Degrees of grammaticalization of Kazakh nominal relators

Lars Johanson

The paper deals with degrees of grammaticalization of nominal relators in Kazakh, suggesting a new kind of systematic classification. Grammatical relators are understood as free or bound elements occurring after nominals to mark their syntactic functions. Four levels of relators are distinguished: (i) case relators, (ii) opaque postpositional relators, (iii) less grammaticalized relators, and (iv) complex postpositional relators. The four levels represent decreasing degrees of grammaticalization.

Lars Johanson, Institut für Slavistik, Turkologie und zirkumbaltische Studien, Johannes Gutenberg-Universität, Hegelstraße 59, D-55122 Mainz, Germany.
E-mail: johanson@uni-mainz.de

Introduction

The present paper deals with degrees of grammaticalization of nominal relators in Kazakh, suggesting a new kind of systematic classification.[1] Grammatical relators are understood as free or bound elements occurring after nominals to mark their syntactic functions. Nominals are parts of speech that can be inflected for number, possession, and case: nouns, nominalized adjectives, pronouns, and numerals. An extended nominal is a nominal group larger than a single nominal.

The left-branching syntax of Kazakh has a rich system of grammatical relators functioning as postposed case markers and postpositions. The relators form relator phrases with their complements, linking these to other elements in the clause, in particular to verbs.

Level 1: Case relators

At level 1, the level of maximal grammaticalization, the relators are bound case markers. They combine morphologically with their hosts, forming phonological words with them and often showing sound harmony.

Kazakh cases cover semantically broad, less differentiated relational concepts. As is well known, five oblique cases are expressed by accentable suffixes, genitive

---

1 The paper is based on a talk presented at the *International Symposium on Kazakh Philology*, 26–27 October 2013, Minzu University, Beijing.

{-NIŋ}, accusative {-NI}, dative {-GA}, locative {-DA}, and ablative {-DAn}. The markers are slightly different after possessive suffixes. The accusative is marked with {-n} after third-person possessive suffixes. A so-called "pronominal *n*" appears in locative {-ndA} and ablative {-nAn} < {-ndAn}. Personal and demonstrative pronouns have a different declension, displaying some irregular forms, e.g. *men* 'I', genitive *meniŋ*, accusative *meni*, dative *maɣan*, locative *mende*, and ablative *menen*. The final *-l* in the demonstratives pronouns *bul, ol, sol* is replaced by *-n* in oblique cases except in the dative and in some ablative forms such as *bu-dan.* The old compound suffix {-Iŋ-KI} exhibits a variation, e.g. *men-iŋ-ki ~ men-i-ki* 'mine'. The *i* of the latter form looks like an accusative marker but may be a shortened genitive or an oblique stem form.

There have been speculative attempts to trace core case suffixes back to originally independent words, particularly nouns. Gerhard Doerfer (1977: 208–214) has suggested that the Orkhon Turkic dative marker {-KA}, which exhibits adlocational (allative, directive, terminative) functions, might go back to a Proto-Turkic root **ḳa* 'near(ness)', present in words such as *ḳat* 'side', *ḳat-* 'to join (tr.)' and East Old Turkic *ḳa:* 'family' (Clauson 1972: 578a). The markers may have developed from early level 2 postpositions (see below), though this is impossible to prove.

Kazakh relators of level 1 do not govern other cases. There are no complex forms of pronominal declension of the type documented in East Old Turkic, e.g. locatives such as *biz-in-dä* 'in us' instead of *biz-dä*. There are no traces of composite case suffixes, so-called 'double declension' markers, which are common in Mongolic and Tungusic and also occur in Yakut and Dolgan (e.g. comitative + accusative). It is highly improbable that the dative marker {-KA} goes back to two Proto-Turkic elements, as has been suggested by some scholars. The alleged second element has been compared to the Mongolic dative-locative marker {-A} (Tekin 1968: 130). The original dative suffix is most likely {-KA}, not {-A}.

There are other level 1 relators which are not core case markers, e.g. the unstressed bound equative marker {-šA}, developed from {-čA} and mostly expressing comparison ('like', 'as'), e.g. *ḳazaḳ-ša* '(in) Kazakh'.

Level 2: Opaque postpositional relators

Kazakh has, as other Turkic languages, a rich system of grammatical relators functioning as postpositions. Though they have functions similar to those of case markers, they form a distinct grammatical category. The number of cases is very limited, whereas there are numerous simple and complex postpositions. Postpositions mostly occur as free wordforms, but may also have bound variants, sometimes showing sound harmony. Postpositions may contribute to various kinds of complement phrases. Unlike case markers, they can even be separated from their complements by other intervening words.

Postpositions may determine the case of their complements. Some may even govern more than one case, depending on their meanings. Postpositions are fre-

quently classified according to such criteria. Here we will only deal with them as representing certain degrees of grammaticalization.

Postpositions express more differentiated relational concepts than cases do. Kazakh has numerous postpositions that can be used to express a wide range of spatial, temporal, and other semantic relations. They mostly locate their complement in space or time, but also express comparison, cause, purpose, instrument, etc.

Level 2 consists of highly grammaticalized, opaque postpositions of high age, including the four primeval Turkic postpositions described by Grönbech (1936: 35): *bi:rlä:* 'with', *tä:g* 'like', *üčün* 'for', *üzä:* 'over'. It is unknown what lexical elements they may go back to.

In modern languages, postpositions of this type constitute a closed class. They do not occur as other parts of speech, for instance as adverbs. They are not derived productively from words of other categories. They normally do not carry case markers. They are not used attributively, for instance by taking on the adjectivizing suffix {-GI}. They cannot be immediately preceded by quantifiers.

The East Old Turkic postposition *tä:g* 'like', e.g. *täŋri tä:g* 'god-like', is represented by {-DAy} in Kazakh, e.g. *ḳar-day appaḳ* 'white as snow, snow-white'. In the southwestern branch of Turkic, it has been replaced by the type *kib-i*, which belongs to level 3, a noun *ki:b* 'mould, model' plus a third-person singular possessive suffix, e.g. Ottoman *ḳar gibi* 'like snow'. In nominalized use, modern Kazakh {-DAy} can be followed by case markers in forms of pronouns and proper names, e.g. *mun-day* 'like this', genitive *mun-day-dïŋ* 'of something/someone like this', *Ḳanat-tay-dïŋ* 'of those like Kanat'.

Modern Kazakh *üšin* has prodessive, causal, and purposive functions: 'for', 'for the sake of', 'because of', 'in order to'. Its Old Uyghur equivalent *üčün* 'for' can govern the genitive or nominative case of personal and demonstrative pronouns in singular, e.g. *sen-iŋ üčün ~ sän üčün* 'for you'; cf. Uzbek *sen-iŋ üčün ~ sen üčün.* Some modern languages employ the genitive, e.g. Turkmen *θen-iŋ üčü:n* ‹seniň üçin›, Turkish ‹sen-in için›. Certain languages attach an accusative-like suffix, certainly a shortened genitive marker, to the complement, e.g. Karachay-Balkar *sen-i üčün.* Kazakh or Noghay *üšin* does not use the genitive case of pronominal complements, e.g. *sen üšin* 'for you'. The equivalents in other Turkic languages mostly occur as independent words, but they may also have bound variants, e.g. {-čUn}. Chuvash grammarians reckon with a causal or purposive case in {-šĭn}.

The old postposition *bi:rlä:* 'with' survives as modern Uyghur *birlä ~ bilä ~ birlän ~ bilän*, Tuvan *bilä*, Uzbek *b(i)làn,* Turkmen ‹bile› ~ ‹bilen› etc. Its shape is rather different from that of the Kazakh and Kirghiz instrumental-comitative postposition *menen* 'with', which has equivalents in several languages, e.g. Bashkir *mĭnän* ‹менəн›. Kaare Grönbech (1936: 35) analyzed *bi:rlä:* as *bi:r* 'one' plus an otherwise unknown "emphatic particle" *-la:*. Sir Gerald Clauson (1972: 364b) analyzes it as an abbreviated converb of *bi:rlä:-* 'to unite', which would mean that it was used as an adverb. The word also came to be used as an adverb meaning 'even', e.g. Ottoman *bilä*. Kazakh *menen* is, however, not used adverbially, and clearly belongs to level 2.

Kazakh *menen* and its Kirghiz counterpart tend to occur in bound forms, as the non-harmonic enclitic {-Men}, e.g. *ušaḳ-pen* 'by air'. Neither *menen* nor {-Men} are case markers, since they themselves govern a case, namely an oblique case of singular personal and demonstrative pronouns, e.g. *men-i menen* 'with me'. The equivalents of *bi:rlä:* behave in similar ways, e.g. Uzbek *meniŋ* (genitive) ~ *men* (nominative) *blàn,* Turkish ‹benim ile› ~ ‹benim-le› 'with me' (Deny 1921, § 437), the latter containing the harmonic enclitic {-(y)lA}.

The marker attached to the Kazakh complement gives the impression of being an accusative suffix (Geng Shimin et al. 1999: 205). It can, however, be a shortened form of the genitive marker; cf. *biz-di-ki* 'ours' and *men-iŋ-ki* ~ *men-i-ki* 'mine' (Menges 1959: 469). Zhang Dingjing (2004: 212) states that a high unrounded vowel is inserted before {-Men}, when it is attached to pronouns such as *men*, *sen*, and *ol*. This is a correct synchronic description. Diachronically, however the element *i* attached to the complement, e.g. in *men-i menen* 'with me', may reflect an old stem-final oblique segment (Menges 1959: 467). As noted, singular personal and demonstrative pronouns exhibit oblique stems ending in a "pronominal *n*" in many Turkic languages. East Old Turkic thus shows locatives such as *biz-in-dä* ← *biz* 'we' rather than *biz-dä*. Kazakh has oblique pronominal stems in locatives such as *bun-da* 'in this'. In Uzbek, which has lost the "pronominal *n*" in the nominal declination, the case markers {-gà} (dative), {-dà} (locative), {-dàn} (ablative), {-čà} (equative), and a few other suffixes are attached to the oblique stem of certain pronouns, e.g. *bun-dà* 'in this', *bun-dày* 'like this'. Oblique pronominal stems of this kind also occur in Mongolic, Tungusic, and Japanese.

Level 3: Less grammaticalized relators

Level 3 comprises less grammaticalized postpositions. They are formally indistinguishable from other parts of speech such as adverbs, adjectives, or nouns.

Postpositions sharing their shape with adverbs are often petrified converbs, e.g. *ḳara-y* 'toward' ← *ḳara-* 'to look', *gör-i* 'in comparison with, rather than' ← *kör-* 'to see', *bol-a* 'for' ← *bol-* 'to (be)come', *say-ïn* 'each' (< 'counting' ← *say-* 'to count'), e.g. *kün sayïn* '(counting) each day'. Other types contain adverbializing markers, e.g. *bas-tan* 'from ... on' ← *bas* 'head'. The type *boy-ïn-ša* 'according to', provided with the bound equative marker {-šA}, rather represents level 4 (below).

Postpositions sharing their shape with adjectives include the type *basḳa < bašḳa* 'other (than)', 'except (for)', 'apart (from)', 'excluding', 'without' ← *baš* 'head', governing the ablative case, e.g. *bu-dan basḳa* 'other than this'. Relator phrases at this level may be formally identical to adjective phrases with a complement that modifies a noun, e.g. *üy-ge žaḳïn* 'close to the house' < *yaḳïn* 'near' ← *yaḳ-* 'to approach, to be near'. Another type, containing adjectivizing markers, is represented by *sïyaḳ-tï* 'like' and *tuwra-lï* 'about', both governing the nominative, e.g. *sen tuwra-lï* 'about you'.

A number of postpositions consist of plain nouns without any adverbial endings. For example, Kazakh *burun* ‹бұрын› 'before' has clearly nominal properties. It goes back to the noun *burïn* 'nose', 'beak', which was also used for other protruding objects, generalized in the sense of 'place in front (of)'. In the temporal sense, it did not mean 'time still to come' (time following the moment of speaking) but, on the contrary, 'earlier (than)', 'prior (to)', e.g. Karakhanid *bu kün-dä burun* (this day-ABL before) 'before today'. According to Maḥmūd al-Kāšγarī, it was used for 'the first (*al-awwal*) of anything' (Clauson 1972: 366b–367a); cf. postpositions of the type *äv:äl* in several Turkic languages. Kazakh *burun* is regarded as an adverb that is also used as a postposition. It basically seems to be a noun meaning 'time or place before'. It can take on case markers and be used attributively, provided with the suffix {-GI}, *burïn-γï* 'previous'. It can also be immediately preceded by quantifiers, e.g. *keški tamaktan eki saγat burun* (dinner-ABL two hour before) 'two hours before dinner'. Both features are impossible in the case of postpositions of level 2.

In the same way, Kazakh *keyin* is basically a noun meaning 'time or place after' with the adjectival form *keyin-gi*. It goes back to *ke:-dïn*, a denominal noun ← **ke:* 'place behind', 'time after'. The East Old Turkic marker {-dÏn} forms denominal place nouns, e.g. *Säläŋä ke:-dïn* 'the place behind (west of) the Selenga'. It is more or less synonymous with the old word *ke:n* 'after', an allegedly instrumental form of **ke:*, used as an adverb or postposition, meaning 'place behind', 'time after (later)', e.g. *män-dä ke:n* 'after me'. It can be declined as a noun, which, according to Clauson 1972: 724b, might mean that its origin had been forgotten. It rather seems to be a noun which, like *ke:-dïn*, could occur in both adverbial and postpositional functions.

Kazakh *soŋ* 'end' is a noun that can be used as a postposition without an adverbial ending. Like other nouns, it can take on the adjectivizing suffix {-GI}, i.e. *soŋ-γï*. The older *soŋï-ra* 'afterwards', 'after' is a form in {-rA} ← *soŋ* 'end'. Though the East Old Turkic marker {-rA} is mostly said to form locational adverbs, it rather forms nouns that can be used both locationally and adlocationally, e.g. *ič-rä* 'in, inward' ← *ič* 'interior', *taš-ra* 'out, outward' ← *taš* 'exterior', *öŋ-rä* 'in front of, forward' ← *öŋ* 'front', *as-ra* 'below, downward'. The derived forms are not primarily adverbs, but place nouns, e.g. Ottoman *taš-ra* 'the outside', 'the provinces'. Similarly, *soŋï-ra* had nominal properties, forming nouns meaning 'time or place after'. Compare East Old Turkic *üzä-rä*, one of the four primeval postpositions that was also employed as an adverb. It replaced *üzä:* 'over', originally an adverb 'above' and also was used as a postposition meaning 'above, upon, on', governing the nominative, occasionally locative or genitive of pronouns (Clauson 1972: 280b).

Kazakh *beri* 'since' is a deictic space- and time-related postposition which also occurs as an adverb 'here', 'hither', 'on this side'. It originally contained {-rU} an alleged directive suffix that rather is a marker of place and time nouns. It was often used in antithesis to *aŋaru* 'that place/time', 'that side'. Like other nouns, it could take on an adjectivizing suffix, e.g. *bärü-ki* '(situated) towards this side'. Compare Kazakh *keri* 'backward place' < *ke:rü* ← **ke:* 'place behind', Turkish ‹geri›.

The markers of place and time nouns are comparable to space nouns such as Kazakh *žoγarï*. The old form *yo:ḳaru:* 'upper place' < **yo:ḳ-γaru:* is derived from *yo:ḳ* 'high ground' by means of the suffix {-GArU:}, which is a space-noun suffix rather than a directive marker. In adverbial expressions with verbs expressing motion, such nouns may occur in the meaning 'up', or 'upwards' without an adverbial marker of direction or goal. Thus Turkish ⟨yukarı⟩ 'upper place', can, if employed adlocationally, be synonymous with ⟨yukarı doğru⟩ 'upwards'. Used together with dynamic verbs it does not need a dative marker to express direction or goal (Johanson 2012).

Level 4: Complex postpositional relators

The complex relators found at level 4 are the least grammaticalized, most transparent relators. They provide more differentiated semantic specifications than other relators do, particularly expressing various exact spatial and temporal relations. They are based on auxiliary nouns provided with possessive suffixes and case markers. The possessive suffixes agree with the preceding complement, which is in the genitive or the nominative case, e.g. *üy-diŋ ald-ïn-da* ⟨house-GEN front-POSS3SG-LOC⟩ 'in front of the house', *stol üst-in-de* ⟨table top-POSS3SG-LOC⟩ 'on the table'.

These constructions are similar to free morphological combinations, but they are petrified ("frozen") grammaticalized items; compare the status of the complex English preposition *on top of* as against the locative phrase *on the top of*. The possibility of inserting elements between the complement and the relator is also very limited. The boundary between complex postpositions and free combinations is sometimes less distinct. It may be difficult to distinguish the postpositions from the phrases they derive from, particularly as most auxiliary nouns function as regular nouns.

The complex relators generally refer to place and time. Most of them go back to nouns denoting body parts. The relevant notions are mapped onto various spaces that may be conceived of as bodies or body parts. Most postpositions express relations in space: attachment ('at'), inclusion ('in'), superposition ('on', 'over'), subposition ('under'), exteriority ('outside'), laterality ('beside'), proximity ('near'), circumposition ('around'), contraposition ('against', 'opposite to'), intermediacy ('between', 'among'), etc. Spatial relators may describe dynamic situations, implying goal, direction, delocation ('motion to', 'motion toward', 'motion away from'). Relations in time are mostly expressed by the same postpositions. Some complex postpositions express comparison, cause, purpose, instrument, etc.

Kazakh employs auxiliary nouns such as *ald* 'front', *ara* 'interval', *art* 'back', *ast* 'bottom', 'lower part', *iš* 'interior', *orta* 'middle', *sïrt* 'exterior', *žan* 'side', *üst* 'top'. Some of them, e.g. *al-d*, *ar-t*, *as-t*, *üs-t*, are petrified relators that have lost their original functions. The auxiliary nouns can take on the adjectivizing suffix {-GI}, e.g. *art-ḳï* and *sïrt-ḳï*.

Notations

The following brackets and arrows are used for notations.
Graphic forms are given in angle brackets of the type ‹ ›
Glosses are given in angle brackets of the type ⟨ ⟩
Morphophonemic transcriptions are given in curly brackets of the type { }
← means 'is derived from'
 < means 'has developed from'

References

Clauson, Sir Gerald 1972. *An etymological dictionary of pre-thirteenth-century Turkish.* Oxford: Clarendon.

Deny, Jean 1921. *Grammaire de la langue turque (dialecte Osmanli).* (Bibliothèque de l'École des langues orientales vivantes 5.) Paris: Éditions Ernest Leroux.

Doerfer, Gerhard 1977. Zu türkisch *bana* 'mir', *sana* 'dir'. *Central Asiatic Journal* 21, 208–214.

Geng, Shimin et al. 1999. *Ḳazirgi ḳazaḳ tili.* [Modern Kazakh]. Beijing: Minzu Publishing House.

Grönbech, Kaare 1936. *Der türkische Sprachbau.* Kopenhagen: Munksgaard.

Johanson, Lars 2012. Pyramids of spatial relators in Northeastern Turkic and its neighbors. In: Suihkonen, Pirkko & Comrie, Bernard & Solovyev, Valery (eds.) *Argument structure and grammatical relations. A crosslinguistic typology.* Amsterdam & Philadelphia: Benjamins.191–210.

Menges Karl H. 1959. Die aralo-kaspische Gruppe. In: Deny, Jean et al. (eds.) *Philologiae turcicae fundamenta* 1. Wiesbaden: Franz Steiner. 434–488.

Tekin, Talât 1968. *A grammar of Orkhon Turkic* (Indiana University Publications. Uralic and Altaic Series 69.) Bloomington: Indiana University; The Hague: Mouton & Co.

Zhang Dingjing 2004. *Xiandai hasakeyu shiyong yufa* [A practical grammar of Modern Kazakh]. Beijing: The Publishing House of the Central University of Nationalities.