URALIC NUMERALS

For Vladimír Skalička (*Aug 19, 1909)

§1. The following are the (proto)forms of the numerals of the first decad in Fenno-Ugric and Samoyed languages:

<table>
<thead>
<tr>
<th>Numeral</th>
<th>Fenno-Permic</th>
<th>Ugric</th>
<th>Samoyed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>*ükti</td>
<td>Ma Tj ük, KU äk &quot;</td>
<td>? SeT ukkir, okkir</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kh V ëj, Kaz. I, I ii, Hu ëgy</td>
<td>*oj-/*äj-</td>
</tr>
<tr>
<td>2</td>
<td>*kakta</td>
<td>*kektä</td>
<td>*op</td>
</tr>
<tr>
<td>3</td>
<td>*kolmi</td>
<td>Kh *käälem</td>
<td>*kitä</td>
</tr>
<tr>
<td>4</td>
<td>*ti/neljä</td>
<td>Hu három // Ma *kuurem</td>
<td>*näkur</td>
</tr>
<tr>
<td>5</td>
<td>*wi(i)t(i)</td>
<td>*witi</td>
<td>*tet,td / *tetti / *tejt,tö</td>
</tr>
<tr>
<td>6</td>
<td>*ku(u)t(t)i</td>
<td>*kotti</td>
<td>*sämpä / *sämpälänkö</td>
</tr>
<tr>
<td>7</td>
<td>*seščemä</td>
<td>*šęptš</td>
<td>*mäktut</td>
</tr>
<tr>
<td>8</td>
<td>FV *kaktekan &lt; *kakta-ikak-s3</td>
<td>ObUg *ňňľęy // Hu nyolc</td>
<td></td>
</tr>
<tr>
<td>Pe</td>
<td>*kiktames &lt; *kakta-ikam-s3(s)</td>
<td>*kitšľętš &lt; *kitä tetš</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SeT ššiti čaŋg köt</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>FV *ūkteš &lt; *ūkte-ikš-šc</td>
<td>Ma *änt-tšš-ššy</td>
<td></td>
</tr>
<tr>
<td>Pe</td>
<td>*ökt-tem-es &lt; *ūkte-ikšem-es(e)</td>
<td>Kh *šš-čšť-ššy</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hu kilenc</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>*šmäšššššš</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MTK *op-šššššš</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SeT ukkir čaŋka köt</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>FiMd *kämäni</td>
<td>Ma lšy</td>
<td></td>
</tr>
<tr>
<td>LpMr</td>
<td>*luka</td>
<td>Kh *šš-čšť-ššy</td>
<td></td>
</tr>
<tr>
<td>Pe</td>
<td>*das</td>
<td>Hu šš</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>*wüt</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MT ššen</td>
<td></td>
</tr>
</tbody>
</table>
§2. Comparative-etymological analysis:

1.1. FP *ükti “1” (UEW 81; S 552) can be of the FU age, if Ma TJ ük, KU ük” id. is related (Szinnyei 1910: 108). Honti 1993: 75f adds Kh *ejs and Hu ëg(y) & eg(y) deriving them from *ëy/*õy” and *iy3/*üy” respectively, and reconstructs FU *ikel/*üke (= *uki after S) — see 1.2. below. The internal FU etymology is uncertain, cf. perhaps Md vejs(e) “together” < *üke (Honti 1993: 80); Komi ękan / etnam “ich allein” (Honti 1993: 82); FV *wikta > Fi vihdoin “endlich, yhden kerran”, Md *(v)ukta > Erzja udalo, Mokša fiäl “hinter” (SKES 1734–35; Keresztes 1986, n.516). A similar semantical development is known e.g. from IE languages: Welsh cyntaf “der erste” : Old High German hintana, hintar “hinter” and Old Church Slavonic is-koni “von Anfang”: konec “Ende” (Pokorny 1959: 564). The same semantical way out can be seen in SeT ukkjr, okkjr “1” : ukj “vorderes Ende; Nase; Vorderteil”, ukon, ukot “früher, vormals” (Joki 1975: 729–30) < Sm *ukš “Enge, Vorderteil” (SW 30) < U *uks “Kopf” (UEW 542–43). The initial sequence *vi- in BF can probably be explained through contamination with FU *wiši “Ende” > Fi viimein “endlich”, viime “letzten, jüngst, vorig” etc. // Hu vég “Ende, Schluss, Zweck” (UEW 575). There is a promising external parallel in YkN øxte, ax “only, alone” (Collinder 1940: 104). The difference in initial vocalism is perhaps comparable with FU (U ?) *piši “grouse” (UEW 383; S 547–548) : YkS pange “kulik” (Krejniovč 1982: 80). On the other hand, there are also some possible Altaic cognates here: Tk *ük “upper, superior” // OKor (Paekče) *oko, MKor 'ux “upper part, top, surface” // OJp ökör- “to arise”, ökos- “to lift up” (Menges 1984: 291).

1.2. Hu ëgy and Kh *ejs “1” are probably of pronominal origin (UEW 67; MSzFE 1: 141–142; Majtinskaja 1979: 174–175), although it is tempting to speculate about their relationship to their Sm counterpart (1.3.).

1.3. Sm *oj-/ *oj- “1” (Helimski 1986, 136; Janhunen does not separate it from Sm *op id., see SW 28) has promising Altaic counterparts in Kor oi, ö “only, a single”, oi-nun-thóti “one-eyed person”, compared by Ramstedt (PKE 134) with Tg *ugi “few” > Evenki uyi, uvi, Orok oji etc. (TMS 2: 246), and IE *oý- in *oý-no-/*oý-wó-/*oý-ko- “1”, probably an old derivative of the IE anaphora *eý (Hamp 1992: 904). There is a more probable Tg parallel in Oroč ojoke “some, one” (TMS 2: 9: Oroč + Mong oira “near”) and perhaps pJp *uja “the same” (Starostin 1972: 74: Ḷp + U *ükki “1”).

1.4. Sm *op “1” (Helimski 1987: 77; Janhunen reconstructs *oap and derives it from *o(š)-, see SW 28) is comparable with possible Altaic parallels: KorN obun “all”, S on “voll, vollständig, all” // Tg *up-ka- “whole, every, all” // Mong ogtalu “fully” (SKE 177; PKE 133; TMS 2, 281).

2.1. U *kektä “2” (S 537; UEW 118–119: *kakta / *käktä) is the only numeral reliably reconstructed for the Uralic proto-language including the semantical identity of all attested forms. There are hopeful external cognates in Yk: Omok tkit “2” = *kit ?, cf. kit kim nel “20” (Collinder 1957, 118; Tailleur 1959, 84 compares the Omok forms with YkN kín / kíd “two” < *kiji-nt with
the genitive marker *-nt, cf. Nikolaeva 1988: 18) and in Tg *gagda “one of a pair” (TMS 1: 135: Tg + Mong gagča /ganča “one, single, only”; see also Dolgopol’skij 1969: 299: U + Tg), perhaps also in OJp kata “one (of a pair of set), single (member of a pair of set)” (Starostin 1972: 74; Street 1985: 641 derives it from Alt *kalta “half”). Ankeria 1951: 137 and Dolgopol’skij, Voprosy jazykoznanija 1964/2: 57 also add ČK data to U “2”, cf. Itelmen “2”: Tigil River (Billings) kaxan, Ukän (Pallas) kaza, W (Radloff) kaxs, S (Krašennikov) káss, kaáž etc. (Anderson 1982: 31).

Note: There is an alternative attempt to find the external relationship for U “2” in IE *kʷet- “pair” > Sl *četa id. and Ossetic caxd(æ) “a pair of bulls in yoke” (Abaev 1: 293). Erhart 1970: 95 also sees the same root in *kʷet(e)-H-o-r “4” = “2x2” (?). If related to U *kekta, it excludes the Tg form and allows to analyze the form *kekta < *ketV “2” + suffix of dual *-kal/*-kä by way of a metathesis caused (at least in FU) by *ükti “1” (more about the numeral “2” and dual in U – see Helimski 1982: 114–17).

3.1. FP *kolmi & Kh *kāaləm “3” stand in opposition to Hu három & Ma *kuurem (S 543; UEW 174). The -r-forms can be original, if we accept the influence of the following numeral *neljä “4” (Collinder 1965: 145). A hypothetical FU reconstruction could be *kurmi. If we isolate the suffix *-mi expressing perhaps the abstract noun “Dreiheit”, we get a hypothetical root *kur-, compatible with Sm *nä-kur “3” (Castrén 1854: 194; Helimski 1987: 77; SW 99: *näkk(j)r). The same demonstrative (?) marker *nä- perhaps appears in some Sm postpositions: *näŋ “zu” (dat.sg.), *nänä “bei” (loc.sg.), *nätä “von” (abl.sg.), *nämmländä (< *nän-mända) (prosecutive sg.; see SW 99).

The protoform of the Altaic numeral “3” can be reconstructed as *gu[r]-: Mong gurbən “3”, gurmusun “dreistährig, dreistähriger Strick” (Collinder 1965: 145: FU + Mong), gutagar “third”, gučin “30” < *gu[r]tin (> Tg *gutin, cf. Benzing 1955: 31); Mong (Khitan ?) > Jurchen *nuor-xuan “13” (Miller 1975: 146) // Kor *ku in MKor nirkup “7”, KorN nirgup, KorS ilgop = *[n]yərh “10 ... *ku *“3” ...öps “be nonexistent”, the negative of iss- “to be, exist, have”, cf. MKor ’yətrp “8” = “10–2”, ’ahop “9” = “10–1” (Miller 1971: 244; Menges 1984: 278; SKE 167: *il “3” & *kop “bending”, cf. Tg *ilan “3”) // ? OJp kōkōnö- “9”, besides nogono- (Pallas) < *kōnakono = “3x3”? (Miller 1971: 236; Anderson 1982: 42), the Tg counterpart *xuyğin “9” after Benzing 1955: 101, or *xegün according to Starostin 1991: 141 (correctly probably *xùnyığin “9”), is perhaps compatible with OJp, but not with Alt *gu[r]- “3”.


4.1. FU *neljä */neljä “4” (S 546; UEW 315; Honti 1993: 92–93) has no evident cognates in Sm, but there are at least hypothetical parallels in Yk: Čuven (Matjuškin) njagon, (Boensing) nāgane, nāxane “4” are derivable from *na[l]ga-, cf. Čuven (Matjuškin) xanbamegej, (Boensing) xambo megii “10” =
“double hand” vs. YkNW *malg-andklon “4” = “double two”, further YkN attrib./predic. malejli/-malajla- “6” vs. ja^-ljalo- “3” (cf. Tg *ilan id., but “Tonsugu-Konni” by Strahlenberg (1730) yelan! — see Anderson 1982: 46) and maleluku/-malajlaka “8” vs. jeluku/-jalakla- “4”, besides YkS melhal/mlhalo “6” vs. ja^-ljalo- “3” and malhileku / molhileqlo “8” vs. ileku / ileqlo “4”, cf. YkN malhur “on both sides” and YkS malhi “joint” (Tailleur 1962: 72; Krejnovič 1982: 114–118; Jochelson 1905: 113). On the other hand, the Čuvan “4” can be borrowed from Koryak, cf. (Krašennikov) nišken “4”, (Pallas) niyax, (Bogoras) gašax etc. (Anderson 1982: 30).

Traces of the root related to FU *neljä “4” can be also found in Altaic. MKor nayh “4” (Lee 1977: 174; about KorN ndujìn “4” quoted in SKE 277 see Miller 1971: 219) can be derived from a hypothetical protoform *nari-h, cf. OKor (Silla) *narih “river” > MKor nayh or *murih “mountain” > MKor moyh (Lee 1977: 80), similarly MKor sàyh “3” vs. syårhin “30” (Lee 1977: 174). A hopeful cognate can be identified in the Tg numeral “6”, usually reconstructed *nögün (Benzing 1955: 101; TMS 1: 647–48) and interpreted as “2x3” (Miller 1975: 147). The pattern “2x3” is evident for Mong ʒirgu-gan “6”, but doubtful for the Tg numeral in spite of Poppe’s attempt to reconstruct a development *nirgun < *nirgun < *ʒirgun (Poppe 1960: 130). On the other hand, a series of very archaic numeral stems is preserved in Jürchen names for the teens. So, for “16” the form ni-χun, earlier nül-χon, corresponding with Manchu niolxun “16th day of the 1st month”, while for “6”, Jürchen has *niŋ-ʒu (Miller 1975: 147). On the basis of these data the Tg prototype for “6” can be modified into the form *niol-ʒu(n) and interpreted as “10” (Tg *ʒuwan) (minus) “4” (*nöl-). This subtractive model forming the numerals 6 — 9 is well-known from Ainu (Lauffer, JAOS 37[1917]: 192–208) and probably from Korean (Miller 1971: 243–44 analyzes only 7, 8, 9 in this fashion, but MKor ʒeɕis “6” reveals a presence of ʒeɕh “10” similarly as ʒeɕtም “8”). On the other hand, a series of very archaic numeral stems is preserved in Jürchen names for the teens. So, for “16” the form ni-χun, earlier nül-χon, corresponding with Manchu niolxun “16th day of the 1st month”, while for “6”, Jürchen has *niŋ-ʒu (Miller 1975: 147). On the basis of these data the Tg prototype for “6” can be modified into the form *niol-ʒu(n) and interpreted as “10” (Tg *ʒuwan) (minus) “4” (*nöl-). This subtractive model forming the numerals 6 — 9 is well-known from Ainu (Lauffer, JAOS 37[1917]: 192–208) and probably from Korean (Miller 1971: 243–44 analyzes only 7, 8, 9 in this fashion, but MKor ʒeɕis “6” reveals a presence of ʒeɕh “10” similarly as ʒeɕtም “8”). Mong nayiman “8” could also belong here; it is possible to derive it from *nayid-man where *nayid- represents a regular plural to the sg. *nayil-, and -man together with -ban / -ben / -gan are suffixes forming most numerals of the first decad (Blašek 1997: 48–49). Besides hypothetical Yukaghir and Altaic cognates, there is a perfect correspondent in the Dravidian numeral *nāl “4” (DEDR #3655; cf. Tyler, Language 44[1968]: 807: FU + Dr; Menges 1984: 248: FU + Kor + Dr). Bouda (Anthropos 55[1960]: 358) opines that Nivkh root *nV of the numeral “4” also appears in ni-mar-i “quarter” < *n-nar- and minr- “8” = “2” (me) x “4”. This more archaic form *n(a)r “4” corresponds well with the hypothetical Korean *n*[r]i-h “4” (Krejnovič 1955: 135: Nivkh + Kor) and probably also with FU *nëljä “4” (Bouda, I.e. explains the change */ --> *r by the unacceptability of the expected cluster *nl- in Nivkh).

4.2. Sm *tet,tâ /*tetši/*tejt,tâ “4” (SW 159), of course, does not correspond with FU *nëljä (in spite of Menges 1975: 92). The attempts to find an internal Sm etymology are also unconvincing: (i) Joki 1975: 730 connected Sm “4” and Nganasan tatù-ame “sehr, viel”; (ii) the formally similar word
"Band" (SW 158) could represent the same root *tet- extended by various suffixes; but the semantical motivation is not clear (a similar semantical development was proposed for Ug *nala3 "8": Hu nyalâb "Bündel, Bund", KhV nula "zusammen" — see UEW 875 and below 8.3.).

On the other hand, Sm "4" could be a borrowing. The most probable candidates for a source are Tk forms of the type Čuvaš tâvattâ, Volgaic Bulgarin *tîät "4" (Benzing 1959: 730) or Lobnor döjt, tâjt, döt, Salar döt, döt, Ujghur dial. tööt etc. (Sevortjan 1980: 284–286), all from Tk *dört (Dybo). A source of the Čuvaš-Bulgarin type seems to be the most probable, because the same source of borrowing is evident for Sm *jûr "100" (SW 50; Janhunen, MSFOu 158[1977]: 125). The Sm + Tk comparison was presented already by A. Trombetti (1923: 395), but in a genetic framework.

5.1. FU *wi(i)t(t)i "5" corresponds to Sm *wiît "10" (S 541; UEW 577). The semantical relationship is certainly plausible, cf. e.g. Sîtulmen (Pallas) kûmnaka "5": kumextuk "10" (Anderson 1982: 33). For U *wi(i)t(i) "5" (Janhunen, JSFOu 77[1981]: n.124 reconstructs *wi(t)(i)*wxît) Joki (1975: 729) supposed the original meaning "grosse Anzahl, grosse Menge, viel", cf. Fi viittä vaivainen vaila "dem Armen fehlt viel". There are no evident external parallels, the only hypothetical exception would be OJp itu "5", OKor (Kogurjo) *utu (Murayama) or *uć (Lee) — see Miller 1971: 241; Kazâr 1980: 60: Jp + U.

5.2. Sm *sömë or *sömpšäkša "5" (SW 133) can be convincingly etymologized on the basis of Nenets sampa “in den Händen schaukeln, schwingen” (Joki 1975: 730).

6.1. FU *ku(u)t(t)i or *kotti "6" (S 544; UEW 225) and Sm *mäktut "6" (Helimski 1987: 77; SW 85) are evidently unrelated. But both the numerals were probably formed on the basis of the same semantical pattern: “six” = "beyond five", cf. (i) U *kutt3 "Rücken" (UEW 225) > MaN čûtł “hinter” etc.; (ii) U *mukâ “back” (S 538; SW 85) > Selkup Taz moqoqjít “hinter, hinteren” etc. If this etymology is correct, the Sm numeral represents a compound *mäkt- & -*ut. The latter component is probably identical with Sm *wiît "10", preserving the original meaning "5". Alternatively, Sm *utâ “hand” (SW 30) could be identified here, cf. IE *Ksweks "6" < *ģ's-weiks "hand-overgrowing". There is a suggestive parallel in Ossetic farast "9" = “beyond 8” (Winter 1992a: 14). The traditional interpretation of Sm *mäktut based on the comparison with Kamasin mäkt- "schreiten" etc. < U *mukča- “über das Wasser fahren” (so Joki 1975: 730; cf. UEW 284–85) is perhaps plausible semantically (cf. Eskimo of SWAlaska ar’Firtoa “I cross over to” vs. arFinligin "6" — see Thalbitzer, JSFOu 25/2[1908]: 13), but less probable for phonological reasons (UEW 284).

7.1. FP *sejćcem "7" (UEW 773; S 553: *se(e)/š/cVni; Honti 1993: 100 reconstructs *sejćcém comparing it with Sm *sejt,wš "7", cf. 7.3.) is probably borrowed from some Baltic dialectal form of the type OLithuanian sëkmas "7th" (Fraenkel 2: 772). Napolskikh 1995: 126 prefers some early Slavic
source (certainly preceding East Slavic *sents, perhaps a hypothetical *sedents), but this solution entails serious difficulties in phonology and chronology.

7.2. Ug (Hu & Kh) *stäpt(3) / Ma *sät "7" (UEW 844; Honti 1982: 138) has usually been derived from some Indo-Iranian source (Korenchy 1972: 70–71; MSzFE 284) preserving the original *s-, contrary to the characteristic Iranian change *s > *h (FU *s- gives regularly Ug *h-, while Ma *s- reflects FU *s-). Joki 1973: 313 judges that Ma *sät can be borrowed from a different Indo-Iranian source (cf. Dardic: Šina sat, sät, Prasun sete etc. and Mittani Ar-yan (Kikkuli) šatta !), rather than from Toch (cf. Toch A šäptänt- "7th"). On the other hand, Napolskikh 1995: 124 supposes just the Tocharian origin of the Ug "7". Winter (1992b: 109–110) reconstructs common Toch *ṣapta continuaing in Tocharian A and transforming in Tocharian B *swat > *swät/śwät (further leading to the historically recorded forms sukt, šukt, šuk, assuming the influence of the following numeral *aktu > okt "8"; but the change *-pt- > Tocharian B -kt- can be regular, cf. Napolskikh 1995: 120). A form resembling *swät could have been the source of pre-Mansi *sät, although the chronological correlation remains uncertain.

7.3. Sm *sejt,wä "7" (SW 139–40), with a more archaic variant in Sayan Sm *sejptä, are probably of Tocharian origin too (Janhunen, MSFOu 185[1983]: 119; Napolskikh 1995: 119 prefers to reconstruct Sm *sejkwä derivable from a source of the type Toch B suk(t)).

8.1. FV *kakteksan "8" has been analyzed in *kakta "2" + Negationsverb *e- + modal-reflexives Konjugationssuffix *k + Px3 Sg *sä + Dualsuffix *n, hence "zwei existieren nicht" (UEW 643; Honti 1993: 108). Napolskix (1997) mentions that the hypothetical form *e-k-sä-n is artificial without any support in real languages (e.g. in Fi it would give *eise, but really ei ole is used). The negative auxiliary *e- (UEW 68–69) is never independent. He proposes his own solution based on the abessive suffixes (a) for nouns *-tkak ~ *-tkek (Fi -ttal-ţă, Udm -tek) and (b) for adjectives & adverbs *-tkVm (Fi -ton/-tdn, Md -toma, Udm -tern). The final *-s3/*-sc is probably a nominal derivational suffix (e.g. Udm nil-ês "forest" vs. nįl-pu "Abies"). This approach allows us to recognize parallel structures in both Fenno-Volgaic and Permian numerals "8" and "9", differing only in the use of the abessive suffixes (a) and (b) respectively:

<table>
<thead>
<tr>
<th>Fenno-Volgaic</th>
<th>Permian</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>*kakta-tkak-s3 &gt; *kakteksa(-n)</td>
</tr>
<tr>
<td>9</td>
<td>*ükte-ikek-sc &gt; *ükteksä</td>
</tr>
</tbody>
</table>

8.2. Pe *kikjamis "8" has been derived from *kikja min(3)s "8" (Honti 1993: 156), and analyzed as the compound *kik "2" + frozen suffix of dual *ja + *min(3)s, the ablative from Pe *mīn, forming tens: Komi komiň, Udmurt
kvamjn “30” etc. (Honti 1993: 117f). Besides the irregular depalatalization, this solution means two different ways of forming of the numerals “8” and “9” in Fenno-Volgaic and Permian, in contrast to the consistent solution of Napoleonskikh (see 8.1.).

8.3. Ug *kvls-(kɔ-) “8” > ObUg *nǐl(j)γ // Hu nyolc (UEW 875; Honti 1982: n.462 and 1993: 111–15) probably represents the dual of Ug *nilji “4” (Szemerényi 1960: 145; Gulya 1976: 314). This internal structure formally corresponds to IE *H₂oko(u)H, “8” = the dual of the form continuing in Avestan ašt- “breadth of four fingers” (Winter 1992a: 13). The main objection against this solution (Honti 1993: 112, 253, fn. 112) consists in the varying root vocalism of the numerals “4” and “8” (front vs. back). But the same opposition appears in the numeral “2” between the Fenno-Permic and Ugric Numerals.

The other etymologies, e.g. those based on Hu nyalab “Bündel, Bund”, KhV ŋula “zusammen” or ObUg *náI // *nöl “nose” (UEW 875; Majtinskaja 1979: 166–67; MSzFE 488) are semantically rather vague. Given that in some Mansi dialects the protoform *ñaláljγ “8” consisting of *lγ “10” is reconstructible (cf. Honti 1993: 114), the “nose” etymology could be applied — although, of course, not with the primary meaning “Nasenzehn” (so Nyíri, Juhász — see Honti 1993: 113). It is possible to imagine the expression “8” = “[two] before ten” where “before” could represent the result of grammaticализation of the ObUg “nose” (cf. e.g. Tzotzil nil ‘(il) “nose; in front of”).

8.4. Sm *kitantett3 / *kitšntett3 “8” < *kitā tett 3 “2 x 4” (SW 71–72).

8.5. The subtractive pattern 8 = 10 - 2 appears in SeT šittį čąγ köt (Honti 1990: 105).

9.1. FV *ükteksā “9” is evidently of the same structure as FV *kakteksan “8” (see 8.1.). Traditionally it has been analyzed as *uküti “1” + negation verb *e + medial reflexive conjugation suffix *k + Px3 Sg *sā (UEW 807; Honti 1993: 110, 116). The final -n in Fi yhdeksān appears probably under the influence of kahdeksan “8” where -n represents the dual suffix. The alternative and evidently more convincing solution of V. Napoleonskikh based on the abessive suffix *-tkak ~ *-tkek for FV, and *-tkam ~ *-tkem for Pe numerals “8” and “9”, was discussed above (8.1., 8.2.).

9.2. Pe *ökmiš “9” (KESK 211; Honti 1993: 159) is of the same structure; see 8.2.


9.6. Sm *ämājšm/n ś “9” is derived from the indefinite pronoun *ämāj “anderer, zweiter” (SW 19), while the second component corresponds perhaps to Sm *ton “Zahl” (SW 165), cf. Kamasin amithun “9” vs. “Abakan” thun “100”.

9.2. Pe *ökmiš “9” (KESK 211; Honti 1993: 159) is of the same structure; see 8.2.


9.6. Sm *ämājšm/n ś “9” is derived from the indefinite pronoun *ämāj “anderer, zweiter” (SW 19), while the second component corresponds perhaps to Sm *ton “Zahl” (SW 165), cf. Kamasin amithun “9” vs. “Abakan” thun “100”.

9.2. Pe *ökmiš “9” (KESK 211; Honti 1993: 159) is of the same structure; see 8.2.


9.6. Sm *ämājšm/n ś “9” is derived from the indefinite pronoun *ämāj “anderer, zweiter” (SW 19), while the second component corresponds perhaps to Sm *ton “Zahl” (SW 165), cf. Kamasin amithun “9” vs. “Abakan” thun “100”.
9.7. MTK *op-ţ3-jągk3 "9" = op “1” + Px3 Sg *t3 + *jągk3 “is absent” (Helimski 1987: 77, 99; cf. SW 41).

9.8. SeT ukkjr čągka kót “9” = “one is absent [from] ten” (Sm *jągk3 > Se čągka, see SW 40).

10.1. FiMd *kümneni “10” (Keresztes 1986, 54; UEW 679) is comparable with Yukaghir *kümne- “10” > Omok kimnel, YkNW kune, YkS chuniella (Billings), kunol(en) (Klička), attrib. / predic. kuni- / kune’o-, YkN kunil’i / -kunali- (Tailleur 1959: 84; Krejnovič 1982: 117–118). Ankeria 1951: 138 also adds ItelmenS (Radliński) kumhtuk “10” (vs. koomnak “5”).

On the other hand, Bouda (UJb 20[1940]: 89) assumed that the Yukaghir forms may be borrowed from Eskimo, cf. SWAlaska qoln, EGreenland qulin, WGreenland qulit “10”, originally **“the top ones”, cf. quleq “the uppermost” (Thalbitzer, JAOS 25/2[1908]: 8–9, 12). But Omok kimnel practically excludes this explanation and Yk *kümne- represents a very hopeful cognate to the FiMd “10”.

Čop, Linguistica 13[1973]: 147, 148 reconstructs FiMd *kümneni “10” interpreting *-n as the dual suffix. The root *kümen- is supposed to correspond with IE *ǵēp isolated from *dekqep “10” = “2x5” (or “1 x 10”, cf. Erhart 1970: 93–94; Winter 1992a: 17). But there is an alternative etymology based on the verb *dek- “to reach, accomplish”, ruling out Čop’s comparison.

There is also a possibility of internal etymology (cf. Majtinskaja 1979: 176) analyzing *kümneni as a compound of the interrogative particle *ku (UEW 101) + *m3n3 “quantity, many” (UEW 279), cf. Komi ki’mjn; Kamasin k’u’men, Koibal kumine “how many” (KESK 152; SW 75) and YkS xamun id. (Angere 1957: 251). The component *m3n3 forms tens (e.g. Komi nel’amjn, Ma naltin; Hu negyven “40”), and so it is natural to expect that it also forms the numeral “10”. The semantical development “quantity/many” → (“number”) → “ten” is plausible, cf. Semitic *t̥āšar- “10” vs. Egyptian ṣis “many, numerous; quantity, multitude”. The shift “number” → “ten” appears in the following case (10.2.) too.

There are promising Altaic parallels: Tk: Čuvaš mōn “big” // MKor mänha- “many, much”, mKor mantha “to be many”, män “size, amount, number” (PKE 105–106) // OJp mane-si “much, many” < *mania-/ *manai (Starostin 1991: 94–95, 145, 276) and /or Even *mian “10” (TMS 1: 534) // Kor män “hand” (PKE 105). A hopeful cognate also appers in IE *men-/ *mon- “many, much”, cf. Lithuanian minià “Menge” (Skalička, UAJb 41[1969]: 341).

The same component *m3n3 could also form the numeral “20” attested in Md kom(e)š “20”, if the derivation *koj3 + *m[3n]3š is correct (cf. KESK 150), and also perhaps Pe *kauž, Ug *kuš “20” (S 544: FU *kuusi; UEW 224), accepting the same internal structure. In the first component U *koji “man” has been identified (S 543; UEW 166–167; 224). But if the second member meant “10”, *koj3 could represent another word for “2”. Alternatively, the Pe-Ug isogloss could be explained as *koj3 “2” + plural marker preserved in Udm -eš, Komi -ęš (Majtinskaja 1979: 83). The same pattern forming tens, i.e. the
cardinal of the first decad + plural marker, is known in some Semitic languages, e.g. Hebrew *šešer “10” vs. *šešārīm “20” etc. or Arabic *fāšar “10” vs. *fisrīna “20” (Gray 1934: 69-70).

There are promising external parallels: Yk *kuj-* > OYk *kijön : (Bilings) purchion, (Klička) purkijen “7”; Čuvan (Boensing) kujen, kujun “2” vs. imoxanbo kijon “7”, (Matjuškin) kuen “2” vs. emganbagu kuek “7”; YkN *kij- “2” (Tailleur 1962: 70) // Mong qojar / qogar “2”, qorin “20” (Anderson 1982: 44).

10.2. FU *luki “10” has been derived from the root *luki- “to count” (S 545) via the meaning “number” (Majtinskaja 1979: 175; UEW 253; Honti 1993: 121). Its traces can be seen in YkN/S (attrib.) jeluku-/ileku- “4” vs. ja- “3”, indicating that “4” = “add 3” or sim. (Krejnovič 1982: 118; Jochelson 1905: 113).

Ankeria 1951: 137-38 compares FU *luki- with Itelmen liux- “Zahl, Rechnen”, Ćukči lg- “rechnen, zählen; Finger”.

FU *luki also has been connected with IE *leğ- “sammeln, zählen, lesen” (Joki 1973: 278-79 with older lit.). Illič-Svityč, Étimologija 1965[67]: 366 adds Mong toga(n) “Zahl, Rechnung”, cf. also Manchu ton “Zahl” (TMS 2: 161-162).

10.3. Pe *das (or *dās) and Hu tīς “10” are borrowed from some (M)Iranian source(s) (Joki 1973: 329-30, 257).

10.4. Kh *jǎŋ “10” has no convincing etymology. The hypothetical parallels or sources of borrowing quoted by Honti 1990: 103 (e.g. Tg *šuwan “10” or rather some of later forms as Evenki jǎŋnǎ “ten objects” — TMS 1: 248) can be supplemented by Turkic dial. forms, like OKipčak ong, cf. Osman dial. Ṝp, ḍp, Ćuvāš vun, vun(n)ǎ “10”, etc. (Sevortjan 1974: 455). The unclear initial *j- can be of later origin. A similar prothesis appears more frequently before front vowels in Ob-Ugric, cf. Kh *jil “vorder” < U *eđe “das Vordere” (Honti 1982: 71; UEW 71-72). The hypothetical source of Kh “10” could be the form of the type *ön. But the front vocalism does not apper in any historically attested Turkic language.

10.5. MT ùuen “10” (Helimski 1987: 77) is undoubtedly of Tg origin, cf. Nanai zɔan, Sibo žuаn, etc. < *šuwan “10” (Benzing 1955: 101; TMS 1: 248).


§3. Conclusion

The most promising cognates of the first decad and the related words in Uralic and other Northern Eurasian language families are distributed as follows:
Note: The symbol (+) means that the meaning is not a simple numeral, e.g. Kor oi, ö “only, a single”, oi-nun-thoji “one-eyed person”. The symbol * is used for non-corresponding meanings of numerals, e.g. FU *nil neljä “4” vs. Tg *nöl) ()ju- “6” (= “10 — 4” ?).

This cognate set reflects the East (North) Nostratic level, approximately corresponding to J.H. Greenberg’s Eurasian. The material correspondences among numerals indicate a closer relationship among Uralic, Yukaghir and Altaic (including Korean and Japanese) within (East / North) Nostratic.

The internal evidence and external parallels allow us to reconstruct the proto-Uralic numeral system consisting of the numerals 1–5. Although there are no evident Sm cognates to FU *neljä “4”, the probable foreign origin of the Sm “4” (< Tk) and the external evidence justify projecting this numeral onto the proto-Uralic level. The FU and Sm numerals “6” differ etymologically, but are formed on the basis of the same semantical pattern. The numeral “7” is borrowed from various IE branches (Baltic, Indo-Iranian, Tocharian). The numerals “8” and “9” represent the subtractions “10–2” and “10–1” respectively, perhaps with exception of the Ug “8” analyzable as the dual of the numeral “4”. Among various denotations of the numeral “10”, the FiMd *kümeni seems to be the most archaic, given proto-Yukaghir *küme- “10”.

The counting system with firmly established numerals 1–5 and 10 looks perhaps as illogical to Europeans, but it is well-known e.g. in Bantu languages. It does not mean the numerals 6–9 did not exist in proto-Uralic; they could be formed (and were formed) through the existing numerals 1–5 & 10 and elementary arithmetic operations.

Abbreviations: Alt Altaic, BF Balto-Fennic, ČK Čukči-Kamčatkan, Dr Dravidian, E East, Fi Finnish, FP Fenno-Permic, FU Fenno-Ugric, FV Fenno-Volgaic, Hu Hungarian, IE Indo-European, Jp Japanese, Kh Khanty (= Ostyak; with dialects: I Irtyš, Kaz Kazym, V Vach), Kor Korean, Lp Lappic, m modern, M Middle, Ma Mansi (= Vogul with dialects: KU lower Konda, TJ Tavda of Janyčkova), Md Mordvin, Mong Mongolian (written), Mr Mari, MT(K) Mator-Taigi(-Karagas), N North, O Old, p proto, Pe Permian, S South, SeT
Selkup Taz, Sl Slavic, Sm Samoyed, Tg Tungusian, Tk Turkic, U Uralic, Udm Udmurt, Ug Ugric, W West, Yk Yukaghir.

References:


UÁj *Ural-Altaische Jahrbücher*.


UJb *Ungarische Jahrbücher*.
