

Khurshid, the vizier of *Qilij-Arslan IV*, at a feast given in honour of *Baiju* in *Ab-i Garm*, the present day *Irgin*, S.E. of *Akşehir*.<sup>55</sup>

#### 15. Xul

A more correct form of the name — *Fuli* — occurs in *Kirakos*. It is of course the Turkish *Qulī* («His slave»).

*Qulī*, the second son of *Jochi's* eldest son *Orda*, the founder of what was later known as the White Horde, was sent by his father to assist *Hulegū* in his operations against the *Ismā'īlis* and the Caliph. He arrived in Persia by way of *Dihistān*, i.e. the district north of the *Atrek* in the present-day Turkmenistan, and *Mazandaran*. *Grigor* implies that he met a violent end, but *Rashid al-Din* says simply that his death occurred after those of *Balayai* and *Tutar*.

#### 16. Xurum'ī Nuin

*Xurum'ī* is probably a scribal error for an original *Xurumšī*,  $\zeta \acute{\epsilon}$  and  $\zeta \grave{\delta}$  being readily confused; and in fact the correction has been made in the index to *Blake and Frye's* text.

As *Pelliot* remarks in the passage quoted by *Cleaves*,<sup>56</sup> there were many bearers of this name, which means «*Khwārazmian*.» Perhaps this is the third son of *Orda*, whom *Pelliot* identifies with the Mongol chieftain *Corenza*, encountered by *John de Plano Carpini* on the *Dnieper*.<sup>57</sup>

<sup>55</sup> *The Successors of Genghis Khan*, pp. 104 and 123.

<sup>56</sup> *Mongolian Names*, pp. 433—434.

<sup>57</sup> *Loc. cit.*, also *Notes sur l'histoire de la Horde d'Or*, p. 9.

RAL 1975

## THE FOREIGN ELEMENTS IN EARLY TURKISH

BY

Sir GERARD CLAUSON (London)

A hundred years ago scholars believed that there were two kinds of Turkish: Western Turkish, primarily *Osmanli*, and Eastern Turkish, primarily *Chagatai*, and that when the obvious Arabic and Persian, and in the case of *Osmanli* a sprinkling of Greek and *lingua franca*, loan words had been eliminated from the vocabulary of these languages, all that was left was pure Turkish.

The situation has now entirely altered. Extensive texts in much earlier Turkish languages have been discovered on monuments in Mongolia and the Soviet Union, by excavations in Sinkiang and Kansu, and in some ancient libraries, mainly in Turkey and the Soviet Union. In addition Sogdian and Tokharian manuscripts have been discovered, languages that no-one had even heard of a hundred years ago. It is now obvious that even the earliest Turkish languages were no more free from loan words than the language of any people which has been for centuries in close contact with peoples speaking other languages. The problem is not one of proving that loan words exist in the early Turkish languages but one of identifying and classifying those loan words and seeing what historical deductions can be drawn from them.

So far as identifying them is concerned, there are a few simple phonetic laws which enable us to identify some, but perhaps not very many, words as loan words.

There are two sounds which occur in pure early Turkish only in strictly defined contexts, but otherwise only in loan words.

The unvoiced velar fricative *x* occurs only in two consonantal clusters -*xt*- in words like *axtar*- «to roll on one's back» and -*xš*- in words like *oxša*- «to resemble». If it occurs in any other contexts, for example in *xayun*, *xatun* and *tarxan*, it proves that they are loan words. I shall return later to the question of what kind of loan words these particular words are. But not all words containing -*xt*- or -*xš*- are necessarily pure Turkish; *axšam* («evening», for example, is a Sogdian loan word.

The voiced palatal sibilant *ž* occasionally occurs as a secondary form of *z* in the immediate vicinity of *č*, as for example in the *Xakani* verbs *čiz-* and *čž-*, which are secondary forms of *čiz-* and *čž-*, otherwise it occurs

only in loan words. As *z* is a very common sound in Sogdian, it is generally safe to assume that words which contain it are Sogdian loan words. But we are in this difficulty, that the vocabularies of the surviving Sogdian, and also Tokharian, texts, which are mainly translations of Buddhist scriptures, are rather restricted. We can be certain that *a:zu:n* «a state of existence» is a loan word because it frequently occurs in Sogdian texts, but we are not in the same position regarding *üzmd:* «a mulberry», because the Sogdian word for «mulberry» has not survived.

There are two other sounds, *m* and the unvoiced palatal sibilant *š*, which are common in the medial and final positions but occur as initials in pure Turkish words only as the secondary forms of other sounds.

In almost all Turkish languages initial *-b*, if followed by a nasal, has become *m-* by nasal attraction. When the nasal in question is the dental nasal *n* or the guttural nasal *ŋ*, what has happened is quite clear. For example *bän* «I» and *bängü:* «eternal» have become *män* and *mängü:* in almost all languages except 8th and 9th century Türkü and the Oğuz languages. But when the nasal is the palatal nasal *ñ* the position is more complicated because this sound was beginning to disappear in the 8th century and has become *y* in nearly all languages and *yn* or *ny* only in a few. Thus there are a few words with an initial *m-* followed by a *y*. Fortunately there is sufficient evidence to show that this *m-* must originally have been *b-*. For example one of the rather numerous words for «dung» is *mayak* in Uyğur, Xakani, Kıpçak and one or two modern languages, but Kāšyari says that the Oğuz form was *baynak*, which proves that the word was originally *bañak*. Again the word for «brain» occurs in a bewildering variety of forms ranging from *mī* in Khakas to *miñ/miñi* in Sinkiang, but the Osmanli form *bäyin* and Türkmen *bäymi* prove that it must originally have been *bäñi*.

All other words with initial *m-* are loan words. The most interesting is perhaps «honey», *mīr*, in Uyğur. Immediately this is a loan word from the Chinese word *mī* (in Middle Chinese something like *myit*, pronounced *mīr* in North West China in the middle of the first millennium A.D.). This Chinese word is itself a loan word from Tokharian *mit*, a good Indo-European word cognate to Russian *мед*, Greek *methu*, Sanskrit *madhu*.

Of the remainder some like *mākkū:* «ink» and *mīnyuy* «paste» are Chinese, and some like *midik* «clayman» and *mīškūš* «wild cat» are Sogdian loan words. Of one or two like *maši:* «cat» the origin is still uncertain.

There are not many early Turkish words with initial *š-*, perhaps about forty in all, if the obvious Sanskrit words in translations of the Buddhist scriptures are left out of account. These include one or two onomatopoeics

like *ša.b ša.b*, the sound of a slap. In nearly all the rest the initial *š-* is a secondary sound due to phonetic attraction. In words like *šiš-* «to swell» (*šiš-*) it represents dental *s*, in words like *šayla:* — «to shout» (*šayla:-*) the denti-palatal affricate *š-*, and in one word, *šišāk* «two-year-old sheep» (*išāk*) dental *t*.

When these are eliminated, we are left with one or two Chinese loan words like *šik* «a measure for grain», conventionally a little less than a hectolitre, one or two Sogdian loan words like *šimnu:* «devils», which, like Persian *ahrimān* goes back ultimately to Avestan *arəmānyu*, a word *ša.bū:k* «quickly» which is no doubt connected with Persian *šābūk* and one or two more.

Of these the most interesting is *šad* «prince». Ultimately this goes back to Avestan *zāyadhīya*; it can hardly be Sogdian *šēš*, perhaps Old Bactrian (Kushan) *šao* is the nearest.

Another word is *šalu:* «a ladder (or staircase)». It is first found in Uyğur, occurs in the *Kutabyu:* *Bilig* in association with *bayna:* «a rung (or step)», and survives in one or two modern languages. As the early Turks lived in tents they are not very likely to have had their own words for such things, and both words are almost certainly loan words, but so far as I know their origin has not yet been discovered.

There are three more sounds, *l*, *r* and *z* which are common in the medial and final positions, but never occur as initials of pure Turkish words.

Although initial *l-* was strange to the Turks, they do not seem to have had any difficulty in pronouncing it. There are very few early words with this initial and all are certainly or almost certainly loan words. One or two are Sanskrit loan words like *lakšan* which are found only in Buddhist texts. Of the rest, about a dozen in all, most are Chinese loan words like *lu:* «dragon» and *la:v* «sealing wax», one *lešp* «mucus» is Tokharian, and one *la:yu:n* «a drinking mug» is a widely distributed international word probably derived ultimately from Assyrian *laginnu* which was probably brought east by the Manichaean missionaries. Finally there are two animal names, *la:šin* «a falcon» and *layzin* «a pig», the origin of which is uncertain. They cannot be Sogdian as *l* is not a Sogdian sound.

There are no pure Turkish words with initial *z*, and very few in all. Kāšyari records two onomatopoeics *zep zep* and *zak zak*, two plant names, *zarjūnēmud* «sweet vasil», which is certainly, and *zaranza:* «safflower» which is probably, a Sogdian loan word, *zūngūm* «a kind of Chinese brocade», which is certainly, and *zanbi:* or *zenbi:* «field cricket» which is probably a Chinese loan word.

Initial *r-* is a different matter. The early Turks had such difficulty in pronouncing it that they habitually put a prosthetic vowel before loan

words which contained it. Thus in the *Kutab-yu: Biliq* Sogdian *ryž*, pronounced *rēž*, cognate to Persian *rāy* «paradise» is spelt *ārdž*, and Arabic *raṣl* «easy-going, slacks» *draḍl*. The Sanskrit word *ratna* «a jewel» was spelt *rtny* in Sogdian and appears in that form in one Manichaean Uyyur text but is otherwise spelt *ārtini*: or *ārdini*. It became a Mongolian loan word as *erdeni* and still survives in that language.

The only word with initial *r*- listed by Kāšyari is the Ganjak word *rabat* «unpaid forced labour», which is obviously Indo-European and cognate to Russian *рабо́та* «works».

Finally there is the odd case of initial *n*-. The only pure Turkish words in which it occurs are *nā* «what?», various words like *nāḷā*, *nāḷāk* derived from it, and *nāḡ* which has two meanings: «at all» in negative sentences and «a thing». It too is probably connected etymologically with *nā*-. There are in Kāšyari and other early texts about half a dozen loan words beginning with *n*-. Three *nevašigi*, «good spirits», *nīday*, «a whetstone», and *namīza*-, «one's wife's sister's husband», are Iranian, probably Sogdian. One *nag*, «crocodiles», is ultimately derived from Sanskrit *nāga* and one *nom*, «doctrines» and the like, which formed several derived words *nomla*-, *nomliḡ* etc. is ultimately derived from Greek *nomos* and was brought into Turkish by the Manichaean missionaries. It still survives as a loan word in Mongolian. In the 13th century, which is beyond the scope of this paper, several Mongolian words with initial *n*- like *nōkōr* «a servant» were introduced into Turkish.

So much for the phonetic laws. The other rules for identifying loan words are less categorical. The most helpful one can be formulated as follows: if an early Turkish word corresponds more or less precisely in form and meaning to a known word in Chinese, Tokharian, Sogdian or occasionally some other Iranian language or Sanskrit, it can be taken to be a loan word from the language concerned, unless it has a good Turkish etymology, since a reverse borrowing by the foreign language is unthinkable at this period.

Although the Turks were in contact with the Tibetans from an early period and the «Xayan of Tūpūt» sent a representative to Kül Tegin's funeral there are no Tibetan loan words in Turkish earlier than those which occur in late (13th or 14th century) translations of Tibetan Buddhist texts. Most editors of the Kül Tegin inscription credit the representative with the proper name *Bōlōn*; in fact of course this is the Tibetan word *blon* «ministers», spelt *bōlōn*, since the Türktü could not pronounce an initial *bl*-. There are numerous Turkish loan words in early Mongolian and even one in Kitañ, *taolai* «shares», a distortion of *tavišya:n* in its L/R form \**tavi-*

*ya:n*, but there are no Mongolian loan words in Turkish earlier than those which entered the language after Chinggis conquered various Turkish tribes in the late 12th, early 13th century.

There are so many examples of the application of this rule that it is not possible to give more than a few in a short lecture. The most interesting are those from Tokharian or Proto-Tokharian. In a recent paper *Chinese and Indo-Europeans* (*Journal of the Royal Asiatic Society*, 1966 pp. 9ff.) Professor Pulleyblank demonstrated that the Indo-European tribes who were the ancestors of the Tokharians who still survived in Sinkiang in the first millennium A.D. must have reached the frontiers of China in the early second or even late third millennium B.C. They must therefore have been in contact with Turkish-speaking peoples for many centuries. One loan word *lešp* «mucus», probably rather a late one, has already been mentioned. Another is *kūncit* «sesame». Another is *tūmān* «ten thousand». This word which exists in slightly different forms in both Tokharian dialects has no plausible Indo-European etymology. Professor Pulleyblank has told me that he is sure that it is in fact a Chinese loan word. «Ten thousand» in Chinese is now *wan*: in Old Chinese it was something like *myan*, but Professor Pulleyblank believes that at a much earlier date, before Chinese and Tibetan drifted apart, Proto-Chinese, like Classical Tibetan, had initial consonantal clusters and «ten thousand» was something like *tman* or *dman*. Another is *ōkuz* «ox», *okās* in Tokharian-A, *okso* in Tokharian-B.

Another probable example is a common, and at first sight typically Turkish, word *arpa*: «barley». We do not unfortunately know the Tokharian word for «barley», but the Proto-Indo-European form has been reconstructed as \**albhi* and P.I.E. \**bh* habitually becomes *p* in Tokharian.

Another rule for identifying loan words can be formulated as follows: if the language is found to contain more words for slightly different articles of the same kind than its speakers might have been expected to require before they made contact with foreign peoples it is likely that some of them are loan words. This rule must obviously be applied with discretion. It is not surprising that a pastoral people like the Turks have an elaborate vocabulary for livestock of various ages and sexes and the colour of horses' coats. Indeed it is much more significant that they have only two words for pig, no doubt wild pig, and that one of them, *layzin*, is a loan word which soon became obsolete, while the Mongols have a most elaborate vocabulary for pigs, different words being used for wild and domesticated pigs. But when we find that there are over a dozen Turkish words for silk and other fabrics, it is hard to believe that most of them are not loan words.

It is generally agreed that *bāz* «cotton fabric» is ultimately derived from Greek *bassos* «linen»; other words which occur in Kāšyari or earlier texts are *āḡḡiti*-, *ḡit*, *ḡiranḡi*-, *ḡikin*, *ḡuz*, *torḡu*-, *kaḡyar*, *kaḡaḡ*, *kuḡay*, *xuḡiḡ*, *loḡtay*, *ḡalaḡu*: and *zūḡḡūm*. There may be others. Some are actually described by Kāšyari as Chinese fabrics, and it is certain that those which begin with *x*-, *l*-, *k*- or *z*- or contain medial *-f*- or *-x*- are loan words, most of them, but not all, Chinese.

The question of royal titles and titles of office is the most difficult of all. The first Türkü Empire, if that is the right term to employ, took shape in the third quarter of the sixth century A. D.; there had been previous similar «Empires», the earliest of which we have any substantial knowledge being that of the Hsiung-nu, which took shape in the last quarter of the third century B. C. It seems to have been the regular practice of each «Empire» to take over the titles of its predecessor. Some of the early Türkü titles are unquestionably loan words, the question whether there are any which are not depends on the extent to which we regard the rulers of these earlier «Empires» as ancestors of the Turks. The Chinese undoubtedly thought that the Hsiung-nu were ancestors of the Turks, but Professor Pulleyblank in an article called *The Hsiung-nu language*, an appendix to another paper in *Asia Major IX* (1963) pp. 239 ff. has assembled evidence which leads him to believe that that language was related not to Turkish but to the Yeniseyan group, Arin, Asan, Kettish, Kottish. This view is not yet shared by many other scholars, but on balance it would be unwise to assume that the Hsiung-nu did talk Turkish. No-one has yet produced any coherent theory about the language talked by the Juan-juan (Jou-juan etc.), whose «Empire» immediately preceded that of the Turku.

In these circumstances it seems more sensible to assume that all these inherited titles are loan words. Two of them certainly are. *Tarzan* and *tegin* form the wholly un-Turkish plurals *tarzant* and *tegiḡ*. This has long puzzled scholars but the explanation has recently been discovered in a most unexpected way. Some years ago a monument was discovered in Mongolia which has now been identified by Klyashtorny and Livshitz as dating from the reign of a *xayan* of the first Türkü Empire, probably Taspar, T'a-po in the Chinese histories, who reigned from A.D. 572 to 587. The inscriptions on it have been very severely damaged. What survives of the Sogdian inscription on three sides has been published by these two scholars in *Strany i Narody Vostoka VIII*. It contains a list of titles of Türkü dignitaries in the plural including *tarzant* and *toḡunt*. It does not, of course, follow that because these are Sogdian plurals the titles are themselves Sogdian; indeed this is

very unlikely, but they are not Turkish either. Mr. Klyashtorny has told me that the inscription on the fourth side of the monument is in Brahmi not Chinese as stated in the published article. It has not yet been read, but I have suggested to him that it is probably a tribute to Taspar by the famous Indian monk Jinagupta, who took refuge with the Türkü when he was expelled from China by the Northern Chou Emperor.

*Xayan* is traceable back through the Juan-juan, and T'u-yu-hun probably to the late Hsiung-nu, and can safely be regarded as a loan word, so too can *Xatun*, but in this case there is a possible Sogdian etymology. I have already referred to the Iranian title *šah*. In the Šine-usu inscription there is a title *švara*: with a plural *švaraz*; this can hardly be anything except the Sanskrit word *śvara* «lord» which has precisely this plural.

Finally I come to the interesting question whether the loan words in early Turkish, and particularly those which can be put into groups, throw any light on the prehistory of Turkish peoples. There is nothing particularly exciting about the fact that so many of the words for silk fabrics are Chinese loan words; we knew already that the Turks got such things from China.

The Sanskrit title *śvara* is a little more interesting; it was probably obtained through Khotan, which is mentioned once, under the name *Kordan*, as the enemy in the west in line 14 of the inscription of Toñukuk. But by far the most interesting point is the cluster of Tokharian loan words relating to agriculture and animal husbandry, *kunḡit*, *ḡkuz*, probably *bokarsī*: «a wooden plough» which occurs in the *Irḡ Biliḡ* and Kāšyari, and seems to be the Tokharian word *pyākā* «a pointed stick», which is basically what a wooden plough is, and possibly *arpa*. We know that the Proto-Tokharians were pastoralists and believe that they were also agriculturalists to the extent that the physical conditions allowed, as they did in the oases of Sinkiang in the first millennium A.D. and probably over a wider area in the steppes at an earlier date. If *arpa*: really is a Tokharian word this is particularly significant, since palaeobotanical research has recently shown that barley was the first bread grain to be domesticated, that is deliberately grown for human consumption, and that for a long time it was a more important food-stuff than wheat. It is exactly the kind of seed that the Proto-Tokharians would have taken with them when they moved east from their original home somewhere in the Ukraine or the steppes east of the Volga. Nothing is more natural than that the early Turks should have learnt agriculture and animal husbandry from them and borrowed some of their words for things connected with those occupations in the process, and that in fact seems to be what they did.