# Notes on the Turkic loanwords in Kangjia

Hans Nugteren\*

#### 1. Introduction

This paper will discuss the approximately 25 Turkic loanwords that can be identified in the Mongolic language Kangjia. It will touch upon three aspects. Firstly, the lexical fields will be established in which words were adopted. Secondly, I will compare the particular set of Turkic items found in Kangjia to those found in related languages. As will be shown, the set of Turkic words is very similar to that documented for Baoan and Dongxiang. Thirdly, I will attempt to determine the specific sources of the Turkic lexemes. The only Turkic language in the vicinity is Salar, but in view of the word shapes this is usually not the obvious donor language.

Kāngjiā [kɔdʒia] is the most recently documented Mongolic language of the Qīnghǎi-Gānsù region in China (Sečenčogt 1999). It is spoken by only a couple of hundred people in Jainca county in eastern Qīnghǎi. The Kangjia speakers officially belong to the linguistically diverse Huí ('Chinese muslim') nationality.

Kangjia belongs to the Shirongol subgroup of Mongolic. Together with Baoan (Bonan) and Dongxiang (Santa) it constitutes the Baoanic branch of that subgroup. It further shares genetic and areal features with the Monguor branch of Shirongol (consisting of Mongghul and Mangghuer), as well as with Eastern Yugur, which is not part of Shirongol. All these Mongolic languages also belong to a larger *Sprachbund* which in addition includes the Turkic languages Western Yugur (Sarïg Yugur) and Salar, as well as varieties of Amdo Tibetan and Northwest Mandarin. The members of this *Sprachbund* share phonological, morphological and syntactic features, in addition to lexical similarities like those discussed here.

Apart from a healthy percentage of native Mongolic vocabulary, and many words of obscure etymology, the Kangjia lexicon contains loanwords from Amdo Tibetan, Northwest Mandarin, Arabic, Persian and Turkic. Many of the loanwords have become difficult to recognise due to internal phonetic changes after the period of borrowing.

## 2. Some notes on Kangjia

As mentioned above, Kangjia is closely related to both Baoan and Dongxiang. Sečenčogt noted that it shared some features with the former and others with the latter, and concluded that both Baoan and Dongxiang have contributed to make Kangjia what it is today, i.e. a language of mixed ancestry in his view (1999: 277). Stephen Kim, in his contribution on Dongxiang in *The Mongolic Languages*, seems to consider Kangjia a dialect of Dongxiang (2003: 347–348). Similar sentiments may have played a role when it was

\* Dordrecht.

decided that there would not be a Kangjia chapter in *The Mongolic Languages*. The fact that Kangjia possesses many features that are shared by Baoan and Dongxiang, but tends to lack the defining traits of either, may also indicate that Kangjia is the result of an independent isolation event out of Baoanic stock. To demonstrate that, at the very least, Kangjia is not 'just Dongxiang', the following differences may suffice.

Kangjia	Dongxiang	CM	
ders	zieroŋ	*dörben	four
ixgə	oćin	*ökin	girl
eter	oćielu	*ötel	to age
пөхдн	nokie	*nöker	friend
gər	fugie	*hüker	bovine
χэr ~ χuar	oqo	*okar	short
tɔ(r)mɔ	ćiemaŋ	*teermen	mill
vəis <del>u</del> n	osuŋ	*ebesün	grass
kuar	kuaŋ	*köl	foot
səʉki	sumuğe	*sibüge	awl

#### 3. Turkic words in Mongolic

After centuries of contacts between Turkic and Mongolic languages, there are no modern Mongolic languages without Turkic loanwords. At least two layers of Turkic words are documented from the earliest Middle Mongolian onwards, and in view of the present distribution of these words, both layers must have been present in the Common Mongolic lexicon. Instead of layers one may speak of 'types' since the layers represent loans from different Turkic languages but do not necessarily belong to different periods, given that the modern Turkic subgroups had already diverged before the first Middle Mongolian documents emerged.

The following examples illustrate some of the different phonetic correspondences of the two main types. The oldest layer is typified by, among other things, the familiar correspondences CM \*r = CT \*z, CM \*l = CT \* $\check{s}$ , CM  $\emptyset$  = CT \*p, whereas the youngest layer involves only slight adaptations to Mongolic phonology.

'Layer' 1:	Mongolic	Turkic	
	*hüker	*höküz	bovine
	*jalau	*ya:š	young
	*jaidaŋ	*yapïtak	without saddle
'Layer' 2:	Mongolic	Turkic	
	*bös	*böz	fabric
	*jimis	*ye:miš	fruit
	*tobarag	*toprak	earth

'Layer' 3 pertains to the loanwords that are present in certain subgroups or individual languages, and probably entered after the split-up of Mongolic. Nevertheless some were adopted early enough as to be documented in Middle Mongolian, e.g. *jarqaq* 'hide' in the Secret History, *qarqulaq* 'tiger cat' in the Huá-Yí Yìyǔ, *čulbuq* 'gum in the eyes' in the Muqaddimat al-Adab.

Among the modern Mongolic languages, Kalmuck and Moghol abound in recent Turkic loanwords. They are more rare elsewhere, but can be found, e.g. Khalkha *evšges* 'female sable'  $\leftarrow$  *\*evči kiš*, *sonbas* 'stag head mask'  $\leftarrow$  *\*sïgun bašï*; Ordos *yalman* 'gerbil'  $\leftarrow$  *\*yamlan*, *bušxaq* 'skin of the legs of a skinned animal'  $\leftarrow$  *\*bïčgak*; Buriat dialects *saaža*, *šaža* 'braid'  $\leftarrow$  *\*sač*.

In the Qīnghǎi and Gānsù area, perhaps a couple of hundred Turkic words are used by Eastern Yugur; about fifty (so far) have been found in all of Shirongol combined.

#### 4. Lexical fields

Many of the Turkic words in Kangjia describe cultural concepts related to agriculture and trade. All loanwords that I could identify so far are nouns. Most words denote tangible concepts, from the following lexical fields. **Natural environment**: *cïrima* 'hail' < \*ya:rma 'groats', *dur* ~ *durpa* < \*toprak 'earth', t(i) $\chi asi ~ ta\chi asi < *tas 'stone'.$ **Animals**:*kəkəuliu*< \*keklik 'partridge', gugučiɔ < \*kökürčgün 'pigeon', s(i)Jiğɔ < \*sičgan 'mouse'.**Agriculture**& horticulture:*arpa*~*appa*< \*arpa 'barley', bašiğ(ï) < \*bašak 'ear (of grain)', bəġdi ~ baġdi < \*bugday 'wheat', kənčir < \*kendir 'hemp', urma < \*orma 'harvest', turma ~ tumar < \*turma 'turnip', jaŋgəy < \*yaŋak 'nut'.**Human environment and material culture**: ağa < \*agil 'village', bağa ~ bəğa ~ buğa < \*bakir 'copper > money', ba(r)mɔ ~ bɔrmɔ < \*batman 'unit of weight',*čügu*~*čigu*< \**čöki* $'chopsticks', gebde ~ gibde < *kegde 'paper', <math>\chi de ~ \chi te < *kitañ 'Chinese', dan < *tam 'wall'.$ **Kinship terms**: ana < \*ana 'mother', ade 'grandfather'? < \*ata 'father', aʒï < \*eze 'elder sister'. The latter category is notoriously problematic because it contains*Lallwörter*, and both Chinese and Tibetan use a number of forms that are phonetically close to their Turkic counterparts.

#### 5. Distribution of Turkic words in Shirongol

At first sight the distribution of the Turkic words across the Shirongol languages is very chaotic. Many words are only attested in a single language, while others are shared by several sets of languages. But in case of more widely attested items, some patterns do emerge. As established in Nugteren 1998, there is a set of Turkic loanwords that is shared by both Shirongol branches, while other sets are either restricted to Baoanic or to the Monguor languages (the few Turkic loans in Mangghuer, not discussed in that article, agree with Mongghul). Mongghul shares several items with Eastern Yugur, but as the latter has a far larger number of Turkic words, this may be of little significance for classification purposes.

Based on other shared features one expects the Turkic loans in Kangjia to largely agree with those found in Baoan and Dongxiang rather than with those in Mongghul or

#### Hans Nugteren

Eastern Yugur. This is generally the case. Kangjia has several words that were already known to occur in both Shirongol branches, e.g. \*bakïr 'copper > money', \*ta:š 'stone', \*turma 'turnip'. In other cases Kangjia indeed fits in specifically with the Baoanic branch: \*a:gïl 'village', \*arpa 'barley', \*bugday 'wheat', \*kegde 'paper', \*sičgan 'mouse', \*ta:m 'wall'. In case of \*yaŋgak 'walnut', Kangjia shares the shape \*jaŋgak with Baoan and Dongxiang, as opposed to \*jak in Mongghul, which, like Middle Mongolian (Huá-Yí Yìyŭ) ji'aq, represents a different development, likely a different Turkic source form (cf. also çaq in the Dahe dialect of Western Yugur).

However, several words found in Baoan and Dongxiang, and therefore expected in Kangjia, are missing, e.g. *\*čelpek* 'mucus (in the eye)', *\*erük* 'apricot', *\*kaymak* 'cream', attested in both Baoan and Dongxiang, and *\*közegü* '(charred stick used as a) poker', which is attested in all the other Shirongol languages, including Mangghuer.

There are only few Turkic words in Kangjia that are not attested so far in Baoan or Dongxiang, e.g. \*bašak 'ear (of grain)'.

The following list shows the Turkic words borrowed by Kangjia and their distribution in the remaining Mongolic languages of the region. (NB: the reconstructed Turkic source forms need not be identical to Old Turkic. Common Turkic vowel lengths are omitted.)

Some cases (marked +?) are uncertain due to a phonetic or semantic discrepancy between the attested form and its assumed Turkic source form.

Kangjia *ade* 'grandfather' differs semantically from \**ata* 'father' (cf. Dongxiang *ada* 'father'). Sečenčogt derives Kangjia *čauši* ~ *čoši* 'ladle' from Turkic without mentioning a form. Perhaps he was thinking of \**čomiš*, which is an unlikely match phonetically. Kangjia *urma* 'harvest' is a good match for Turkic \**orma*, but Sečenčogt marks it as a Tibetan word (also without providing a source form). Two other uncertain correspondences involve the following widespread borrowings. Kangjia  $\chi aya$  'handful' closely resembles Turkic \**haya* 'palm of the hand', but its Mongghul counterpart  $\chi ayag'$  inexplicably ends in *g*, casting doubt on this etymology. Turkic \**örtgün* 'threshing floor' may be the ultimate origin of Kangjia *torğo* ~ *tarğo* ~ *toğo* 'id', although Kangjia and remaining Shirongol suggest a form \**ütergen*. Hán Jiànyè (1992: 61) derives Kangjia (*y*)*ila* 'to cry' from Turkic, but this form corresponds well to Mongolic \**uila*. Hán (ibid.) views Turkic \**baka* as the source of *bab*(*a*)*ğa* 'frog', but it, and perhaps even the Kangjia form *paka* given by Sečenčogt, may be alterations of Amdo Tibetan dialect forms (Literary Tibetan *sbal-pa*; cf. Mongghul *sba:vag*, Nantoq Baoan *mbawa*).

	Turkic source form	L		Shirongol				
			E. Yugur	Mongghul	Mangghuer	Baoan	Kangjia	Dongxiang
1	*agïl	camp	Ø	Ø	Ø	+	+	+
2	*ana	mother	+	+	+	+	+	+
3	*arpa	barley	Ø	Ø	+	+	+	+
4	*ata	father	Ø	+	+	Ø	+?	+
5	*aze < *eze	elder sister	Ø	Ø	Ø	+	+	+
6	*bakïr	copper	+	+	Ø	+	+	+
7	*bašak	ear of grain	Ø	Ø	Ø	Ø	+	Ø
8	*batman	unit of weight	Ø	+	+	+	+	Ø
9	*bugday	wheat	+	Ø	Ø	+	+	+
10	*čomïš	ladle	+	Ø	Ø	Ø	+?	Ø
11	*čöki	chopsticks	Ø	Ø	Ø	Ø	+	+
12	*haya	handful	Ø	+	+	Ø	+	+
13	*kegde	paper	Ø	Ø	Ø	+	+	+
14	*keklik	partridge	+	+?	+?	Ø	+	+
15	*kenčir *kentir</td <td>hemp</td> <td>+</td> <td>Ø</td> <td>Ø</td> <td>+</td> <td>+</td> <td>+</td>	hemp	+	Ø	Ø	+	+	+
16	*kïtay <*kïtañ	Chinese	Ø	Ø	+	+	+	+
17	*köküšgün < *kökürčgün	pigeon	+	+?	+?	+?	+	+
18	*orma	harvest	Ø	Ø	Ø	Ø	+?	Ø
19	*örtgün	threshing floor	Ø	+	+	+	+	+
20	*sïčgan	mouse	Ø	Ø	Ø	+	+	+
21	*tam	wall	Ø	+	Ø	+	+	+
22	*taš	stone	+	+	+	+	+	+
23	*toprak	earth	Ø	Ø	Ø	Ø	+	Ø
24	*turma	turnip	+	+	+	+	+	+
25	*j(aŋg)ak < *yaŋak/*yagak	nut	Ø	+	Ø	+	+	+
26	*jarma < *yarma	hail	Ø	+	Ø	+	+?	+

# 6. Word shapes and source languages

Most of the Turkic words in Kangjia cannot be derived from a specific source form (which also applies to many loans in the other Shirongol languages and Eastern Yugur). The following items *could* stem from (an earlier stage of) Salar, but lack the informative phonetic features needed for a certain match.

Kangjia		Salar	W. Yugur	CT
ağa	village	ağïl	ağil	*a:gïl
bašiğ(ï)	ear (of grain)	(baš)	paşaq	*bašak
bağa/bəğa/buğa	money (copper)	vaxïr	pa <sup>h</sup> qir	*bakïr
kəkəuliu	chukar (bird)	keklik	kekilik	*keklik
t(ï)xaši/taš	stone	daš	tas	*ta:š
čügu ~ čigu	chopsticks	čugu	Ø	*čöki
dan	wall	dam	Ø	*ta:m
χde/χte	Chinese	χïdi	q <sup>h</sup> ïti	*kïtañ
turma	turnip	turma	turma	*turma

In other cases modern Salar has a deviating form that cannot be the source:

Kangjia		Salar	W. Yugur	CT
arpa ~ appa	barley	ahra/arfa	harva	*arpa
ba(r)mɔ ~ bərmə	weight; scales	omïn	Ø	*batman
bəġdi/baġdi	wheat	boğji	Ø	*bugday
dur ~ durpa	earth	torïx	tur(w)aq	*toprak
gʉgʉčiɔ	pigeon	guryunjux	kükişkïn	*kö:kürčgün
tarğɔ ~ tə(r)ğə	threshing floor	yürdïn	Ø	*örtgün

The Turkic word, or at least a convincing source form, may not be attested in Salar at all. In case of 'hemp' and 'paper' the Western Yugur form is closer.

Kangjia		Salar	W. Yugur	CT
s(ï)jiğə	mouse	Ø	şïğan	*sïčgan
jaŋgəy	walnut	Ø	[çaq]	*yaŋak
kənčir	hemp	gemdur	k <sup>h</sup> emćir	*kendir
urma	harvest	Ø	Ø	*or-ma
gebde/gibde	paper	[xaxït]	k <sup>h</sup> egti	*kegde
cïrima (sic)	hail	yarma	Ø	*ya:rma

The phonetic shapes of the Turkic words in Kangjia and remaining Shirongol are diverse to a degree that suggests that the source languages belong to various periods and subgroups of Turkic. For instance, intervocalic \*g was preserved in \* $ag\ddot{i}l$  'village', but VgV was contracted in \* $k\ddot{o}:zeg\ddot{u}$  'poker' (found in all of Shirongol except in Kangjia). The latter development is an innovation also found in Kypchak. Postconsonantal \*g was preserved in \* $s\ddot{c}gan$  'mouse', but lost in \* $k\ddot{o}:k\ddot{u}r\ddot{c}g\ddot{u}n$  'pigeon'. The latter development is normally found in Oghuz. Initial Turkic \*y is represented by an affricate (suggestive of Kypchak) in \*ya:rma and \* $ya\eta ak$ .

Some variable developments cannot be ascribed to Turkic source languages, e.g. final \*k was preserved in \*bašak 'ear' and \*yaŋak 'nut', but lost in \*keklik 'chukar (partridge)' and \*toprak 'earth'. Likewise \*r was preserved in \*kendir 'hemp' but lost in \*bakïr 'copper'. Similar inconsistencies can be observed in native words. They are the result of the ongoing 'sinification' of Kangjia syllable structure. Some Kangjia forms are idiosyncratic. The inexplicable *c* of *cirima* 'hail' would have been reason to doubt the identification, if it were not for the presence of this word in the other Baoanic languages (in shapes close to Kypchak *\*jarma*). The *b* of *gibde* 'paper' is unexpected as well.

On the other hand, the affricate in *kənčir* 'hemp' is mirrored by the other Baoanic languages, and also found in Eastern Yugur and Kalmuck. This suggests that a Mongolic cognate *\*kenčir* had developed quite early.

In short, the Kangjia word shapes indicate that most of the Turkic loanwords are nonrecent and/or non-local. Moreover, not all words can be derived from a single Turkic language. There are a couple of conceivable explanations which are not mutually exclusive.

The Kangjia speakers may have adopted the Turkic words elsewhere, before they settled in their present location. The words may have come from an earlier stage of Salar or from a Turkic substrate language whose precise affinities are unclear. They may have entered via other Baoanic languages, without any direct contact between Kangjia and Turkic.

These scenarios depend on the assumption of historical circumstances that are actually unknown. An alternative explanation is the following.

As many word shapes are consistent with phonetic developments found in Chaghatai and Kypchak Turkic, such words may have been carried from Turkestan to Qīnghǎi by Chinese-speaking Muslims, or by an actual Turkic-speaking component of such a migrating group.

Some of the Turkic items may have been handed down in a 'package' of cultural terminology typical for the Huí, and largely unknown to non-Muslim speakers of North Western Mandarin dialects. This package also contained the Persian and Arabic words found in Kangjia, such as d # nya 'world', *meiči* 'mosque', *tesbixar* 'prayer beads', from Arabic *dunyā*, *masjid*, *tasbīħ*, and *axuŋ* 'imam', *činčiri* 'chain', d # z e e 'hell', *pere* 'spirit' from Persian 'a x w nd, *zinjīr*, *dozax*, *parī*. Several of these words show adaptations to Chinese phonology, notably with regard to syllable structure.

The following Arabic or Persian words must have entered via Northwest Mandarin, because their Kangjia shapes do not simply represent phonetic adaptations of the Arabic or Persian words. Judging from the peculiarities they contain (the added i of *asmani*, the n in *manla*, etc.) they are based on the forms these words took on in Northwest Mandarin.

Kangjia	N. W. Mand.	Modern Uygur	
asmani	a sĩ ma ńi	asman	sky
manla	man la	molla(h)	mullah
naimaʒï	nɛ ma cï	namas	namas prayer
nikaχar	ni k <sup>h</sup> a xa	nika(h)	wedding pronouncement

Northwest Mandarin as spoken by the Huí also adopted loanwords of Turkic origin, including the following items of cultural vocabulary (tones omitted):

Hans Nugteren

N. W. Mandarin	Modern Uygur	
a t <sup>h</sup> a	ata	father
a ta śi	adaš	friend
pa la	bala	child
wu ma şï	umaš (umač)	porridge
ć <sup>h</sup> ia ma kur	čamğur	turnip
k <sup>h</sup> æ̃ t <sup>h</sup> u mæ̃	kätmän	mattock
ya k' śi	yaxši	good

On the other hand the number of Turkic words which are attested both in Kangjia and in Huí dialects is small. For instance the names of animals and cultivated plants that Kangjia took from Turkic are not attested in Huí. This can only partly be blamed on the small size of the Kangjia corpus, as there is a considerable overlap where the Persian and Arabic words are concerned.

Evidently the matter of the entry route(s) of the Turkic words in Kangjia cannot be resolved in this brief contribution. The possibility of Chinese intermediation could be confirmed when more Chaghatai and Kypchak-like forms attested in Kangjia are also found in Northwest Mandarin.

### Sources:

- Bökh et al. 1983. Düngsiyang kelen-ü üges / Dōngxiāngyŭ cíhuì [Vocabulary of the Dongxiang language]. Hūhéhàotè: Nèimĕnggŭ rénmín chūbănshè.
- Bökh & Chén Năixióng 1981. Tóngrén Băo'ānhuà gàiyào [Outline of the vernacular of Tongren Bao'an]. Mínzú Yŭwén 2, 61–75.
- Chén Năixióng et al. 1985 [1986]. Boo An kelen-ü üges / Băo'ānyŭ cíhuì [Vocabulary of the Bao'an language]. Hūhéhàotè: Nèiměnggŭ rénmín chūbănshè.
- Chén Năixióng et al. 1986 [1987]. Boo An kele ba Mongyol kele / Băo'ānyŭ hé Měnggŭyŭ [Bao'an and Mongolian]. Hūhéhàotè: Nèiměnggŭ rénmín chūbănshè.
- Chén Năixióng 1995. Gānhétān Băo'ānhuàde yŭyīn hé cíhuì [Speech sounds and vocabulary of the Ganhetan Bao'an vernacular]. Chén Năixióng Lùnwénjí. Hūhéhàotè: Nèiměnggŭ rénmín chūbănshè. 99–157.
- Čeremisov, K. M. 1973. Burjatsko-russkij slovar'. Moskva: Sovetskaja ènciklopedija.
- Clauson, G. 1972. An Etymological Dictionary of Pre-thirteenth-century Turkish. Oxford: Clarendon Press.
- Hán Jiànyè 1992. Chū tán Kāngjiāhuà yŭyīn xìtŏng jí cíhuì de gòuchéng [Preliminary remarks about the phonological system and the composition of the lexicon of the Kangjia vernacular]. Qīnghǎi mínzú yánjiū 1, 52–62.
- Hangin, G. 1986. A Modern Mongolian-English Dictionary. Bloomington: Indiana University.
- Janhunen, J. (ed.) 2003. The Mongolic Languages. London, New York: Routledge.
- Janšansin, J. 2009. Kratkij dungansko-russkij slovar'. Moskva: IPB.

- Khasbaatar et al. 1985 [1986]. Mongyor kelen-ü üges / Tŭzúyŭ cíhuì [Vocabulary of Monguor]. Hūhéhàotè: Nèiměnggŭ rénmín chūbănshè.
- Kim, S. S. 2003. Santa. In: Janhunen, J. (ed.) The Mongolic Languages. London, New York: Routledge. 346–363.
- Lín Liányún 1992. S*ālā Hàn Hàn Sālā Cíhuì* [Salar-Chinese, Chinese-Salar Vocabulary]. Chéngdū: Sìchuān mínzú chūbănshè.
- Liú Lìlĭ 1989. *Huímín Wūlŭmùqí yŭyán zhì* [Description of the language of the Hui of Ürümči]. Wūlŭmùqí: Xīnjiāng dàxué chūbănshè.
- Mă Guózhōng, & Chén Yuánlóng (eds.) 2000. Dunxian kielien khidei kielienni lugveqi / Dōngxiāngyŭ Hànyŭ Cídiăn [Dongxiang-Chinese dictionary]. Lánzhōu: Gānsù mínzú chūbănshè.
- Mostaert, A. 1941-1944. Dictionnaire ordos I-III. Peiping: Catholic University.
- Mostaert, A. (de Rachewiltz, I. ed.) 1977. *Le matériel mongol du Houa i i iu de Houng-ou (1339) I.* Brussels: Institut belge des hautes études chinoises.
- Nugteren, H. 1998. On some Turkic Loanwords in Monguor, Bao'an, and Dongxiang. In: Johanson, L. & Csató, É. Á. et al. (eds.) The Mainz Meeting. Proceedings of the Seventh International Conference on Turkish Linguistics.. Wiesbaden: Harrassowitz. 683–696.
- Poppe, N. 1955. The Turkic loan words in Middle Mongolian. *Central Asiatic Journal* 1, 36–42.
- Potanin, G. N. 1893. *Tangutsko-tibetskaja okraina Kitaja i central'naja Mongolija*. Sanktpeterburg: Imperatorskoe russkoe geografičeskoe obščestvo.
- Sečenčogt 1999. *Kāngjiāyŭ yánjiū* [Kangjia language research]. Shānghǎi: Shānghǎi Yuǎndōng chūbǎnshè.
- Sečenčogt 2002. Kāngjiāyŭ gàikuàng [Brief account of Kangjia]. Mínzú Yŭwén 6, 66-67.