Once Again on the Comparison of Personal Pronouns in Proto-Languages

The article discusses the importance of comparison of personal pronouns for studies dealing with long-range linguistic relationship. It is recognized that individual similarities between pronouns reconstructed for various low-level protolanguages may be ascribed to chance resemblance or borrowing, but the same is much less probable for comparisons between entire paradigmatic systems of pronouns. Based on a brief survey of the principal pronominal forms and markers in most of the protolanguages reconstructed for the world’s language families, it is asserted that many of the current hypotheses on macrofamilies, such as Nostratic, Sino-Caucasian, Austro, and others, are indeed supported by such paradigmatic similarities.

For those scholars who engage in long-range comparison of the world’s languages, personal pronouns have been a viable source of genetic data. Indeed, as was already noticed by the first Indo-Europeanists nearly two hundred years ago, personal markers, both syntactically independent and bound with verbs or nouns, are among the most stable elements of the language’s morphology. The Ukrainian pronominal form мене ‘me’ and the Irish form tú ‘thou’ both sound very much the same as they must have sounded some 6000 ago in the speech of Indo-Europeans, and bear the same meanings. Three personal pronominal forms are deservedly placed in the Swadesh list of the one hundred most stable lexical items in human vocabulary.

Moreover, personal pronouns are relatively rarely borrowed between languages, especially from a genetically different language stocks. It would be an extreme thing to say that they cannot be borrowed at all — since they can and are, as shown by a number of cases — but this happens much more rarely than with nominal or verbal lexical items. Instances of the first person singular pronoun ‘I’ being a loanword in a language are extremely scarce, probably no more than ten reliable examples known so far, and almost all of them can be explained by some peculiar sociolinguistic (or grammatical) conditions.

These two unique characteristics of personal pronouns — internal stability and resistance to external borrowing — make the analysis of personal pronouns, both lexical and morphological, extremely important for comparative linguistic studies.

Can personal pronouns serve as solid proof or, at least, a considerable heuristic argument in proposing distant genetic relationship between languages? This question has long been a subject of discussion in various fields of linguistics. Personal markers were a strong argument in favor of the Altaic theory advocated by numerous authors starting from J.-F. STRAHLENBERG ([1730]). They were one of the keystones in postulating the Nostratic macrofamily by H. PEDERSEN, and the Amerind family in the Americas by J. GREENBERG ([2000]). In Papuan linguistics, personal markers arguably remain as the most important factor in positing linguistic relationship and elaborating classifications, most of which are not yet supported by conventional lexical comparative analysis ([FOLEY 1986: 3]). Pronouns are widely drawn upon elsewhere to prove genetic relationship, and, therefore, methodologically a positive answer to the above question would help us to validate a number of macrofamily hypotheses still under discussion.

Unsurprisingly, opponents of long-distance relationship claims have their own reaction. A number of respectable linguists have challenged the argument that personal pronouns may be essential for long-range genetic comparison. The most widespread and easily understandable objections are quoted as follows:

1. Personal pronouns can be borrowed. Well known examples of this kind are drawn mainly from closely related languages, like English, where they is a Scandinavian loanword, or Dravidian, where Kolami is claimed to have borrowed the 2nd person singular niv from Old Telugu ([KRISHNAMURTI 2001: 91]). More distant languages also demonstrate sporadic examples: thus, Micronesian Chamorro borrowed Spanish yo ‘I’ ([TOPPING 1973: 107]), while Ambonese Malay adopted ose ‘thou’ from Portuguese Você ‘you (Sg.)’. Languages of New Guinea tend to “exchange” certain personal pronouns quite often ([LAYCOCK—Z’GRAGGEN 1975: 732, 737]), though this phenomenon has not yet been sufficiently studied.
2. Personal pronouns are, in fact, mostly very short lexical items usually consisting of CV roots which are hard to compare, because the chance of coincidence is very high ([CAMPBELL 2003]). Moreover, it has been tentatively shown that pronominal patterns tend to crystallize an opposition of a nasal to a stop consonant as the basic phoneme of the pronominal root ([NICHOLS 2003]). Even if this hypothesis is not entirely correct, this pattern works in a great number of the world’s languages.

In our opinion, the most responsible answer to both of these points is not to try and prove the opposite; rather it is to use personal pronouns for language comparison not just as isolated lexical items, but as a paradigmatic subsystem of morphology.

In the majority of the world’s languages, pronouns constitute a balanced paradigm. The types of such paradigms are very well presented by M. CYSOUW ([CYSOUW 2003]), the most basic being the opposition of the two pronouns of the 1st and 2nd persons singular: the pronominal system is limited by this binary opposition in a number of languages, e.g. Golin and Salt-Yui in New Guinea ([FOLEY 1986: 70]). Other structural samples include the number opposition, with four basic members like I — we, thou — you. Still other systems are augmented even further, including 3rd person pronouns, usually adopted from the demonstrative domain; the introduction of the category of clusivity (inclusive / exclusive pronouns of the 1st person plural); more number categories (dual / trial / paucal etc.). Finally, personal pronouns can be declined for case, as in most Indo-European languages, adding even more forms to the pronominal paradigm. Nominative and oblique forms are frequently derived from suppletive roots, e.g., the Indo-European 1st person singular nominative *egHiot(ŋ) vs. oblique *me-.

It seems clear that in order to prove that the languages are genetically related, it is the entire paradigm that should be compared rather than isolated personal pronouns as disconnected lexical items. Adopting the latter approach, we may certainly encounter loanwords or mere coincidences, like Indo-European *me ‘me’ = Sumerian (Eme-sal) me ‘id.’. However, comparison of entire paradigms drastically decreases the probability of either chance similarity or borrowing.

In those languages where personal pronouns constitute paradigms, there are no known cases of borrowing the whole paradigm from another language, provided exceptional situations with pidgin and mixed languages are excluded ([THOMASON–EVERETT 2005]).

One of the most recent claims ([CAMPBELL–POSER 2008: 214]) is that the paradigmatic approach to comparison is equally invalid for long-range comparison purposes. The main reason here is that, as the authors claim, personal pronouns are essentially iconic “phonosymbols”, an assertion that is as insufficiently grounded as it is simply untrue; if it were so, pronominal comparison would be useless altogether, since pronouns would tend to share the same phonological features in all language families — which is clearly not the case, as can be easily seen from the chart below. Campbell & Poser also claim that pronouns are subject to analogical reformations, which is obviously correct (we state the same in [БАБАЕВ 2008: 209]). However, analogical reformations usually have internal reasons and should not be expected to make one language’s paradigm closer to that of another.

Paradigm comparison is also important for historical linguistics, because in long-range studies any comparison of materially short cognates with a CV or VCV structure, where the consonant is the only reliable sound to be compared, is dubious by definition. Therefore, comparison between proto-languages, aimed at reconstructing a 10 000-year-old or older linguistic situation, should be made between reconstructed cognate paradigms of personal pronouns rather than between personal pronouns as lexical roots.

Such attempts have, in fact, been made, some of them quite successful. It was noticed long ago that the pattern exhibited by personal pronoun paradigms in a number of language families in Northern Eurasia is remarkably similar; this pattern can be described as *mV for the 1st person vs. *tV for the 2nd person. The Nostratic/Eurasian theory proposes a genetic link between languages that present this pattern, namely, Indo-European, Altaic, Uralic, Yukaghir and, most probably, Chukchi-Kamchatkan. Indeed, (a) the linguistic probability of borrowing this pattern is practically zero; (b) the statistical probability of a chance similarity in such a geographical cluster is also close to zero; and (c) a pattern like this is not attested for any other proto-language of any other family on Earth. This makes the “pronominal factor” one of the strongest bits of grammatical evidence for the Nostratic hypothesis.

Another example is the American pronominal pattern *na- (1st pers.) vs. *mi (2nd pers.), first noted by A. Trombetti ([1905]) in a great number of the languages of both Americas and later de-

---

1 One well known case is that of the Thai language, which uses Chinese and English personal pronouns in specific sociolinguistic contexts ([COOKE 1968: 11–12]). However, they do not replace the common Thai personal pronouns.
veloped by J. Greenberg into one of the key arguments of his Amerind theory. Again, the spreading of the pattern is extremely wide, and the chance of borrowing or coincidence is lower than low.

The present article is inspired by a tentative list of reconstructed pronominal forms for the first two persons in the world’s languages presented by M. Ruhlen ([1994])². However, we find it necessary to continue Ruhlen’s effort by substantially adding to the sketch pencilled by his pioneering work. Our task will be to present an updated and accurate comparative chart of personal pronoun paradigms, reconstructed for proto-languages of most of the world’s language families, to be subject to further analysis.

To make our results as precise as possible, we will rely on the most recent data collected by scholars, especially concerning languages of Africa and the Pacific region. Although many families still lack thorough proto-language reconstructions, most of their pronominal systems have been at least tentatively sketched out since Ruhlen’s work was published fourteen years ago. Apart from independent pronominal forms, we will try to include bound verbal and nominal affixes denoting person as well, provided they differ from the former.

Finally, it is necessary to give the appropriate reference links to etymological dictionaries and/or comparative grammars for each reconstructed form that we quote for a proto-language. Unfortunately, no such references are given in Ruhlen’s list, which makes it harder to resolve all the problems that may emerge during one’s analysis of it. We have attempted to find as many links as possible for the reconstructions, bearing in mind that for a number of language families (African for example) no comparative work exist at all, and for these cases we make our own tentative reconstruction based on the available dictionaries and word lists (these reconstructions are marked by a double asterisk).

The chart below will provide sufficient material for such an analysis. Along with the usual pronominal forms reconstructed for the family’s proto-language, we also mark a number of peculiar forms from separate languages (or language groups) of the family in case they differ from the main family pattern — e.g., Svan forms are placed alongside the reconstructed Proto-Kartvelian ones, because Svan is considered the most morphologically archaic dialect of the family, and a number of judgments on Proto-Kartvelian pronouns has been made based exclusively on very peculiar Svan forms.

We have attempted to make our chart below as taxonomically correct as possible. The majority of the families listed go back to proto-languages which should have existed at roughly the same chronological period: 6 to 8 thousand years ago, considered by many of today’s mainstream linguistics as the limit for comparative reconstructions. And, although a number of macro-families have also been recognized in the last decades, for some of them it is not always easy to reconstruct a single system of personal pronouns. Thus, the Afrasian pronominal reconstructions are still under serious discussion. Niger-Congo and Austro-Asiatic may be taken as even better examples: these two macrofamilies seem to be well established and widely accepted, but the personal pronoun systems for their proto-languages have not yet been verified. Therefore, we will only demonstrate pronominal forms for their subfamily proto-languages: Mande, Gur, Dogon, &c. and Munda and Mon-Khmer respectively. The Australian macrofamily is so far an unverified hypothesis (although widely accepted), and the New Guinean ‘phyla’ are even less so ([Foley 1986]). The most complicated situation lies in the Americas: the classification of solidly reconstructed families includes more than 60 taxonomic units ([Campbell–Mithun 1979: 39–46]), and, although it has been reduced to around a dozen macrofamilies by Greenberg and Ruhlen ([2007: 6]), their theory is not universally accepted and still needs further verification.

The above chart, with its schematic information, cannot even pretend to be perfect, since each form in this matrix is subject to long and deep analysis per se. But even this sketchy picture still allows to make a number of interesting conclusions, based on the patterns of personal pronoun paradigms.

1. The languages of northern Eurasia show a pattern “M—T” which is found nowhere else in the world on the family proto-language level. Namely, these are Indo-European, Uralic, Yukaghir, Altaic and Chukchi-Kamchatkan languages. This phenomenon has been identified long ago and constitutes a part of the morphological evidence for the Nostratic hypothesis.

2. A similar pattern “M—S” is present in Indo-European, Kartvelian, Altaic and, arguably, Chukchi-Kamchatkan, where the form *ɣəš ‘thou’ is of an unclear origin. In Indo-European, the distribution of

² Editors’ note: A similar list of 1st and 2nd person pronouns has been earlier produced by S. L. Nikolaev during his talk on the Nostratic Seminar in Moscow in April of 1985. The list was, however, much shorter, including only those major families across Eurasia, Africa, and North America for which at least some sort of protolanguage reconstruction had already been established or sketched.
“M—T” vs. “M—S” is morphological (the former is found in the pronominal system, the latter in the verb system), while in Altaic the variability is dialectal (Mongolic vs. Turkic and Tungus-Manchu). In both Indo-European and Altaic, however, the two systems are projected onto the proto-language level.

It is necessary to add that Nostratic languages also share the case paradigm for the 1st person pronoun: the form *mVn(V)- may be reconstructed for the Nostratic proto-language on the basis of Indo-European, Altaic, Uralic and Kartvelian languages ([ИЛЛИЧ-СВИТЫЧ 1971: 6]; [БАБАЕВ 2008]).

<table>
<thead>
<tr>
<th>Language Family</th>
<th>'T'</th>
<th>'Thou'</th>
<th>'We'</th>
<th>'You (Pl.)'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indo-European³</td>
<td>n *eg'ho(m)</td>
<td>*iū, (t)i(e)</td>
<td>n *me(s), *we(i)-</td>
<td>n *iū-</td>
</tr>
<tr>
<td></td>
<td>obl *me(ne) v st *-H-</td>
<td>v *-s-</td>
<td>obl *ne/o-</td>
<td>obl *we- v *-te</td>
</tr>
<tr>
<td>Uralic⁴</td>
<td>n *mE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>obl *mENV v st *-k</td>
<td>n *tE obl *tENV</td>
<td>*mE</td>
<td>*tE</td>
</tr>
<tr>
<td>Yukaghir³</td>
<td>mit</td>
<td>tit</td>
<td>met</td>
<td>tet</td>
</tr>
<tr>
<td>Altaic⁶</td>
<td>n *bi obl *mi-n-, ? *ŋa</td>
<td>n *si, *t'i obl *si-n-, ? *na</td>
<td>*ba / *bu obl. *ma-n-</td>
<td>*t'a, *su obl. *su-n-</td>
</tr>
<tr>
<td>Turkic</td>
<td>n *bi obl *mān</td>
<td>n *si obl *sān vprt *-ŋ</td>
<td>*biŋ</td>
<td>*siŋ</td>
</tr>
<tr>
<td>Mongolian</td>
<td>obl *nama-</td>
<td>či &lt; *t'i</td>
<td>obl man-</td>
<td>obl ran-</td>
</tr>
<tr>
<td>Korean</td>
<td>na &lt; *ŋa, čé</td>
<td>ne &lt; *na</td>
<td>ne-huy</td>
<td>-huy</td>
</tr>
<tr>
<td>Japanese</td>
<td>*a &lt; *ŋa, wa(re)</td>
<td>na(re)</td>
<td>-huy</td>
<td>-huy</td>
</tr>
<tr>
<td>Kartvelian⁷</td>
<td>*me(n) vsubj *(x)w-</td>
<td>*si vsubj *x- vobj *g-</td>
<td>*čwe(n)</td>
<td>*tkve(n)</td>
</tr>
<tr>
<td>Svan⁸</td>
<td>excl naj / nāj vincl subj l- vincl obj gw-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dravidian⁹</td>
<td>n *nyān obl *yan- vabs *-kV</td>
<td>n *nīn obl *nin- vabs *-ti</td>
<td>n exci *nyām obl exci *yam- n incl *nām obl incl *nam-</td>
<td>n *nīm obl *nim-</td>
</tr>
<tr>
<td>Chukchi-</td>
<td>*ŋa-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kamchatkan¹⁰</td>
<td>*ŋa-š obl *ŋa-n</td>
<td>muri</td>
<td>turi</td>
<td></td>
</tr>
<tr>
<td>Eskimo (W)¹¹</td>
<td>rel -ma abs -qa</td>
<td>rel -pa(k) abs -n</td>
<td>rel -m-ta abs -(x)put/-wut vdu -(x)puk</td>
<td>v rel –posi vabs -(x)si vdu -(x)tok</td>
</tr>
<tr>
<td>Elamite¹²</td>
<td>v (h)u obl un abs –k</td>
<td>nu, ni v -(a)ti(i)</td>
<td>incl ela, elu excl nika, niku</td>
<td>num</td>
</tr>
<tr>
<td>Chadic¹³</td>
<td>I *?ran-i</td>
<td>masc I *ka, II</td>
<td>*ku fem *ki(m)</td>
<td>incl *muni excl *na, *?yna</td>
</tr>
</tbody>
</table>

³ [Беехес 1995: 207–21].
⁵ [Николаева–Хелимский 1997: 166].
⁶ [EDAL 225, 1237]; [Иващи 1998: 130–135].
⁷ [Климов 1998], [Шмидт 1984].
⁸ [Тутт 1997: 18–26].
⁹ [Сердюк 2006].
¹¹ [Бергсланд 1986]; [Меновщик 1997].
¹² [Дьяков 1979].
¹³ [Блахек 1991: 37].
<table>
<thead>
<tr>
<th>Language Family</th>
<th>T</th>
<th>'Thou'</th>
<th>'We'</th>
<th>'You (Pl.)'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berber[^14]</td>
<td>subj *ṭnakkw</td>
<td>subj masc *kayy, fem *kamm</td>
<td>*na-</td>
<td>*-tum-</td>
</tr>
<tr>
<td></td>
<td>obj *ṭy, *ṭi(w)</td>
<td>subj *ok, ik, fem *am</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cushitic[^13]</td>
<td>subj *ʔani</td>
<td>subj *ʔan-til</td>
<td>subj *muni</td>
<td>*(ʔan)-tínV</td>
</tr>
<tr>
<td></td>
<td>obj *ʔi</td>
<td>subj *ʔa</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>N *ʔantu, S *yV</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>fem *ʔan-tina</td>
<td></td>
</tr>
<tr>
<td>Egyptian[^13]</td>
<td>('in)-k, -j</td>
<td>fem tm(t), -t</td>
<td>('in)-n</td>
<td>tn</td>
</tr>
<tr>
<td>Sino-Tibetan[^5]</td>
<td>*ŋa</td>
<td>*nay, *kIV-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Caucasian[^16]</td>
<td>I (n) *ẓo (E *ẓō-, W *sa)</td>
<td>I *wo (E *wō- / *g w, WC *wa)</td>
<td>EC incl *Lā:</td>
<td>EC *zwē</td>
</tr>
<tr>
<td></td>
<td>II (obl) *nV / *mV</td>
<td>WC fem *ba</td>
<td>EC excl *(n)Vži</td>
<td>WC *ḥa</td>
</tr>
<tr>
<td></td>
<td></td>
<td>II *yu- / *yV-</td>
<td>WC *štV</td>
<td></td>
</tr>
<tr>
<td>Hurro-Urartian[^17]</td>
<td>*(j)es- / *so-</td>
<td>*we- / *u/-o-</td>
<td>*(j)es- / *so-</td>
<td>*we-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yeniseian[^18]</td>
<td>I *ʔaŋ</td>
<td>I *ʔaw</td>
<td>*ʔawŋ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>II *b-</td>
<td>II *ʔV-</td>
<td>*ʔawŋ</td>
<td></td>
</tr>
<tr>
<td>Na-Dene[^19]</td>
<td>*šwī / *xwī</td>
<td>*ni, *wī (Tlingit)</td>
<td>*na-, *nu-, *ta-, *tu-</td>
<td>*ʔuX-</td>
</tr>
<tr>
<td>Burushaski[^20]</td>
<td>ʒa poss a-</td>
<td>un obl gu- / gō</td>
<td>mi</td>
<td>ma</td>
</tr>
<tr>
<td>Basque[^21]</td>
<td>ni v trans *-da-</td>
<td>(h)i v trans masc *-ga-</td>
<td>gu</td>
<td>zu</td>
</tr>
<tr>
<td>Sumerian[^22]</td>
<td>nae poss -ṇu *mV (?)</td>
<td>zae poss -zu</td>
<td>mēdē poss -me</td>
<td>mēdē, -ze</td>
</tr>
<tr>
<td>Nahali[^23]</td>
<td>ʒo obl en-</td>
<td>nē</td>
<td>ʒo, tye-ko</td>
<td>nē(ko)</td>
</tr>
</tbody>
</table>
| Nivkh[^24] | ʔi | ċi | excl ʔi n du mê- | excl ċi, -as-
| | | | incl subj u- incl obj ʔi- |
| Ainu[^25] | ku- obj 'en- / -'en- | 'e- | excl abs ċi-, -as-
| | | | incl subj u- incl obj ʔi- |
| | | | | 'eci-, es- |
| Thai-Kada[^26] | *k(ā)u | *m(a)u | rau, *tu | miŋ, suu |

[^15]: [Бенди́кт 1972: 93]; [Старос́tin 1984].
[^16]: [Старос́tin 1984]; [Nicholls 2003: 293].
[^18]: [Старос́tin 1984].
[^19]: [Бенди́кт 1992].
[^20]: [Бергер 1974: 24].
[^21]: [Трасск 1997: 218].
[^23]: [Ша́фер 1941: 363]; [Пиннов 1966: 189].
[^24]: [Грузде́ва 1998].
<table>
<thead>
<tr>
<th>Language Family</th>
<th>Τ'</th>
<th>'Thou'</th>
<th>'We'</th>
<th>'You (Pl.)'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austronesian27</td>
<td>*(w) anu / *ku</td>
<td>*kaw, *(ka)šu</td>
<td>excl *(k)ami incl *(k)ita</td>
<td>*(ka)mu</td>
</tr>
<tr>
<td>Mon-Khmer28</td>
<td>*(V) ko / *ki, *?Vo / *?Vy</td>
<td>*mi? / *mVj</td>
<td>*hi? incl *bon du *?a?</td>
<td>*(b)na?, *pi?</td>
</tr>
<tr>
<td>Munda29</td>
<td>*iy</td>
<td>*me</td>
<td>excl *le incl *bi(n)</td>
<td>*pe</td>
</tr>
<tr>
<td>Miao-Yao30</td>
<td>*ko(u), *waŋ</td>
<td>*koŋ, *kam</td>
<td>*m(p)V</td>
<td>Miao du *mu \ Yao *ni-V-</td>
</tr>
<tr>
<td>Kusunda32</td>
<td>tsi</td>
<td>tok</td>
<td>nu</td>
<td>nok</td>
</tr>
<tr>
<td>Papuan (general types)33</td>
<td>I nV / *-Vn, v *-Vk v obj *-Vy</td>
<td>*.Vm(?), *-Vt v obj *-Vš</td>
<td>*-nān</td>
<td>*-wāw</td>
</tr>
<tr>
<td>West Papuan29</td>
<td>*nV</td>
<td>*nV</td>
<td>I ni</td>
<td>I ki</td>
</tr>
<tr>
<td>Keresiouan</td>
<td>*(V)nā-</td>
<td>*ni-</td>
<td>*ni</td>
<td>*ki</td>
</tr>
<tr>
<td>Hokan</td>
<td>*n'a / *n'i</td>
<td>*mi / *ma</td>
<td>*m- / *p-</td>
<td>*y-</td>
</tr>
<tr>
<td>Penutian</td>
<td>*n-</td>
<td>*mV</td>
<td>III na</td>
<td>III mi</td>
</tr>
<tr>
<td>Aztec-Tano40</td>
<td>*neʔ</td>
<td>*peme</td>
<td>*ta-</td>
<td>*peme</td>
</tr>
<tr>
<td>Chibcha</td>
<td>*na-sV</td>
<td>*mue-ya(nV)</td>
<td>*na-m</td>
<td>*mato</td>
</tr>
<tr>
<td>Quechua</td>
<td>*nu-qa</td>
<td>*qa-m</td>
<td>*mi</td>
<td>*no</td>
</tr>
<tr>
<td>Panoan</td>
<td>*mi</td>
<td>*no</td>
<td>incl *ānā excl *ām</td>
<td>*wēm</td>
</tr>
<tr>
<td>Nilo-Saharan41</td>
<td>*āgā</td>
<td>*āh</td>
<td>*ini</td>
<td>incl *-r(i)- incl *-n(a)- *-p(a)-</td>
</tr>
<tr>
<td>Kordofanian42</td>
<td>*pi</td>
<td>*ŋa</td>
<td>excl *-r(i)- incl *-n(a)- *-p(a)-</td>
<td></td>
</tr>
<tr>
<td>Atlantic43</td>
<td>*mi</td>
<td>obj *ne</td>
<td>*mo</td>
<td>*ti(V) (cəv.) *sV?</td>
</tr>
<tr>
<td>Mande (S)44</td>
<td>*N, *mā</td>
<td>*i / *č, *bV</td>
<td>excl *yũ/o incl *koa</td>
<td>*kā</td>
</tr>
</tbody>
</table>

---

30 [Benedict 1975: 203, 205]; [Peiros 1998].
32 [Watters 2006: 44].
33 [Wurm 1975: 191].
34 [Capell 1975: 678].
35 [Laycock 1975: 768–770].
36 [Foley 2000].
37 [Blake 1979: 31–69].
38 [Blake 1988: 7].
40 [Steele 1979: 447].
41 [Ehret 2001: 225–231].
43 [Posdniakoff – Segerer 2004].
44 [Yıldırı̇m 2006].
Ones Again on the Comparison of Personal Pronouns in Proto-Languages

nouns in North Caucasian should be reconstructed as *nV suppletive pronominal paradigms.

It is interesting that the Sino-Tibetan proto-language paradigm does not fit this pattern. This cannot be interpreted as anything other than a trace of genetic kinship between the languages of Eurasia.

3. Another common paradigm, shared by the languages grouped within the Nostratic/Eurasian hypothesis, is the “K—M” pattern for the 1st person stative (intransitive) vs. relative (transitive). The Uralic intransitive opposition of the 1st person markers *-k (intransitive) and *-m (transitive) ([ХЕЛИМСКИЙ 1982]) is usually compared with the Turkic preterite (“pl”) *-k and present *-m, and the Eskimo absolutive *-qa together with relative *-ma. The Hittite “series” of verb conjugation in *-hi and *-mi, usually regarded as the ancient Indo-European opposition of stative and transitive, might also be derived from the Nostratic type, since there are clues for the velar (or postvelar) origin of the Indo-European laryngeal phoneme(s). The Nostratic postvelar form was proposed by Илич-Свityч ([ИЛИЧ-СВИТЫЧ 1971: 149]).

The syntactic function-related “K—M” opposition of the 1st person forms finds a similar correlation in the 2nd person, where the stative counterpart of “K” is usually “T”. The pattern “K—T” in its stative aspect is found in Dravidian appellatives ([ZVELEBLI 1990: 35–36]), in Indo-European (where *-t is the 2nd person perfective affix in the same paradigm as the 1st person *H), as well as in Elamite and Afrasian families, which are usually considered as the closest relatives of Nostratic or Eurasian by long-range comparativists.

Again, it should be noted that this particular system is never reconstructed for any other proto-language around the world. This uniqueness, hardly representing the result of cross-borrowing, cannot be interpreted as anything other than a trace of genetic kinship between the languages of Eurasia.

4. Further evidence for another reconstructed macrofamily, Sino-Caucasian, or even broader, Dene-Caucasian, said to include North Caucasian, Sino-Tibetan, Yeniseian, Na-Dene languages and possibly some ancient tongues of Europe and the Near East (Hurrian, Urartian, Hatti, Basque) as well, is also found in the chart above. The pattern “S/Z—W” is easily traced in North Caucasian, Yeniseian, Na-Dene, and Hurro-Urartian languages. The primary consonant phoneme in the 1st person is not that clear, but it is definitely a coronal sibilant/affricate. Nahali data might be considered here as well.

It is interesting that the Sino-Tibetan proto-language paradigm does not fit this pattern. This discord might, however, be explained by the hypothesis that Proto-Sino-Caucasian, in fact, had two suppletive pronominal paradigms. Starostin ([1984]) explicitly showed that the second set of pronouns in North Caucasian should be reconstructed as *nV / *mV (1st person sg.) – *yu / *ywV (2nd person sg.), which correlates well with such Sino-Tibetan forms as, respectively, *ŋa and *kwV, as

---

Language Family | 'I' | 'Thou' | 'We' | 'You (Pl.)'
---|---|---|---|---
Ijo | *a, *i/c | *i | *w3 | *o
Dogon | mi | u- / uwV | be, ene | e
poss ebV
Gur-Adamaua | *mi / *ma | *mo | *-a(n) | *u(n) / *-u(n)
**Gur | *ma | *u / *bV | *tV | *i, *nV
Ubangi | *mi | *mc | *rele | *pene
**Kwa | subj *N | *o, *wo | *ne ? | *mo / *mu
obj *mc
**Benue-Congo | subj *N | *o, *u | excl *tV | obj *mi
obj *mc
**Bantoid | *N | *a, (obj?) *wo | *se | *ne, *we
obj *mc
Bantu | subj *Ni | *tu | *nV | subj *u / *we
obj *mc
obj *ku |
Khoisan | N | *mV, S *yV | excl *e | *-u
C | *ti- / *ta-

---

45 [WILLIAMSON 2004].
47 [BOYD 1989: 207].
49 [MONIÑO 1995].
50 [SEGERER 2002–2007].
52 [GREENBERG 1963: 70–74]; [STAROSTIN 2003]; [GÜLDEMANN 2005].
well as Yeniseian prefixed pronouns *b- and *kV-. The Burushaksi forms might also fit here, as proposed by Bengtson ([1997]) in his cognate list for a theory of Burushaski as a Sino-Caucasian language. It should be noted that, typologically, both Yeniseian and Burushaski share the two-paradigm system of personal pronouns with North Caucasian.

5. Languages of South-Eastern Asia and the Pacific seem to share the pattern “K—M” (for the opposition between the 1st and 2nd singular persons), as shown by Tai-Kadai, Miao-Yao, Mon-Khmer, and Austronesian data. These are the families grouped by P. Benedict into the hypothetical Austric macrofamily. Additional similarities can be seen between single forms of pronouns in Austric languages (like *Vŋ ‘I’); some of them are shared by Ainu, which is also considered to be related to Austric ([Bengtson—Blazek 2000]). The protoforms of pronouns for the Austro-Asiatic protolanguage are still foggy.

It is curious that the same “K—M” pattern also characterizes a number of languages in New Guinea: this is the type identified by Wurm ([1975]) as “Set II” of the Papuan personal pronouns. It is emphasized that set II is supposed to mark the most archaic strata of human language in New Guinea; nowadays, it is used in languages of the Torricelli phylum and those languages of the Trans-New-Guinea phylum which show a strong substratum influence. Although genetic relations between Austric languages of Asia and the Pacific have not yet been thoroughly studied, this coincidence might become a hint for any further studies in this regard.

6. Languages viewed under the Niger-Congo macrofamily umbrella show distinct similarities in the pronominal system. A pattern “M—W” for the 1st and 2nd person singular pronouns is displayed by most of the language families of Niger-Congo. Moreover, at least one 1st person plural pronoun can probably be reconstructed as *tV based on Bantu, Atlantic, and Gur languages. It seems to us that the Niger-Congo system actually consisted of two series of pronouns, one subject, one object (oblique), just as it looks like in numerous languages of practically all the families. The other 1st person singular pronoun *N (representing a variable syllabic nasal) most probably denoted subject forms opposed to the object (oblique) *m.

Niger-Congo forms tend to show some basic similarity in the pattern (N[any nasal] — W), in the line of the hypothesis of their common genetic roots ([Blench 2007]). Kordofanian and Ijoid languages in this regard lie closer to Nilo-Saharan than to common Niger-Congo patterns. But there is still obvious lack in solid basis for the proto-forms, so these clues should be used carefully before reliable reconstructions are made for all sub-families of Niger-Congo.

7. Another striking similarity (“T—N”, basically 1st pers. *tV vs. 2nd pers. *nV, with plural forms in *-k) lies between Kusunda, a language isolate of a forest gatherers’ tribe of central Nepal, and the West Papuan languages whose speakers are thought to represent the latest wave of Papuan migration to the island of New Guinea ([Wurm & al. 1975: 947]). More anthropological and cultural similarities can be found between Papuans and Kusunda, so the two languages do deserve deeper linguistic comparison ([Whitehouse & al. 2004]).

It goes without saying that the more we delve into language prehistory, the more observations like these will be formulated. Non-evident cognates may well be hidden behind phonetic changes, impossible to uncover without a thorough analysis.

However, some more general comments might be even more valuable from the point of view of linguistic typology.

It would be incorrect to call the choice of phonemes for personal pronouns a linguistic universal. Neither in proto-languages nor in living ones does there seem to be any inclination towards a “resonant — stop” phonematic distribution like the one proposed by J. Nichols. From a statistical point of view, it is obvious that there are plenty of proto-languages with both the 1st and the 2nd person pronouns based on nasals (even such a phenomenon as sharing the same nasal sounds for both persons is quite common in living languages, e. g. Korean), or both formed by stop consonants (Miao, Afrasian, Austronesian, Kusunda, &c.). There are also clear cases when both basic pronouns are built on the basis of sibilants. Although this statistical survey is relatively small, from a certain point of view it is more precise than one that would include all the living languages of the world, since the disproportionality in absolute language numbers among the various families would spoil the picture in the latter case. Thus, several hundred Austronesian languages will show phonetic values that will decisively overwhelm those of the single Basque language, but only because they are closely interrelated, while Basque is an isolate.

Subsequently, it is also incorrect to claim that “chance resemblance” can play an important part in pronominal comparison between languages of different families. There are absolutely no coinci-
idences in paradigm patterns between the languages which are not thought to be genetically related by modern long-range comparativists. One can certainly find similarities that may be hard to explain (like the above-mentioned Basque and Sumerian, or Papuan and Nivkh which share the “N—K” pattern). However, we are dealing here not just with single isolates instead of family proto-languages, but also with vastly diverse chronological periods: Basque and Sumerian data are separated by a time gap of at least 3000 years. It is only the comparison between chronologically close proto-languages of large families that may show us some clues in how to reconstruct deep linguistic prehistory.

References


45


RENISON 2000 — R. RENNISON. Koromfe–English/French/German Dictionary [Forthcoming].


WILLIAMSON 2004 — K. WILLIAMSON. Proto-Ijoid Reconstructions [Forthcoming].
лает вывод, что многие из ныне существующих гипотез о дальнем родстве (таких, как ностратическая, сино-калых местоименных парадигм. Основываясь на кратком обзоре основных форм личных местоимений и местоимений случайным сходством или как заимствования, такая ситуация гораздо менее вероятна при сопоставлении це-
риний и местоименных показателей для исследований в области дальнего родства языков. Утверждается, что, хотя


Дьяконов 1967 — И. М. Дьяконов. Языки древней Передней Азии [Languages of the Ancient Southwest Asia]. М.


Канева 2006 — И. Т. Канева. Шумерский язык [The Sumerian Languages]. СПб.: «Петербургское Востоковедение».


Резюме

Статья посвящена вопросу о том, насколько значимым является сопоставление данных личных местоиме-
ний и местоименных показателей для исследований в области дальнего родства языков. Утверждается, что, хотя отдельные, бессистемные сходства между местоимениями в пражах нелгобукого уровня могут быть объяснен-
ны случайным сходством или как заимствования, такая ситуация гораздо менее вероятна при сопоставлении це-
лых местоименных парадигм. Основываясь на кратком обзоре основных форм личных местоимений и местоименных показателей, реконструируемых специалистами для подавляющего большинства пражков, автор де-
лает вывод, что многие из ныне существующих гипотез о дальнем родстве (таких, как ностратическая, сино-кавказская, австрическая и т. п.) подтверждаются наличием как раз такого рода парадигматических сходств.