

Anatolian **meyu-* '4, four' and its cognates

The article reexamines the reconstruction of the Proto-Indo-Hittite counting system. Using comparative and typological arguments, I attempt to show that the original Indo-Hittite system was structured as '1', '2', '3', '>3/several/many'. Words for a new numeral, '4', were introduced in Anatolian and Narrow Indo-European branches separately after their split. Anatolian **meyu-* '4' (Hittite *meyu-* ~ Luwian *mauwa-*) is unetymologizable within Indo-European. I suppose that it is an archaism and propose an external (Proto-Altaic) non-numeral cognate for this Anatolian stem. Anatolian **meyu-* is one of the rather numerous Nostratic roots that were retained in Anatolian, but lost in the Narrow Indo-European branch.*

Keywords: Linguistic reconstruction, Numerals, Counting system, Nostratic, Indo-European, Anatolian, Hittite, Khoisan, Amerind

§1. The archetypal human counting systems appear to belong to the following types:

- 'this one', '>1, several/many'
- '1', '2', '>2, several/many'
- '1', '2', '3', '>3, several/many'
- or perhaps '1', '2', '3', '4', '>4, several/many'

Nowadays such systems are attested in traditional cultures of Amazonia, South Africa, Papua etc. (see Greenberg 1990 w. lit. for examples). When the new economic situation leads to the expansion of the counting system, an increase in the number of numerals (i.e. the integers which can be expressed in this language) is achieved in three ways:

- using lexemes other than numerals in the sense of numerals (e.g., metonymy based on body parts;¹ rare "fraternal" strategy²);
- borrowing numerals from neighboring languages or *lingua franca*;
- forming new numerals by mathematical operations with the original numerals, such as addition and (less frequently) subtraction and multiplication (2-, 3-, 4-, 5-, 6-cycle counting systems are attested cross-linguistically, see Lean 1992 w. lit., Greenberg 1990 w. lit.) or, as in Eskimo languages, with the help of a word/affix meaning 'in addition', 'of the other hand' etc.

It goes without saying that these methods can be combined within one language:

* I am grateful to Harald Hammarström (Göteborg), George Starostin (Moscow), Ilya Yakubovich (Chicago) and Mikhail Zhivlov (Moscow) for their useful and important remarks on the preliminary version of this paper. Naturally, all the infelicities are the author's only. Language names and their genetic attribution are given according to Ethnologue 15, unless otherwise mentioned.

¹ 'Eye(s)' > '2' (Epps 2006 for Dâw and Hupdë [= Hup]) or the common shift '(left) hand' > '5'.

² The expression "has a brother/sibling/companion" for even numerals above '3' and "has no brother/sibling" for odd numerals above '3' in several languages of Maku (Nadahup) family and in Tucano (Epps 2006).

- as one counting system. For example, Gagadu (< Gunwingguan < Australian) has un-derived '1' and '2', '3' is '2+1', while '5' is based on the word for 'hand' (Harvey 1992: §8.5.6). Naro (< Central Khoisan) has a decimal system, where '1'–'3' represent un-derived forms, '4'–'10' are borrowed from Nama, and '5' and '7' can be alternatively expressed by, respectively, such proper words as 'hand' and 'index finger'; the latter can be used alongside loanwords from Nama (Visser 2001; for the more archaic Naro system see Schniz 1891: 545);
- as parallel counting systems. E.g., Haruai (< Trans-New Guinea) has four numeral systems, as described in Comrie 1999:
 - a) abstract system '1', '2', '2+1', '2+2', '>4, many';
 - b) body-part based system, going from '1' up to '18' (or even higher, up to '36');
 - c) clear decimal counting, borrowed from Tok Pisin (English-based creole);
 - d) English numerals, adapted to the local pronunciation, i.e. the contemporary Tok Pisin system.

A question that deserves special investigation is what kind of lexemes other than numerals can be used to form new (larger) numerals. Although metonymy (e.g., based on hand and fingers) is the most widespread method, other scenarios are also possible.

§2. According to the Tower of Babel Project, it is possible to reconstruct numerals '1' and '2' for the Nostratic proto-language.³

Nostratic *VrV '1' >
 Uralic: Finno-Ugric *erä 'part, single, 1';
 Kartvelian: *ert- '1';
 Dravidian: *or- '1'.

Nostratic *UdV '1'(?) >
 Indo-European: *edh- (> Slav. *ed-inъ '1');
 Altaic: *iude 'to imitate, to simulate';
 ? Uralic: Finno-Ugric *ükte '1' (metathesis kt > tk?);
 Kartvelian: *od- 'only, yet';
 Dravidian: *ud- 'match, pair'.

Nostratic *tu?V (~ *tuqV) '2' >
 Indo-European: *dwo- '2';
 Altaic: *(i)ubu '2';
 ? Uralic: Balto-Finnic *toinen / *toise- 'second'?⁴
 Kartvelian: *tqu-b- 'twins'.⁵

³ The Tower of Babel Project unites Indo-European, Altaic, Uralic, Kartvelian and Dravidian proto-languages as *Eurasiatic* macrofamily, provisionally reserving the term *Nostratic* for the next level, where *Eurasiatic* is linked to the Afroasiatic (Semito-Hamitic) macrofamily as its closest relative. Since *Eurasiatic* is not a common term thus far, in my paper I use the more usual *Nostratic* as the name of the proto-language that yielded IE, Altaic, Uralic, Kartvelian and Dravidian proto-languages. The genealogical tree of Nostratic (<http://starling.rinet.ru/images/globet.png>), based on 35-word lists and certain morphological data, shows that the family diverged ca. 10–9 millenium B.C., i.e. in the Mesolithic or Early Neolithic period.

⁴ Problematic; the derivation from deictic *to 'that' is more probable from the phonetic point of view.

⁵ If not a Northwest Caucasian loan; cf. Proto-West *tq̄l:^wA '2', having reliable East Caucasian cognates, NCED: 924.

Nostratic *jVrV '2, pair' >
 Altaic: *ǒre (~ *ǒri) 'other, one of two'
 Uralic: *jOrtV 'friend'
 Kartvelian: *jor- '2'
 Dravidian: *ir- '2'

No common Nostratic form for the numerals '3', '4' or higher can be reconstructed, but it is important that both '3' and '4' are reconstructable for the Dravido-Uralo-Altaic branch of the Nostratic macrofamily:

Nostratic *ηü- '3' >
 Altaic: *η[iu] '3';
 Dravidian: *mū- '3'.

Nostratic *ńVIV '4' >
 Altaic: *ńöl- '4' (Tungus-Korean isogloss, Blažek 1999: 92, 130)
 Uralic: *ńeljä (~ *neljä) '4';
 Dravidian: *nāl- '4'.

Thus we must accept the following counting system for Proto-Nostratic: **"1, 2, >2/several/many"**. The Dravido-Uralo-Altaic protolanguage increased its numeral row up to **"1, 2, 3, 4, >4/several/many"**, while IE and Kartvelian expanded their systems separately. Note that none of the common IE numerals have any external Nostratic etymologies except for IE *duo- '2'.

§3. It is well-known that some IE languages show evidence for an old base-4 counting system (Luján Martínez 1999; Гамкрелидзе—Иванов 1984: 846, 849 ссл. [= Gamkrelidze & Ivanov 1995: 744, 746 ff.]; Winter: 1992; Blažek 1999; etc.). As a matter of fact, internal IE reconstruction points to a simple system **"1, 2, 3, 4, >4/several/many"** for the early stage of the IE proto-language, which later developed a base-4 counting system, and, still later, the decimal system that is well attested in known IE languages (except for Anatolian, see below).

The following facts are important for the hypothesis outlined above.

1. Numerals '2'—'4' are underived, while '5'—'10' give grounds for speculations concerning their underlying meanings and root connections. E.g., *penk^{ue}- '5' can mean 'fist', *dek^m(t) '10' possibly means 'right hand' or 'two-hands', *sept^m '7' is, without a doubt, a Semitic loan (Иллич-Свитыч 1964: 7). One should also mention the impossibility of reconstructing a single proto-form for IE '6' and '9', despite the similarity of their attested phonetic shapes in known IE languages. Further see Luján Martínez 1999; Mallory & Adams 2006: 308 ff.; Blažek 1999: 234 ff.; Bomhard 2008.
2. In some IE languages the numerals '1', '2', '3', '4' are declinable and grammatically agree with the noun, while numerals '5'—'10' are indeclinable and stay in apposition to the noun. To be sure, it is not uncommon in languages around the world for the lower numerals to show a specific syntactic and/or morphologic behavior, and some have attempted to explain this in terms of frequency and "cognitive reference points" (see Hammarström 2004 w. up-to-date literature and discussion). Two points, however, must be stressed. First, we do not know any language where numerals show syntactic unification at an early stage and syntactic individualization of lower numerals at a later one. Second, it is entirely unclear why the high frequency of lower numerals should provoke syntactic and/or morphologic complication of the system.

3. Iranian **ašti-* '(breadth of) four fingers' (measure of length), retained in Avestan compounds, see Henning 1949; ЭСИЯ 1: 252; Emmerick 1992: 300, Edelman 1999: 226–227. Goes back to IE **okt-*.
4. The dual, not pl. ending in nom.-acc. of IE **oktō* '8' (OInd. *aṣṭā(u)* [RV+; obliques with pl. endings], Grk. ὀκτώ, Lat. *octō*, Goth. *ahtau* etc.) points to '8' as 'doubled 4, 4×2'. This pattern has reliable typological parallels.
5. Proto-Kartvelian '4', reconstructed as **otxo-* in Klimov EDKL: 145–146 and early as **o(s₁)tx(w)* in Климов ЭСКЯ: 150 (Georgian *otx-* ~ Svan *wošt(xw)*). Be it an early IE loan in Proto-Kartvelian (with metathesis⁶) or even vice versa, this form cannot be separated from Iranian **ašti-* and IE **oktō*. Kartvelian **otxo-* strongly indicates that before the appearance of the enigmatic **k^uetuor*, the proto-IE root for '4' was **okt-*.
6. It is possible to derive **neu(e)n* '9' from the root **neu-o-* 'new' (cf. especially the prefixal Grk. form ἐννέα and Arm. *inn* where the first part can be identified as **en* 'in'), which points to 'nine' as an 'onset of the new (set of numbers)'. This etymology is supported by exclusive Iranian forms for '9' collected in Edelman 1999: Osset. *far-ast* 'beyond 8' (**pāra-ašta*), Khwarezmian {š'δ} 'increased' (**frad-*) or 'beyond 8' (**fra-ašta*), Pashto dial. *terai* 'passed', Waziri *tār pa wōta* 'passed beyond 8'. Edelman claims that the above-mentioned Iranian forms for '9' are innovative, since they are not used in words for '19', '90' or '900', derived from Iranian **nawa* '9'. There are, however, no octimal counting systems in the region, from which the Iranian dialects could borrow or calque the numeral '9' as 'next to 8'. On the other hand, it is natural that words for larger numerals '19', '90', '900', seldom used in natural speech, could be borrowed from the neighboring languages that possessed the stem **nawa* for '9' (more about IE '9' see in fn. 13 below).

§4. Greenberg 1990: 276 formulates the following universal (No. 6): "The largest value of L in system with only simple lexical representation is 5 and the smallest is 2".⁷ His example of L=5 is the Guana (Arawakan) system: "**1, 2, 3, 4, many**", and he especially notes that "the most common values for L are 3 and 4".

As was claimed in §3 above, we should reconstruct "**1, 2, 3, 4, many**" for the early IE protolanguage. But when we proceed deeper, reaching the Indo-Hittite stage (using E. H. Sturtevant's term, see, e.g., Lehrman 2001 and other works by Lehrman on the issue of the Indo-Hittite proto-level), it turns out that a common word for '4' cannot be reconstructed: Anatolian has **meyu-* as '4' and no traces of IE roots **okt-* and **k^uetuor* in the known lexicon.⁸

As is typical of cuneiform orthographies, numerals in Hittite and Luwian texts are almost always written logographically, while their phonetic attestations are scanty. The following Anatolian cardinal numerals (or numeral stems) are currently known in phonetic realization (Eichner 1992; Goedegebuure 2006; Hoffner & Melchert 2008: 153 ff.; Neumann 2007; HEG; HED):⁹

⁶ Accessive clusters (like *KT*) are very unusual in Kartvelian languages and Proto-Kartvelian, as opposed to the frequent decessive clusters (*TK* pattern).

⁷ "L" is "the next largest natural number after the largest expressible in the system".

⁸ Hitt. *kutruwan-* c. 'witness' can hardly be compared with **k^uetuor* '4'. For plausible Baltic cognates see HED K: 298 ff.

⁹ Note that several other Lycian numerals are known by now, all of them without an exact translation and/or transparent etymology: *aitāta*, *mupñme* ('4' if to Luw. *mauwa-*?), *nuñtāta* ('19'/'90' if to HLuw. *nuwa(n)*?), *qññā-kba* ~ *qññā-tba* ('12'?), *sñta-* (not '100'), *tupñme*, and perhaps *kñma-*.

	Hitt.	CLuw.	HLuw.	Lyc.
'1'	šia-		s... (?)	
'2'	ta-		tuwali-	(A) kbi-, (B) tbi-
'3'	teri-		tara/i-	(B) tri-
'4'	miyu-/meyu-	mauwa-	mi... / mu... (?)	
'7'	šiptam-	šaptam-		
'9'			nuwa(n?)	

All of these forms, except for Hitt. *šia-* '1' and Hitt. *meyu-* '4' (with a Luw. cognate), are matched by IE numeral stems (**d̥yo-* '2',¹⁰ **tri-* '3', **septm̥* '7', **new̥* '9').

The Hitt. stem *šia-* '1' corresponds to the IE pronominal demonstrative stem **so-*, fem. **sā/sī* (Pok.: 978–979), cf. especially the OInd. variant *syā*, fem. *syā* 'jener, der'. This numerical meaning for *šia-* seems unparalleled in the other IE languages (cf., however, EDHIL: 751 for doubtful Grk. *ἄ* f. 'one'), but the semantic shift 'this' > '1' is quite common.

The common Anatolian numeral '4' remains without any clear IE cognates. The paradigm of Hitt. *miyu-/meyu-* is collected in HED M: 116 ff. Although the graphic representation of its attested forms is usually ambiguous, its Anatolian proto-form can only be **meyu-*, in view of Cuneiform Luwian *mauwa-* '4'.¹¹

So far, three controversial IE etymologies have been proposed for Anatolian **meyu-* '4' (see HED M: 116 ff. for references):

- 1) a *u*-stem from IE **mei-* 'to diminish' (Pok.: 711), i.e. '4' as 'little (hand)', 'hand without thumb'. This etymology is morphologically possible, but semantically unparalleled in any of the world's languages;
- 2) a *u*-stem from the root **mei-*, same as in Hitt. *mai-* 'to grow; to increase; to prosper' (rather to Lat. *mātūrus*, not to IE **mē-* 'to measure' [Pok.: 703]), i.e. '4' as 'grown' ('bigger than 3'). This semantic development also seems to be unparalleled;¹²
- 3) a *u*-stem from IE **mē-* 'to measure', with a complicated semantic development: 'measure' > '(firmly) measure' > '(four)square'. This proposal is equivalent to attempts to connect **meyu-* with any other IE root that has the shape *mV(i)-*; meaning shifts proposed by J. Puhvel are unprovable.

I want to stress that the authors of these three etymologies implicitly proceed from the same schema as I have advocated above: Indo-Hittite had the counting system "1, 2, 3, >3/several/many", while words for '4' appeared in Anatolian and Indo-European branches of Indo-Hittite already after their separation. This conclusion stems from the fact that only two ways to eliminate old words for (lower) numerals are attested cross-linguistically, as far as I know:

¹⁰ Hitt. *ta-* '2' cannot be separated from IE **d̥yo-*, although loss of *u* is inexplicable. The same development (Indo-Hittite **d̥u* > Proto-Hittite **d*) can be seen in Hitt. *idalu-* 'evil' ~ Luw. *attuwal* 'evil' (: Grk. *ὀδύνη* 'pain of body?'), but a sonorant metathesis *u-l* in Hittite or Luwian is an equivalent solution for this case. Bomhard 2008 proposes that the proper Indo-Hittite stem for '2' was **do-*, while the form **d̥yo-* was latter borrowed from West Caucasian **tq̄l:ʷA* '2' (cf. fn. 5 above); this is theoretically possible, but not very likely.

¹¹ There are no other reliable examples illustrating the development of **aīu* in Luwian, but analogy with the more or less regular loss of intervocalic *i* in other combinations with non-front vowels (cf. Melchert 1984: 164) is not risky.

¹² The model 'big(ger than) X' for 'X+1' is attested, but the simple model 'big(ger)' is not.

- 1) borrowing from neighboring language or *lingua franca*;
- 2) using mathematical operations. E.g., Lake Miwok (< Penutian) has lost its old word for '4' and now uses $\text{ʔot-ót-a} = '2+2'$ (for more details on the Lake Miwok situation see fn. 24 below); Old Irish has a word *mór-feiser* 'семеро, group of 7 (people)', literally 'big 6', 'bigger than 6 [by 1]'

The appearance of IE **septm* '7' and **newn* '9' in Anatolian numeral systems should not confuse us. Undulatory divergent-convergent processes between closely related dialects frequently take place after the definite split of their proto-language. I propose that the Anatolian stem **septam-* (Hitt. *šiptam-* ~ CLuw. *šaptam-*) was borrowed from neighboring Proto-IE dialects, immediately after the latter had borrowed **sept-* from early Semitic. The situation with '9' is more interesting: we cannot reconstruct one IE form for '9', since known languages point to **newen*, **newn*, **e(n)-nun* or **en-neun* (Pok.: 318 f.; Blažek 1999: 277). Phonetically none of these proto-forms could directly yield Hieroglyphic Luwian *nuwa(nʔ)*, with *u* instead of expected *a*. Therefore, HLuw. *nuwa(nʔ)* is either a loanword from closely related IE dialects or a calque from the same source.¹³

§5. Although one can hypothesize that Anat. **meyu-* '4', likewise, represents a loanword, no plausible sources for such a borrowing can be found among the neighboring Afroasiatic, North Caucasian or Kartvelian languages. I propose the following external (Nostratic) etymology for Anatolian **meyu-* '4' — Proto-Altaic **móju* 'all (totus), whole': Tungusic **muja-* 'whole, all (totus)', Japanese **múina* 'all (totus), all (omnis)', Korean **māin* 'most, extremely, very' (EDAL: 939 f. w. lit.).¹⁴

Phonetically, the comparison of Anatolian **meyu-* (virtual Indo-Hittite **meju-*) and Alt. *móju* is exact,¹⁵ but the semantic shift 'all' > '4' requires typological confirmation.

In §1 above I have mentioned that, when a counting system begins to expand, other scenarios than metonymy based on body parts (or "fraternal" strategy, etc.) are also possible for the formation of new numerals.

Example 1. Proto-Arawakan (Amazonia) had **pa-* for '1' and **(a)pi-* or **yama-* for '2'; it is unlikely that '3' can be reconstructed at the Proto-Arawakan level (see Payne 1991). The Proto-Arawakan counting system must have been as follows: "1, 2, >2/several/many".

Let us now look at the Inland subbranch of Arawakan as listed in Ethnologue 15 (except for Resígaro, where the numerical system is influenced by Witotoan). Data are given according to the following sources: Native American Number Worksheets; Aikhenvald 1998: 298–299;

¹³ The last supposition is very natural. First, the Nostratic stem **nVwV* 'new' is present in the Anatolian lexicon (Hitt. *newa-* 'new, fresh'). Second, the loan of a *concept* rather than an *exact form* of a new (larger) numeral is well attested, especially between related idioms. E.g., Arawakan languages use 'hand' as '5', but words for 'hand' may differ among dialects, so the common Arawakan or common Proto-"Inland" stem for '5' cannot be reconstructed (see my Ex. 1 in §5); Proto-Maku (Amazonia) must be reconstructed with the counting system "this one, a few, many", some modern Maku languages introduce "2" as 'eye(s)', but the roots and morphological structures of these expressions are not identical (Epps 2006). Some Iranian words for '9' show the concept 'beyond 8', which is, however, expressed differently in individual languages (see §3 above). The same is true of Proto-IE '9' in Greek and Armenian, which show the prefix **en-*, unlike the other branches of IE.

¹⁴ Anat. **meyu-* is one of the numerous Indo-Hittite stems that were lost in Narrow IE branch, but having external Nostratic parallels. See Kassian in print for ca. 40 reliable Anatolian—Nostratic etymologies, such as body parts, kinship terms, etc.

¹⁵ The secondary Indo-European (Indo-Hittite) apophony **ole* is irrelevant for the Nostratic vocalic reconstruction.

Aikhenvald 2003: 217–218; Schauer 2005. In the table below “loan” indicates a Spanish or Portuguese source.

L-ge	'1'	'2'	'3'	'4'	'5'	'6' and higher
Achagua	<i>báque</i>	<i>chámai</i>	<i>matálii</i>	<i>kuátru</i> (loan)	<i>abakáahi</i> (< <i>káahi</i> 'hand')	base-5 compounds, being superseded by loans
Baniwa	<i>apada</i>	<i>dzamada</i>	<i>madariwa</i>	<i>wadaka /</i> <i>ri-kwa-da-ka</i> (‘it is enough’)	<i>apamawakaapi</i> (< <i>kaapi</i> 'hand')	base-5 compounds, being superseded by loans
Curripaco	<i>pada</i>	<i>yamada</i>	<i>madalida</i>	<i>li-kua-da-ka</i> (‘it is enough’)	<i>pemapakapi</i> (< <i>kapi</i> 'hand')	base-5 compounds, being superseded by loans
Guarekena (Warekena)	<i>pe-/pa-</i>	<i>-naba</i>	<i>te.lefi</i> (loan)	<i>kwatru</i> (loan)	<i>sinku</i> (loan)	loans
Piapoco	<i>abéiri</i>	<i>pucháiba</i>	<i>máisiba</i>	<i>báinúaka</i> (un- clear struc- ture)	<i>abéemàa wa-</i> <i>káapi</i> (< <i>káapi</i> 'hand')	base-5 compounds, being superseded by loans
Tariana	<i>pa: / páada</i>	<i>ñama / yama /</i> <i>yamáda</i>	<i>madali /</i> <i>manalída</i>	<i>kephunipe</i> (< <i>ka-iphu-nipe</i> 'somebody who has a companion', calque of Tucano numeral)	<i>pa-kapi /</i> <i>peénkapi</i> (< <i>kapi</i> 'hand')	base-5 compounds, being superseded by loans
Yucuna	<i>pahluwá(ha)</i>	<i>iyamá</i>	<i>wéehi kéele</i>	<i>pa'ú kéele</i> (unclear structure)	<i>pahluwa té'ela</i> <i>kéele</i> (< <i>yáte'ela</i> 'hand'), <i>sínku</i> (loan)	base-5 compounds, being superseded by loans

Proto-“Inland” ‘3’ can be reconstructed (sounding approximately like **mata-*; Yucuna *wéehi* is unclear), while ‘4’ cannot. The numeral ‘5’ as ‘hand’ is surely a common cross-cultural innovation, where roots may differ, cf. ‘5’ in Yucuna, derived from *yáte'ela* ‘hand’, not from Proto-“Inland” **kāpi* ‘hand’ as in other languages. Thus the Proto-“Inland” counting system must have been as follows: “1, 2, 3, >3/several/many”.

The new numeral ‘4’ can be formed in different ways. It is the morphologically transparent form ‘it is enough’ in the neighboring and closely related languages Baniwa and Curripaco that is important for our further conclusions.

Example 2. A typical Khoisan (South Africa) counting system consists of three or four words:

- ‘1’, ‘2’, ‘>2, many’
- ‘1’, ‘2’, ‘3’, ‘>3, many’

Of course, in the immediate past many languages were forced to expand their counting systems. Let us look at some Peripheral Khoisan systems. I adopt the Khoisan glottochronological classification by G. S. Starostin (Starostin G. 2008; Starostin G. KhoiEDB), which is generally consistent with the Ethnologue's classification (Ethnologue 15). In the table below North Khoisan languages are cited according to the following sources:

!Kung Doke — Doke 1925: 154, 158, 159¹⁶; !Kung Wilhelm — Wilhelm 1922; !Kung Ekoka — König & Heine 2001; !Kung Weich — Weich 2004¹⁷; !Kung Angolan — Snyman 1979¹⁷; Ju!'hoan Snyman — Snyman 1970; Ju!'hoan Dickens — Dickens 1994; ||Kx'au||'ein (‡Kx'au||'ein, ||Au||en, Auen) — Bleek 1956.

L-ge	'1'	'2'	'3'	'4'	'5'	'>3, several, many'	'all (omnis)'
!Kung Doke	<i> nʔe</i>	<i>sā:</i>	<i>!!kao / !kao</i>	<i>sā: sa sā</i> (‘2-and-2’)	(no data)	<i>!kao</i>	
!Kung Wilhelm	<i> nē</i>	<i>sa / sām</i>	(not used)	(not used)	(not used)	<i>!gau</i>	
!Kung Ekoka	<i>n è'è / n è</i>	<i>tcā</i>	<i>!āō</i>	<i>k ām</i> (loan?)	<i>g āō n è'è</i> (‘one hand’)		
!Kung Weich	<i> nehe</i>	<i>tsa</i>	<i>!ao,</i> <i>dri</i> (loan)	<i>firi</i> (loan)	<i>faiḑi</i> (loan)		
!Kung Angolan	<i> nè'è</i>	<i>tšā / tsā</i>	<i>!āo</i>	(no data)	(no data)		
Ju!'hoan Snyman	<i>n e'e</i>	<i>tsā</i>	<i>n eni</i> (loan?)	(not used?)	(not used?)	<i>!xare, gesī</i>	
Ju!'hoan Dickens	<i>n è'è</i>	<i>tsàn / tsàqn</i>	<i>n ànì</i> (loan?)	<i>tsánkútsán</i> (‘2+2’)	(no data)		<i>wàqnsì</i> (also ‘whole’, i.e. ‘totus’), <i>wècè</i>
Kx'au 'ein	<i> né / nēe</i>	<i>tšā / tsa</i>	<i>!nan(n)i</i> (loan?)	<i> kai / kei</i>	<i>!gou</i> (‘hand’)		<i>!kau</i>

NB: Grammarians stress the fact that even though certain languages have numerals ‘3’ and larger, numerals larger than ‘2’ or ‘3’ are seldom used in natural speech (König & Heine 2001: 70 for !Kung Ekoka; Bleek 1956: 473 s.v. *!nanni*).

Since the semantic shift ‘2’ or ‘3’ > ‘several/many’ is typologically incredible, the only possible North Khoisan counting system that can be reconstructed based on the above table is the following one:

‘1’ **ŋ|eʔe*

‘2’ **cā*

‘>2, several/many’ **!au* (cognates are highlighted in the table above)¹⁸

¹⁶ Doke’s click transcription has been adapted to modern notation, with the omission of some vowel diacritics that are irrelevant for us.

¹⁷ Same as !O!Kung!/O!ung language according to G. Starostin, KhoiEDB.

¹⁸ Strictly speaking, it is unclear how we should reconstruct the initial click in North Khoisan **!au* ‘>2, several/many’. Doke gives *!kao* with an alveolar click and a by-form *!!kao* with a retroflex one. An unstable articulation of the retroflex *!!* and its confusion with the alveolar *!* is not so rare in North Khoisan dialects, retaining retroflex clicks; on the other hand, Doke’s materials are not very reliable in general, and, therefore, the form in *!!-* can simply be erroneous. It is important that !Kung Angolan has the alveolar onset *!-* in *!āo*, not the lateral *||-* (regular correspondences are Doke *!-* ~ Angolan *!* and Doke *!!-* ~ Angolan *||*), thus the proper !Kung Doke form is *!kao* rather than *!!kao*. More about retroflex clicks in North Khoisan see in Старостин Г. 2005.

The reconstructed system coincides with the attested !Kung Wilhelm forms.

When counting systems begin to expand, we find a variety of forms for new numerals ('3', '4' etc.). These can be loans from Afrikaans, as in !Kung Weich (Ju|'hoan and ||Kx'au||'ein *n!eni / n!àni / !nan(n)i* '3' seems to be a loan from Central Khoisan dialects; the same may apply to !Kung Ekoka *k||ãm̃* < Central Khoisan?¹⁹), mathematical operations ('2+2'), body parts ('hand') or the old word for '>2, several/many' *!au used for the next new member of the numerical sequence, i.e. '3'.

An intermediate situation is probably observed in !Kung Doke: (!)k_o shifted to '3', '4' is '2+2', however, sometimes (!)k_o retains its old meaning 'many'.

In ||Kx'au||'ein *!ao shifted to 'all (omnis)', cf. Bleek's example: "we are all sleepy".

As a possible external cognate for North Khoisan *!au '>2/several/many' I can propose !au '4' in South Khoisan²⁰: Masarwa (Kakia dialect of !Xóõ) !kau '4', quoted in Bleek 1956: 412 with the example: "they are four". If the Masarwa form is reliable, the Proto-Masarwa counting system was "1, 2, 3, >3/several/many". One has to assume a subsequent semantic shift: 'several' > 'next new numeral', i.e. '4', similar to the one discussed above. It is strange that simultaneously several other Masarwa forms for '4', taken from the same source as !kau, are quoted in Bleek 1956: ||gai '4' and ||kē '3; 4'.

Example 3 is similar to Ex. 2, but concerns the form ||kai ~ ||kei '4' in the ||Kx'au||'ein dialect of North Khoisan (quoted in the table from my Ex. 2 above).

This form can be compared with the South Khoisan numeral *||ai '>2, several/many' reconstructed according to the following data from the Taa branch of South Khoisan:

Masarwa ||kai '3; 5', ||gai '4' and ||kē '3; 4' (quoted in Bleek 1956) probably represent different transcriptions of the same word ||kai '>2, several, many' with occasional diphthong contraction. Cf. Bleek's examples for ||kai: "the flies are three", "in five days"; for ||gai: "four children"; and for ||kē: "they are three", "the stones here are three", "they are four";

The |Nu|en counting system shows the same situation (as listed in Bleek 1956 s.v. ||gai): !oe '1', !num '2', ||gāĩ '>2' (examples for ||gāĩ: "three people", "the hoofs are four");

!Xoong (according to Traill 1994) has three numerals: †'ûã '1', †nûm '2', ||âe '3'. Traill gives no words for '4' or '5', therefore, it remains uncertain whether ||âe means exactly '3' or rather '>2, several'. Cf. Traill's examples: "we three", "three times", "make a threesome" (s.v. ||âe), "the three of them" (s.v. ûh).

An additional example comes from the !Wi branch of South Khoisan: ||Xegwi (Batwa) ||kai 'much' (Bleek 1956 with an example: "much food").

If so, ||Kx'au||'ein ||kai ~ ||kei '4' shows the same semantic development as illustrated in Ex. 2: 'several' > 'next new numeral', i.e. '4'.

¹⁹ The Proto-Central Khoisan word for '4' is *haka (Starostin G. KhoiEDB), but in some languages this form can be superseded by mathematical '2+2', e.g., archaic Naro (Schniz 1891: 545) had |gam '2' and |gam tscha ra hi |gam tscha ra hi '4', where |gam regularly goes back to Proto-Khoe *|ám̃ '2' (the modern Naro dictionary — Visser 2001 — gives only the Nama loan hãka for Naro '4'). It is theoretically possible that !Kung Ekoka borrowed the first morpheme from this Central Khoisan collocation to denote '4'. On the other hand, G. Starostin rejects this etymological solution for !Kung Ekoka *k||ãm̃* (Starostin G. KhoiEDB: "cannot be easily identified as a borrowing from any known source") in view of the phonetic irregularity: lateral click || in Ekoka vs. dental | in Central Khoisan. This is a serious counter-argument.

²⁰ South Khoisan joins North Khoisan at the next level, which is called Proto-Peripheral Khoisan or Proto-Bushman by G. Starostin (the term Ju-Taa is also sometimes found in literature).

Alternatively, Bleek's transcription of $\|Kx'au\|ein \|kai \sim \|kei$ '4' can be compared with Bleek's $\|Kx'au\|ein \|kai$ 'all together', which goes back to North Khoisan $*\|kxae$ 'together': Jul'hoan Snyman $-\|x'ae$ 'together'; Jul'hoan Dickens $\|káé$ 'to be together'; !Kung Weich $\|kae$ 'together (with); simultaneously'; !Kung Ekoka $\|x'āē$ 'to meet, find', $-\|x'āē$ verbal suf. 'together'. These forms can be further compared with South Khoisan $*\|kxáé$ 'to join' (Starostin G. KhoiEDB).

G. Starostin (Starostin G. KhoiEDB) proposes Peripheral Khoisan $*!_1a-$ '3' on the grounds of North Khoisan $*!au$ '3' and Proto-Taa (South Khoisan) $*\|âe$ '3'. This connection, however, is not very probable, since the counting systems of Proto-North and Proto-South Khoisan did not have the numeral '3' (as shown above) and the diphthong correspondence North Khoisan $*au \sim$ South Khoisan $*ae$ appears to be irregular (of course, vowel correspondences between Khoisan sub-branches deserve additional investigation).

Example 4. Traces of old ternary or senary (base-6) counting systems are typical of Penutian languages spoken in western North America (Beeler 1961; Callaghan 1990). For the Proto-Costanoan levels (Penutian sub-branch) the numerals '1', '2', '3' and '6' can be reconstructed. The word for '6' ($*\|sak\cdot e-n$) is unetymologizable, however, some Costanoan dialects show another root for '6': Mutsun *nakči*, Soledad *heme-noksi*. These forms are probably related to Proto-Miwok (closest relative of Costanoan) $*nake(\|s)$ 'end, edge' (in this case Soledad *heme-noksi* '6' has a transparent structure and means '1 edge', cf. Soledad *himič-?a* '1'); see Callaghan 1990 for details.

Example 5. For the common Dravidian root $nāl-$ '4' (Starostin G. DravEDB; DEDR: #3655) a parallel meaning 'a few/several/many' is attested in a number of Dravidian languages. Some scholars (e.g., Blažek 2009: 71) claim that it is a late semantic development. As a matter of fact, polysemy '4' \sim 'a few/several/many' for $*nāl-$ is attested in all subbranches of the Dravidian family except for the Central Dravidian group (Kolami-Gadba and Gondi-Kui), where $*nāl-$, according to the available sources, means only '4', and Brahui, where this proto-root has been lost:

North Dravidian: Kurukh *nākḥ* '4; a few' (Grignard 1924: 512);

South Dravidian: Tamil *nālu* '4; many, manifold; a few' (Tamil lexicon), Malayalam *nāl*, *nālu*, *nān* '4' \sim *nālu* 'several', *nānā* 'manifold';

Telugu: *nālugu* '4; an indefinite number, a few, several, many' (Gwynn 1991).

Since North Dravidian languages do not contact with Telugu and South Dravidian dialects, it is impossible to invoke a shared innovation in this case, and so one must reconstruct the polysemy '4' \sim 'a few/several/many' for $*nāl-$ for the Proto-Dravidian level.

As was shown in §2 above, the numeral '4' is assuredly reconstructed for the Dravidio-Uralo-Altaic branch of Nostratic macrofamily: Altaic $*\|nöl-$ '4' \sim Uralic $*\|neljä$ '4' \sim Dravidian $*nāl-$ '4'. Of course, it is possible to suppose that the meaning 'several' represents a secondary development in Proto-Dravidian. But the opposite hypothesis is hardly weaker: if the primary meaning of Nostratic $*\|nIV$ was 'several', then later, in the Dravidio-Uralo-Altaic proto-language, this root was used to form the next new numeral, i.e. '4'. The second solution is supported by the fact that the semantic shift '3'/'4'/'5' $>$ 'several' would be typologically unparalleled. If so, the situation with Dravidian dialects, retaining the polysemy '4' \sim 'several', is the same as with !Kung Wilhelm (see my Ex. 2 above), where *!kau* means both 'several' (old sense) and '3' (new sense).

§6. Conclusions. In §1–4 above I have shown that Nostratic (ca. 10–9 millenium B.C.) counting systems consisted of three elements: “1, 2, >2/several/many”. Indo-Hittite (ca. 5 millenium B.C.) expanded it to “1, 2, 3, >3/several/many”. For the next newly formed numeral **meyu-* '4' Anatolian used the Nostratic stem that yielded Proto-Altaic **móju* 'all (totus)'. In §5 I have given several examples illustrating the typological plausibility of such a semantic development.

Ex. 1 (Arawakan family) illustrates that the word 'enough' can be used for the expansion of the numerical system; the shift in meaning between 'enough' and 'all (omnis)' or 'all (totus)' is natural.²¹

Exx. 2–3 (Khoisan family) illustrate that the last member of a primitive counting system (>X/several/many) can become the next new numeral; the semantic shift between 'many' and 'all (omnis)' or 'all (totus)' is also natural.²²

Ex. 4 (Penutian family) illustrates that the base of a primitive counting system can be expressed via the word for 'edge, end'. The semantic development from 'edge, end' to 'all (omnis)' or 'all (totus)' is well attested.²³

Ex. 5 (Dravidian family) illustrates the synchronic polysemy between '4' and 'several', but, strictly speaking, the direction of semantic development cannot be established in this case.

As for the Indo-European branch of Indo-Hittite, I propose that the early IE word for '4' was **okt-*, which was latter superseded by **k^uetuor* and retained only in the expression for '8' **oktō* ('doubled 4')²⁴ and in Iranian **ašti-* '(breadth of) four fingers'. When the 4- and 5-threshold in the row of numerals was overcome, the IE counting system became base-4 (in addition to '8', lit. 'doubled 4', traces of a base-4 system are retained in some Iranian dialects, where '8' is understood as a round number).

The origin of IE **okt-* and **k^uetuor* remains unknown.²⁵ From the typological viewpoint, the most plausible scenario would be to assume that **okt-* is the old word for 'palm' (cf. Iranian **ašti-*) or 'index finger', while the later **k^uetuor* is a foreign loanword. There is, however, no positive evidence that could support this theory.

Literature

- Гамкрелидзе—Иванов 1984 — Т. В. ГАМКРЕЛИДЗЕ, Вяч. ВС. ИВАНОВ. *Индоевропейский язык и индоевропейцы*. Тбилиси, 1984
- Иллич-Свитыч 1964 — В. М. ИЛЛИЧ-СВИТЫЧ. Древнейшие индоевропейско-семитские языковые контакты // *Проблемы индоевропейского языкознания*. М., 1964
- Климов ЭСКЯ — Г. А. КЛИМОВ. *Этимологический словарь кавказских языков*. М., 1964.

²¹ E.g., the Russian exclamation *всё!* 'enough, stop it' < neut. sg. of Slav. **vъsv* 'all (omnis/totus)'.

²² E.g., ||Kx'au||'ein *!kau* 'all (omnis)', quoted in the table in Ex. 2, or Tungus-Manchu **gere-n*, meaning 'many' and 'all (omnis)' among the daughter languages (Цинциус, 1: 182; Tower of Babel Project).

²³ E.g., Tamil *muṭi* 'to end' ~ *muṭiyal* 'all, whole'.

²⁴ The situation when a language acquires a new form for '4', retaining the old root '4' in the word for '8', has typological parallels. E.g., Lake Miwok (Western Miwokan < Utian < Penutian; California) has a new word '4' as *ʔoṭ-óṭ-a*, ('2+2') while '8' is *ʔóṭ^h-aja* < **ʔóṭ-haja* ('2×4'), where *haja* corresponds to *huja* '4' in other Western Miwokan languages (Callaghan 1994: 167).

²⁵ The occasionally proposed comparison between IE **k^uetuor* '4' and Uralic **kutte* '6' is semantically unconvincing.

- Старостин Г. 2005 — Г. С. СТАРОСТИН. Некоторые аспекты исторического развития кликсов в койсанских языках // *Аспекты компаративистики [Aspects of Comparative Linguistics]*. Vol. 1. Moscow: RSUH Publishers, 2005. P. 281—299 [available online at Tower of Babel Project]
- Цинциус — *Сравнительный словарь тунгусо-маньчжурских языков* / Отв. ред. В. ЦИНЦИУС. В 2 тт. Ленинград, 1975—1977.
- ЭСИЯ — В. С. РАСТОРГУЕВА, Д. И. ЭДЕЛЬМАН. *Этимологический словарь иранских языков*. М., 2000—.
- Aikhenvald 1998 — A. AIKHENVALD. Warekena // D. C. DERBYSHIRE, G. K. PULLUM (eds.). *Handbook of Amazonian languages*. Vol. 4. Walter de Gruyter, 1998. P. 225—440.
- Aikhenvald 2003 — A. AIKHENVALD. *A Grammar of Tariana, from Northwest Amazonia*. Cambridge University Press, 2003.
- Beeler 1961 — M. S. BEELER. Senary counting in California Penutian // *Anthropological Linguistics* 3 (1961). P. 1—8.
- Blažek 1999 — V. BLAŽEK. *Numerals: Comparative-Etymological Analyses of Numerals Systems and Their Implications*. Brno: Masarykova univerzita, 1999.
- Blažek 2009 — V. BLAŽEK. Dravidian numerals // *Вопросы языкового родства [Journal of language relationship]*. No. 1. 2009. Moscow. P. 69—80.
- Bleek 1956 — D. F. BLEEK. *A Bushman dictionary*. New Heaven, 1956.
- Bomhard 2008 — A. BOMHARD. Some thoughts on the Proto-Indo-European cardinal numbers // J. D. BENGSTON (ed.). *In Hot Pursuit of Language in Prehistory. Essays in the four fields of anthropology. In honor of Harold Crane Fleming*. Amsterdam—Philadelphia, 2008. P. 213—222.
- Callaghan 1990 — C. A. CALLAGHAN. Proto-Costanoan Numerals // *International Journal of American Linguistics*. Vol. 56, No. 1 (1990). P. 121—133.
- Callaghan 1994 — C. A. CALLAGHAN. Proto-Miwok Numerals // *International Journal of American Linguistics*. Vol. 60, No. 2 (1994). P. 161—176.
- Comrie 1999 — B. COMRIE. Haruai numerals and their implications for the history and typology of numeral systems // J. GVOZDANOVIĆ (ed.). *Numeral Types and Changes Worldwide*. Mouton de Gruyter, 1999. P. 81—94:
- DEDR — T. BURROW, M. B. EMENEAU. *A Dravidian etymological dictionary*. 2nd ed. Oxford, 1984.
- Dickens 1994 — P. DICKENS. *English—Ju'hoan, Ju'hoan—English dictionary*. Köln, 1994.
- Doke 1925 — C. M. DOKE. An Outline of the Phonetics of the Language of the Chū: Bushmen of North-West Kalahari // *Bantu Studies and General South African Anthropology*. 1925, vol. 2 (3). P. 29—65.
- EDAL — S. A. STAROSTIN, A. V. DYBO, O. A. MUDRAK. *Etymological Dictionary of the Altaic Languages*. Brill, 2003.
- EDHIL — A. KLOEKHORST. *Etymological Dictionary of the Hittite Inherited Lexicon*. Brill, 2008.
- Edelman 1999 — DŽ. I. EDELMAN. On the history of non-decimal systems and their elements in numerals of Aryan languages // J. GVOZDANOVIĆ (ed.). *Numeral Types and Changes Worldwide*. Mouton de Gruyter, 1999. P. 221—243.
- Eichner 1992 — H. EICHNER. Anatolian // J. GVOZDANOVIĆ (ed.). *Indo-European Numerals*. Mouton de Gruyter, 1992. P. 29—96.
- Emmerick 1992 — R. EMMERICK. Iranian // J. GVOZDANOVIĆ (ed.). *Indo-European Numerals*. Mouton de Gruyter, 1992. P. 289—346.
- Epps 2006 — P. EPPS. Growing a numeral system: The historical development of numerals in an Amazonian language family // *Diachronica*. Vol. 23(2). 2006. P. 259—288.
- Ethnologue 15 — R. G. GORDON, JR. (ed.). *Ethnologue: Languages of the World*. 15th ed. Dallas: SIL International, 2005. Online version: www.ethnologue.com/
- Gamkrelidze & Ivanov 1995 — TH. V. GAMKRELIDZE, VJAČ. VS. IVANOV. *Indo-European and the Indo-Europeans*. Mouton de Gruyter, 1995.
- Goedegebuure 2006 — P. M. GOEDEGEBUURE. A New Proposal for the Reading of the Hittite Numeral '1': šia- // *The Life and Times of Ḫattušili III and Tuḫaliya IV: Proceedings of a Symposium held in Honour of J. de Roos*, 12–13 December 2003, Leiden / Ed. by T. P. J. van den Hout & C. H. van Zoest. Leiden: 2006. P. 65—88.
- Greenberg 1990 — J. H. GREENBERG. Generalizations about numeral systems // *On Language: Selected writings of Joseph H. Greenberg* / Ed. by K. DENNING and S. KEMMER. Stanford University Press, 1990. P. 271—309. [Originally published in: *Universals of Human Language* / Ed by J. H. GREENBERG, Ch. A. FERGSON, & E. A. MORAVCSIK. Vol 3. Stanford University Press, 1978. P. 249—295]
- Grignard 1924 — A. GRIGNARD. *An Oraon-English Dictionary*. St. Gabriel-Mödling, 1924 [reprint: 1986].

- Gwynn 1991 — J. P. L. GWYNN. *A Telugu-English dictionary*. Delhi; New York: Oxford University Press, 1991 [available online: dsal.uchicago.edu/dictionaries/gwynn/]
- Hammarström 2004 — H. HAMMARSTRÖM. Properties of Lower Numerals and their Explanation: A Reply to Pawel Rutkowski // *Journal of Universal Language* 5 (2), 2004. P. 1—20.
- Harvey 1992 — M. HARVEY. *The Gaagudju people and their language*. PhD at the University of Sydney, 1992.
- HED — J. PUHVEL. *Hittite Etymological Dictionary*. Vol. 1—. Berlin—NY—Amsterdam, 1984—.
- HEG — J. TISCHLER. *Hethitisches etymologisches Glossar*. Innsbruck, 1977—.
- Hoffner & Melchert 2008 — H. A. HOFFNER, JR., H. C. MELCHERT. *A grammar of the Hittite language*. Part 1: Reference grammar. Eisenbrauns, 2008.
- Henning 1949 — W. B. HENNING. Oktō(u) // *Transactions of the Philological Society* 1948. London, 1949.
- Kassian in print — A. S. KASSIAN. Anatolian lexical isolates and their external Nostratic cognates.
- Klimov EDKL — G. A. KLIMOV. *Etymological Dictionary of the Kartvelian Languages*. Walter de Gruyter, 1998.
- König & Heine 2001 — CHR. KÖNIG & B. HEINE. *The !Xun of Ekoka*. Germany, 2001.
- Lean 1992 — G. A. LEAN. *Counting systems of Papua New Guinea and Oceania*. Thesis submitted toward fulfilment of the degree of doctor of philosophy. December, 1992. With four appendices. Papua New Guinea, University of technology. [www.uog.ac.pg/glec/thesis/thesis.htm]
- Lehrman 2001 — A. LEHRMAN. Reconstructing Proto-Indo-Hittite // *Greater Anatolia and the Indo-Hittite language family*. Papers presented at a colloquium hosted by the University of Richmond, March 18—19, 2000 [*The Journal of Indo-European Studies Monograph Series* No. 38] / Ed. R. Drews. Washington, 2001. P. 106—130.
- Luján Martínez 1999 — E. U. LUJÁN MARTÍNEZ. The Indo-European Numerals from '1' to '10' // J. GVOZDANOVIČ (ed.). *Numeral Types and Changes Worldwide*. Mouton de Gruyter, 1999. P. 199—219.
- Mallory & Adams 2006 — J. P. MALLORY & D. Q. ADAMS. *The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World*. Oxford University Press, 2006.
- Melchert 1984 — H. C. MELCHERT. *Studies in Hittite historical phonology*. Göttingen, 1984.
- Native American Number Worksheets (American Indian Numerals) (www.native-languages.org/numbers.htm) // "Native Languages of the Americas" project (www.native-languages.org) [last visited: 01.2009]
- Neumann 2007 — G. NEUMANN. *Glossar des Lykischen*. Überarbeitet und zum Druck gebracht von J. TISCHLER. Harrassowitz, 2007.
- Payne 1991 — D. L. PAYNE. A classification of Maipuran (Arawakan) languages based on shared lexical retentions // D. C. DERBYSHIRE, G. K. PULLUM (eds.). *Handbook of Amazonian languages*. Vol. 3. Walter de Gruyter, 1991. P. 355—499.
- Pok. — J. POKORNY. *Indogermanisches etymologisches Wörterbuch*. Bern—München, 1959.
- Schauer 2005 — S. & J. SCHAUER, E. YUCUNA, W. YUCUNA. *Diccionario bilingüe yukuna-español, español-yukuna*. Bogotá, 2005.
- Schniz 1891 — H. SCHINZ. *Deutsch-Südwest-Afrika*. Forschungsreisen durch die deutschen Schutzgebiete Gross-Nama und Hereroland, nach dem Kunene, dem Ngami-See und der Kalahari 1884—1887. Oldenburg: Schulzesche Hof-Buchhandlung und Hof-Buchdruckerei (A.-Schwarz), 1891.
- Snyman 1970 — J. W. SNYMAN. *An introduction to the !Xū (!Kung) language*. Communications from the School of African Studies, No. 34. Cape Town, 1970.
- Snyman 1979 — J. W. SNYMAN. The relationship between Angolan !Xū and Žu|hōasi // *Bushman and Hottentot linguistic studies*. Ed. by J. W. SNYMAN. University of South Africa (UNISA), Pretoria, 1979. P. 1—58.
- Tower of Babel Project — *Etymological database project "The Tower of Babel"*, headed by S. A. STAROSTIN. In affiliation with the Evolution of Human Languages (EHL) Project at the Santa Fe Institute. <http://starling.rinet.ru/> [last visited: 01.2009].
- Starostin G. 2008 — G. S. STAROSTIN. From Modern Khoisan Languages to Proto-Khoisan: The Value of Intermediate Reconstructions // *Аспекты компаративистики [Aspects of Comparative Linguistics]*. Vol. 3. Moscow: RSUH Publishers, 2008. P. 337—470 [available online at Tower of Babel Project].
- Starostin G. DravEDB — G. S. STAROSTIN. Dravidian etymological databases // *Tower of Babel Project* [last visited: 01.2009].
- Starostin G. KhoiEDB — G. S. STAROSTIN. Khoisan etymological databases // *Tower of Babel Project* [last visited: 01.2009].
- Tamil lexicon — *Tamil lexicon*. University of Madras, 1924—1936 [available online: dsal.uchicago.edu/dictionaries/tamil-lex/].

Traill 1994 — A. TRAILL. *A !Xóǀ dictionary*. Köln, 1994.

Visser 2001 — H. VISSER. *Naro Dictionary*. 4th ed. Gantsi, Botswana, 2001

Weich 2004 — F. WEICH. *San dictionary*. Johannesburg, 2004.

Wilhelm 1922 — J. H. WILHELM. Aus dem Wortschatz der !Kuñ- und der Hukwe-Buschmannsprache // *Zeitschrift für eingeborenen Sprache*. Bd. XII. 1922. S. 291—304.

Winter 1992 — W. WINTER. Some thoughts about Indo-European numerals // J. GVOZDANOVIĆ (ed.). *Indo-European Numerals*. Mouton de Gruyter, 1992. P. 11—28.

В статье подвергается пересмотру традиционная реконструкция индоевропейской системы числительных. Прибегая к компаративистским и типологическим аргументам, автор показывает, что первоначальная (индо-хеттская) система состояла из четырех членов: '1', '2', '3', '>3/несколько/много'. Числительное '4' появилось независимо в анатолийской и собственно индоевропейской ветвях после распада индо-хеттского праязыка. Анатолийское **теуи*- '4' (хет. *теуи*- ~ лув. *таиша*-) не находит себе этимонов внутри и.-е. языков. Автор считает эту основу архаизмом, обладающим внешней (алтайской) этимологией. Анатолийское **теуи*- представляет собой один из многочисленных ностратических корней, сохраненных в анатолийской, но утраченных в собственно индоевропейской ветви.