooden club holding the side of a yoke, a long stick (seme: 134a 7-8)

Index of Pahlavi Glosses[4]

ābā (131b 11-12)
ardihāla (132a 9-10
āseme (132a 7-8)
āšufte (133a 1-2)
āteš (132a 9-10)
āteš-dān (132a 9-10)
āteš-engīz (132a 9-10
āteš-kāv (132a 9-10)
aχker (134a 3-4)
āyīn (132a 7-8)

bāδe (132b 12-13)
balaχ ? (132a 3-4)
bānū (132a 13-14)
barχāš (131b 5-6)
bāšgūne (132a 13-14)
bāver (131b 5-6)
bāz (132a 13-14)
bazaγ (131b 7-8)
bāzbīḥ ? (132a 5-6)
bezdum ? (132a 13-14)
biγāz (133b 13-14)
bulbulī (132a 1-2)
bunduz (131b 3-4)

čafāla (131b 3-4)
čāvāle ? (131b 1-2)
čazdere (132b 3-4

čeklīz (132b 5-6)
čere ? (132a 5-6)
činīne (132b 1-2)
čīzū (132b 5-6)

dānze (132b 7-8)
dekihān (133a 7-8)
demdeme (132b 5-6)
derχāl (132a 3-4)
destān (132b 5-6)
dezār (132b 7-8)
dihqān (132b 7-8)
dilād (132b 5-6)
dīvāne (133a 1-2)
dūle (133b 13-14)

efsūs (132a 7-8)
efzāh ? (132a 13-14)
esīb (132a 7-8)
ėy besā (132a 11-12)
ežend (132a 7-8)

fer (131b 5-6)
ferāšā (133a 7-8)
ferhānǰ (131b 5-6)
feršeng (132a 3-4)
firāχ (132a 13-14)
furūǰe (133a 3-4)

γabγab (133a 3-4
γarbūn (134a 5-6)
γāvšeng (132a 1-2)
γūk (131b 7-8)
γınk (131b 5-6)

gān (133a 5-6)
geš (134a 1-2)
girizmān (131b 13-14)
gītī (133a 7-8)
gurīǰ (133a 3-4)
hav (133b 3-4)
hazzār (132b 3-4)
hemdem (133b 9-10)
hemvāre (131b 11-12)
henguft (133b 1-2)
hīle (133b 13-14)
hinduvāle (133b 3-4)
hir (133a 11-12)
hiyam (132a 5-6)
hiδ (133b 9-10)
hīzīne (133b 3-4)
hūš (133b 3-4)
imsāl (132a 11-12)
isperiš (132a 3-4)
išim (131b 13-14)

[4] I would like to express my thanks and regards to Prof. Oktor Skjaervo of Harvard University for his very valuable help in reading and rendering certain Pahlavi words in the text.

ǰehende (132b 3-4)
ǰehīn (132b 3-4)
ǰere‘ (131b 7-8)
ǰiger (132b 1-2)

kāčī (132a 9-10)
kākā (133b 7-8)
kālāne (133a 5-6
kālim (131b 3-4)
kālūǰ (133b 7-8)
kāšāne (133a 3-4)
kavak (133b 7-8)
kaχša ? (133a 5-6)
kelāže (133b 11-12)
kelāšekere (133b 11-12)
kelendür (132a 1-2)
kemīn (133a 5-6)
kengāǰ (133a 5-6)
kerunǰ (133a 7-8)
keš (134a 1-2)
kešef (131b 9-10)
keyfe (133a 7-8)
keyfer (131b 9-10)
kihīn (133b 7-8)
kilk (133a 5-6)
kiyā (133b 11-12)
kūh (131b 3-4)
kurde (132b 1-2)
kuṣṣa (131b on margin)
kuzāyiš (133a 7-8)

mālīz (133a 9-10)
māγ (133a 9-10)
māγče (133b 5-6)
māšū (133a 9-10)
māχ (134a 7-8)
mebīn ? (133b 13-14)
mevīzāb (134a 9-10)
mīšū (131b 11-12)
mivlāleng (133b 5-6)
mol (133b 9-10)
mustemend (132a 1-2)
mul (132b 12-13)
mūye (133a 3-4)
muždegānī (133a 11-12)
nā-guzīr (133a 11-12)
naγz (133a 11-12)
nāk (133b 7-8)
nāšitā (133a 11-12)
nešīmen (133a 11-12)
nežīm (133a 11-12)
nižāz (134a 7-8)
nuhuft (133b 1-2)
nuhunbīn (133b 1-2)
numūdār (133a 13-14)
numūne (133a 13-14)
nūse (131b 13-14)
nuvāχt (133a 13-14,
nuvāz (133b 1)
pāhu (133b 5-6)
pār (132a 11-12)
pek ve lek (134a 5-6)
pelek (134a 3-4)
perestār (133b 9-10)
perχū (134a 1-2)
pervā (134a 1-2)
pervīzen (133a 9-10)
pežīr (134a 3-4)
pīh (132b 3-4)
pīrār (132a 11-12)
pīšterm (133b 13-14)
poš (134a 1-2)
pūt ? (133a 9-10)
pūyā (134a 5-6)

qāb (132a 1-2)
qoš (131b 13-14)

rānīn (131b 13-14)
renǰ (134a 5-6)
reste (131b 9-10
riḥš (131b 3-4)
rustī (132b 5-6)
rūze (132b 3-4)

sāl (132a 11-12)
sāl-i āyende (132a 11-12)
sefeǰ ? (134a 3-4)
sekāle (131b 5-6)
seme (134a 7-8)
senāb (132b 11-12)
senāber (132b 12-13)
send (132b 11-12)
seng-pošt (131b 9-10)
sevīšem (131b 5-6)
siperī (134a 11-12)
sikāl (132b 11-12)
sitān (133b 9-10)
sitāmī (132a 3-4)
sitūh (133a 1-2)
siyegī (132b 12-13)
supurz (132b 1-2)
surfa (132b 1-2)
surūš (133b 5-6)
sūsmārā (131b 9-10)
sutrug (133b 7-8)
šāδ-kām (132a 7-8)
šāl (133a 1-2)
šān (131b 3-4)
šast (134a 1-2)
šehd (131b 3-4)
šenā (132b 11-12)
šenāγ (132b 11-12)
šenāver ((132b 12-13)
šešχanǰ (134a 7-8)
šeydā (133a 1-2)
šikem (132b 3-4)
šilek ? (134a 3-4)
šināh (132b 11-12)
šīrīn (133a 1-2)
šīrīnī (133a 3-4)
šīrīn-kār(133a 3-4)
šūγ (134a 7-8)
šuš (132b 1-2)

tabār (132a 3-4)
tabīra (133b 9-10)
tāk (133b 5-6)
tārmīχ (133a 11-12)

tek ? (133b 5-6)
teng (132a 13-14)
tīrāže (131b 13-14)
tīve (131b 11-12,
tūy ? (134a 5-6)

uqru'a ? (132a 5-6)

vālā (132a 5-6)
vazaγ (131b 7-8)
verziš (133b 1-2)
vīž (133b 1-2)
vīže (133b 3-4)

χafa (132b 1-2)
χālū (132a 13-14)
χāmī (132a 5-6
χāne (132b 3-4)
χarāšīn (134a 1-2)
χarbuze (133b 3-4)
χarǰenk (131b 7-8)
χārpošt (131b 7-8)
χasvīte (131b 11-12)
χilālūš (133b 11-12)
χišl (131b 3-4)
χorsend (133b 11-12)
χ^{w}onǰīze (133b 13-14)
χurrem (132a 7-8)
χuvāze ? (132a 1-2)

yāδ (132b 7-8)
yāδigār (132b 8-9)
yārsīn (132b 7-8)
yāver (132b 8-9)
yeksān (131b 11-12)
yešk (131b 11-12)
yūγ (134a 7-8

zāǰ (132a 5-6)
zālū (133a 9-10)
zāvilāne (132b 8-9)
zehre (132b 1-2)
zelek (134a 3-4)
zeme (132b 8-9)
zerde-zemīn (132b 8-9)
zīger (132a 3-4)
žāla ? (131b 1-2)
žem (133b 9-10)

Dictionaries and works consulted

Anonymous I. *Farsça-Türkçe lügat.* Istanbul: Suleymaniye Library, Darülmesnevi, No: 581. (Unpublished)

Anonymous II. *Lügat-i Pehlevī.* Istanbul: Suleymaniye Library, Fatih, No: 5163, Hicri 895 [1489 or 1490] (Unpublished)

Bagdādī, Abdülkadir: *Lugat-i Şāh-nāme-i Firdevsī.* Istanbul: Suleymaniye Library: Reisülküttab, No: 1112. (Unpublished)

Burhan-ı Katı' tercümesi. (Mütercim Asım Efendi: Kitab-i Tibyan-i Nafi-yi Tercüme-yi Bürhan-i Katı [by Tabrīzī]), Kostantiniye [Istanbul]: Darü't-tıbaatu's-Sultaniye, 1214 [1799].

Clauson, Gerard 1972. An etymological dictionary of pre-thirteenth-century Turkish. Oxford: Clarendon Press.

Dictionary of most common Avesta words. *Avesta-Zoroastrian Archives* 02.03.2005. http://www.avesta.org/avdict/avdict.htm

Dihkhudā, Alī Akbar 1372-1373 (1993-1995). *Lughatnāmah.* 15 vols. Tihran: Dānishgāh-i Tihrān.

Dilçin, Cem (ed.) 1991. *Mes'ud bin Ahmed: Süheyl ü Nevbahar: İnceleme, metin, sözlük.* Ankara: Atatürk Kültür Merkezi, 1991.

DS = *Derleme sözlüğü.* (Türkiye'de halk ağzından derleme sözlüğü) Ankara: Türk Dil Kurumu. 1-12. 1963-1982.

Ebeling, Erich 1972. Das aramäisch-mittelpersische Glossar Frahang-i-pahlavik im Lichte der assyriologischen Forschung. Osnabrück: Zeller. [reprint of original from 1941].

Eren, Hasan 1999. *Türk dilinin etimolojik sözlüğü.* Ankara: Bizim Büro Basım Evi.

Farahvashī, Bahrām 1352 [1973]. *Farhang-i Pahlavī.* Tihrān: Intishārāt-i Danishgāh-i Tihrān. Glossary of Zoroastrian terms. *Avesta-Zoroastrian Archives.* 02.03.2005.http://www.avesta.org/zglos.html

Halīmī, Lutfullah bin Ebi Yusuf. *Lugat-i Halīmī: Bahru'l-Garāib.* Istanbul: Suleymaniye Library, Husrev Paşa 734. (Unpublished)

Hasan Şu'ūrī 1314 (1897). *Ferheng-i Şu'ūrī, Fārsīden Türkçeye mükemmel lügat.* İstanbul: Cemal Efendi Matbaası.

Inju Shirazi, Husayn ibn Hasan. *Ferheng-i Cihāngīrī.* Istanbul: Suleymaniye Library, Ayasofya, No: K. 4710. (Unpublished)

Kaşgari, Mahmud. *Divanü lûgat-it-Türk tercümesi.* 1-3. Tercüme: Besim Atalay, Ankara: Türk Dil Kurumu. 1941.

Lane, Edward William 1984. *Arabic-English lexicon.* Cambridge, England: Islamic Texts Society.

MacKenzie, David N. 1971. *A concise Pahlavi dictionary.* London, New York: Oxford University Press.

Mashkur, Muhammad Javad 1346 [1967]. *Farhang-i Huzvārishhā-yi Pahlavī.* [Tihrān]: Bunyād va Farhang-i Irān.

Mehmed Es'ad Efendi 1216 (1801). *Kitābi lehcetü'l-lugat.* Kostantiniye: Darü't-Tıbaatü'l-Ma'muret-ül-Sultaniye.

Mo'īn, Muhammad 1371 (1992). *Farhang-i Fārsī.* 6 vols. Tihrān: Amīr Kabīr.

Nakhjivānī, Hindushāh bin Sanjar 1366 (1987). *Sihah al-'Ajam: Nuskha-'i Gazan,* (ed.) *Ghulam Husayn Begdili.* [Tihran]: Vahid.

Oružov, Ä. Ä. (ed.) 1966-1987. *Azärbaycan dilinin izahlı lüğäti.* 1-4. Bakı: Elm Näşriyyatı.

Redhouse, Sir James W. 1890. *A Turkish and English lexicon: Shewing in English the significations of the Turkish terms.* Constantinople: A. H. Boyajian.

Steingass, Francis J. 1957[4]. *Persian-English dictionary.* London: Routledge & Kegan Paul Limited.

Tekin, Talat et al. (eds.) 1995. *Türkmence-Türkçe sözlük.* (Türk Dilleri Araştırmaları 18.) Ankara: Simurg.

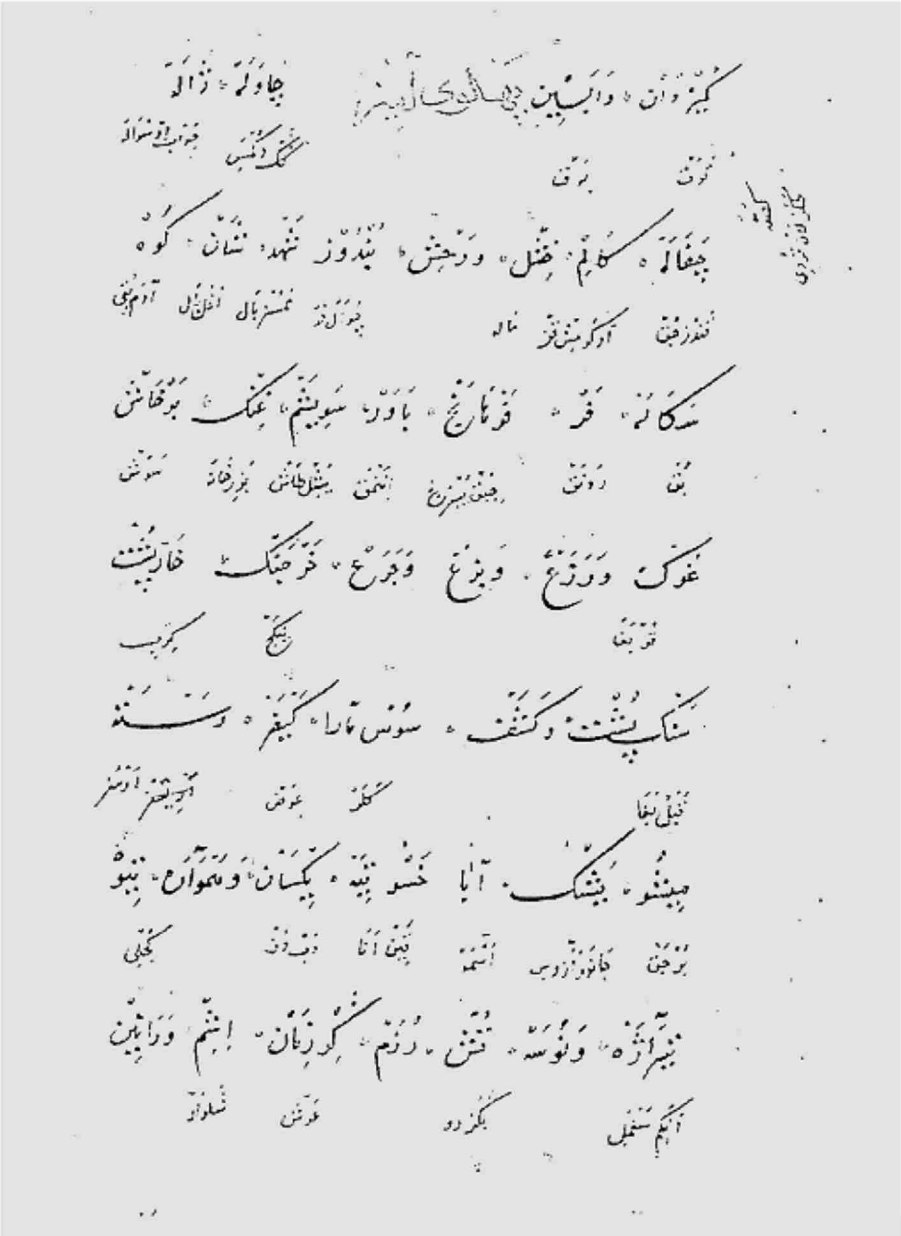
Tietze, Andreas 1999. *Wörterbuch der griechischen, slavischen, arabischen und persischen Lehnwörter im anatolischen Türkisch.* (Türk Dilleri Araştırmaları 26.) İstanbul: Simurg.

TS = *XIII. yüzyıldan beri Türkiye Türkçesiyle yazılmış kitaplardan toplanan tanıklariyle tarama sözlüğü.* 1-8. Ankara: Türk Dil Kurumu. 1963-1977.

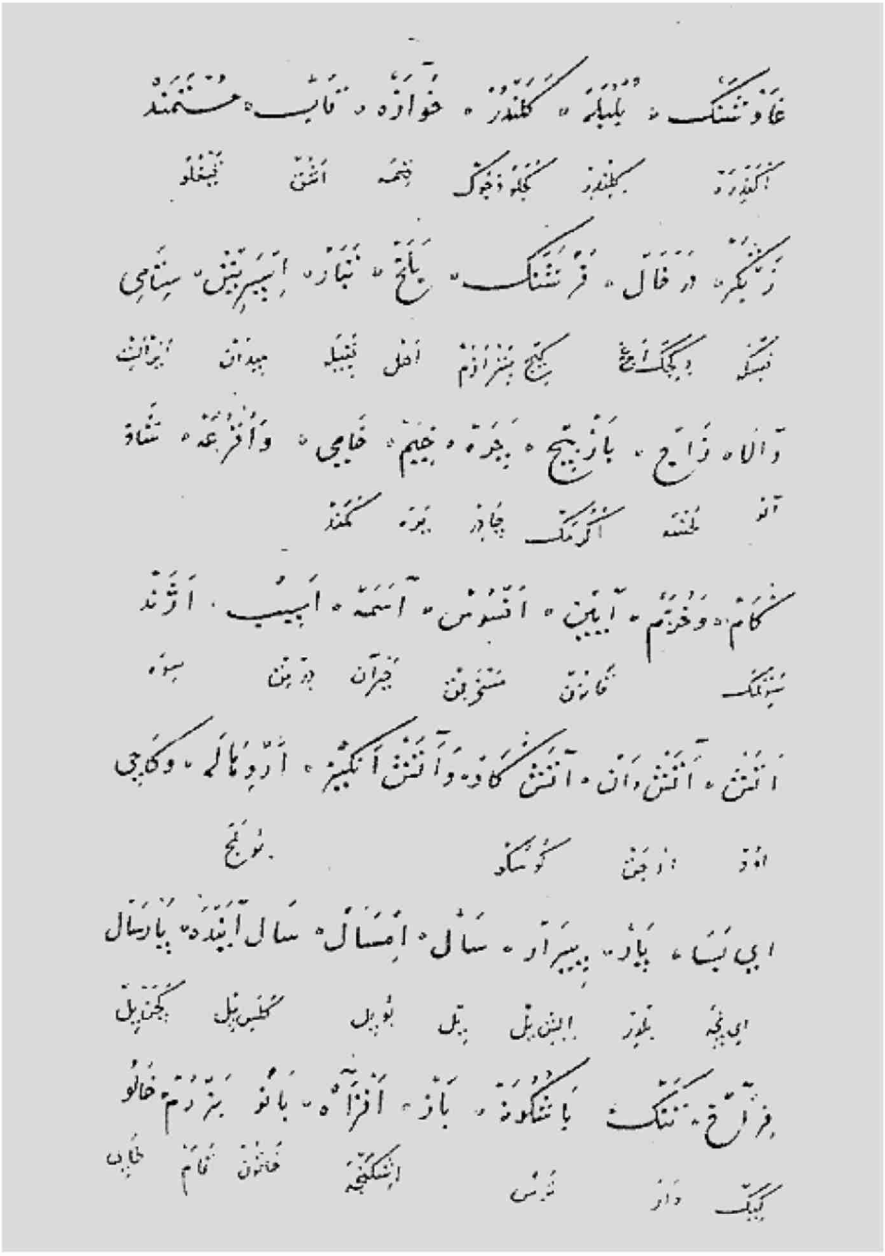
Turan, Fikret 2001. *Eski Oğuzca sözlük: Bahşayiş lügati.* İstanbul: BAY.

Vefik Paşa, Ahmed 1306 (1888). *Lehçe-i Osmānī.* İstanbul: Mahmud Bey Matbaası.

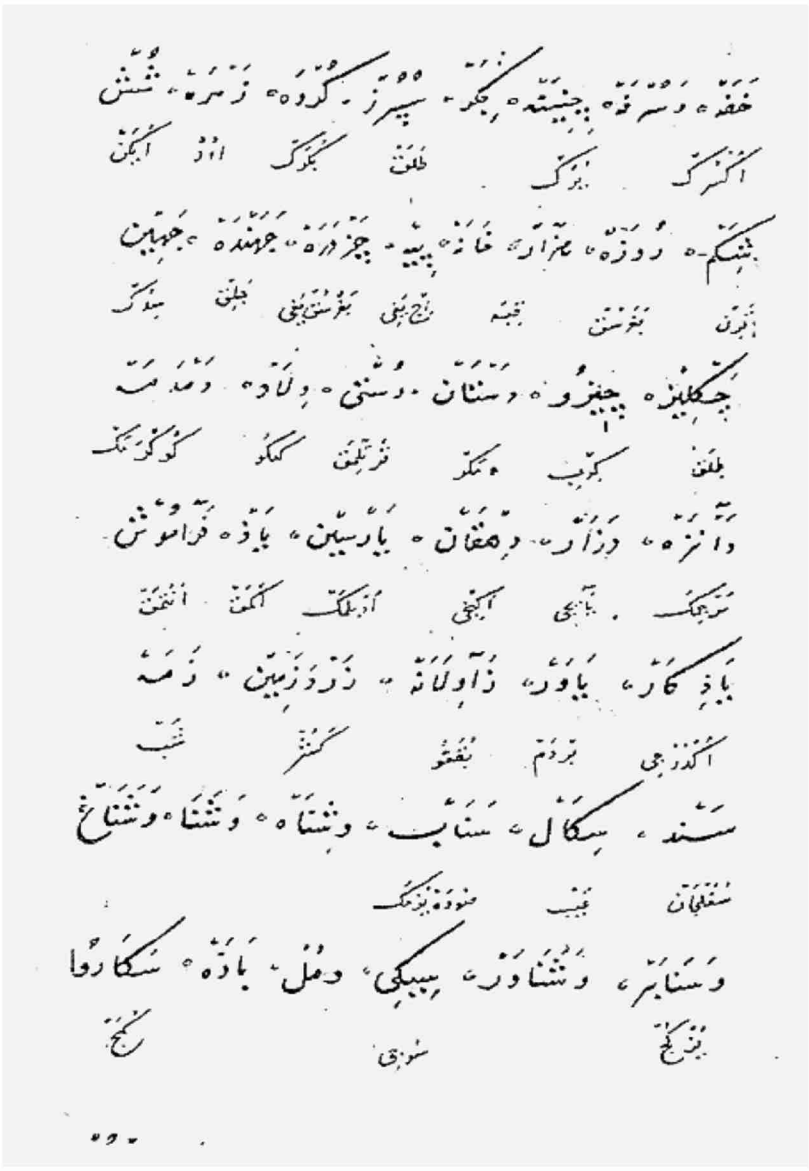
Pahlavī-Āmīz, Suleymaniye Library, Carullah 2043 : 131b



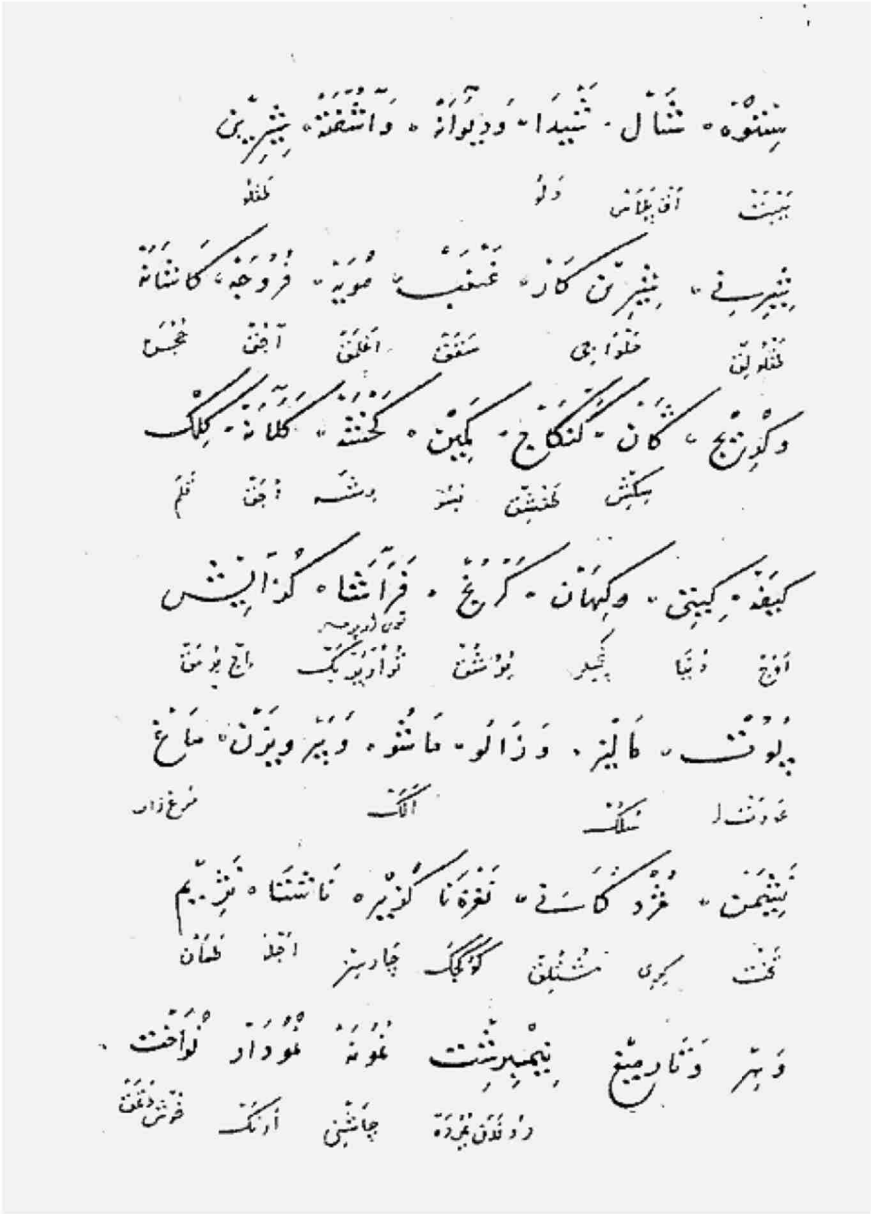
Pahlavī-Āmīz, Suleymaniye Library, Carullah 2043 : 132a



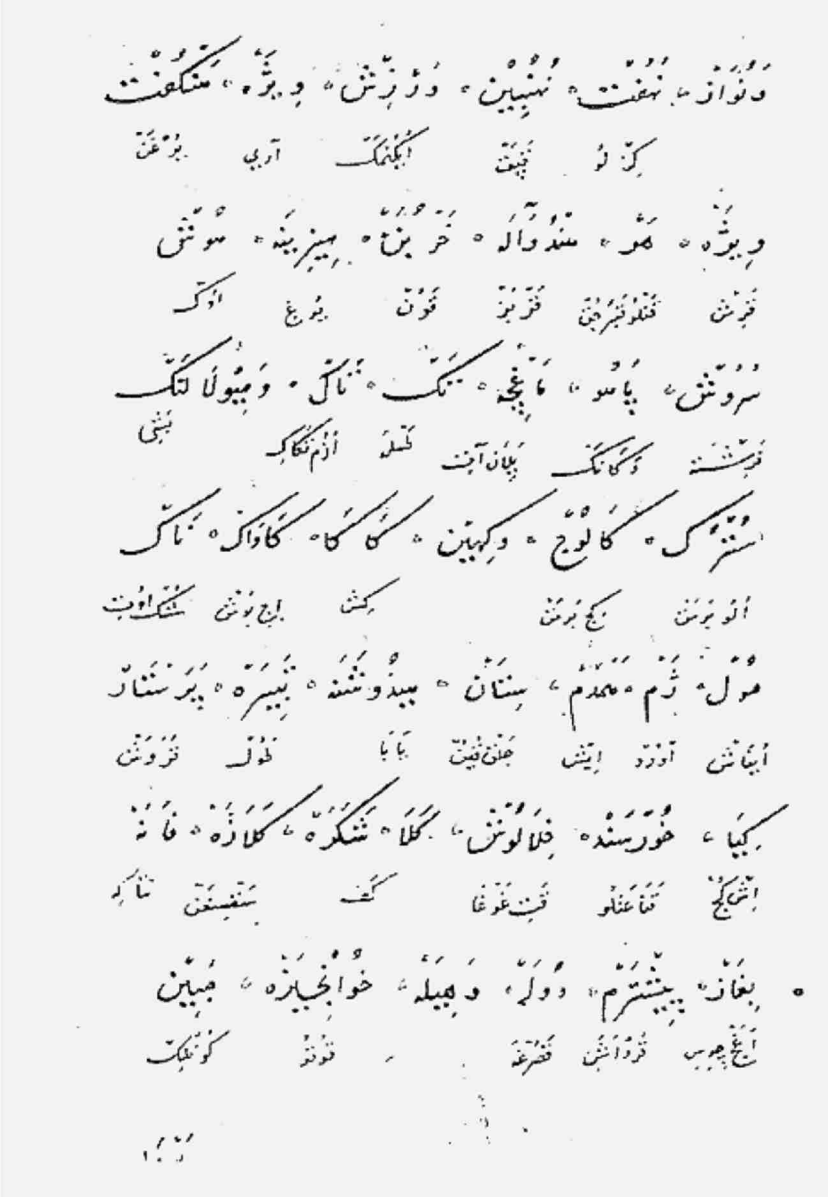
Pahlavī-Āmīz, Suleymaniye Library, Carullah 2043 : 132b



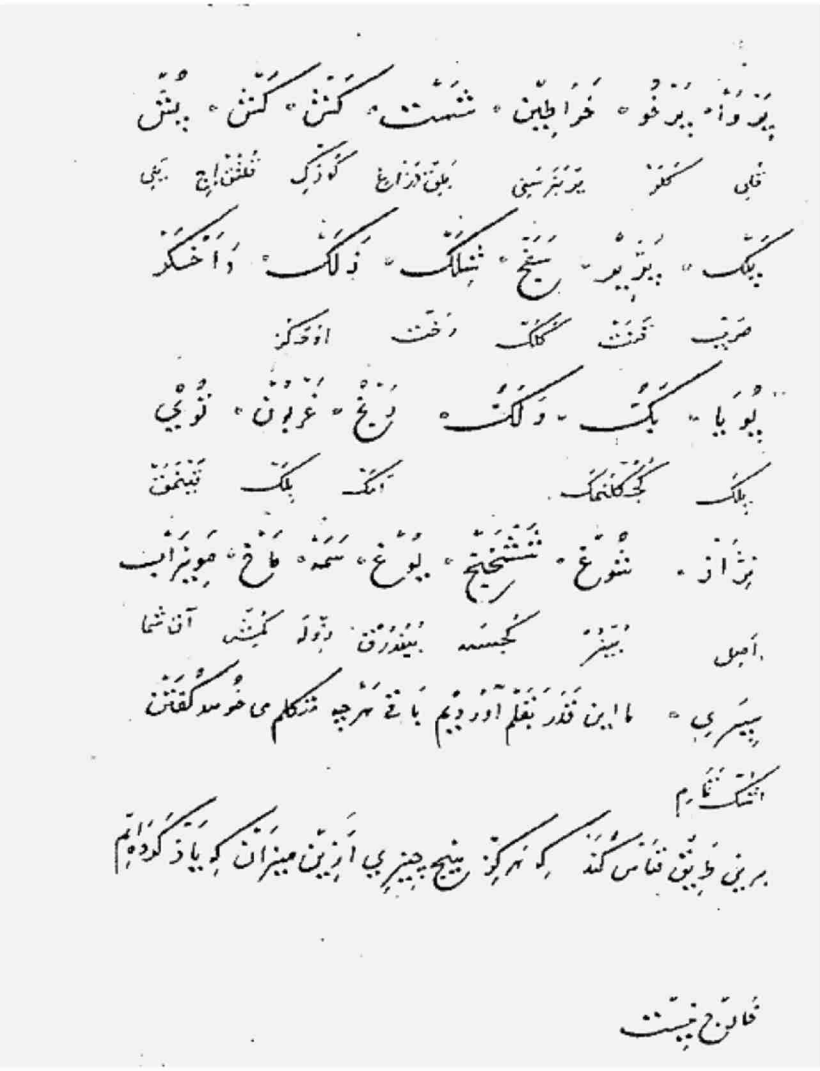
Pahlavī-Āmīz, Suleymaniye Library, Carullah 2043 : 133a



Pahlavī-Āmīz, Suleymaniye Library, Carullah 2043 : 133b



Pahlavī-Āmīz, Suleymaniye Library, Carullah 2043 : 134a



A hierarchical explanation against the distinction of nominal copular sentences in Turkish

Nadir Engin Uzun

Uzun, Nadir Engin 2006. A hierarchical explanation against the distinction of nominal copular sentences in Turkish. *Turkic Languages* 10, 111-130.

A philosophically-based distinction which divides nominal copular sentences into two types, "attributives" and "equatives", and which has been considered in linguistics for a long time has never been studied in Turkish. Only Tura (1986) suggests that "equative-attributive" distinction is also observed in Turkish. In this paper, I will claim that this distinction is not valid for nominal copular sentences, at least in Turkish, and show that the inversion observed in such sentences is not subject to a distinction of sentence types, but to a semantic hierarchy among noun phrase types. Drawing on the data of personal pronouns in copular sentences from Turkish and various other languages, I will also show that the semantic hierarchy seems to be typologically formed.

Ankara University, DTCF Dilbilim Bölümü, Sıhhiye 06100 Ankara, Turkey.
uzun@humanity.ankara.edu.tr

1. Introduction

Various aspects of copular sentences in Turkish have received a great deal of attention over the last decade since the comprehensive work of Lees in 1972 (see Kornfilt 1996, Johanson 2000, Göksel 2001, Kelepir 2002). However, a philosophically-based distinction which divides nominal copular sentences into two types, "attributives" and "equatives", and which has been considered in linguistics for a long time has never been studied in Turkish, except by Tura (1986).

According to Tura (1986), the following examples support a distinction between attributive nominal copular sentences and equative ones:

(1)	*Bu*	*adam*	*bir*	*hırsız*
	this	man	a	thief
	/iyi	*bir*	*doktor*	
	good	a	doctor	
	/çok	*iyi*	*bir*	*doktor-(dur).*
	very	good	a	doctor-(EMM)[1]

'This man is a thief/a good doctor/a very good doctor.'

[1] EMM stands for Emphatic Mood Marker and is used for marking predicative phrases in this paper.

(2) Bizim dekanımız Ahmet Dalmaç(tır).
we-GEN dean-POSS1PL Ahmet Dalmaç-(EMM)
'Our dean is Ahmet Dalmaç.'

While the attributive copular sentence illustrated in (1) cannot be inverted, the equative one in (2) can be inverted easily, as seen in (3) and (4), respectively:

(3) **Çok iyi bir doktor bu adam-(dır).*
very good a doctor this man-(EMM)

(4) *Ahmet Dalmaç bizim dekanımız-(dır).*
Ahmet Dalmaç we-GEN dean-POSS1PL-(EMM)
'Ahmet Dalmaç is our dean.'

In this paper I will claim that 'equative' and 'attributive' distinctions are not valid for nominal copular sentences and show that the inversion observed in such sentences is not subject to a distinction of sentence types, but to a semantic hierarchy among noun phrase (NP) types. Drawing on the data of personal pronouns in copular sentences from Turkish and various other languages, I will also show that the semantic hierarchy seems to be typologically formed.

2. Nominal copular sentences and inversion

Nominal copular sentences (NCSs) consist of two NPs and a copular element. The copular element in NCSs can be located in different syntactic positions cross-linguistically. For example, in the basic word order, the copular verb follows the two NPs in Turkish, whereas this type of verb occurs between the two NPs in English. However, the Irish copular element, which is not a verb, is normally placed first in the sentence. The formulas in (5) show these possibilities and the sentences in (6) exemplify them.

(5) NP_1 NP_2 cop *Turkish*
NP_1 cop NP_2 *English*
cop NP_1 NP_2 *Irish*

(6) Ali doktor <u>idi.</u>
John <u>is</u> the doctor.
<u>Is</u> é Seán an dochtúir.
(The copular elements are underlined.)

Another peculiarity of NCSs is that they can be classified into two types based on the semantic properties of NPs in the first and second positions. The NCS whose first NP is definite/referential but whose second NP is indefinite/nonreferential is called an "attributive copular sentence". The copular sentence in which both NPs are definite/referential is called an "equative copular sentence". (7) shows these relations.

(7)	*Attributive copular Sentence*		*Equative copular Sentence*	
	NP_1	NP_2	NP_1	NP_2
	def-ref	indef/non-ref	def-ref	def-ref

In attributive copular sentences, the secondary NP forming a unit with the copular element is considered a predicative NP. In such sentences the secondary NP attributes describe or classify the first NP. In equative copular sentences, referential sources of both NPs are equalized and the predicative element is the copular element.

The main difference between attributive copular sentences and equative ones is that the former cannot be inverted:

(8) a John is a teacher. *Attributive*
b. *A teacher is John.

In this respect, attributives are like adjectival copular sentences:

(9) a. John is honest. *Adjectival*
b. *Honest is John.

However, equative sentences invert easily:

(10) a. John is the teacher. *Equative*
b. The teacher is John.

In Turkish, where the copular element is the verb *–i* which appears when the tense, mood and agreement suffixes are needed, the picture is the same as in English at first glance. The examples in (11), (12) and (13) show the acceptability of the inverted forms of copular sentences in Turkish:

(11) a. *Ali bir öğretmen-(dir).*
Ali a teacher-(EMM)
'Ali is a teacher.'
b. **Bir öğretmen Ali-(dir).*
a teacher Ali-(EMM)

(12) a. *Ali dürüst-(tür).*
Ali honest-(EMM)
'Ali is honest.'
b. **Dürüst Ali-(dir).*
honest Ali-(EMM)

(13) a. *Ali o öğretmen-(dir).*
Ali that teacher-(EMM)
'Ali is that teacher.'

b. *O öğretmen, Ali-(dir).*
that teacher Ali-(EMM)
'That teacher is Ali.'

Note that the copular verb is absent when the subject is in the third person and the tense is present in Turkish.

On the other hand, when it comes to the presence of personal pronouns in so-called "equative copular sentences", a clear distinction between English and Turkish is apparent. It is not possible to invert a copular sentence in English when the first NP is a personal pronoun:

(14) a. *The teacher am I.
b. *The teacher are you.
c. *The teacher is he.

However this type of sentences inverts easily in Turkish:

(15) a. O öğretmen ben-im.
that teacher I-COP1SG
'That teacher is me.'
b. O öğretmen sen-sin.
that teacher you- COP2SG
'That teacher is you.'
c. O öğretmen o-(dur).
that teacher he-(EMM)
'That teacher is him.'

At this point, again note that the copular verb is not present in (15c) since Turkish has no agreement suffix for the third person singular.

Italian, like Turkish but unlike English, allows inversion in NCSs whose subject is a personal pronoun, a fact observed by Moro (1995) and (1997):

(16) a. Io sono il re della Francia.
'I am the king of France.'
b. Il re della Francia sono io.
'I am the king of France.'

The conjugated form of the copular verb (*sono*) for agreement in (16b) shows that the second NP, not the first one, must be the subject because the morphology of the copular verb belongs to the first person, not to the third.

As we see in (15) and (16), Turkish and Italian display different properties from English when it comes to inversion in equative copular sentences which have a personal pronoun in the first position.

On the other hand, the inversion is not always free in Turkish and Italian. In the examples of Turkish in (17)-(20) where the second position is also filled with a pronoun, inversion is blocked, except for the third person pronoun:

(17) a. Ben sen-im.
I you-COP1SG
'I am you.'
b. *Sen ben-im.
you I-cop1sg

(18) a. Sen ben-sin.
you I-COP2SG
'You are me.'
b. *Ben sen-sin.
I you-COP2SG

(19) a. Ben o-yum.
I he-COP1SG
'I am him.'
b. O ben-im.
he I-COP1SG
'He is me.'

(20) a. Sen o-sun.
you he-COP2SG
'You are him.'
b. O sen-sin.
he you-COP2SG
'He is you.'

In Italian, inversion is blocked whenever the subject is a personal pronoun, including the third person.

(21) a. *Lui sono io
he be-COP1 I
'I am him.'

Moreover, in Turkish, it is impossible to construct a sentence with the third person pronoun in the subject position when the first or second person pronoun is predicative.

(22) a. *O ben(-dir)
he I-(EMM)
b. *O sen(-dir)
he you-(EMM)

Overt regular Turkish NPs also behave similarly to the third person pronoun in copular sentences in this respect.

(23) a. *O öğrenci ben(-dir)
that student I-(EMM)

b. *O öğrenci sen(-dir)
that student you-(EMM)

The picture is the same in Italian with overt NPs. Inversion in this language is constrained if there is an overt NP in the subject position and a pronoun in the predicate position, as in (24):

(24) *Il re della Francia è me
ART king GEN France be-COP3 me
attempted reading: 'The king of France is me.'

However, a pronoun in the predicate position seems to cause no problem in English:

(25) That teacher is me.

The examples given in (14)-(25) make us think that inversion is not a valid criterion to observe the difference between equative and attributive sentences. What is more, some of the equative sentences cannot be inverted at all, and inversion seems to be affected by the type of NPs, including personal pronouns and overt NPs, not the type of sentence in which it occurs.

3. Towards a hierarchical explanation

The first structural analysis of NCSs within the generative framework comes from Tim Stowell. Stowell (1978) and then Stowell (1984) take the copular verb as a raising verb, a verb that does not assign any theta-role and case but selects small sentences as its complement. The first NP takes the external role from the second one and raises to the initial position of the sentence as shown in (26) and (27):

(26) be [John the teacher]

(27) John_i is [t_i the teacher]

This approach, however, does not take any distinction between attributive and equative sentences into account. So, in his analysis the first phrase of the sentence should raise from the subject position to the sentence initial position as seen in (28):

(28) The teacher_i is [t_i John]

Heggie (1987), accepting the presence of equative sentences like *John is that teacher,* says that because *the teacher* in (28) also has an indefinite/non-referential reading, this phrase must have been raised from the predicate position of an attributive NCS. According to her, indefinite/non-referential reading is possible only with a phrase originally located in the predicate position. The most original evidence Heggie (1987) presents is that the focus in the basic order can be either on the first or the second NP, whereas in the inverted structure the focus can only be on the second NP, as seen in (29):

(29) a. JOHN was the teacher.
b. John was the TEACHER.
c. The teacher was JOHN.
d. # The TEACHER was John (with the non-referential reading of *the teacher*)

Based on some tests used in Heggie's work, whose details will not be dealt with here, *the teacher* in (28) does not show any features of an argument because it is originally predicative.

Edwin Williams, who accepts that copular sentences are divided into two types, adjectivals and equatives in Williams (1983), later (see Williams 1997) claims that there are actually no equative sentences. He states that there is an epistemic difference, thus an asymmetry between the two NPs. According to Williams (1997), the idea that the subject is referential but the predicate is not is false, and the predicative phrase should only be considered less referential than the subject phrase.

In fact, it was noted by Heggie in 1988, long before Williams (1997), that there is an asymmetry between the two NPs of NCSs, and semantically the NPs can either be in exact balance or the second one is in a lower position in a semantic hierarchy, as she proposed for English in (30). "Deixis" (actually demonstratives) being at the top of the list are followed by (proper) "names", "definite descriptors" and finally "indefinites", respectively (see Heggie 1988: 106):

(30) deixis > names > definite descriptors > indefinites

Within such a hierarchy, when we reconsider Heggie's (29c), we realize that the non-referential use of the first NP can only be possible when a proper noun is used in the secondary NP. Otherwise, if we reconstruct this sentence as in (31), this reading will disappear:

(31) The teacher is a woman.

This obviously shows that the semantic features of NPs in NCSs affect inversion and related readings.

4. A semantic classification and hierarchy for Turkish NPs

In this section, following Williams' (1997) idea, I propose to put attributives and equatives, that is all NCSs, into the same category of nominal copulars by considering only adjectival copular sentences in (9) and (12) as attributives. According to this proposal, there are no "attributive NCSs" anymore.

And, applying Heggie's hierarchy to Turkish, I claim that whether nominal copulars can be inverted is determined by a semantic hierarchy which is formulated and defined based on semantic properties of NPs (such as definiteness, specificity, genericity) in this language.

Diagram (32) represents a typology including four types of NPs with three distinctive semantic features, and (33) represents the semantic hierarchy derived from this typology.

(32) *A semantic classification for Turkish NPs*

	definite	specific	generic
1st NP- type	+	+	-
2nd NP-type	-	+	-
3rd NP-type	+	-	+
4th NP-type	-	-	+

(33) *The hierarchy for Turkish NPs*

1. 1st and 2nd personal pronoun (*ben, sen*) >
2. definite specific NPs > (including proper nouns and the 3rd personal pronoun *o*)
3. indefinite specific NPs >
4. definite generic NPs >
5. indefinite generic NPs

As can be seen, if one compares the NP-types in (32) and the levels in the hierarchy of (33), 1st and 2nd personal pronouns in Turkish are separate from the rest of the definite-specific NPs, including proper nouns and the 3rd personal pronoun.[2]

Now let's examine the hierarchy studying each NP type in (32) with Turkish NCSs.

(34) a. Pelikan kuş(tur).
Pelican bird(EMM)
'The pelican is a bird.'
b. Kuş pelikan(dır).
c. Pelikan bir kuş(tur).
Pelican a bird(EMM)
'The pelican is a bird.'
d. *Bir pelikan kuş(tur).

[2] For the preliminary discussions of the classification in (32) and the hierarchy in (33), see Uzun 2003.

The sentences in (34a) and (34b), both of whose NPs are definite generic, and the sentence in (34c), whose secondary NP is lower than *pelikan* in the hierarchy, are fine, but if in the first position we put an NP which is hierarchically lower than the secondary NP, the sentence becomes incorrect, as in (35d). This observation states that definite generic NPs are higher in the hierarchy than indefinite generic ones, and the higher one must come first in copular sentences. The indefinite generic NP *bir pelikan* cannot stand at a higher position than *pelikan*, which is a definite generic NP. So, the hierarchy blocks the string with *indefinite generic NP + definite generic NP*.

The difference between the sentences in (35), on the other hand, shows that indefinite specific NPs are at a higher position than definite generic NPs. A sentence with *gizli polis* in the first position and an indefinite specific NP in the second position becomes ill-formed when the first NP takes a definite generic reading, as illustrated in (35b):

(35) a. Bu okulda bir öğretmen, gizli polis(tir).
this school-LOC a teacher secret police-(EMM)
'In this school one of the teachers is a secret police officer.'
b. *Gizli polis, bu okulda bir öğretmen(dir).
c. *Bir gizli polis, bu okulda bir öğretmen.

Of course, the blocking still works in sentence (35c) with an indefinite generic NP in the first position.

However, the incorrectness of (35b) and (35c) does not seem to be only a matter of hierarchy. When we look at (36), we find a more general restriction in Turkish copular sentences: Secondary NPs in nominal copulars can not be specific if there is a specific NP in the first position of a sentence.

(36) *Çocukların dün bahçede yakaladığı bir kuş,
children-GEN yesterday garden-LOC chase-PART-AGR a bird
bu kuş(tur).
this bird(EMM)
'This bird is a bird that the children caught in the garden yesterday.'

Before moving on to examples of first level NPs in the hierarchy, let's consider the second level NPs together with definite generic ones. The examples in (37) illustrate this relation:

(37) a. Pelikan bu kuş(tur).
Pelican this bird-(EMM)
'The pelican is this bird.'
b. Bu kuş pelikan(dır).
'This bird is the pelican.'

The definite generic NP *pelikan* violates the hierarchy standing over the definite specific NP *bu kuş* in (37a), but the sentence is still perfectly grammatical. So, this means that (37a) is the inverted form of (37b).

The fact that both of the NPs can be focused in (37b) while only the secondary NP in (37a) can, strongly supports this claim:

(37') a. *?PELİKAN bu kuş(tur).
Pelikan BU KUŞ.
b. BU KUŞ pelikan(dır).
Bu kuş PELİKAN(dır).

Let us now focus on an example for uninvertible pairs: The indefinite generic NP in (38) violates the hierarchy and as a result, the sentence is ungrammatical. No sort of focusing can make (38) acceptable (see (38')). Skipping two levels of hierarchy should also be an effective factor.

(38) *Bir pelikan bu kuş(tur).
(38') *BİR PELİKAN bu kuş(tur).
*Bir pelikan BU KUŞ(tur).

Here I claim that there is no inversion between NPs having the same status in the hierarchy. At this point, if we return to (34), we can say that neither (34a) nor (34b) is the inverted form of the other. We can observe this better when we compare the definite specific NPs that occur in (39) and (40) to the examples given before:

(39) O öğretmen, Ali(dir).
Ali, o öğretmen(dir).
(40) O, Ali(dir).
Ali, o(dur).

Crucially, we can focus all the NPs freely as illustrated in (39') and (40'), and this proves that there is no inversion among the pairs given:

(39') O öğretmen, ALİ(dir).
O ÖĞRETMEN, Ali(dir).
(40') ALİ, o(dur).
Ali, O(dur).

What is more, (39) and (40) indicate that definite specific NPs, including reference expressions, proper nouns and third person pronouns, do not stand in an asymmetric relationship to each other. This explains perfectly why proper nouns can stand in secondary position in small clauses in Turkish:

(41) a. Ben Ali'yi okuldan bir arkadaşın sandım.
I Ali-ACC school-ABL a friend-POSS2SG consider-PAST-1SG
b. Ben Ali'yi/seni sen/Ali sandım.
I Ali-ACC/you-ACC you/Ali consider-PAST-1SG
c. Ben Ali'yi o avukat sandım.
I Ali-ACC that lawyer consider-PAST-1SG

In contrast, as Rappoport (1995) observed earlier, the secondary position of small clauses in English is reserved only for non specific NPs:

(42 a. *I believe Jones a certain friend of mine.
b. *I proved our professor Riki.
c. *I consider Tali that woman over there.

Let us now consider the NPs at the top of this hierarchy and study this observation. First of all, let us make a quick statement: If the two NPs or two pronouns have different personal features, then the phrase whose feature is higher in the hierarchy determines the agreement. We shall now check this statement with Turkish examples:

(43) a. Ben o öğretmenim(dir).
I that teacher-COP1SG-(EMM)
'I am that teacher.'
b. O öğretmen benim(dir).
that teacher I-(EMM)
'That teacher is me.'
c. *O öğretmen ben(dir).

In (43a), based on the basic word order facts in Turkish, the first person pronoun which occupies the first NP position is the subject of the sentence and the copula must be inflected with the first personal agreement morphology. However, if we put the NPs in reversed order, as in (43b), we see that this time the copula agrees with the NP in the second position, since this inverted phrase is still the subject. The sentence in (43c), on the other hand, illustrates the unacceptability of the third person agreement morphology.

Again, we see that the top-most element determines the agreement whatever its position is in the sentence. The unacceptability of the focus on the first phrase in (43b') shows that this sentence is inverted:

(43b') O öğretmen BENİM(dir).
*O ÖĞRETMEN, benim(dir).

The facts in (43) indicate that *o*, which is traditionally considered the third person pronoun, can not be a genuine personal pronoun. Therefore, I claim that personal

pronouns are divided into two groups, 1st and 2nd personal pronouns, and there is no third personal pronoun in Turkish.

Here is the supposed personal pronoun 'o' and the relationships derived from "the one at a higher position determines the agreement" principle:

(44)	a.	Ben	oyum.	O	benim.
		I	he/she-COP1SG	He/She	I- COP1SG
		'I am him/her.'		'He/She is me.'	
	b.	*Ben	o(dur).	*O	ben(dir).
		I	he/she-(EMM)	He/She	I-(EMM)
(45)	a.	Sen	osun.	O	sensin.
		You	he/she-COP2SG	He/She	you- COP2SG
		'You are him/her.'		'He/She is you.'	
	b.	*Sen	o(dur).	*O	sen(dir).
		You	he/she- (EMM)	He/She	you- (EMM)
(46)	a.	Ben	senim.	Sen	bensin.
		I	you-COP1SG	You	I-COP2SG
		'I am you.'		'You are me.'	
	b.	*Ben	sensin.	*Sen	benim.
		I	you-COP2SG	You	I-COP1SG

5. Cross-linguistic data

I believe that the semantic hierarchy for NPs in copular sentences is cross-linguistically configurated. The peculiarity of languages, such as presence of articles and demonstratives or the subdistinction of personal pronouns and their relations with demonstratives, may vary the hierarchical ordering of NPs. Therefore, it is to be expected that this hierarchy may affect the agreement relations between NPs and copulas differently in different languages.

Heycock and Kroch (1998) explain how the copular verb in Italian agrees with the NP in the second position in Moro's examples by pointing to the scrambling features of this language. They see a connection between the fact that the NP in the second position of copular sentences is the subject and the fact that the subject of a sentence in Italian comes after the verb, unless it is emphatic. This special connection could also be made for Turkish, which has scrambling features as well. However, Heycock and Kroch (1998) would be correct only if no other language that has parallel structures but lacks typical scrambling features existed. In addition, the scrambling feature is not a cline. Some languages, like Turkish, allow scrambling freely; some others, like Italian, are less free, and yet English shows this feature very poorly. How much scrambling is sufficient for inversion?

When we deal with inversion, which has many language-specific features, we are forced to state that the facts must be motivated by lexical sources of languages. At this point, I believe we can explain unexpected inversion aspects of languages through the configuration of lexical categories in a semantic hierarchy.

Let us look at Serbian, which shows parallel behavior with Italian, and Urdu, which shows parallel behavior with Turkish in respect to the personal pronoun hierarchy.[3]

In Italian, personal pronouns are divided into three as in (47), and the copular verb is conjugated in three different forms with these three personal pronouns as in (48):

(47) io ('I')
tu ('you')
lui ('he')

(48) sono ('be' for 1st person singular)
sei ('be' for 2nd person singular)
è ('be' for 3rd person singular)

The personal pronoun which comes first, including also the 3rd person pronoun, determines the agreement in Italian, as seen in (49):

(49) a. Io sono te.
I am you
'I am you.'

b. Tu sei me.
you are me
'You are me.'

c. Lui è me.
he is me
'He is me.'

If we invert the examples in (49), the pronoun in the second position which will be in the first position after the inversion, does not allow the copular verb to agree with the pronoun which will be in the second position after the inversion, contrary to the fact that Italian easily allows this type of configuration with the overt NPs in (16):

(50) a. *Io sei tu.
I are you

b. *Tu sono io.
you am I

c. *Lui sono io.
he am I

d. *Lui sei tu.
he are you

In Serbian, personal pronouns are also divided into three, as in (51), and the copular verb is conjugated according to these personal pronouns as in (52):

[3] For a deeper discussion of cross-linguistic data, see Uzun 2004.

(51) ja ('I')
ti ('you')
on ('he')
(52) sam ('be' for 1st person singular)
ti ('be' for 2nd person singular)
je ('be' for 3rd person singular)

In this language, the personal pronoun which comes first, including also the 3rd person pronoun, determines the agreement, as in Italian:

(53) a. Ja sam ti.
b. Ti si ja.
c. On je(ste) ti/ja.

Again, no personal pronoun in Serbian allows the copular verb to agree with the pronoun in the second position after the inversion, as seen in (54):

(54) a. *Ja si ti.
b. *Ti sam ja.
c. *On sam ja.
d. *On si ti.

However, Serbian, just like Italian, easily allows the copular verb to agree with the personal pronoun in the second position when the first position is filled by an overt NP, as seen in (55):

(55) Doctor sam ja.
Doctor si ti.
Doctor je(ste) on.

These examples indicate that these two languages give all personal pronouns the same place in the hierarchy.

On the other hand, in Urdu, although traditionally there are three personal pronouns as in (56), subject agreement facts are not the same as in Italian or Serbian.

(56) men ('I')
tum ('you')
vo ('he/she/it')
(57) hun ('be' for 1st person singular)
ho ('be' for 2nd person singular)
he ('be' for 3rd person singular)

Urdu, like Turkish, puts the third person pronoun in a different place from the other personal pronouns. Agreement facts in NCSs are the same as in Turkish in every aspect, as in (58):

(58) a. men tum hun. *men tum ho.
I you am I you are
'I am you.'
b. Tum men ho. *Tum men hun.
you I are you I am
'You are me.'
c. Vo men hun. *Vo men he.
he I am he I is
'He is me'
d. Vo tum ho. *Vo tum he.
he you are he you is
'He is you'

Thus, in my opinion, Urdu has only two personal pronouns for the first and second persons, and the pronoun (*vo*) is not in the same place in the hierarchy as the other pronouns. Urdu has two and Turkish has three demonstrative pronouns to mark different spatial positions. One of these pronouns (underlined in (59)) is also used as a personal pronoun in these languages:

(59) ye 'this' - vo 'that' *in Urdu*
bu 'this' şu 'that' o 'that' *in Turkish*

Urdu and Turkish use one of the demonstrative pronouns as the third person pronoun. At this point, a piece of evidence comes from the demonstrative classifications in the languages discussed here. There are specific demonstrative pronouns independent from the third personal pronoun in Italian and in Serbian.

(60) questo 'this' codesto 'that' quello 'that' *in Italian*
ovo 'this' to 'that' ono 'that' *in Serbian*

Above all, the hierarchy may require some unfamiliar categories. Korean is one of the languages whose hierarchy shows a very special category.

In Korean, we find demonstrative pronouns independent of personal ones, so we expect that the subject NP can not be inverted if it is not higher than the predicative NP in the semantic hierarchy. But the picture is not that clear. First, let us look at the personal pronouns in Korean:

(61) ne ('I')
neo ('you')
gu ('he/she/it')
oori ('we')
dangshin ('you')
gudul ('they')

In addition, this language has independent demonstrative pronouns, as seen in (62):

(62) i ('this')
cheo ('that')

But, Korean has no personal agreement system realized on predicative elements. Therefore, there is no way to observe the inversion relations between the NPs in an NCS:

(63) a. Neka sonsengida.
I-NOM teacher:COP
'I am a teacher'
b. Neoka sonsengida.
You- NOM teacher: COP
'You are a teacher'
c. Guka sonsengida.
He/She/It-NOM teacher: COP
'He/She/It is a teacher'
d. Oorika sonsengida.
we- NOM teacher: COP
'We are a teacher'
e. Dangshini sonsengida.
You(PL)-NOM teacher: COP
'You are a teacher'
f. Guduli sonsengida.
They-NOM teacher: COP
'They are a teacher'

As seen in (63), the copular element (*i-da*) never changes due to different pronouns in the subject position, so personal agreement relations say nothing about inversion. At this point, evidence for inversion may come from another area of agreement in this language: "honorifics".

In Korean, when the subject is an honorific, the suffix *–sshi* is added to the predicative phrase preceding the copular element. Thus, in a copular sentence carrying this honorific suffix, we can say that there is inversion if we see an honorific lexical element in the second position, and a non-honorific element in the first position.

In (64a), the subject is not an honorific, so the copula does not need a suffix for the honorific. Therefore, the honorific suffix makes (64b) ungrammatical:

(64) a. Dodogi neoida
thief-NOM you-COP
'The thief is you'
b. *Dodogi neoisshida
thief-NOM you:HON-COP

But if we keep *dodog* 'thief' in the first position and put the honorifical word *songseng* 'professor' in the second position, we see that the honorific suffix may come before the copular element, as seen by comparing the sentences in (65):

(65) a. Sonsengi dodogi(sshi)da
teacher-NOM thief-(HON-)COP
'The teacher is the thief'
b. Dodogi sonsengi(sshi)da
thief-NOM teacher-(HON-)COP
'The thief is the teacher'

Of course, the sentences in (65) will sound more natural if the copular element is replaced by the elements of exclamation as seen in (66):

(66) Dodogi sonsengi(sshi)-ne/ya/guna
thief teacher-(HON-)EXC
'The thief is the teacher!'

At this point, a problematic case appears when the sentences in (64) and (65) are put together: The sentences in (64) show us that Korean is similar to English in terms of inversion involving the personal pronouns and that this language puts its personal pronouns, including the third person, in a higher position in the hierarchy than the NPs, but the sentences in (65) unexpectedly show us that Korean is similar to Turkish or Italian in terms of agreement facts involving honorific NPs and that this language does not block inversion when the copular sentences have honorific NPs. Korean puts honorific NPs in the hierarchy as a particular lexical category, not only in respect to agreement, but also inversion.

6. Additional evidence for the hierarchical explanation

Lastly, let us see how a unitary approach to nominative copulars —considering all of them under one category— can account for an observation that Özsoy (2001) made following Contreras (1995).

Özsoy (2001) points out that small clauses in raising structures in Turkish behave differently with regard to the type of copular they include. Adjectival copulars in small clauses can construct a sentence while nominatives can construct only complex predicates, not sentences. In the same vein, we can also say that adjectivals can undergo negation and obtain a governing domain while nominals cannot, as exemplified in (67b):

(67) a. Herkes [beni mutlu değil] sanıyor.
everybody I-ACC happy not consider-PRES3SG
'Everybody considers me not to be happy.'

b. *Herkes [beni avukat değil] sanıyor.
everybody I-ACC lawyer not consider-PRES3SG
'Everybody considers me not to be a lawyer.'

Obviously, it is not possible to explain why (67b) is ungrammatical when we accept the traditional distinction for copular sentences as equative or attributive, and the idea that attributives include the ones whose secondary position is occupied by an NP as the embedded clause, as illustrated in (67b). The small clause in (67b) is actually an attributive NCS according to the definition in (7), so it is not different from the small clause in (67a) in respect to copular type. This case brings a question to mind: What makes (67a) grammatical, but (67b) ungrammatical?

At this point, assuming that a widely observed feature of NCSs occurs in Turkish, too, it becomes clear why nominal copulars cannot construct a sentence.

In Russian, attributive sentences have an optional copular verb while equatives have an obligatory pronominal. In Yiddish, copular sentences in the present tense do not have a copular verb, and attributives have an optional agreement pronoun, but this agreement pronoun is obligatorily used in equatives. English has an optional copular verb in small clauses whose secondary position is occupied by an adjective; however, it is obligatorily used when the small clause is equative:

(68) I consider the winner (to be) a good runner.
I consider the winner *(to be) Mary.

All of the above observations indicate that the equative nominal copular sentences have a copular element obligatorily; that is, the inflectional category is obligatory for this type of sentence. Assuming that this obligatory mechanism works in Turkish for "equative" copular sentences, for the NCSs within the framework of this paper, we can say that the NCS in (67b) can not construct a sentence since it does not have an inflectional category.

The ungrammaticality of (67b) also tells us that this copular sentence is still an NCS even if the secondary NP is indefinite, and that all NCSs are monosemic in Turkish.

7. Conclusion

This paper claims that inversion is not a criterion for the distinction of NCSs as attributives and equatives and that inversion facts in NCSs are determined by a semantic hierarchy.

This claim also states that only those grammatically well-formed sentences that violate the hierarchy should be considered as inverted. Therefore, I reject the analysis argued by Tura (1986) in which she proposes an inversional relation between (2) and (4), repeated in sentences a and b in (69):

(69) a. Bizim dekanımız Ahmet Dalmaç(tır).
'Our dean is A. D.'
b. Ahmet Dalmaç bizim dekanımız(dır).
'A. D. is our dean.'

According to my claim, these two sentences are not only in their own basic word order, but also independent from each other, i.e., they are not derived from one another.

Moreover, this paper shows that the unexpected behavior of personal pronouns in terms of inversion can be explained based on how languages rank their lexical categories in this semantic hierarchy, and not on the scrambling features of these languages. This is where cross-linguistic variations such as the Korean data related to honorifics are especially significant since they draw attention to the need for examining more languages.

References

Göksel, Aslı 2001. The auxiliary verb at the morphology-syntax interface. In: Erguvanlı Taylan, Eser (ed.) 2001. *The verb in Turkish.* (Linguistik aktuell = Linguistics today 44.) Amsterdam, Philadelphia: John Benjamins. 151-181.

Heggie, Lorie 1987. A unified approach to copular sentences. In: Borer, Hagit (ed.) 1987. *Proceedings of WCCFL 7.* Stanford, CA: Stanford Linguistics Association. 129-142.

Heggie, Lorie 1988. The syntax of Copular Structures. Unpublished doctoral thesis, University of Southern California.

Heycock, Caroline & Kroch, Anthony 1998. Inversion and equation in copular sentences. In: Alexiadou, Artemis & Fuhrop, Nanna & Law, Paul & Kleinhenz, Ursula (eds.) 1998. *Papers on syntax of clefts, pseudo-clefts, relative clauses, and the semantics of presents perfect.* (ZAS Papers in Linguistics 10.) Berlin: ZAS. 71-78.

Johanson, Lars 2000. Traces of a Turkic copula verb. *Turkic Languages* 4, 235-238.

Kelepir, Meltem 2002. Olmak, var, yok ve değil. Paper presented in the National Conference on Turkish Linguistics (XVI), Hacettepe University, Ankara.

Kornfilt, Jaklin 1996. On copular clitic forms in Turkish. In: Alexiadou, Artemis & Fuhrop, Nanna & Law, Paul & Löhken, Sylvia (eds.) 1996. *Papers on clitics.* (ZAS Papers in Linguistics 6) Berlin: ZAS. 96-114.

Lees, Robert B. 1972. The Turkish copula. In: Verhaar, John W. M (ed.) 1972. *The verb 'be' and its synonyms: philosophical and grammatical studies. 5, Urdu, Turkish, Bengali, Amharic, Indonesian, Telugu, Estonian.* (Foundations of language. Supplementary series 14.) Dordrecht: Reidel. 64-73.

Moro, Andrea 1995. Small clauses with predicative nominals. In: Cardinaletti, Anna & Guasti, Maria Teresa (eds.) 1995. *Small Clauses.* (Syntax and Semantics 28.) San Diego, London: Academic press. 109-132.

Moro, Andrea 1997. *The raising of predicative. Predicate noun phrases and the theory of clause structure.* Cambridge: Cambridge University press.

Özsoy, A. Sumru 2001. On 'small' clauses, other 'bare' verbal complements and feature checking in Turkish In: Erguvanlı Taylan, Eser (ed.) 2001. *The verb in Turkish.* (Linguistik aktuell = Linguistics today 44.) Amsterdam, Philadelphia: John Benjamins. 213-237.

Stowell, Tim 1978. What was there before there was there. *Chicago Linguistics Society 14*, 458-471.

Stowell, Tim 1984. Subjects across categories. *The Linguistic Review 2*, 272-285.

Tura, Sabahat S. 1986. Definiteness and referentiality in Turkish nonverbal sentences. In: Slobin, Dan I. (ed.) 1986. *Studies in Turkish linguistics*. Amsterdam: John Benjamins. 165-194.

Uzun, N. Engin 2003. Türkçenin koşaç tümceleri üzerine. A paper presented at the 17th National Conference on Turkish Linguistics, Anadolu University, Ankara.

Uzun, N. Engin 2004. Inversion and personal pronouns in nominal copular sentences. A paper presented at the ICKL Conference, Çolaklı.

Williams, Edwin 1983. Semantic vs. syntactic categories. *Linguistics and Philosophy* 6, 423-446.

Williams, Edwin 1997. Asymmetry of predication. In: Blight, Ralph C. & Moosaaly, Michelle J. (eds.) *The syntax and semantics of predication*: *Proceedings of the 1997 Texas Linguistic Society Conference*. (Texas Linguistic Forum 38.) Austin, Texas: Department of Linguistics, University of Texas at Austin. 323-333.

The comprehension of humorous texts by Turkish first graders

Nalân Kızıltan

Kızıltan, Nalân 2006. The comprehension of humorous texts by Turkish first graders. *Turkic Languages* 10, 131-148.

This study will be of great help for language acquisition studies since it discusses to what extent children's cognitive development affects their comprehension of humorous texts.
In this study, two groups of 60 Turkish first graders in Turkey have been tested to determine to what extent they are able to understand six Nasreddin Hodja jokes through reading and listening skills. Although the first graders are significantly better at understanding the jokes through reading skills than listening skills, due to the design of the jokes, none of the groups has shown significant success in comprehending the jokes.
In conclusion, humorous texts may not be comprehended well by seven-year-olds because these texts have a structure specific to humor, in which the paradox between superstructure and substructure causes laughter and explosiveness. Since they are beginning readers and at the very beginning of the Concrete Operational Period, they can be said to have failed to comprehend the punch line hidden in each joke. Therefore, the results of this study may be taken into account in first language acquisition studies. It is important to point out that texts chosen for beginning readers should match their current language development.

Nalân Kızıltan, Ondokuz Mayıs University, Faculty of Education, Department of English Language Teaching. Kurupelit/Samsun, Turkey, kiziltan@omu.edu.tr

1. Introduction

Books are a great medium for introducing the world, imagined or real, to children and teaching them how to cope with its problems and dangers. They help children to develop their language skills. Books can expand vocabulary and organize more cohesive and coherent texts. Glazer (1991: 103) holds that "children's language develops best in an environment where mature language is heard, where children have many opportunities to communicate with others, and where language is presented in context".

Children learn language by constructing for themselves the grammar of the language they hear. They need to hear stories and see print in their environments. Children can be exposed to adult language in literature, hearing varied and complex syntactical structures. Therefore, through textbooks, children are exposed to some literary materials. However, the difficulty level of the language in those literary materials must be taken into account. Literary texts may be full of some incomprehensible in-

put that uses unknown and ambiguous vocabulary items and compound complex sentences. In Turkish primary schools, this difficulty level is adjusted according to the Bulletin of Notices (Tebliğler Dergisi) published by the Turkish Ministry of National Education, which puts forth the educational objectives in elementary schools. For instance, in the first grade, teachers aim to help students

1. develop the ability to listen to the teacher or a story for 3-5 minutes,
2. learn to write words, and take dictation what they have listened to,
3. be interested in learning to read,
4. pronounce the vocabulary items accurately according to Turkish phonetic rules. (Kızıltan 1995)

Since reading plays an important role in the development of children's language within the school program, reading texts are chosen according to the criteria given by the Bulletin of Notices (Tebliğler Dergisi). According to the Bulletin of Notices, reading texts are chosen, taking the following criteria into consideration.

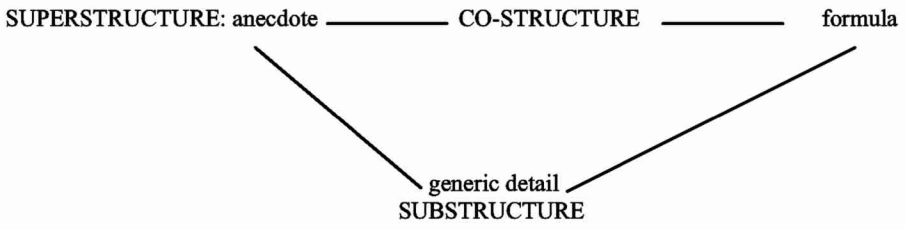
1. Texts must be suitable for children of this age group; they must appeal to children's interests and pleasures.
2. Texts must be good examples of Turkish literature.
3. Texts must be appropriate to the subjects of other courses students take. Appropriateness is valid only for the choice of vocabulary items (Kızıltan 1995).

Although reading texts are chosen according to the above criteria, they may not be appropriate for children's present literacy level or appeal to their immediate interests in reading or listening. For example, humorous texts may be problematic for first graders since humorous texts are different from other narratives in terms of syntactic and semantic structures. Nash (1986: 1-13) points out that humorists compress the soul of wit, word-play and banter; slogans and captions and catchwords; allusion and parody; ironies; satires.

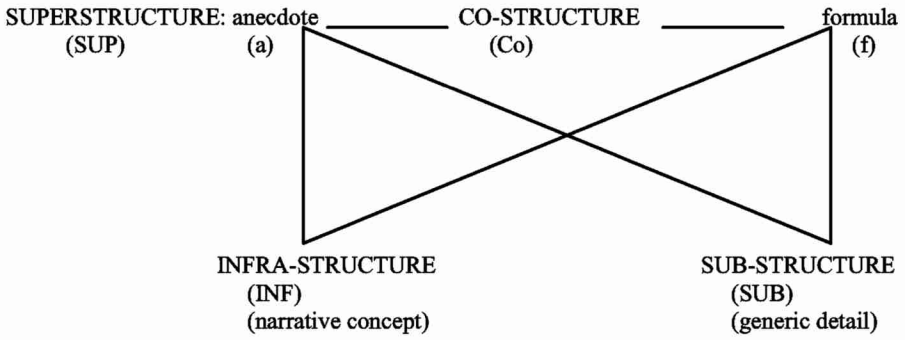
First graders may not comprehend witty compression or comic expansion properly since their reading skills are still in the process of emerging. Listening skills may help them to understand the culture of the joke, the natural facts, logic and language of jokes if their schemata, which is knowledge of the world activated from memory represents the children's knowledge about the underlying concepts, objects, situations, events, sequences of events, actions and sequences of actions.

Özünlü (1992: 194), referring to Raskin (1985) describes "two types of texts one of which is in the superstructure and the other of which is in the substructure. If the elements in both texts do not overlap, the paradox between these texts cause laughter and explosiveness".

Nash (1986: 32) presents a simple structure of humor, assuming that the superstructure is some kind of formula, and that the substructure consists of a few relevant generic details.



Nash (1986: 33) also presents a complex structure of humor:



> "A complex piece of humor involves diverse relationships between what is overt—i.e. the texts—and what is covert—i.e. the generic detail and the narrative concepts" (Nash 1986: 33).

Nash (1986: 34-36) explains design of the joke in two main parts: Pre-location, which is some sort of preparation for the discharge of the joke and location, which clinches or discharges the joke. Pre-location embodies three elements: the *signal* that a joke is intended, the *orientation,* the type of joke, and the *context* in which the joke operates, whereas location embodies locus, which is a word or phrase clinching or discharging the joke. Witty compression is accomplished through foregrounding, parallelism, repetition and deviation (Özünlü 1992: 199).

On the other hand, from a pragmatic point of view, if Grice's cooperative principles such as quantity, quality, relation and manner are violated in the text, they may create locus. Raskin (1985: 56) considers presuppositions, implicature and speech acts as humorous elements (quoted in Özünlü 1999: 140). Besides, if any one of Halliday's (1975) seven basic functions—the instrumental, the regulatory, the interactional, the personal, the heuristic, the imaginative or the representational—is violated, it may cause laughter and explosiveness.

When the afore-mentioned elements are taken into account in a humorous text, children who are said to be in Piaget's (1995) concrete operational period may not

understand the logic and likelihood of a joke. Although children begin to use rudimentary logic and problem solving during the concrete operational period, which covers roughly the ages of seven to twelve, if they are not broadly informed before being exposed to the humorous texts, they are unable to understand the humor. In humorous texts, if the children are exposed to a context in which the punch line is not comprehended, the input may not be comprehensible for them according to Krashen (1985), since it is beyond the learner's current level of competence in the language. The input must be slightly beyond the learner's current level of competence in the language (what Krashen (1985) calls 'i+1') for both comprehension and acquisition to occur (Lightbown & Spada 2000: 39). Therefore, the context should be defined through the students' schemata, their world knowledge. In this way, children may be able to understand the joke regardless of the complex structure of humor in which the superstructure is made up of both the infra-structure and sub-structure of humor.

Recognizing the importance of jokes in the cognitive development of children, textbooks for children include some jokes in their selections. But if children are not able to read closely an unfamiliar literary text of above-average difficulty with good comprehension of its content and literary characteristics, the joke goes nowhere, and it may cause some comprehension problems. Therefore, in this study, the comprehension of the humorous texts by first graders who have just learned how to read has been tested using six Nasreddin Hodja jokes.

The aim of this study is to examine whether the first graders are competent enough to comprehend the humorous texts.

In order to understand to what extent these jokes match the level of the first graders' language development, the study addresses the following questions:

1. Is there a significant difference between the overall comprehension scores of the first graders according to reading and listening skills?
2. Are the first graders successful in comprehending the jokes in terms of the reading and listening skills?
3. Do the first graders show significant comprehension differences between the various jokes?
4. What types of questions are most frequently answered correctly by the first graders in reading and listening comprehension skills?
5. Which jokes are understood better by the first graders?
6. How competent are the first graders in comprehending Turkish humorous texts which conform to their current level of competence?

2. Method

2.1. Sources of Data

The research is based on sets of data from 120 first graders in Samsun Atatürk Elementary School, whose students come from middle class families.

2.2. Data Collection

As humorous texts, six Nasreddin Hodja jokes in the first graders' Turkish textbooks were chosen: "The quilt is gone and the fight is over", "The cauldron died", "Do you believe a donkey's word rather than mine?", "What day is it today?", "Let the quilted turban read", "One who gives the money".

During the summer semester of the academic year, the first graders were divided into two groups. Each joke was given to both groups in the respective Turkish lesson. To assess listening skills, the researcher read aloud the above given jokes to 60 first graders twice, and the 60 first graders were then asked to read these jokes twice silently. After the first graders had read the jokes, the written texts of the jokes were collected. Comprehension of each joke was tested by a series of questions designed to detect to what extent the first graders had understood the jokes with regard to reading and writing skills. Although both groups were given the written questions prepared by the researcher on separate sheets beforehand, each question was read out loud and the students were given enough time to write their answers on the sheets. Then the answer sheets were collected. 41 questions were given in total.

Computerized evaluation of the data received from the questions was conducted research assistant Mustafa Caner and the Department of Statistics at Ondokuz Mayıs University. The programs used for this purpose were SSPS 10-0, Minitab 13-0 and Microsoft Excel.

T-tests and proportion tests were run to explore possible differences between the comprehension scores of the first graders according to reading and listening skills.

3. Findings

The findings are given in the order of the research questions.

1. Comprehension scores of the first graders according to reading and listening skills.

To test whether there is a significant difference between the comprehension scores of the first graders according to reading and listening skills, a t-test was conducted. The results are as follows:

Table 1: Difference-of-means between comprehension scores of the first graders according to reading and listening skills.

Skill	N	Mean	t-Test	P-Value
Reading	60	0,495		
Listening	60	0,432	2,15	0,034

$P=0.034<\alpha=0.05$

The first graders' comprehension scores show a significant difference between reading and listening skills. According to these scores, the first graders' reading skills are better than their listening skills.

2. The first graders' comprehension scores indicate that they have not understood the jokes well in terms of either reading or listening skills. The mean comprehension score for each skill is as follows:

$\text{Mean}_{\text{Listening}} = 0.432$
$\text{Mean}_{\text{Reading}} = 0{,}495$

According to these results, the first graders are better in comprehending jokes in terms of reading than in terms of listening. But none of them indicated success in fully comprehending the jokes.

3. The comprehension of each joke according to reading and listening skills.

a) "The quilt is gone and the fight is over"

In Table 2, the t-test results of the first graders according to reading and listening skills have been given.

Table 2. Difference-of-means between comprehension scores of the first graders according to reading and listening skills

Skill	N	Mean	t-Test	P-Value
Reading	60	0,551		
Listening	60	0,468	2,12	0,036

$P=0.036<\alpha=0.05$

As can be seen, the difference between the comprehension scores of the first graders according to reading and listening skills is significant. It may be concluded that the first graders are more successful in reading skills than they are in listening skills.

b) "The cauldron died"

In Table 3, the t-test results of the first graders according to reading and listening skills have been given.

Table 3. Difference-of-means between comprehension scores of the first graders according to reading and listening skills

Skill	N	Mean	t-Test	P-Value
Reading	60	0,490		
Listening	60	0,438	1,33	0,185

$P=0.185>\alpha=0.05$

As can be seen, the difference between the skills in this respect is insignificant.

c) "Do you believe a donkey's word rather than mine?"

In Table 4, the t-test results of the first graders according to reading and listening skills have been given.

Table 4. Difference-of-means between comprehension scores of the first graders according to reading and listening skills

Skill	N	Mean	t-Test	P-Value
Reading	60	0,406		
Listening	60	0,464	-1,63	0,106

P=0.106>α=0.05

According to the above results, it may be said that the difference between the skills is not significant.

d) "What day is it today?"

In Table 5, the t-test results of the first graders in terms of reading and listening skills have been shown.

Table 5. Difference-of-means between comprehension scores of the first graders according to reading and listening skills

Skill	N	Mean	t-Test	P-Value
Reading	60	0,457		
Listening	60	0,342	3,18	0,002

P=0.002<α=0.05

The first graders seem to have comprehended the joke better by means of their reading skills. The difference is significant.

e) "Let the quilted turban read"

In Table 6, the t-test results of the first graders in terms of reading and listening skills have been presented.

Table 6. Difference-of-means between comprehension scores of the first graders according to reading and listening skills

Skill	N	Mean	t-Test	P-Value
Reading	60	0,475		
Listening	60	0,369	2,57	0,011

P=0.011<α=0.05

As the results indicate, the first graders have comprehended the joke significantly worse through listening skills than reading skills.

f) "One Who Gives the Money"

In Table 7, the t-test results of the first graders in terms of reading and listening skills have been given.

Table 7. Difference-of-means between comprehension scores of the first graders in terms of reading and listening skills

Skill	N	Mean	t-Test	P-Value
Reading	60	0,587		
Listening	60	0,514	1,73	0,086

P=0.086>α=0.05

As can be seen, the first graders have performed more highly in getting the meaning of the joke through reading skills. The difference between the comprehension skills is significant.

4. Tables 8-11 illustrate comprehension scores in listening and reading skills for each question. (black indicates wrong answers, grey indicates incomplete answers and white indicates correct answers).

Table 8. Frequency of the comprehension scores according to the questions of the first three jokes in reading skills

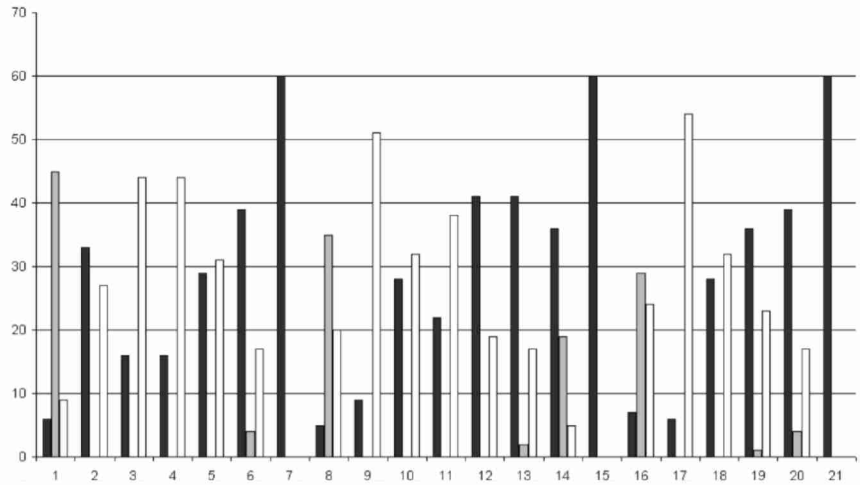


Table 9. Frequency of the comprehension scores according to the questions of the last three jokes in reading skills

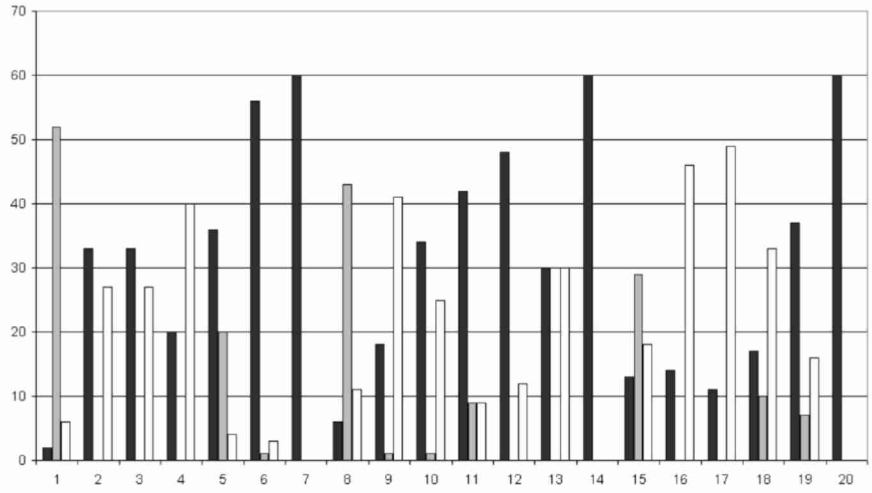


Table 10. Frequency of the comprehension scores according to the questions of the first three jokes in listening skills

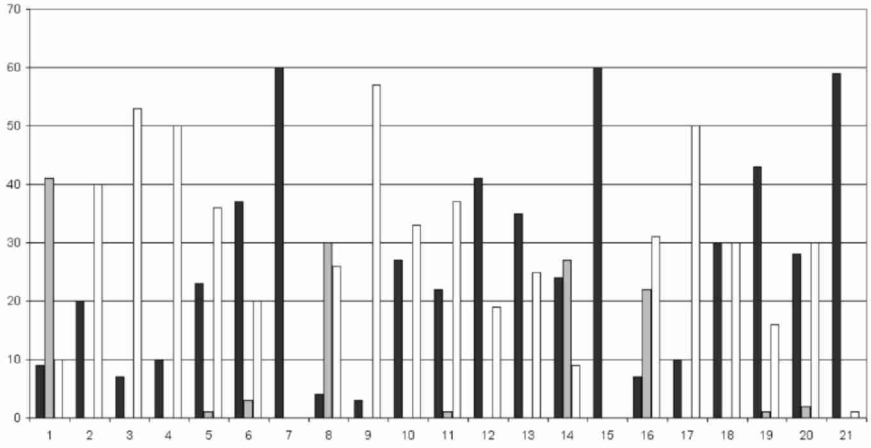
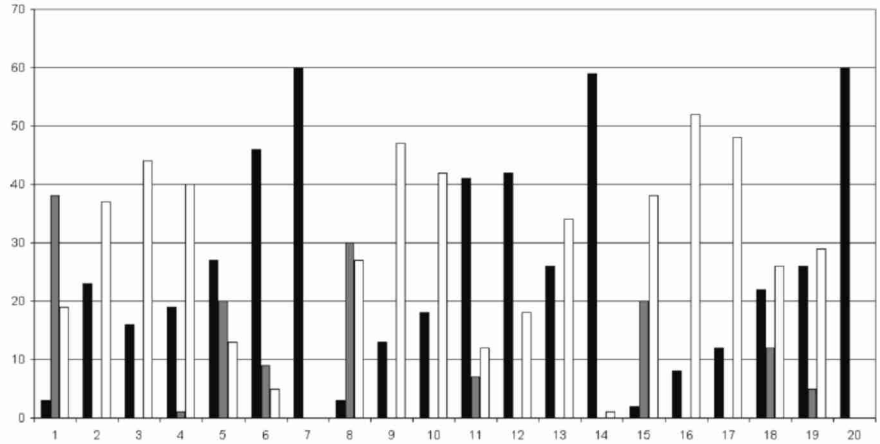


Table 11. Frequency of the comprehension scores according to the questions of the last three jokes in listening skills



Tables 8-11 indicate that the correct answers are less in number when they are compared with the wrong answers. Additionally, incomplete correct answers have also been given by the first graders.

The following comprehension questions were used to check to what extent the first graders were able to understand Nasreddin Hodja jokes. Additionally, Tables 8-11 show the frequency of the comprehension scores in listening and reading skills on the following questions for each joke .

Questions

"The quilt is gone and the fight is over"

1. Who are the characters of the joke?
2. Why was Hodja curious?
3. What did Hodja take on his back when he went out?
4. Why did Hodja go out with his quilt wrapped around himself?
5. Why did Hodja go inside helplessly?
6. What utterance of Hodja makes us laugh?
7. Why is this joke so funny?

"The cauldron died"

8. Who are the characters of the joke?
9. What did Hodja borrow from his neighbour?
10. What did Hodja do while he was giving the cauldron back to his neighbour?
11. How many times did Hodja borrow the cauldron from his neighbour?
12. What did Hodja's neighbour do when Hodja said that his cauldron had died?
13. What did Hodja say when his neighbour said "Cauldrons don't die!"
14. What utterance of Hodja makes us laugh?

15. Why is this joke so funny?

"Do you believe a donkey's word rather than mine?"
16. Who are the characters of this joke?
17. What did Hodja's neighbour ask Nasreddin Hodja?
18. What was Nasreddin Hodja's answer about where the donkey went?
19. What did the neighbour say when he heard the donkey braying?
20. What utterance of Nasreddin Hodja makes us laugh?
21. Why is this joke so funny?

"What day is it today?"
22. Who are the characters of the joke?
23. Where does Hodja live?
24. Where does Hodja go?
25. What does the man who approaches Hodja ask?
26. How does Hodja answer the man's question?
27. What utterance of Nasreddin Hodja makes us laugh?
28. Why is this joke so funny?

"Let the quilted turban read"
29. Who are the characters of the joke?
30. Why did the man give his letter to Hodja?
31. What did Hodja do when he took the letter?
32. What did the man say to Hodja?
33. What utterance of Nasreddin Hodja makes us laugh?
34. Is the joke funny?
35. Why is this joke so funny?

"One who gives the money?"
36. Who are the characters of the joke?
37. What did the children of the neighborhood asked Hodja to buy in the marketplace?
38. Did Hodja buy all of them a whistle?
39. Whom did Hodja buy a whistle? Why?
40. What utterance of Nasreddin Hodja makes us laugh?
41. Why is this joke so funny?

5. When the findings in Tables 2, 3, 4, 5, 6, 7 are scrutinized, the first graders have understood the jokes "The one who pays the Money" and "The Quilt is Gone and the Fight is Over" better in both reading and listening skills than the other jokes in the study.

3. Discussion

According to the findings, the first graders seemed significantly incompetent in understanding the jokes. The problematic jokes in terms of reading skills are "Do you believe a donkey's word rather than mine" and "What day is it today?". In listening

skills the problematic jokes are "Let the quilted turban read" and "What day is it today?". In explaining the first graders' comprehension of the texts, their linguistic competence, needs, interests, and knowledge on the subjects in the given jokes must be taken into account.

The students have not comprehended the witty compression and comic expansion properly since they are "unnamed and unacknowledged" (Brady 1999) in the superstructure of the jokes, but they are given in the substructure of the joke, which contains a few relevant generic details. These details can only be understood if the meaning is activated by means of the learners' schemata. Therefore, most of the first graders had difficulties in comprehending the jokes due to their lack of experience and background knowledge.

The first graders' reading comprehension seems better than their listening comprehension in this study. This may be because of the usual aids such as pictures illustrating the jokes and written verbal material (Bloom 1974: 89). According to Bloom's classification of comprehension, the first graders scored better on the questions of jokes 1 (2,3,4), 2 (8, 9, 10, 11); 3 (16, 17, 18); 4 (23, 24, 25); 5 (29, 30, 31, 34); 6 (36, 37, 38, 39, 40) in both reading and listening skills since these questions relate to interpretation behavior the evidence of which may be found in the inferences, generalizations, or summaries produced by the individual. Incomplete answers were given to the questions of jokes 1(1); 2(8, 14); 3(16); 4(22, 26); 5(29); 6(36, 39) since they are related with extrapolation behavior that involves making inferences with respect to implications, consequences, corollaries, and effects which are in accordance with the situations described in the jokes given. The answers to the questions (1, 8, 16, 22, 29, and 36) that tested to what extent the first graders were able to understand the orientation were incomplete, as the first graders were unable to name the other characters in the Nasreddin Hodja jokes. Most of them wrote Hodja's name, but could not identify the other characters in the jokes. The last two questions for all the jokes tested to what extent the first graders were able to understand the location, which clinches or discharges the joke. As seen in Tables 8-11, the first graders gave wrong answers since they were unable to comprehend the location. Due to the orientation, Nasreddin Hodja, they gave answers such as "The joke was funny because it was about Nasreddin Hodja; it was funny because it was a joke; since the joke was about Nasreddin Hodja; I laughed at each single utterance of Nasreddin Hodja". They were able to grasp the funny utterance better in "The Quilt is Gone and the Fight is Over". This may be due to the comic rhyme of the two past tense verbs in "Yorgan Gitti, Kavga Bitti".

Paralinguistic features are significant in understanding the content, as Nash (1986) indicates:

> In comic versifying, rhyme and rhythm have, potentially, a dual function. We may regard them as merely decorative applications, providing a setting and, so to speak, a lighting for the humour, or we may assign to them a more significant role as directive elements, features that organize the comedy and are essential to it. (154)

When the jokes in this study are taken into account in terms of listening and reading skills, significant differences in comprehension are detected for "The quilt is gone and the fight is over"; "What day is it today?"; "Let the quilted turban read", and "One who gives the money" (Tables 2, 5, 6, 7). The first graders seemed to have comprehended "The quilt is gone and the fight is over", "What day is it today?", "Let the quilted turban read" and "One who gives the money" better from reading them than from listening to them. Paralinguistic features did not help the first graders in this study.

In the test of reading skills for "The quilt is gone and the fight is over", the first graders were able to give the correct answer to question (3), which asked about the quilt. This may be due to the title of the joke. In listening skills they might have missed the title. In reading skills they were able to give the reason for what made them laugh, with more incomplete answers than in listening skills.

In the listening comprehension test for "What day is it today?" the first graders were unable to understand the allusion asked for in question (26). Although social questions such as "What Day is it Today?" are generic knowledge , when listening, the first graders were unable to catch Hodja's answer consisting of an indirect remark to the ridiculous question: "I'm a foreigner here. I don't know the days of this village. Ask someone who lives here". Additionally, they misunderstood question (25), giving as an answer the name of the day on which they were given the joke during the experiment. Whereas Hodja had violated the maxim of relevance, the first graders did not violate it. Since the first graders gave the wrong answer to the question (26), they also failed in questions (27) and (28).

In "Let the quilted turban read" the first graders were able to get question (32) regarding the signal of the joke, which introduces the joke, on the reading skills test, providing an incomplete answer to that question. The utterance "Shame on you!" helped them to find the answer. However, they were unable to get the hidden meaning of "quilted turban". Within the design of the joke "the quilted turban" was the locus. It was used as a hyponym for "learned person". Since in the schemata of the first graders the subordinate term of the subordinate concept was missing, they were unable to answer questions (33), (34), (35) properly. Besides, this hyponym seems to be beyond the first graders' literacy level. Therefore, Krashen's (1985) Input Hypothesis has been violated here.

In listening skills, the first graders confused the word "kavuk" (the quilted turban) with "tavuk" (hen). This might be due to their lack of knowledge of the meaning of "quilted turban".

About half of the first graders in both groups did not find the joke very funny since they were unable to interpret the function of Hodja's utterance "Let the Quilted Turban read".

Although according to the findings, the first graders demonstrated better comprehension in both listening and reading skills for the joke "One Who Gives the Money", question (41), which tested the locus, was left unanswered and question

(40) was answered incompletely and incorrectly. The students could not grasp Hodja's entire utterance,which was supposed to make them laugh.

An analysis of the first graders' common answers to the questions asking where the punch line of the jokes was hidden makes the following results plausible:

1. "The quilt is gone and the fight is over"

The punch line in this joke is that one of the fighters pulled Hodja's quilt off his back and fled although the fight had nothing to do with the quilt. The first graders misunderstood that the fight was over since one of the fighters had taken Hodja's quilt away.

2. "The cauldron died"

The first graders understood the 'logic' factor concerning how many times Hodja had borrowed his neighbor's cauldron, returning it with a stewpot inside and the explanation "Your cauldron gave birth".

Additionally, they comprehended that Hodja did not return the cauldron after he had taken it again. They remembered that Hodja snapped at his neighbour: "Well, you believed that it gave birth, so why don't you believe it died?!". However, they did not understand the 'likelihood' factor requiring the acceptance of an absurd proposition. They did not give complete answers to question (14). They were unable to reproduce Hodja's full utterance. In the last question, (15), they were unable to explain the violated 'truth conditions' in the joke which function as a locus.

3. "Do you believe a donkey's word rather than mine?"

In this joke, although Nasreddin Hodja's neighbour wants to borrow his donkey, Hodja, who does not want to give it to the neighbour, lies that it is not there. Upon hearing the donkey braying in the shed, the neighbour accuses him of telling a lie, to which Hodja retorts "Is it good manners to believe a donkey's word rather than that mine?"

The first graders, as in other jokes were unable to name both Hodja and his neighbour as the characters of the joke. Almost half of them gave only Hodja as a character of the joke. They understood well what the neighbour wanted to borrow from Hodja, but they could not get Hodja's full utterance after being accused of telling a lie. They did not get the comic unlikelihood compelling the logic in question (21). According to the truth conditions, animals do not speak. A donkey is an animal and it cannot speak. In this joke, truth conditions are violated, so it is a logic-twisting joke. At the same time Hodja violates the Maxim of Manner, which makes the joke funny.

4. "What day is it today?"

In this joke, Hodja leaves his village for Konya. While he is strolling in the market a man approaches him and asks: "What day is it today?" Since to Hodja this is a ridiculous question, he answers "I'm a foreigner here. I don't know the days of this village. Ask someone who lives here".

The first graders were unable to understand why the joke was funny. They were unable to find the signal (of the intention of the joke) and they seemed to have missed the locus which clinches the joke.. Since Hodja is a foreigner in Konya, he

cannot define the days there. The joke violates the generic knowledge of the truth in which the days of the week are the same all over the world. The first-graders were unable to give this answer.

5. "Let the quilted turban read"

In this joke, a man asks Hodja to read a letter he has received since he is unable to read cursive handwriting. Hodja looks at the letter, but he cannot read it. Upon hearing this, the man gets mad and says "You should be ashamed of yourself. You are wearing a Quilted Turban, but you can't read the letter".

The first graders were unable to get the idea about the man's visit to Hodja in terms of the substructure of the joke, where the generic detail of the visit was hidden. The man sought out Hodja, and nobody else since Hodja was the wise man to be consulted. They were unable to relate "the quilted turban" to being a literate human being.

6. "The one who pays the money"

In this joke, some neighborhood children gather around Nasreddin Hodja when he is on his way to the marketplace and ask him to buy each one a whistle. But only one of them gives him the money for the whistle. When Hodja returns from the marketplace, the children surround him and ask him if he has bought each of them a whistle. Hodja gives one whistle to the child who had given him the money and says "The one who pays blows the whistle".

The punch line of this joke was almost understood by the first graders. The parallelism helped them here. However, they were unable to explain why the joke was so funny. Here they did not relate the formula to the infrastructure, where the narrative concept was hidden. The narrative concept here is to buy something with the help of money, not with the help of somebody.

The findings of this study indicate that the first graders may not comprehend the humorous texts well because the humorous texts have two types of texts, one of which is in the superstructure and the other of which is in the substructure. The paradox which causes laughter and explosiveness may not be understood by children since they are unable to see the two types of texts in one humorous text.

3. Conclusion

Our findings and analysis indicate that the first graders may not comprehend the jokes properly in reading and listening skills. They appear to do better with written humour than oral humour. The written words and illustrations accompanying the jokes may have helped the first graders to develop memory strategies.

Although they are able to understand the superstructure of the jokes, they are unable to get the punch line in the substructure of the jokes due to their lack of world knowledge. The signal in the pre-location may not help them to get the locus which clinches or discharges the joke. They may laugh at the orientation, the type of joke rather than the punch line. Nasreddin Hodja in their schemata is a comic person and his name enables them to dissolve into laughter.

In order to comprehend, the beginning readers may have the interpretation behavior, that is, they may reorder the ideas, make inferences and generalizations implied in the jokes, but they may not have extrapolation behavior, which includes the making of estimates or predictions, implications in the jokes.

This research may be useful to both first and second language acquisition studies, since neither cognitive language development nor socio-psychological factors in language acquisition can be ignored. As Turkish first language acquirers are at the beginning of Piaget's concrete operational period , they have difficulties in understanding the humorous texts, which have a specific text structure in terms of text linguistics. This result can be taken into account in second language acquisition studies, and cultural factors as a part of language development, therefore, cannot be underestimated, since they clinch or discharge the joke in humorous texts. Linguists who conduct studies in first language acquisition may take into account the statistical results of this study to understand how cognitively mature Turkish first graders are in comprehending humorous texts. Additionally, the results indicate that the acquirers' general knowledge of the world makes it easier to understand the texts. Thus, the first graders have understood the jokes 'One who gives the money' and 'The quilt is gone and the fight is over' better than the other jokes in the study. Their responses indicate a better understanding of these two jokes' infrastructure, where setting, events and responses are given.

This study may bring about a profile of the first graders. Therefore, the results of this study should be considered by the Turkish Ministry of National Education. The primary school curriculum prepares the first graders for being good readers in Turkish by introducing them to selected informative texts, jokes, poems and short stories. However, as is seen in the research, the level of difficulty for the first graders is ignored. Since this study is limited to jokes in Turkish text books for the first grade, it has tested the jokes' level of difficulty for the first graders' current literacy level. Since jokes are humorous texts that make readers laugh when they are comprehended, they help the first graders enjoy their reading experience; but if they are incomprehensible to them, they may discourage their reading skills.

Our findings and discussion indicate that the first graders may not comprehend the jokes properly in listening skills. They understand written humor better than oral humor. Written visual letters and pictures illustrating the jokes must have helped the first graders to develop memory strategies. The first graders may have missed the humor in all the jokes (tested in questions 7, 15, 21, 28, 35 and 41). Thus, they appeared unable to understand the punch line in the substructure of the jokes due to their lack of world knowledge.

The findings also clearly indicate that the first graders are not cognitively mature enough to be able to engage in problem solving, deduction and complex memory tasks. Their general knowledge of the world is not extensive enough for some of the humorous texts given in this study. Therefore, they would not seem to be competent enough to understand humorous texts.

Furthermore, these jokes perhaps should not be included in the first-year curriculum, but instead in the second-year curriculum since beginning readers first need to develop some literacy. To develop the first graders' literacy level and introduce them to the culture of jokes, such humorous texts may be presented through simulations which will help the children to understand written discourse by incorporating verbal and non-verbal language. Paralinguistic signals, such as gestures, motions and tone of voice make both the denotative and connotative meanings of the vocabulary items in written discourse clearer. Additionally, some anaphoric, cataphoric, and exophoric items missed by children in reading may be understood better. Maxims of conversation cannot be violated by children during the first language acquisition period. When their exposure to humorous texts is accompanied by supportive and relaxed discussions, the first graders are able to extend their world view, and develop important critical thinking skills. Research also reveals that jokes must be chosen according to the first graders' interests and experiences.

The results clearly indicate that the first graders' linguistic competence, which is their mental representation of linguistic rules, may not be sufficiently developed for them to comprehend the pragmatic meaning hidden in the humorous texts. Therefore, this study emphasizes the importance of the child's experience and cognitive development in the comprehension of humorous texts. Children at this age may not understand the paradox hidden in the humorous texts which causes laughter and explosiveness.

Finally, similar studies should be conducted for each type of text in elementary schools textbooks in order to test to what extent students are able to understand the various texts. Linguists interested in first language acquisition should take the above results into account when determining the significance of cognitive development for the comprehension of certain types of texts.

References

Bakırcıoğlu, Rasim & Yalcin, Güner 2001. Yorgan gitti kavga bitti. *İlköğretim Türkçe ders kitabı 1.* Ankara: Özgün Matbaacılık Sanayii ve Ticaret AŞ.

Bloom, Benjamin S. (ed.) 1974. *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive dominion by a committee of college and university examiners.* New York: David McKay Company, Inc.

Brody, Roberto 1999. Buoys in the gap: Satire, sarcasm and irony as sense-making strategies. Paper presented at the annual meeting of the International Communication Association. San Francisco, California, May 27.

Glazer, Joan I. 1991. *Literature for young children.* New York: Macmillian Publishing Company.

Hacıoğlu, Hasan Selim & Çelebi, Celile & Çelik, İbrahim Halit 2001. Bugün günlerden ne? *İlköğretim Türkçe ders kitabı 1.* İstanbul: Prizma Yayıncılık AŞ. 153.

Karaca, Aliyar & Kapulu, Ahmet 1999. Kazan öldü. *İlköğretim Türkçe 1.* Ankara. Koza Yayın Dağıtım. 81-82.

Kaya, Dilek & Sarıdağ, Fevzi 2001. Eşeğe mi inanıyorsun bana mı? *İlköğretim Türkçe ders kitabı 1.* Ankara: Bir Yay Yayınevi. 17.

Kızıltan, Nalân 1995. A linguistic study of the language in Turkish children's literature. Unpublished Doctoral Dissertation. Hacettepe University.

Krashen, Stephen D. 1985. *The input hypothesis: Issues and implications. London: Longman.*

Lightbown, Patsy M. & Spada, Nina 2000[2]. *How languages are learned.* Oxford: Oxford University Press.

Nash, Walter 1986. *The language of humour: Style and technique in comic discourse.* London: Longman.

Özünlü, Ünsal 1992. Dil, Gülmece, Deyiş. *20. Yıl Yazıları: 20th Anniversary Year Book.* Ankara: Karaca Dil Kursu.

Özünlü, Ünsal 1999. *Gülmecenin dilleri.* Ankara: Doruk Yayımcılık İnceleme. 191-203.

Piaget, Jean 1995. *The language and thought of the child.* New York: Meridian Books.

Tortop, Ramazan & Barışcan, Hasan 2001. Parayı veren. *İlköğretim Türkçe ders kitabı 1.* İzmir: Top Yayıncılık. 32.

Tortop, Ramazan & Barışcan, Hasan 2001. Kavuk okusun. *İlköğretim Türkçe ders kitabı 1.* İzmir: Top Yayıncılık. 93.