

CAMBRIDGE
UNIVERSITY PRESS



The Turkish Numerals

Author(s): Gerard Clauson

Reviewed work(s):

Source: *Journal of the Royal Asiatic Society of Great Britain and Ireland*, No. 1/2 (Apr., 1959), pp. 19-31

Published by: [Cambridge University Press](#)

Stable URL: <http://www.jstor.org/stable/25202205>

Accessed: 23/08/2012 08:02

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Cambridge University Press and Royal Asiatic Society of Great Britain and Ireland are collaborating with JSTOR to digitize, preserve and extend access to *Journal of the Royal Asiatic Society of Great Britain and Ireland*.

<http://www.jstor.org>

THE TURKISH NUMERALS

BY GERARD CLAUSON

THE TURKISH LANGUAGES¹ are unusually rich in numerical series, and there are peculiarities in their numerical system which are, I believe, unparalleled elsewhere. The main facts are well known, and it might be thought that there was nothing more to be said about them ; but several interesting points have never been properly discussed. The purpose of this paper is to call attention to them.

There are four distinct numerical series in Turkish :—

(1) Cardinals : one, two, three, etc.

(2) Ordinals : first, second, third, etc.

(3) Distributives : one each, two each, three each, etc.

(4) Collectives : one by itself, two together, three together, etc., as well as a few other words with a numerical basis, such as noun/adjectives for “ twin, triangle, quadrilateral ”, etc., and verbs for “ to do a thing twice ”, etc. On the other hand there are no numerical series either for the numerical adverbs, once, twice, thrice, etc., phrases composed of cardinals followed by a noun which varies from language to language being used for this purpose, or for the

¹ I use this term to include all languages of the Turkish family from eighth-century Türkü, the language of the “ Orkhon Inscriptions ” written in “ Runic ” scripts, which are the earliest substantial remains of Turkish, down to the modern languages of this family still spoken in Turkey, Persia, the Soviet Union, and N.W. China. By “ early Turkish ” I mean Türkü and Uyğur and the Manichæan dialects, which are very close to Türkü and, in their earliest known forms, practically contemporary with it. In an article, “ The Turkish Y and Related Sounds,” in *Studia Altaica, Festschrift für Nikolaus Poppe* (Ural-Altäische Bibliothek, Wiesbaden, 1957), I explained at some length what languages I have covered in my studies and the terminology employed. I use here the system of transcription explained in that article, roughly the Turkish Official Alphabet, with a few added letters and signs to represent sounds not adequately represented by it, notably **z** for the unvoiced velar fricative, the closed **é**, distinguishable from open **e** in early Turkish, and the use of an attached colon to indicate long vowels (**a:** is long **a**, and so on).

Roman numerals indicate a century of the Christian era from VIII onwards.

References to the Türkü inscriptions are to the texts published in H. N. Orkun's *Eski Türk Yazıtları*, T.D.K., İstanbul, 1936–1941.

Kaş. is an abbreviation for Mahmüd al-Kāşğari's (XI) *Divanü'l-Luğati'l-Türk* ; references by volume, page, and line are to Besim Atalay's Turkish translation published by the Türk Dil Kurumu, Ankara, 1939 and foll.

References to the (XI) *Kutadğu Bilig* are to the critical edition by R. R. Arat, T.D.K., İstanbul, 1947.

References to the (XV/XVIII) *Sanglax* are to the MS. belonging to the Gibb Memorial Trust, a reproduction of which will, it is hoped, be published shortly.

fractions which are expressed by such phrases as *üçte bir* "one third".

The Cardinals. The earliest known forms of the cardinals are as follows:—

Units : *bir*, *ékki*(?), *üç*, *tö:rt*, *bé:ş*, *altı*, *yétti*(?), *sekkiz*(?), *tokkuz*(?).

Tens : *o:n*, *yégirmi:*, *ottuz*(?), *kırk*, *ellig*(?), *altmış*, *yétmiş*, *sekkiz*(?)*-o:n*, *tokkuz*(?)*o:n*.

100 yü:z ; *1,000 bin* or *biç* ; *10,000 tümen*.

These are the VIII forms in *Türkü*, *Uyğur*, and the Manichæan dialects. By XI, at any rate in *Xakani*, 80 and 90 had been contracted to *seksö:n* (*Kaş. I 437, 21*), and *tokso:n* (*Kaş. I 437, 16*). In each case *Kaş.* says specifically that these words were crases of the old longer forms.

Some of these words end in consonants and some in vowels. This affects the formation of the other series, and to avoid unnecessary verbiage I shall in future refer to them as C-words and V-words respectively.

It will be noticed that the words for 2, 7, 8, 9, 30, and 50, which contain double inter-vocalic consonants, are followed by a query. In the case of 2 there is a double doubt about the original form. VIII *Türkü* and some Manichæan texts consistently spell the word with an initial *e-*, and there are sporadic spellings with *e-* or *é-* in some modern languages. In the other Manichæan texts, *Uyğur*, and the medieval texts in Arabic script the word is spelt with letters which normally represent *i-* but could also represent *é-*. In nearly all modern languages it is pronounced with *i-*. This is what normally occurs when a word originally contained an *é*, but the point is not free from doubt.

The other doubt in this and the other five similar numerals relates to the (single or double) intervocalic consonants. It is notorious that all Turkish languages are allergic to double consonants, except when a suffix with an initial consonant is attached to a word with a final congruous consonant, and no ordinary basic Turkish word contains a double consonant. It is therefore very remarkable that there is a steady tradition of spelling all these six numerals with a double consonant. This does not, on the face of it, occur in *Türkü*, *Uyğur*, or the Manichæan dialects, but there is no real proof that double consonants would have been written differently from single consonants in the alphabets used for these texts and

some evidence to the contrary. For example, in the "Runic" script words in which a suffix is attached to a root ending in a congruous consonant are spelt with one, not two, consonants. In *I E 7*; *II E 7* **ittı:** (for **ıdtı:**) "gave up, abandoned" is spelt **ittı:**. When the Arabic script, with its convenient device for marking a double consonant with a *taşdîd*, was adopted for writing Turkish the position becomes quite clear. In *Kaş.* the spelling **ıkkı:** (or **ékki** ?) occurs some seventy times, and **iki:** (**éki:** ?) and **ıki:** (**éki:** ?) less than half a dozen times each, and then probably inadvertently. **Yétti:** (*III 27, 9*) is entered under the heading *fa'lal*, which requires a double consonant, although the *taşdîd* is not actually written in the MS. **Sekiz** (*I 365, 14*) is specifically described as an abbreviation (*taxfîf*) of **sekkiz**. **Ottuz** is so spelt (*I 142, 24*); so too is **ellig** (*I 143, 10*). Only **tokuz** on the two occasions on which it is mentioned (*I 437, 16*; *III 127, 14*) is spelt with a single **-k-**, but there is no main entry of the word itself, and the spelling in *Kaş.* of words which are not main entries is notoriously less meticulous.

The subject is also referred to in Mirzâ Mahdî Xân's *Sanglâx*, an XVIII handbook of XV Çağatay (*folio 20 v. 9 ff.*; pp. 107-8 of Sir Denison Ross's edition of the Preface, *The Mabâni'l-Lughat, being a Grammar of the Turki Language in Persian, by Mirzâ Mehdi Khân*, Bibliotheca Indica No. 1225, Calcutta, 1910). It says that there are only a few *hurûf-i muşaddada* in Çağatay, and then mentions only the four numerals **ıkkı**, **yétti sekkiz**, **tokkuz**, adding that they can also be spelt with a single consonant (*taxfîf*). However **éllig**, though not mentioned here, is so spelt in *fol. 113 r. 7*; only **otuz** appears (*fol. 62 v. 10*) with a single **-t-**.

There is another important piece of evidence for the early existence of these double consonants. Modern Chuvash is the direct descendant of the language of a tribe, probably the Proto-Bulgars, which broke away from the main mass of Turks and moved west of the Urals, certainly before VIII and perhaps as early as IV. Chuvash became a written language only recently, and in its modern form shows wide phonetic differences (such as the substitution of **l** for **ş** and **r** for **z**) from standard Turkish. But in spite of these far-reaching changes the language still retains a tendency which must have come down from before VIII, to spell these words with double consonants. The only form of 50 is **allâ**¹; and there are alternative spellings, with and without the double consonants, for 2 (**ıkké**, **ıké**),

¹ See N. K. Dmitriev and others, *Russko-Chuvashskiy Slovar'*, Moscow, 1951.

7 (s'ıçë, s'ıççë), 8 (sakkâr, sakâr), and 9 (täxxâr, täxâr). Only 30 (vâtâr) has lost its double consonant.

In other modern languages *ellig* has retained its double consonant nearly everywhere, but in the other five words there are only sporadic survivals of the longer form. We can, however, confidently say that in the earliest period these six numerals alone among all Turkish basic words were pronounced with a double consonant. There is no obvious explanation of this curious phenomenon.

It will also be noticed that there is a very unusual relationship between the units and the corresponding tens. In most languages the words for one to ten are etymologically independent from one another, but there is a direct etymological connection between the units and the corresponding tens. This applies for example to the Indo-European and Semitic languages (except that in the latter case twenty is usually the dual of ten), and also to Mongolian and the Tungus languages.¹ In Turkish, however, the etymological connection between the units and tens does not start till sixty; there is no connection whatever between two (or ten) and twenty, three and thirty, four and forty, or five and fifty. The question why the connection did not start till this point is perhaps more one for anthropologists than philologists. One possible explanation is that in the remote times when the Turks evolved and stabilized their language they were a pastoral people living in small groups and owning small herds, and that, while they were constantly concerned in counting up to fifty, they never had occasion to talk of higher figures. Another possible explanation is that it is evidence of the superposition of a quinquial on a decimal system. Generally speaking the Turkish is a pure decimal system evolved at a time when things were counted on the fingers (including thumbs) of both hands. It is therefore basically different from the two other known systems of numeration. The sexagesimal system, of which we have survivals in the English system of counting in dozens, having sixty seconds in a minute, and so on, was of course evolved at a time when things were counted up to twelve on one hand by touching each of the three joints of each of the four fingers with the tip of the thumb and using the fingers (and thumb) of the other hand to count up to sixty (five dozens). The third known system is the vigesimal one, of which we have a survival in the English

¹ This fundamental difference between these languages and Turkish is an important argument against the theory that they are genetically connected.

numeration by scores. We seem to have got this from the Continent, where French-speakers, except in parts of Eastern France, Belgium, and Switzerland, call seventy *soixante dix* and have similar expressions for eighty and ninety. It is generally accepted that the French inherited this system from the Basques, who have independent words only for one to ten, a score, and a hundred, and build up 11 to 19 with compounds of ten and the units and 21 to 99 with combinations of one score (two, three, and four score), and the words for 1 to 19.¹ So far as I am aware, no one has ever explained how this system came into existence. My own flippant suggestion is that it was evolved in a dry climate where people could sit down to count and use their toes as well as their fingers.

Another interesting feature of the Turkish system is the difference between sixty and seventy, **altmış** and **yétmış**, and the phrases **sekkiz(ş)on** and **tokkuz(ş)on** for eighty and ninety. Except in these two words the suffix **-miş/-miş** occurs in Turkish only as a verbal suffix forming a kind of participle, and no satisfactory explanation of its use in these two numerals has ever been produced. They go back to an early stage in the language since in Chuvash 60, 70, 80, and 90 are **utmäl**, **s'itmël**, **sakärün**, and **täxärün**. On the other hand, they may not go back to a very remote past, since the North-Eastern languages (Khakas, Tuvan, Mountain Altay, etc. to give them the names by which they are now known in the Soviet Union), which have not diverged to any very great extent from standard Turkish, seem to preserve traces of an earlier stage of evolution. In them 60 and 70 are **aldan/altan** and **çeden/çeton** which are obviously modern forms of **altı: on** and **yéti: on**. It is of course possible that these are not archaic survivals but neologisms, but this is not very probable.

It should be added that in one language which preserves some very archaic traits, that of a Turkish tribe in Kansu, N.W. China who call themselves **Sarığ Yuğur** (Yellow Uyğur),² all the tens above thirty, and sometimes twenty and thirty also, are formed by appending **on** to the appropriate unit; but this language has been so much under Chinese influence that this practice may well be a recent imitation of Chinese, not an archaic survival. **Per yuz** for 100 is certainly a Chinese and not a Turkish idiom.

Apart from those scholars who still hold by the Altaic theory

¹ See W. J. van Eys, *Outlines of Basque Grammar*, London, 1883, p. 27.

² See S. E. Malov, *Yazyk zheltykh Uyğurov*, Alma Ata, 1957, p. 178.

and contend that Turkish, Mongolian, and Tungus are descended from a common ancestor, no one has ever suggested that the words for the units, tens and 100 are anything but pure Turkish. Equally, starting at the other end, no one would now seriously contend that 10,000, **tümen**, is other than a loan word from "Tokharian", either "Tokharian A" (Agnean) *tmām*, or, more probably, "Tokharian B" (Kuchæan) *tumane/tmāne*. The word first appears in VIII Türkü and Uyğur inscriptions. In *II S. 1 bir tümen artuk: yéti: bir süg* "17,000 troops", and *II S. 8 üç tümen süg* "30,000 troops", its numerical significance is precise; but in *I N. 12 bir tümen ağı*: "10,000 precious things" and *Şine Usu E. 9 bir yıllık tümen күnlük* "a thousand years and ten thousand days" (both phrases with a very Chinese flavour), it hardly means more than "an indefinitely large number". This is certainly how *Kaş.* understood it for in *I 402, 5* he translated **tümen** = *al-kaṭīr* "a great many" and **tümen miñ** "a thousand thousand". Thus it looks very much as if the word had first been taken into Turkish as one for a larger number than any hitherto expressed and was only by degrees quantified in its true meaning of 10,000.

So far as I am aware, no one has ever yet suggested that the word for 1,000 is other than pure Turkish, but there do seem to be some indications to the contrary. In the first place it is not unusual for primitive peoples to borrow foreign words for their higher numerical denominations. The Basques, for example (see van Eys, loc. cit.), have no native word for anything higher than 100; *milla*, 1,000, is obviously a loan-word, presumably from Latin. In Turkish **tümen** is a loan-word acquired after they had come into contact with the "Tokharians" (**öktiz** "ox" from *okso* is another); by parity of reasoning 1,000 might also be a loan-word acquired at a rather earlier period. Secondly there is a suspicious resemblance between the Turkish word for 1,000 and the Chinese word for 10,000, in modern Chinese *wan*, in the VIII Tibetan transliterations *hban/hbun* (i.e. *mban/mbun*) (see F. W. Thomas and G. L. M. Clauson, "A Chinese Buddhist Text in Tibetan Writing," *JRAS.*, 1926, p. 518), in Karlgren's Ancient Chinese *mǐ²wm*, and Archaic Chinese *mǐ²ǎn* (*Grammata Serica*, Stockholm, 1940, p. 197). The Chinese character for this word, in its earliest form, is a picture of a scorpion and it is very likely that the word originally meant "a very large number", as many as the insects, a meaning which it still retains in a good many phrases, and was only later quantified

as 10,000, the next denomination higher than the then existing highest denomination, 1,000 (*ch'ien*). Finally there is some uncertainty about the pronunciation of the word itself in VIII Türkü. In *II S. 1* and *Şine Usu, line on edge*, it is **bi:ŋ**; in *Şine Usu E. 9* and perhaps *E. 11* it is **bi:ŋ**; and in *Irk Bitig 48 (XXXII)* it is **mi:ŋ** or **mi:ŋ** (the two are indistinguishable). It can of course be argued that **bi:ŋ** is a misspelling (misspellings do occur in these inscriptions), and the sound change **b** > **m** in words containing nasals is a common phenomenon in Turkish. But it is equally true that variations in pronunciation are particularly common in loan words.

It is, therefore, not impossible that **biŋ/bi:ŋ**, as well as **tümen**, is a loan word. This leaves **yü:z** poised rather precariously between the words which are certainly pure Turkish and those which may be loan words. It looks good Turkish; it is certainly not Chinese or "Tokharian", and there is nothing to suggest that it is a loan word from any other language; but it may not originally have meant specifically 100. If in the earliest period the Turks habitually counted only up to 50, they would certainly have wanted a word for "an indefinitely large number" and that may originally have been the meaning of **yü:z**. The other meaning of **yü:z** is "a face", but it would probably be unduly fanciful to suppose that they got their first indefinitely large figure from the concept of "all those faces".

Another very odd characteristic of the Turkish numerical system, so far as I am aware not precisely paralleled elsewhere (certainly not in Mongolian or Tungus), is the method used to express "broken tens", that is 11, 12 . . ., 21, 22 . . ., and so on. This is done in early Turkish by placing a unit *before the next highest ten*; for example 11 is **bir yégirmi**, 23 **üç ottuz** and so on. This is of course fundamentally different from the subtractional method seen in the Latin idiom for 18 "twenty less two". It looks very much as if this method must have evolved at a time when the Turks did not count above 50; **bir alti: o:n** would have been very clumsy and **ékki: yü:z** frankly ambiguous, since it might be either 98 or 200. This was still the standard practice in Türkü, the Manichæan dialects, and Uyğur, at any rate for numbers up to 89, but for higher figures the higher denomination precedes the lower. In Türkü the two were connected by **artuki**: "and in addition to it", see the phrase in *II S. 1* quoted above, but later this word was omitted. The standard Uyğur method is exemplified in the numeration

of the pages of manuscripts. In the MS. of which parts were published in F. W. K. Müller's *Uigurica IV* (S.P.A.W., 1931) the pagination runs up to **üç yüz beş otuz** 325; it is to be noted that in *Uigurica III* (A.P.A.W., 1922) 79 is **tokuz sekiz on** (p. 51), but 95 is **tokuz on beş** (p. 57). In late Uyğur the latter method is used also for lower figures, e.g. **otuz bir** 31, and by XI Xakani this is the standard method for all "broken tens". The only language later than early Uyğur in which the older method still survives is that of the Sariğ Yuğur (see Malov, op. cit., p. 178) where 11 is **per yığırma** and 21 **per otuz**, but 31 is **uçon per**, which seems to show that the older method survives only up to 29.

Finally it is interesting to consider what light the cardinals throw on the fundamental structure of the Turkish language, since they are all basic words not susceptible of further analysis into component parts. There are seven (or eight) monosyllables: **bir**, **üç**, **tört**, **beş**, **on**, **kırk**, **yüz** (and **biç/biç**); seven dissyllables: **ekki**, **altı**, **yetti**, **sekkiz**, **tokkuz**, **ottuz**, and **ellig**; and one trisyllable **yégirmi**. This is probably a fairly characteristic pattern of early Turkish as a whole, except for the high proportion of monosyllables. It will be noted that one monosyllable and three dissyllables end in **-z**; this again is fairly characteristic, though on the high side. It has from time to time been suggested (see, for example, C. Brockelmann, *Osttürkische Grammatik*, Leiden, 1954, *para. 120a*) that **-z** is a suffix for, or at any rate an indication of, a dual; the suggestion does not stand up to careful analysis, and is certainly not supported by these words, of which one **tokkuz** designates an odd, not even, figure, and the other three cannot possibly be explained as duals of their halves.

Ordinals. These are all, in principle, formed the same way. In early Turkish this was done by adding **-nç** to V-words, and **-nç** preceded by a euphonic vowel, **-u/-ü-** for words containing rounded vowels and **-ı/-i** for the rest, to C-words. In practice, as *Kaş.* points out (*III 373, 9*) **birinç** is rare, being usually replaced in Uyğur by **başınkı**, and after XI by the Arabic equivalent *awwal*. Exceptionally the oldest form of "second" is **ékinti**, which is the only form in Türkü, the Manichæan dialects, and Uyğur. **İkinç** (or **ékinc**?) first appears in XI *Kaş.*, as an alternative to **ikindi**: (*sic*); this latter word still survives in several modern languages, but now only in the meaning "afternoon prayers", or more vaguely "afternoon". The remaining ordinals are all regular, for example

üçünç, törtünç, béşinç in *Kaş. III* 448-9. It is not impossible that **ékinti:** is a survival, the only one, from an earlier period when the suffix was **-ntı/-nti:**, and that this had become abbreviated except in this case to **-nç** by VIII; the sound change **-tı > ty > ç** is a common phonetic phenomenon.

Be that as it may, **-nç** did not long remain as the standard ordinal suffix; by about XV it was becoming obsolete and being replaced by **-nçı/-nçi**, or in some languages **-nçı/-nçi**. The shorter form is the only one known to *Kaş.*, and is generally speaking the standard form till nearly XV; but the longer form appears as an alternative as early as the *Kutadğu Bilig* (almost exactly contemporary with *Kaş.*) no doubt *metri gratia*, and in XV Çağatay the two forms are used indifferently, with preference in prose texts for the longer one. In the *Sanglāx* the relationship between the two forms is rather neatly reproduced, **ikınç** being translated *duyūm* and **üçünç siyyūm**, while **ikinci** and **üçüncü** are translated *duyūmīn* and *siyyūmīn*. The Persian shorter and longer forms are completely synonymous. At the present day the shorter forms in **-nç** seem to survive only in Sarığ Yuğur.

Theoretically an ordinal can be formed from any cardinal, and in the list of Chapters in the Vienna MS. of the *Kutadğu Bilig* the numeration runs up continuously to **yétmiş ikinci**, but it would be difficult to find anything higher.

Distributives. These too are in principle all formed the same way, by adding **-ar/-er** to C-words and **-ar/-er** preceded by a euphonic consonant to V-words. In early Turkish this was **-r-**, but this was later replaced by **-ş-**. In theory a distributive can be formed from any cardinal, but in the nature of things distributives are rarer than numerals of either of the two preceding classes and in most cases no occurrences, or at any rate no very early occurrences, can be traced; indeed in some modern languages the whole series seems to have disappeared. **Birer** is recorded as early as Uyğur (see Bang and von Gabain, *Analytischer Index zu den fünf ersten Stücken der Türkischen Turfan-texte* (S.P.A.W., 1931), p. 15). **İkirer** (**ékirer**?) occurs in Uyğur in such phrases as **tört öđ içinde yana ikirer öđ ađrılır** "again, in each of the four seasons (of the year) two (sub-)seasons are distinguished" (*Türkische Turfan-texte VI* (S.P.A.W., 1934), line 325), and survives at any rate until XV Çağatay, since in the *Sanglāx* (*folio 109 r. 23 and 26*) both **ikirer** and **ikişer** are listed, the first with a quotation from Nawā'ī,

and both translated *dū tā* "two shares (each)". **İkkişer** is listed in a XV (before A.D. 1426) Kıpçak vocabulary (*Et-Tuhfet-üz-Zekiyye*, T.D.K., İstanbul, 1945), and **ikişer** occurs in a letter from Uluğ Muḥammad Xān of the Golden Horde to Sultan Murad II dated A.H. 831/A.D. 1428 (A. N. Kurat, *Altın Ordu, Kırim ve Türkistan Hanlarına ait Yarlak ve Bitikler*, İstanbul, 1940, p. 9). **Üçer** is used in a late Uyğur contract (No. 34 in W. Radloff, *Uigurische Sprachdenkmäler*, Leningrad, 1928); **üçer böz aldımız** "we (two persons) have each received three lengths of cloth". "Six each" is another word of which both earlier and later forms are known. **Altırar** is used in the contract just quoted; **altırar bözni köni bérirbiz** "we undertake to give six lengths of cloth each". **Altısar** is recorded in XIX as occurring in the N.W. (Kazan, Kırim, etc.) and S.W. (Osmanli, etc.) languages and no doubt existed earlier. It is mainly in these language groups that distributives still survive.

Distributive numerals seem to be found in very few languages other than Turkish. They do exist in Mongolian and perhaps the Tungus languages, but the suffixes used are entirely different. In Classical Mongolian (see N. Poppe, *Grammar of Written Mongolian*, Wiesbaden, 1954, p. 55) the suffix attached to the ordinals (usually abbreviated) is traditionally spelt **-ğad/-ged**. I do not know how old these forms are; none seem to occur in the Secret History (the oldest substantial Mongolian text, mid-XIII), but this may be fortuitous. In the Tungus languages (see J. Benzing, *Die tungusischen Sprache*, Wiesbaden, 1955, p. 105) the suffixes used are **-ta/-te** or **-tal/-tel** but these seem to be merely plural suffixes used in a special sense.

Collectives. These too are, in principle, all formed the same way, but there are traces of some uncertainty regarding the V-words. In the case of the C-words the suffix **-ağü:/-egü:** is attached to the cardinal; in the case of the V-words the suffix seems originally to have been **-ğü:/-gü:** similarly attached, but later, in some cases perhaps even in the earliest period, the suffix was **-ağü:/-egü:**, the final vowel of the cardinal being elided. The collective form is certainly as old as VIII *Türkü*, and, as it exists in the N.E. languages, may be even older, but there do not seem to be any indisputable traces of it in Chuvash. In theory a collective can be formed from any cardinal, but in practice only a limited number can be traced. For example, Abu Ḥayyān in his *Kitābu'l-Idrāk li-lisāni'l-Atrāk*

(published in İstanbul, 1931), a handbook of XIV Kıpçak, in a paragraph on this form (*pp.* 114-5) quotes only the following (some with minor spelling errors in the printed edition): **biregü.**, **ikegü.**, **üçegü.**, **dördegü.** (*p.* 22, **dördewü.**), **béşegü.**, **altağu.**, **yédegü.**, **seksegü.**, **toksağu.** This series too has become obsolete in some modern languages, but not the same as those in which the distributive series has, since the latter survives mainly in the N.W. and S.W. groups, while the former survives mainly in the N.E. and N.C. and not at all in the S.W.

The earliest occurrences of **biregü** "one by itself" seem to be in the *Kutadğu Bilig* (*verses* 343, 1238, etc.). By XV, at any rate in Çağatay, the **-g-** had been elided and the word was pronounced **bire'ü** or **birew**. But even at that time the word was no longer current in Osmanli, and in the *Abuşka*, a mid-XVI Osmanli handbook of Çağatay published in V. de Veliaminof-Zernof, *Dictionnaire Djağatai-Turc*, St. Petersburg, 1869, there is an entry (*p.* 32) "**aw** (or **ew**?) *bir kimse* 'a person', also used in the phrase **bir aw** (**ew**?)", with a quotation from Nawā'ī. Substantially the same entry occurs in the *Sanglāx* (*fol.* 53 r. 27), which shows that by XVIII the word had also become obsolete in the South Central group of languages. The word survives, in an abbreviated form, in some North-Eastern and North Central languages, but not apparently elsewhere.

The word for "two together" was originally spelt **ikigü** (or **ékigü**?), which is fairly common in Uyğur. It was still spelt **iki:gü.** (or **éki:gü.**?) in *Kaş.* (*I* 45, 3), but in the MSS. of the *Kutadğu Bilig* beside this spelling we find also **ikegü** (*verses* 331, 875, 1463, etc.), and **ikegü** (or **ékegü.**?) also occurs in some late Uyğur documents. The word still survives in several modern languages, mainly in the North-East and North Central group, none in the South-West, usually in a much distorted form (**ekkü.**, **ikü.**, **ekö.**, **ökö.**, **ikew**, etc.).

Üçegü occurs as early as VIII Türkü (twice in the inscription of Toñukuk), in the *Kutadğu Bilig* (*verse* 802), and, with similar distortions, in much the same modern languages as the preceding word. **Törtegü** occurs in the *Kutadğu Bilig* (*verse* 4502). **Béşegü** is found in Uyğur (see *Analytischer Index* . . . *p.* 15). The earliest known form of "six together" seems to be **altağu.** in XIV Kıpçak (see above) and XIV Rabğuzi's *Qısaşu'l-Anbiyā*. There is unfortunately no evidence for an earlier form **altıgü.** The modern forms are **alda.**, **alto.**, **altu.**, and **altaw**. **Yétegü** is recorded in Uyğur

(*Analytischer Index* . . . p. 57), but the text concerned is a late one and this may not be the earliest form.

Brockelmann in his *Osttürkische Grammatik*, Leiden, 1954, *para.* 130c, alleged that the suffix **-la/-le** was also used to form collectives, quoting as examples **birle**, **ikile**, and other words discussed below. There is no doubt that he was wrong. **Birle**: is one of the four "old Turkish postpositions" discussed in K. Grønbech's *Der Türkische Sprachbau*, Kopenhagen, 1935, p. 35. His theory was that it was made up of **bir** and "the emphatic postposition **-le**". Be that as it may, the word was originally only a postposition, and its only meaning was "with". **İki:le**: is a different matter. It occurs once in *Kaş.* (*III* 244, 20) in the proverb **bir tilkü: teri:sin iki:le: soyma:s** "You cannot flay one fox twice", and also at least once in the *Kutadğū Bilig* (he slept a long time and woke in a fright) **yumulmadı yandru ikile közi** (*verse* 5674); Brockelmann took **ikile közi** to be "his two eyes", but the verse obviously means "his eyes did not close again a second time". The most reasonable explanation is that **iki:le**: is an abbreviated gerund of the verb **iki:le:-** "to do (something) twice", which existed in several Turkish languages and still survives in some.

The later occurrences of **ikile** and the other words quoted by Brockelmann in *para.* 130a are really abbreviated forms of words quoted in *para.* 130c. There was in XIV an outbreak in Turkish languages in the Persian area of collectives formed with the suffixes **-e'üle/-e'ülen** and the like, e.g. **ike'üle/ike'ülen** "two together". These are apparently first recorded in XV Çağatay and survive, in abbreviated form, in some modern languages. In this particular case the modern forms include N.E. **igile**, **ikke:leñ**, **ikkö:leñ**; S.E. **ika'olan**, **iki:le**; S.C. **ikkele**, **ékéla**, **ék(k)éle**; N.W. **ekovian**. These are genuine collectives, but they are not genuine Turkish. The Mongol conquest was followed in XIV, particularly in the area ruled by the Ilkhanids, by a mass Mongolian invasion of the Turkish language. The *Sanğlāx* lists a number of Mongolian loan words in Çağatay described as such, and an even larger number not so described. What was borrowed was not only Mongolian noun/adjectives, but also verbs, which were conjugated as if they were Turkish, and even, in some cases, suffixes, with one of which we are concerned here.

The Mongolian collectives were formed by attaching to the cardinals, sometimes abbreviated, the suffixes which were written

in Classical Mongolian as **-ğulan/-gülen** (Poppe, *op. cit.*, p. 55) but actually pronounced **-'ulan-/'ülen**. These are among the oldest elements in Mongolian and several of them occur in the Secret History (see E. Haenisch, *Wörterbuch zu Manğol un Niuca Tobca'an*, Leipzig, 1939, *passim*). **Ike'ülen** is the Turkish word **iki**: with the Mongolian suffix **-'ülen** and all the other similar collectives are so formed. There is no connection between these collectives, which appear in Turkish only after the Mongol invasion, and only in areas in which the influence of the Mongolian language was overwhelming, and the two old words **birle**: and **iki**:le: which existed long before the conquest and are not collectives.