

The different items can be assigned to the following categories:

- clear cognates: 2, 3, 4, 6, 11, 13, 22, 28, 32, 33, 42, 43, 45, 46, 47, 48;
- possible cognates: 1, 9, 10, 14, 15, 23, 34, 35, 50;
- unclear status: 18, 19, 25, 31, 37;
- unknown words: 5, 8, 12, 16, 21, 24, 27, 29, 30, 38, 39, 40, 41, 49;
- clear or probable loanwords: 7, 17, 20, 26, 36, 44.

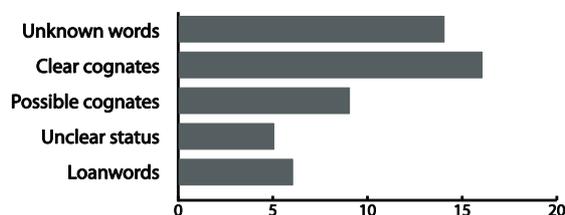


Fig. 1. Situation of Hurrian according to the 50-item word list of the reviewer

According to the reviewer, Greek and Sanskrit score 23 matches on 50, but Proto-Uralic and PIE score between between 12 to 14 matches. This can be compared with Hurrian scoring 16 clear cognates and 25 clear and possible cognates on an incomplete list of only 35 items. Even though the authors do not endorse the method followed or proposed by the reviewer, it is glaringly obvious that it actually supports their theory and completely refutes his own claim that “it should be clear to everybody who is familiar with Hurro-Urartian and IE languages that these two families are genetically unrelated.” (p. 204)

Though the authors reject nearly all of the conclusions reached and criticisms articulated in the review, they are grateful for the reviewer’s time and additional data. They also share his concerns that Hurro-Urartian reconstruction should be carried out and that an up-to-date thesaurus should be made available for further research on these languages.

Alexei Kassian

Russian State University for the Humanities (Moscow)

More about the theoretical foundations of lexicostatistics

0. Since A. Fournet and A. R. Bomhard have presented a reply to my recent unfavorable review of their monograph *The Indo-European Elements in Hurrian* (see Fournet & Bomhard 2010 and Kassian 2010) that has been accepted by the *JLR* editorial board, I am obliged to compile some kind of a “reply to reply”. My text will, however, be shorter than the one by F&B, because this time I will not be touching upon specific Hurrian data and instead confine myself to methodological matters only.

1. As I have previously pointed out (1010: 200 ff.), any pair of languages which are conventionally assumed to be genetically related at a reasonable time depth possesses a significant number of etymological matches with coinciding meanings between the basic vocabularies of these languages, most importantly, between *core* vocabulary, summarized as the 50-item wordlist. One is free to accept lexicostatistics as a working method (like, for instance, the present author) or reject “the theoretical legitimacy of such a wordlist” (as do,

for instance, the authors of the book under review), but the fact is that neither I myself, nor any of my colleagues from the Moscow school are aware of *even a single* reliable exception from the phenomenological rule stated above.¹ On account of this, I consider the presence of etymological matches with coinciding semantics between 50-item wordlists of two languages (or proto-languages) to be a *necessary condition* of recognizing genetic relationship between languages.²

I must emphasize more explicitly that a pair of etymological cognates whose meanings are different

¹ One possible exception could be, e.g., the Na-Dene affiliation of the Haida language. E. Sapir's hypothesis that Haida is genetically related to Athapaskan is not generally acknowledged, but has penetrated into certain basic linguistic *compendia*, e.g., Lewis 2009 (where Haida is classified as a separate branch of the Na-Dene family). As a matter of fact, the 50-item wordlist for Haida demonstrates a factual absence of plausible etymological matches with Na-Dene data (leading me to accept the hypothesis that numerous Haida–Na-Dene isoglosses have contact origin).

In an attempt to challenge this empirical rule, F&B in their reply claim that there are no or almost no matches between Modern English and Bengali 50-wordlists (although both languages belong to the IE family). This is not true. One can count sixteen Modern English–Bengali isoglosses with direct semantic correspondences:

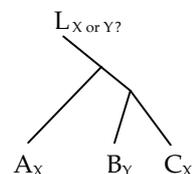
- Eng. *nail* ~ Beng. *nokh* 'id.'
- Eng. *foot* ~ Beng. *pa* 'id.'
- Eng. *horn* ~ Beng. *šing* 'id.'
- Eng. *I* ~ Beng. *ami* 'id.'
- Eng. *name* ~ Beng. *nam* 'id.'
- Eng. *new* ~ Beng. *nob* 'id.'
- Eng. *nose* ~ Beng. *nak* 'id.'
- Eng. *not* ~ Beng. *na* 'id.'
- Eng. *one* ~ Beng. *ek* 'id.'
- Eng. *star* ~ Beng. *tara* 'id.'
- Eng. *tongue* ~ Beng. *jib* 'id.'
- Eng. *tooth* ~ Beng. *dāt* 'id.'
- Eng. *two* ~ Beng. *dui* 'id.'
- Eng. *us* (obl. stem) ~ Beng. *amra* 'we' (dir. & obl. stem)
- Eng. *what* ~ Beng. *ki* 'id.'
- Eng. *who* ~ Beng. *ke* 'id.'

Of course, this list contains a number of phonetically non-transparent cognations, which we can reveal only using our knowledge of Germanic and Indo-Aryan historical phonology. But in fact, a "blind" test of Modern English–Indo-Aryan relationship also yields positive results; see Baxter & Manaster Ramer 2000 on an automated comparison algorithm for 33-item Hindi–English wordlists.

² The Moscow school does not, by any means, regard lexicostatistics as a method that is in any way alternative or opposed to the comparative method and/or internal reconstruction, as F&B state in their reply. Lexicostatistics is complementary to these procedures, and one of the major advantages it offers is that it can work as a safeguard against any possible abusing of the comparative method, which, given enough leeway and freedom of assumption on the part of the researcher, can easily lead to dubious or even absurd results.

between two compared languages constitutes positive evidence for *etymostatistics*, a very different and, at such time depths, much less reliable procedure, but not for *lexicostatistics*. For the latter only *direct* phonetic and semantic correspondences are taken into account. Thus, Lat. *nox* 'night' ~ Grk. *νύξ* 'night' is a positive lexicostatistical pair, whereas Lat. *nox* 'night' ~ Hitt. *neguz* 'evening, nightfall' is not.

A related problem, which I must additionally mention here, is how we reconstruct the basic vocabulary of the *proto-language*. Or, more precisely: in which case and for what reason we assume that the given meaning must have been expressed by the given particular reconstructed root/stem and not by another one in the proto-language. This question is not so trivial, and the problem is regularly omitted in various handbooks of historical linguistics.³ However, there are a number of reliable criteria, some of them internal (e. g., a transparent synchronic level derivation of the word for 'eye' from the verb 'to see' should point that such a term for 'eye' is innovative and should not be projected onto the proto-language level) and some external. Of the latter, the strongest criterion is the distributional one. Let us suppose the following genealogical tree where *L* is a proto-language and *A*, *B*, *C* are its daughter languages.



A certain word in *A* means 'X', its etymological cognate in *B* has a different meaning 'Y', whereas its cognate in *C* also means 'X'. In the absence of additional evidence, the most likely solution is that in *L* this word denoted 'X', not 'Y', since, according to general scientific principles, we should prefer the most economic scenario (one semantic shift 'X' > 'Y' in *B* vs. two independent shifts 'Y' > 'X' in *A* and *C*).

Let us illustrate this with an actual example. The best and obvious candidate for the Proto-IE word for 'bird (in general)' from the distributive viewpoint is **(a)u(e)i-*, because its proto-status is proven by independent attestations in various IE branches (OInd. *vi-*,

³ The only formal approach to semasiological reconstruction known to me is a method of semasiological tables (or semasiological grids) applied by Дыбо А. 1996 to Turkic and Iranian somatic terms of the shoulder girdle. The aforementioned method is somewhat of an elaboration on the concept definition mapping method. Unfortunately, this work by A. Dybo is available only in Russian.

Arm. *haw*, Lat. *avi-s*, all meaning ‘bird (in general)’). In other branches various independently non-matching formations for ‘bird’ are observed. E.g., in Slav. there is an etymologically obscure form **pūt-*. On the contrary, Grk. (Hom.) ὄρνις ‘bird (in general)’ clearly originates from a specific term for ‘eagle’, since the semantics ‘eagle’ (or at least ‘a specific kind of big predatory bird’) for IE **Hor-Vn-* / **Hor-Vl-* is independently proven by data of various IE branches, according to the same distributional principle (Anatolian, Germ., Celtic, Balto-Slavic, all meaning ‘eagle’; also maybe here Arm. ‘raven’ and ‘gull’). It is curious that Proto-Greek rearranged the old IE terms for ‘bird’ and ‘eagle’, cf. αἰετός ‘eagle’, which probably originates from IE **(a)u(e)i-*. Of course, one could also assume that the Greek situation is primary, i.e. **Hor-Vn-* meant ‘bird (in general)’, whereas **(a)u(e)i-* had the meaning ‘eagle’ in Proto-IE, but such a scenario has no advantages over the traditional one, on the contrary, it requires an assumption of a much greater number of independent semantic shifts in various IE dialects.

2. In their reply, F&B propose new IE etymologies for a significant number of Hurrian words and roots from the 50-wordlist (missing from their original monograph), which allow them to claim, thereby, that Hurrian has passed the lexicostatistical test and there are no formal arguments against the Hurrian-IE theory.

Let us now look at these newly proposed etymologies.

2.1. Hurrian ~ Proto-IE.

I have counted only *four* etymologies in F&B’s reply in which the Hurrian meaning coincides with the semantics of its IE correlate, reconstructed with the proper convincing force.

- 1) Hurr. *zur(=)gi* ‘blood’ ~ IE **esHr*, obl. **esHn-* ‘blood’ (its ancestral forms are the default words for ‘blood’ in Anatolian, Tokh., OInd., Arm., Balt. & Lat.; default words for ‘blood’ in Iran., Greek, Alb., Slav. & Germ. do not match each other, representing transparent innovations).
- 2) Hurr. *iš=* [abs., erg.], *šo= / šu=* [obl.] ‘I’ ~ IE **eǵ^(h)-* ‘I’.
- 3) Hurr. *fe=* ‘thou’ ~ IE **tu* [obl.] ‘thou’.
- 4) *šin(i)* ‘2’ ~ IE *d̥u-* ‘2’.

Two of them (pronouns ‘I’ & ‘thou’) are totally unconvincing, because, in order to defend their theory, F&B are forced to arbitrarily segment IE stems into monosyllabic and monovocalic *nuclei* plus “extensions”, i. e. IE **e-ǵ^(h)-* and **t-u*. In the first case IE **e* is compared with Hurr. *i* in *iš=* (the segmentation of the Hur-

rian stem is likewise *ad hoc*). In the second case IE **u* is compared with Hurr. *ulf*. It is empirically clear, however, that using such make-shift segmentations and monosyllabic comparisons, it is easy to provide a solid list of semantically perfect and phonetically trivial etymologies between any two languages among the world. In addition I must note that, as regards IE ‘thou’, its Proto-Indo-Hittite paradigm was **ti* [direct], **tu* [obl.], see Kassian 2009: 172 f. w. prev. lit., so *-u* cannot be the main meaningful element here.

The proposed etymological solution for the word for ‘blood’ is not any more apt. The reduction of a hypothetical Proto-Hurrian-IE initial vowel is unclear. The correspondence Hurr. *ts* (graphical *z*) ~ IE **s* is unique according to F&B themselves.

The same concerns ‘2’, specially for which a new non-trivial correspondence Hurr. *si* (graphical *šī*) ~ IE **d̥u* is postulated by F&B in their reply.

2.2. Hurrian ~ Ancient Greek.

In F&B’s reply there are four Hurr.-Grk. comparisons, where the Hurrian meaning matches the Grk. semantics.

- 1) Hurr. *eradi* ‘bird’ ~ Grk. (Hom.) ὄρνις, gen. ὄρνι-θος ‘bird’.
- 2) Hurr. *iš=* [abs., erg.], *šo= / šu=* [obl.] ‘I’ ~ Grk. ἐγώ ‘I’.
- 3) Hurr. *fe=* ‘thou’ ~ Grk. σύ ‘thou’.
- 4) *šin(i)* ‘2’ ~ Grk. δύο ‘2’.

No. 2–4 are unconvincing (see the previous section for the criticism). No. 1 could be hypothetically acceptable, but note that the Grk. meaning is innovative (see section 1 above for detail).

2.3. Hurrian ~ Old Indian.

F&B’s Hurr.-OInd. comparisons with coinciding meanings are the exact same four that have already been listed in the Hurr.-IE section above.

3. Summing up, the Hurrian-IE theory does not pass the “lexicostatistical test” again. So I must repeat here the conclusion from my first review:

Fournet and Bomhard have not managed to demonstrate the relationship between Hurrian and IE. I suppose that it should be clear to everybody who is familiar with Hurro-Urartian and IE languages that these two families are genetically unrelated.

4. Taking the opportunity, I would like to update my Hurrian 50-word list from Kassian 2010. The Hurr. word for ‘what?’ has been recently revealed by Wegner 2010: Hurr. *aw=* (abs. *au=nni*).

Literature

- Baxter & Manaster Ramer 2000 — W. BAXTER, A. MANASTER-RAMER. Beyond lumping and splitting: Probabilistic issues in historical linguistics // *Time Depth in Historical Linguistics*. McDonald Institute for Archaeological Research, Oxford Publishing Press. P. 167–188.
- Fournet & Bomhard 2010 — A. FOURNET & A. R. BOMHARD. *The Indo-European Elements in Hurrian*. La Garenne Colombes / Charleston, 2010.
- IEW — J. POKORNY. *Indogermanisches etymologisches Wörterbuch*. Bern, 1959.
- Kassian 2009 — A. KASSIAN. Anatolian lexical isolates and their external Nostratic cognates // *Аспекты компаративистики 4* [*Aspects of comparative linguistics 4*] (*Orientalia et Classica: Труды Института восточных культур и античности, вып. XXVIII*). М., 2009. С. 152–186.
- Kassian 2010 — A. KASSIAN. Review of: *The Indo-European Elements in Hurrian*, by Arnaud Fournet & Allan R. Bomhard (2010) // *Journal of Language Relationship*, No. 4 (2010). P. 199–206.
- Lewis 2009 — M. P. LEWIS (ed.). *Ethnologue: Languages of the World*, 16th edition. Dallas, Tex.: SIL International, 2009. Online version: <http://www.ethnologue.com/>.
- Wegner 2010 — I. WEGNER. Das hurritische Fragepronomen *aunni* (*a-ú-u(n)-ni*) = hethitisch *kuit* // J. C. FINCKE (ed.). *Festschrift für Gernot Wilhelm anlässlich seines 65. Geburtstages am 28 Januar 2010*. ISLET, 2010. P. 403–410.
- Дыбо А. 1996 — А. В. ДЫБО. *Семантическая реконструкция в алтайской этимологии. Соматические термины (плечевой пояс)*. М., 1996.