In Defense of the **Comparative Method**, or The **End** of the Vovin **Controversy**

In memory of Sergei Starostin

A. V. Dybo, G. S. Starostin
Moscow, Russian State University for the Humanities

Піскен асқа жеуші көп,
Біткен іске сыншы көп.

When the meal is ready, there'll be plenty of eaters.
When the work is finished, there'll be plenty of critics.

An old Kazakh proverb

It would not be an exaggeration to say that long-range comparison today does not enjoy a good deal of support from historical linguists all over the world. The reasons for this all too well known scepticism are numerous, with both the «long-rangers» and the «scepticists» to be blamed depending on the situation; yet the main reason could probably be defined as a fatal lack of understanding between the two sides, so that «long-rangers» are frequently accused of sins they never committed in the first place, whereas «scepticists» are often treated by «long rangers» as little more than annoying impedimenta for their work despite the fact that their scientific work, more often than not, is important for the purposes of long-range comparison.

Perhaps one of the factors explaining this unpleasant situation has been the relative reluctance of the Moscow school of comparative linguistics — arguably the world's most active proponent of long-range comparison — to engage in detailed theoretical and methodological discussions, preferring instead to concentrate on mostly practical work in such fields as Altaic, Nostratic, Sino-Caucasian, Austric, and others. While such discussions are not generally known to have much practical use, they at least help to lay down one's scientific principles in clear, accessible ways, reducing the degree of misunderstanding and allowing «outside» readers to feel more filled in on the current state of the art. Thus it is only natural that, from time to time, the basic guidelines for the work of the Moscow school be laid out in print, if only to remind the readers of why this work can — and should — be considered scientific, and its results (*some* of them, at least) valid.

A recent review by Alexander Vovin [Vovin 2005] of the newly issued «Etymological Dictionary of the Altaic Languages» [EDAL] by three members of the Moscow school (one of them among the authors of this article), in our view, presents a good opportunity for such a theoretical excursion. First, the review, written in a starkly negative tone (and somewhat boldly subtitled «The End Of The Altaic Controversy»), is not so much a criticism of concrete

Altaic etymologies (although there *is* plenty of concrete criticisms) as it is a general condemnation of the «methodology», presumably employed by the authors in order to arrive at their corpus of comparative data: «…the EDAL authors' methodology is often at odds with what is considered standard methodology in mainstream historical linguistics» [Vovin 2005: 73]. Furthermore, all of the problems with this «non-standard» methodology are carefully spelt out and enumerated, making the review a perfect encapsulation of all the numerous cases of criticism on the part of other opponents of long-range comparison — as well as adding new lines of discussion on its own.

Second, Alexander Vovin (from now on — AV for short) is by no means a stranger to macro-comparative linguistics. Until recently, he himself has been one of the strongest advocates of the Altaic theory, producing a string of works that occasionally offered valuable insights into specific issues in the field of comparative Altaic phonology, morphology, and lexics (e. g. [Vovin 1994], [Vovin 1997a], [Vovin 1999], [Vovin 2001a] and others). His general knowledge of the methodology and achievements of the Moscow school certainly surpasses that of many of its ardent Western critics. Thus, his former willingness to work within that particular paradigm of thought makes the reasons for his subsequent radical rejection of that paradigm particularly interesting, and his arguments well worth replying to in detail.

It is also worth noting that [Vovin 2005] is not AV's first venture into sharp polemics with macro-comparative efforts on the part of Russian scientists. A very similar line of arguing had already been undertaken by him in [Vovin 2002] against the Sino-Caucasian hypothesis (genetic relationship between North Caucasian, Yeniseian and Sino-Tibetan languages; below — SCH for short), scientifically formulated and advanced by the late S. A. Starostin, not coincidentally, also one of the authors of [EDAL][1]. And since that particular article has, for various reasons, until now also remained unanswered, we have deemed it useful to address both at the same time — because, as has already been indicated, AV's criticism goes deeper than simply pecking away at particular Altaic or Sino-Caucasian etymologies; he asks whether, in fact, Altaic or Sino-Caucasian can ever be proven to exist, let alone reconstructed, by means of procedures employed by the main proponents of their existence.

For AV the answer — in both cases — today lies starkly in the negative. The reasons are many and diverse, but most of them could easily be generalized into the following nine points:

[1] [Vovin 2002] is, in fact, just one part of a lengthier discussion, which began with [Vovin 1997b], a critical evaluation of several theories on the origins of the Chinese language, and was later continued by S. A. Starostin's response [Starostin 2002]. However, since most of the critical arguments against Sino-Caucasian, raised in [Vovin 1997b], were already answered in [Starostin 2002], as of now we will concentrate our primary attention on [Vovin 2002], centered — by the author's own admission — around methodological issues more than particular etymologies and correspondences.

1. AV disagrees with the selection of the *lexical* criterion as the major means of establishing genetic relationship, claiming that the best proof of such a relationship should be presented as a series of (preferably paradigmatic) correspondences within the *morphological* systems of the compared language branches. Without such correspondences, lexical evidence cannot serve as definitive proof.

2. Within the frames of lexical comparison, etymologies, in order to be convincing, have to be based on *rigorous phonetic correspondences*. AV claims that, while the authors of EDAL and SCH do admit this necessity, the correspondences proposed by them are, in fact, anything *but* rigorous, easily sacrificed every time they stand in the way of the authors' desperate need of yet one more «Altaic» or «Sino-Caucasian» etymology.

3. A further requirement is that the discrepancy in *semantics* between compared items should lie within reasonably tolerable borders. According to AV, many of the comparisons offered by EDAL/SCH authors are based on extremely unusual or downright impossible shifts of meaning, which suggests coincidental resemblance rather than genetic relationship.

4. Building up on the last two points, in order to be fully convincing, genetic relationship has first to be demonstrated on a *small, compact group of lexical/morphological items* that completely adhere to the formulated set of phonetic correspondences, before proceeding to more «shaky» territory. AV insists that the Moscow school variants of Altaic or Sino-Caucasian allow no such thing, piling up lots of irregular and questionable etymologies instead of setting up a small, «untouchable» core group.

5. AV states that before proceeding to any kind of external comparison for a language or language group, one is first required to perform a meticulous *internal* reconstruction. In the work of EDAL/SCH authors, external evidence always rules supreme, even if it completely goes against historical evidence implied from within the language itself. In addition, the authors overlook the importance of recognizing *borrowings*, frequently inventing etymologies for items that clearly have an outside provenance.

6. A further accusation is that the authors of EDAL/SCH display a peculiar contempt for all sorts of *philological* evidence, procuring their material from dictionaries rather than texts and frequently ignoring the historical, cultural, and ethnic backgrounds of the languages and language groups they are working with.

7. For Altaic at least, its Moscow version, in AV's opinion, reveals that its authors are mostly ignorant about all kinds of valuable research on its branches that have taken place in the last half century. This arrogant attitude towards important *literature* on the subject inevitably leads to an even larger number of factual errors.

8. Also in the case of Altaic (although occasional accusations of the sort are flung at Sino-Caucasian as well), the authors of EDAL are often *irreverent to the*

analyzed material, misquoting the forms and twisting the meanings to fit their subjective vision, occasionally bordering on elementary scientific dishonesty.

9. To conclude, most, if not all of these problems, should be attributed to a certain *religious zeal* on the part of the authors, whose irrational «faith» in macro-families urges them to a speedy «reconstruction» of «Altaic» and «Sino-Caucasian» at any possible cost. This, of course, has nothing to do with true scientific method in historical linguistics.

It goes without saying that all of these accusations have to be taken seriously; being backed up with multiple examples — and AV almost always accompanies them with what looks like supportive evidence — even several of them could be enough to discredit the Altaic and Sino-Caucasian theories in their current incarnation. In fact, AV is fairly confident that this, indeed, is the obvious result of his analysis for both cases. Cf.: «...meanwhile all I can offer in appraisal of his [Starostin's — *A. D., G. S*] 'Sino-Caucasian' hypothesis is an aristocratic funeral in Tibetan *'bum-pa*, an honorary structure which should be erected over this imaginative but futile attempt of the human mind» [Vovin 2002: 167]; «EDAL fails to prove genetic relationship of Tungusic, Mongolic, Turkic, Korean and Japonic, and it is, therefore, another etymological dictionary of a nonexisting language family...» [Vovin 2005: 122].

These and similar statements found in the above quoted works sound harsh indeed, and it is well nigh possible that inexperienced readers, guided by AV as he meticulously exposes all the intricacies of the pseudo-scientific «Sino-Caucasian» and «Altaic» hoaxes, will conclude that both of these theories have about as much value as the average amateurish hypothesis on language relationships of the «Japanese-Basque» or «Russian-Etruscan» variety, usually advanced by people who substitute a serious background in historical linguistics for raw enthusiasm and «inspiration» (which AV would probably call a «religious» attitude).

It may seem somewhat strange, though, that both of these theories are either advanced by or supported by people who not only *do* have a professional background in historical linguists, but also happen to be acknowledged specialists in that field. To err is, of course, human, and not even the most erudite of professionals is immune to occasionally making wrong conclusions; however, it is one thing to err over minor details, and quite a different one if the «error» means conjuring a global theory of potentially crucial scientific importance out of thin air. Such «errors» are definitely less common and require a special explanation.

In [Vovin 2002: 167], AV cautiously suggests that the reason may lie in a certain degree of megalomania on the part of the authors: «It closely reminds of... Hrozny, who after successfully deciphering Hittite, embarked on a long and futile journey to decipher almost all other unknown scripts. SS's «Sino-Caucasian» is, in my opinion, the similar futile enterprise» (we may safely

assume that the same explanation works for Altaic, too). This is a strong statement, perhaps not entirely suitable for an academic discussion, and, in any case, only justified once it has been proven *beyond the shadow of a doubt* that the scientific basis for «Sino-Caucasian» and «Altaic» as established by the Moscow school is, in fact, thoroughly *un*scientific and has nothing to do with the comparative method in its traditionally recognized form.

Has AV, in both of his works, really managed to do that? We assert that he has not. Below, one by one, we are going to show that out of the nine major accusations, listed above, some rest on *false methodological presumptions*; others, while superficially reasonable, are not actually applicable to either Sino-Caucasian or Altaic on a major scale, resting on *insufficient or incorrect evidence*; and still others have very little, if anything at all, to do with the question of proving these macrofamilies' existence, concentrating readers' attention on issues *irrelevant* to these matters.

Moreover, it is also our intention to show that if one were to uncritically accept all of AV's arguments, this would, in the end, make thoroughly impossible *any* research on distant language relationship — along with invalidating quite a lot of research already conducted by linguists on well-accepted language families. Despite his frequent appeals to the principles of the comparative method (e. g. [VOVIN 2002: 166]), AV almost as frequently makes statements that are hardly compatible with these principles.

Our initial plan was simply to go through [VOVIN 2002] and [VOVIN 2005] page by page, providing answers to all the questions; upon consideration, however, that structure was rejected since it resembles too closely the «E-mail citation» form chosen by S. STAROSTIN in his original reply [STAROSTIN 2002], which, according to AV [VOVIN 2002: 154], «allows one to interrupt your opponent's discourse at the convenient points to yourself, frequently distorting or abbreviating the critical points». To tell the truth, this passage strikes us as somewhat surprising, since, in our opinion, it is exactly the «E-mail citation» method that forces the author to answer his opponent's criticisms in a direct manner, where otherwise one could easily get away with the more «uncomfortable» issues by slighting or not noticing them[2].

Nevertheless, honouring our opponent's predilections, we have decided to refrain from the «E-mail style». Instead, we will have to reshuffle

---

[2] Case in point: in [VOVIN 1997b: 310], it is asserted that the Sino-Caucasian reconstruction «with its 150 or 180 consonants does not even remotely resemble a human language». This was, in «E-mail citation form», rightfully contested by S. A. STAROSTIN [STAROSTIN 1997b: 143], who, after checking upon his own reconstruction, was only able to pin down 35 consonant phonemes (80 if clusters are to be counted as phonemes, which is not normally the case in phonology). Had AV's subsequent reply followed the «E-mail form», we might already be aware of the true meaning of his original assertion. As it is, we can only guess.

AV's specific criticisms by assigning them to the nine accusation categories specified above, which, hopefully, will make the proceedings more structured and accessible to those readers who are not well versed in the history of the debate. We sincerely hope that none of the important points will be missed in the process, and that at least in that respect the current article will represent a substantial improvement on the «non-E-mail citation» style of [Vovin 2002]. Along the way, we also hope to address several additional issues, raised by other opponents of long-range comparison (most notably, S. Georg, noted for his thorough criticisms of the Altaic theory both before and after the publication of [EDAL]).

Unfortunately, due to the tragic demise of S. Starostin, his participation in the writing of this article (which we humbly dedicate to his memory) obviously became impossible. The authorship is, therefore, roughly divided between G. S. Starostin (Sino-Caucasian topics, general methodology issues, some of the Altaic problems) and A. V. Dybo (about half of the Altaic sections). We also thank O. A. Mudrak for his helpful advice on several important questions concerning details of Turkic, Tungus-Manchu, and Mongolic reconstructions.

I. *Basic Vocabulary Or Morphology*?

In both of his articles, AV places a particular emphasis on *morphology* as constituting the best basis for proving genetic relationship, and blames the Moscow school for neglecting it in favour of much less reliable lexical comparisons. Cf.:

«In my opinion, the *final proof* is based on morphology... One is only left to wonder why the rich morphology of 'North Caucasian', Yeniseian, and Sino-Tibetan is left out in these attempts to establish 'Sino-Caucasian' macrofamily... The only reason we do accept the etymologies with pandemic irregularity is because we have something else: regular correspondences in morphology as well as comparatively small, but sufficient corpus of lexical etymologies...» [Vovin 2002: 157–158].

«...Any proof of a genetic relationship must be very tight. The best way to make it very tight is to prove a suggested genetic relationship on the basis of *paradigmatic* morphology... and not on the basis of vocabulary, because morphology overall is more stable and more systematic than vocabulary which represents the most unstable part of any language» [Vovin 2005: 73].

The idea that morphology should play a crucial role in establishing language relationship does not, of course, belong to AV. This statement is fairly common among historical linguists (explicitly traceable back to at least A. Meillet), especially those working with morphologically rich language families. One major reason for this is historical: it is no big secret that the Indo-European family was recognized primarily on the basis of the amazing similar-

ity between the paradigmatic systems of Old Indian and classic European languages like Greek or Latin, and, since the general methodology of comparative linguistics grew out of working with Indo-European languages, morphological comparison, by the very force of tradition, is still held in high esteem and frequently suggested as a universal means for establishing relationship.

Another reason lies in the intuitive sphere. Morphology (and grammar in general) is traditionally seen as the «skeleton» of the language, its main constituent which, in comparison with lexics that «comes and goes», is relatively stable and thus far more valid for the first stage of comparison. Thus, if the languages compared do not seem to share much common morphology, but are nevertheless quite close lexically, for many linguists the obvious explanation will be that the languages are not related, but show traces of extensive contacts («convergence»).

From a purely synchronic, structuralist point of view such an understanding of morphology is quite reasonable. And it is most certainly true that regular paradigmatic correspondences in morphology are necessarily indicative of genetic relationship (with the possible exception of creole and pigin languages, whose genetic status is still debatable). But is the reverse also true — that genetically related languages absolutely *have* to share common morphology? And if not, which cases of genetic relationship would be expected to be «morphologically unprovable»?

The first question has already been answered in the negative by AV himself; in [Vovin 2002: 158], he quotes the case of Kadai — and one could certainly add quite an impressive number of isolating languages from the same linguistic region — where morphological comparison is impossible simply because the language happens to lack morphology as such. (Let us at once note that such a situation is, of course, unimaginable in the case of lexicon). This fact alone makes the morphological criterion useless as a *universally applicable* method of establishing relationship. However, this uselessness is by no means restricted to isolating languages.

An interesting test — which we refrain from performing in details since its results are all too predictable — would be to take, for instance, 100 of the most frequent grammatical morphemes reconstructed for Proto-Indo-European and see how well they have been preserved in such modern Indo-European languages as Hindi, French, or English. For the latter in particular, we are afraid, the results would be catastrophic (just a few minor traces, such as the *-s* marker for nominal plurals, *-t* in irregularly formed past participles, etc.). On the other hand, if we take the Swadesh list of 100 most basic lexical items, it seems to fare infinitely better: over 90% of the corresponding English items can be traced back to their Indo-European ancestors, and, what's more important, at least a good third of them can even be shown to have

possessed the exact same meaning in the protolanguage as they do in modern English, including body parts ('eye', 'ear', 'nose', 'tongue', 'foot', etc.), numerals ('one', 'two'), nature terms ('sun', 'star'), pronouns, verbs, etc. In this particular case at least, «morphology» seems to be a far shakier means to ascertain genetic relationship than basic lexicon.

The bottomline here is as follows: we know *for a fact* that a language's system of morphological markers *can* undergo an overwhelming collapse over a relatively short period of time, and frequently does. Thus, it hardly took more than a few hundred years for the elaborate nominal morphology of Classical Latin to be reduced to almost nothing. Chinese, over a period of one millennium, underwent a transformation from an essentially «Sino-Tibetan» type language to an «Austro-Thai» type language, even though genetically its ties certainly lie with the former. The *basic lexicon*, however, of both Chinese and Latin has had a much higher rate of survival [Starostin 2000a: 256]; and, although high-scale borrowings can occasionally speed up lexical replacement, to our knowledge, there is *not a single historically attested case* of the Swadesh 100-wordlist losing even a quarter of its constituents over a one thousand year period.

One could, of course, reasonably object that Indo-European morphology (as well as everything else) is not reconstructed on the basis of modern English or French, but on the basis of archaic languages like Old Indian, Greek, Latin, Gothic, etc., whose morphological systems are, on one hand, far more complex, on the other, much more closely resemble each other than those of modern languages. But it should be noted that, in the general context of historical linguistics, the situation where the researcher has access to a whole number of ancient languages is exceptional: besides Indo-European, the only such family happens to be Semitic. In most cases, however, we have no choice but to work with modern languages, since information on older stages of the respective languages will either be completely unavailable (as is the case with North Caucasian and Yeniseian), or will be limited to, at best, one or two «culturally important» languages with no close relatives of the same chronological depth (Chinese for Sino-Tibetan; most branches of Altaic).

In short, it is fallacious to insist that the reconstruction of Altaic or Sino-Caucasian should follow the «William Jones/Franz Bopp» routine. It is equally fallacious to insist that this routine is the only, or even necessarily the best, way to ascertain genetic relationship when we are dealing with macrofamilies whose ancestors are tentatively dated back to 4,000 BC or higher. A serious reconstruction of Proto-Indo-European morphology, particularly *paradigmatic* morphology, would have been unthinkable without access to language data that significantly decreases the chronological distance from attested languages to their reconstructed ancestor — which is

exactly the case for Altaic and Sino-Caucasian, where the overwhelming majority of information is gained from modern languages.

One must not overestimate the power of intermediate reconstructions when dealing with morphology, either. The case of Romance shows that loss of morphological information in daughter languages can be so drastic that, at best, we can hope to be able to reconstruct bits and pieces of the original system, but little else; there is no way we could arrive at the solid edifice of Classical Latin morphology based on the tentative reconstruction of Vulgar Latin. Even the traditionally accepted Indo-European morphological system, based on a meticulous analysis of Old Indian, Greek, Latin, Germanic, Balto-Slavic etc. data, faces a whole lot of still unsolved problems when confronted with the Anatolian system, such as the three-gender vs. two-gender opposition, the «mystery» of the Anatolian *ḫi*-conjugation, etc.; there is little doubt that, had Anatolian been separated from the rest of Indo-European a thousand years earlier than it was, the lexical criterion would have been the sole serious means of ascertaining its genetic status.

To our knowledge, no paradigmatic morphology has ever been reconstructed for language families older than Indo-European, no matter whether they are commonly accepted (like Afroasiatic) or debatable, like Altaic. What is occasionally dressed up as «paradigms» on close inspection turns out to be a subjectively arranged set of isolated and often questionable comparisons of select morphological markers that prove little, if anything. Such, for instance, is the system proposed for «Elamo-Dravidian» by David MCALPIN [MCALPIN 1981], whose credibility — or, to be more exact, importance in proving the validity of Elamo-Dravidian — is much undermined by lax semantics, paucity of material, and non-exclusive isoglosses (i. e. the proposed morphological markers are not really indicative of a specific Elamo-Dravidian relationship, being quite widespread in other language families of Eurasia as well); for a detailed critical analysis of the reconstructed nominal paradigm, see [G. STAROSTIN 2002: 148–150].

Not much better fares the particular version of Uralic-Eskimo comparison as presented in [SEEFLOTH 2000], much lauded by AV in [VOVIN 2002] as «the remarkable advance over the last two years... the theory demonstrating on the basis of morphology (and not basic vocabulary) that the closest relative of Uralic is not Indo-European as put forward by Nostraticists[3], but Eskimo...

---

[3] This statement is somewhat misleading, since there is actually no general consensus among «Nostraticists», whoever AV is grouping under that label, on the closest relatives of Indo-European. As far as we know, none of the authors of tentative Nostratic dictionaries (which would include V. M. ILLICH-SVITYCH, A. B. DOLGOPOLSKY, and A. BOMHARD) have specifically argued for an explicit «Indo-European-Uralic» subgrouping within Nostratic. As for lexicostatistical calculations, these normally indicate

the fact that Nostraticists failed to notice this fact with all their manpower, is significant». Slightly jumping the gun, AV seems to present the Uralic-Eskimo hypothesis as a proven «fact», even though, to the best of our knowledge, few of the representatives of «mainstream» linguistics, to which AV repeatedly pledges allegiance, exhibit the same hastiness. In reality, while much of the material adduced by U. Seefloth merits serious consideration, the actual situation is not much different from «Elamo-Dravidian»: a serious reliance on subjective internal reconstruction and various possible, but unprovable assumptions, along with the fact that many of the reconstructed «Uralic-Eskimo» grammemes are not at all exclusive for these particular language families: markers such as *-*m* for the 1st person, *-*t* for the 2nd person, and even *-*t* for plural are all well-known trademarks of Nostratic in general.

One might object that even so, the morphological system of Eskimo still shows an unmistakably «Nostratic» character, which makes AV's irony on behalf of «Nostraticists», ignoring solid morphological evidence in favour of unstable basic lexicon, justifiable. But herein lies the rub: the Eskimo-Nostratic connection, among the Moscow school at least, has actually been a commonplace for about *two decades* now, with O. Mudrak actively working on the data (the first serious results were published in [Mudrak 1984]), and, although no detailed and easily available publications on the issue had been made until now (a large number of Eskimo-Nostratic etymologies, however, are publicly available at http://starling.rinet.ru, the Moscow school's collection of etymological databases), the fact of this connection was explicitly mentioned many times, e. g.: «Several attempts have been made to relate some other linguistic families... to Nostratic. The only probable theory by now seems to be that of including Eskimo-Aleut in Nostratic (Oleg Mudrak)» [Starostin 1989: 43]. It is hard to believe that AV, who at that time had relatively strong ties with the Moscow school, is unaware of Mudrak's work on Eskimo.

Needless to say, the Eskimo-Nostratic connection as advanced by the Moscow school is primarily based on *lexical* evidence — and, as O. Mudrak's article in this volume demonstrates, there is plenty of the latter, which would make any supplementary morphological evidence a useful corroborating addition, but little else.

This, in turn, brings us to the second vital part of the issue. We have shown that it is unreasonable to expect to be able to reconstruct paradigmatic morphology when dealing with macrofamilies; it is even less pos-

---

that if Uralic does have tighter connections with one particular branch of Nostratic than with the rest, it is actually Altaic (41 % of potential matches on the 100-wordlist, according to [Starostin 2000b], as compared to just 26 % with Indo-European), which thus seems to corroborate the traditional Uralo-Altaic hypothesis. It is interesting to note that if O. Mudrak's theory on Eskimo-Altaic relations is correct (see [Mudrak 2008] in this volume), this may make us view Seefloth's comparisons in an entirely different light.

sible to come up with a formal statistical evaluation of various exclusive hypotheses of genetic relationship based on morphological evidence. But is it actually true that the situation is always different with the basic lexicon? In other words, how well-grounded is the claim that basic lexicon is more conservative over long periods of time than morphology?

AV, apparently, does not subscribe to that view: «It is very well known fact that the lexical comparison of English and Hindi will not establish Indo-European, not it will be sufficient ⟨sic! — *A. D., G. S.*⟩ for demonstrating that these languages are indeed related» [Vovin 2002: 158], with a reference to [Hock & Joseph 1996]. If this were indeed a «very well known fact», instead of a questionable hypothesis on the part of Hock and Joseph, this would certainly be bad news for macrocomparative linguistics, because, as it turns out, the chronological distance between English and Hindi (≈ 12,000 years, if the disintegration of Indo-European proper be set around 4,000 ℞) is not that much less than between Proto-Turkic, Proto-Mongolic etc., which means that we essentially *do* have to reconstruct Altaic as if it were an English-Hindi comparison. Fortunately for us, the transition from Proto-Altaic to its daughter languages was generally less «destructive» in the phonological aspect than the transition from Proto-Indo-European to English and Hindi, which makes the procedure of reconstruction easier and its results more credible.

Even more fortunately, the statement made by Hock & Joseph upon analysis turns out to be less than a questionable hypothesis: ever since the excellent reply by W. Baxter and A. Manaster Ramer [Baxter & Manaster Ramer 2000], we have to consider it a downright *wrong* hypothesis. The authors, having devised a simple, but effective probabilistic method of evaluating potential English-Hindi cognates, managed to demonstrate that English-Hindi relationship is quite establishable even without any additional knowledge. It is curious that Baxter & Manaster Ramer's article was not mentioned in [Vovin 2002], since it was published in an easily accessible source and was obviously of high relevance to the issue.

Of course, demonstrating English-Hindi relationship on the basis of morphological evidence (not to mention *paradigmatic* morphology) is an entirely different thing altogether. But with such obvious connections within the 100-wordlist as *name/nām, new/nayā, no(t)/na, two/dō,* and *tooth/ dẫt,* confirmed by phonetically similar cases outside the 100-wordlist, you can hardly go wrong, especially with the aid of a formal methodology. It also helps immensely that English and Hindi are, of course, not isolated languages, and a careful reconstruction of Proto-Germanic and Proto-Indic based exclusively on data from modern languages will inevitably result in us being able to track these back to Proto-Indo-European, establishing a large number of common lexical etymologies based on quite rigorous sound correspondences. On the other hand, reconstruction of Proto-Indo-European paradigmatic morphology, based exclusively on data from modern Germanic and Indic languages, is a futile task.

Let us now turn to other examples quoted by AV where, as he insists, genetic relationship is easily determined through morphological evidence but cannot be ascertained through the lexicon. One such case is Na-Dene: «Na-Dene represents even a more extreme example, where very few lexical cognates are found between Athabaskan and Tlingit, but the morphological correspondences are so impeccable, so that only a person unsophisticated in historical linguistics will try to disprove this relationship» [Vovin 2002: 158].

Without placing the «impeccable» morphological evidence under doubt, we might, however, question the exact meaning of the expression «very few lexical cognates»; unfortunately, AV does not provide any further details, and we are very much in the dark as to his exact views on the lexical comparison between Athabaskan and Tlingit — a pity, since, to the best of our knowledge, there is no general consensus on the issue. Thus, in Heinz-Jürgen Pinnow's groundbreaking comparative work on Na-Dene [Pinnow 1966] the number of potential cognates between these two branches exceeds several hundred (hardly «very few»), and the work itself is concentrated primarily on phonological and lexical comparison rather than morphology. Granted, many of Pinnow's comparisons have been sharply criticized by M. Krauss (e. g. in [Krauss 1973]), which does not, however, mean, that he himself has not offered an equally large, if not larger, share of Na-Dene lexical etymologies.

It is true that the first section in the first serious publication on Na-Dene as a historically valid entity [Sapir 1915] is subtitled 'Morphological features', with a detailed grammatical comparison between the branches. But it is equally true that it is immediately followed by a section entitled 'Comparative vocabulary', with about a hundred potential cognates between Tlingit and Athabaskan; in fact, it is highly dubious that a linguist of E. Sapir's erudition and experience could have embraced a distant relationship hypothesis based exclusively on grammar and with no supportive lexical evidence whatsoever.

Finally, let us simply consider the evidence: a selection of cognates on the 100-wordlist between Tlingit and Eyak (Eyak is not exactly Athabaskan, but very closely related)[4]:

Meaning	Tlingit	Eyak
cloud	-góos'	q'ahs
eat	-x̱aa	xaʔ
feather	t'àw	t'ahł
foot	x̱'òos	k'ahš
hair	-x̱awóo	χuʔ
know	-koo	gaʔ
lie	-tee	-ta
one	tléix'	ł̃hɢ
tongue	l'óot'	laʔt'

[4] Sources are [Krauss 1970] on Eyak, [Naish & Story 1963, 1973] on Tlingit.

These are only the most obvious parallels; altogether there are about 30% potential matches between these languages, and even if some of them eventually turn out to be wrong, the above list is already quite sufficient to suggest genetic relationship without the additional aid of any morphological evidence whatsoever (provided, of course, that it did not exist — normally, any additional evidence is always welcome).

AV may, of course, protest about the lack of rigorous phonetic correspondences (see below), but most of these items *do* follow correspondences (for lack of space, we have to refer the reader to the relevant works of Pinnow and Kraus), and besides, even without correspondences these comparisons easily meet and exceed the probabilistic requirements as set by Baxter and Manaster Ramer's algorithm. The only reassuring news for AV will be that it might, in fact, be more *convenient* to use morphology as the starting point for proving Tlingit-Eyak-Athabaskan relationship, if only to save oneself the hassle of rummaging through dictionaries; but no morphological evidence will suffice for proving relationship if it is not backed up with a serious collection of lexical etymologies. Fortunately, in the case of Na-Dene such a collection is not *that* hard to set up.

The second of AV's «proofs» for the supremacy of morphology is even more instructive. AV claims that, based on lexical criteria, it would be impossible to demonstrate that certain Formosan branches of Austronesian are related to its biggest branch — Malayo-Polynesian, i. e. that the Austronesian family as a whole could not be established. To prove his point, he provides a chart containing the names of basic body parts in five Austronesian languages, four of which belong to various Formosan branches. This chart is so interesting that it is worth being reproduced here in full.

	Squliq Atayal	Tsou	Puyuma	Bunun	Malay
nose	ŋuhuu A	ŋɨcɨ B	ungT-an C	ŋutus B	hidung D
eye	loziq A	mcoo B	maTa B	mataʔ B	mata B
ear	papak A	koeu B	Tariŋa C	taŋiʔa C	telinga C
tongue	hmaliʔ A	umo B	ridam C	maʔmaʔ D	lidah E
tooth	ʔnux A	hisi B	wali C	nipun D	gigi E
hand	qbaʔ A	mucu B	rima C	`imaʔ C	tangan D
foot	rapal A	caphɨ B	kui C	dalapa D	kaki E
bone	qniʔ A	cɨehɨ B	ukak C	tohnað D	tulang D
blood	ramuʔ A	hmueu B	daRah C	hairaŋ D	darah C
heart	kualun A	tʔuhu B	muRduRdu C	haputuŋ D	jantung E

AV: «By SS's standards, Squliq Atayal should not be related to the rest of Austronesian, because it shares no common basic body parts vocabulary items with other four branches» [Vovin 2002: 159]. He then remarks that the evidence

for Tsou, Puyuma, and Bunun is so scarce as well that it might easily be negligible; and concludes that if it were not for the fact that «all five languages... happen to share a significant number of morphological markers unique for Austronesian», recognition of Austronesian as a valid taxon would be dubious.

First of all, let us note that neither S. Starostin, nor, in fact, anybody who has ever worked with Swadesh wordlists, has ever stated that «languages *A* and *B* are not related if they do not share any common basic body parts vocabulary». Such a situation would, indeed, be *unusual*, but not entirely unimaginable; the issue here is not over specific parts of the Swadesh wordlist, but over percentages of matches between languages as a whole, regardless of whether the matches comprise body parts or any other part of the basic vocabulary. Thus, modern English and modern Irish, for instance, out of that entire list have but 'heart' and 'tongue' in common, which does not prevent us from ascertaining their being related, as the overall percentage of matches on the entire Swadesh list is well over 20%[5]. AV, presumably, did not present the entire wordlist for Squliq Atayal for reasons of space; but had he been allocated more of it on the pages of JCL, he would somehow have to explain at least the following matches between Squliq and the other Formosan languages:

	Squliq	Tsou	Puyuma	Bunun
ashes	qbu-liʔ	fuu	ʔabu	qabu
fire	pu-niq	puzu	apuy	sapuz
give	miq	mo-fi		
I	-kuʔ	aʔo	ku	sa-k
moon	bya-ciŋ		buḷaŋ	buan
stone	btu-nux	fatu		batu
two	ʔuša-	ruso	ḍua	dusa
we	s-ami	aʔmi	mimi	s-aam

All of these cases are unmistakable matches that, on an intuitive level, do not even require us to verify them through phonetic correspondences (although they *are* verifiable through phonetic correspondences).

Nevertheless, it must be stressed that any statement of genetic relationship based on lexicostatistical calculations boasts full scientific validity *if and only if* it takes into consideration as much available linguistic data as possible. This means that the best results will be attained if (a) calculations are performed on the material of entire language groups rather than isolated representatives; (b) at least a provisionary set of phonetic

---

[5] For even more evidence that lack of a significant number of matches between «body part» words cannot be judged as a strong anti-relationship argument per se, cf. various examples from Dravidian, Uralic, and Indo-European in [Manaster Ramer, Vovin & Sidwell 1998] (yes, AV *is* listed among the authors).

correspondences has been established for the compared units, based on an extensive study of the compared vocabularies *in their entirety*.

The following statement of AV clearly shows that he is unaware of these restrictions, even though they are usually taken for granted in the work of the Moscow school: «I can well understand that using glottochronology for proving genetic relationships is really compelling: all you have to do is just to compare 100 words taken from dictionaries. Easily done, and the results are overwhelming, especially for the folks from the 'Scientific American'» [VOVIN 2002: 164].

As a matter of fact, it is not easily done at all; in reality the procedure involves a massive etymological analysis of entire dictionaries, establishing correspondences, coming up with a set of intermediate reconstructions, then repeating the first two steps for the intermediate reconstructions — pretty much the standard procedure as prescribed by the classic comparative method — and *finally* doing the calculations while making sure to separate original cognates from loanwords[6]. Neither S. STAROSTIN nor any other members of the Moscow school have ever pretended to resort to anything else. Maybe it is because they were too afraid to underestimate the intellectual capacities of unnamed «folks from the Scientific American».

Now let us proceed to the most interesting part of the «Atayal argument»: analyzing AV's list of «body parts», which AV himself has, unfortunately (no doubt, due to space reasons), presented without any etymological comments, and see if the situation there is really as drastic from a purely lexical point of view as AV insists it is.

a) 'nose': Squliq *ŋuhuu* is presented as a mismatch with Tsou *ŋicɨ* and Bunun *ŋutus*, for reasons that are not altogether understood because in many cases, it can be easily demonstrated that Atayal *-h-* is the result of lenition from *-s-*: cf. *hapuy* 'fire' vs. Bunun *sapuz* id., *hiʔ* 'meat' vs. Kavalan *ʔəsiʔ* id., etc. (cf. also 'tongue' below). It also helps to consider such related forms as Amis *ŋoso* and Paiwan *ŋud́us* (or are these, too, mismatches?). And while this particular type of correspondences might not be as recurrent as we would like to, this may well have to do with insufficient analysis of available etymological data rather than anything else (unfortunately, a detailed etymological comparison of Formosan data is still unavailable to the general public). In any case, there is plenty of evidence to suggest that the resemblance goes far beyond chance.

---

[6] *Preliminary* glottochronology — calculations based on subjectively measured phonetic similarity rather than strict correspondences — is also possible, but only for the sakes of forming «first step» hypotheses about relationship, which then have to be corroborated by the procedure described above. Unfortunately, this preliminary step is occasionally confused in literature with scientific glottochronology, right down to suggesting that glottochronology is a (questionable) *alternative* to the comparative method, when in reality it is a useful *addition* to the comparative method.

b) 'ear': Squliq *papak* is indeed a mismatch. However, several other Atayal dialects have here an entirely different root: cf. Mayrinax *caŋiyaʔ*, Palŋawan *caŋeʔ*, and this, according to [Li 1981: 282], is the likeliest candidate for Proto-Atayal 'ear'. In his reconstruction, the root looks like **caŋiaʔ* and thus offers a perfect match with Bunun *taŋiʔa* and the rest. The situation is analogical to the one described by AV in [Vovin 1999], where he advocates the existence of Altaic by showing that a lexical comparison between Old Japanese and Old Turkic yields more matches than a lexical comparison between their modern equivalents — an argument to which we willingly subscribe ourselves. Thus, transition from just one modern Atayal dialect (Squliq) to Proto-Atayal works for us the same way as transition from modern Chinese and Chechen to, respectively, Old Chinese (and Proto-Sino-Tibetan) and Proto-North Caucasian: it makes lexicostatistics more reliable.

c) 'tongue': again, there is some sort of misunderstanding. First, the final *-liʔ* in Squliq is of suffixal origin (cf. Mayrinax *hmaʔ* id.); Li reconstructs **həmaʔ* for Proto-Atayal [Li 1981: 295]. Second, once again what we are dealing with here is the already mentioned lenition from **s-*: cf. Amis *sma*, Paiwan *səma* id., etc. Third, the same root is present in at least one dialect of Puyuma (Nanwang) as *smaʔ* [Dīng Bāngxīn 1978: 358]; obviously, this is an archaism as compared to the form quoted by AV (*ridam*, taken most probably from the Tamalakaw dialect as described in [Tsuchida 1981], although, inconveniently, AV does not quote his sources). Fourth, it is still way too premature to separate these forms from either the Bunun reduplication *maʔ-maʔ < maʔ* or Tsou *umo*, both of which may well reflect **hm- < *sm-* as well.

d) 'tooth': one more case where it helps to consider the history of Atayal before moving on to further comparisons. While Squliq does indeed have *ʔnux*, Mayrinax, in addition to the related *giʔnux*, also has *gipun*, which agrees with all of the dialects of the closely related Sediq Atayal: Toŋan, Toda *rupun*, Truwan, Inago *gupun* < Proto-Atayal **gipun* [Li 1981: 295]. This root is, of course, a good match with Bunun *nipun*[7].

e) 'foot': the mismatch between Squliq and Bunun is unbelievable, as the Squliq form goes back to Proto-Atayal **dapal* [Li 1981: 293], which is comparable with Puyuma *dapal* 'sole of the foot' and multiple other stems in different Formosan languages; out of these, Bunun *dalapa* is an obvious case of metathesis which should not prevent us from recognizing it as a solid cognate form. There are, moreover, serious reasons to believe that Tsou *caphɨ* is also related, but we will not go into details here.

---

[7] In addition, it should be pointed out that Tsou *hisi* and Puyuma *wali* also represent the same etymon; see [Tsuchida 1976: 147] for details (this extremely valuable source on Formosan languages has obviously not been consulted during the preparation of AV's table).

f) 'blood': Tamalakaw Puyuma does indeed have *daRah*, but most of the other dialects certainly do not, cf. Nanwang *damuk*, Chulu *ðamuk* etc. [DĪNG BĀNGXĪN 1978: 361]. All these forms are hardly separable from Squliq *ramuʔ* < Proto-Atayal **damuʔ* [LI 1981: 279].

Continuing this discussion will probably turn the article into a fullblown treatise on Austronesian reconstruction, but we hope that even those few points that have been raised are enough to show that (a) at least a few of AV's «cognation marks» have to be modified (and when we are dealing with a list of 10 words, even one such modification can make a huge difference); (b) in at least several more cases we have to postpone judgement before a more thorough investigation of phonetic correspondences between the various branches of Austronesian has been performed; (c) most importantly, that the analogy between this «Squliq-Tsou-Malay» lexicostatistics and Sino-Caucasian or Altaic lexicostatistics is inherently flawed in that the former compares isolated forms in modern dialects, whereas the latter compares reconstructed protolanguages — and, predictably, yields better results, as we have shown by retracing our steps from modern day Atayal and Puyuma dialects to Proto-Atayal and Proto-Puyuma. Since AV is apparently aware of this fact («…the number of cognates will increase significantly if older forms of languages in question are compared. The increase… becomes even more significant if we compare the reconstructions of related languages» [VOVIN 1999: 92]), the gist of his «Atayal argument» remains unclear[8].

It would seem that the only thing that remains to be done in this section is to simply repeat, perhaps a little more emphatically this time, the initial postulate: *if genetic relationship between two or more languages can be demonstrated on morphological evidence, it will inevitably show up in the basic lexicon as well*. No examples to the contrary are known, and the ones presented by AV turn out to be false upon thorough analysis. However, the opposite is *not* true: *if genetic relationship can be demonstrated on lexical evidence, it will not necessarily be detected within the compared languages' morphology as well*.

That said, however, the most suitable situation (and certainly the most common one for cases of non-distant relationship at least) is when genetic relationship can be demonstrated on both grammatical *and* lexical evidence. Yes, it would be most imprudent to expect a detailed reconstruction of Proto-Sino-Caucasian or Proto-Altaic nominal or verbal paradigms; but it would be equally disparaging to witness their daughter branches boast paradigmatic

---

[8] For the record, according to the lexicostatistical calculations performed by I. PEIROS with the use of STARLING software, Squliq Atayal has approximately 25% to 30% matches with every other Formosan language in AV's table. The figures with Malay are predictably much lower (around 10%?), but, as we have specified, the proper calculations should be performed between reconstructions, not modern languages.

systems that bear no resemblance whatsoever to each other and find no explanation in the protolanguage, as if they arose independently from nowhere. Fortunately, in either case the situation is nowhere near as bad as it may seem to the general reader upon considering certain passages in AV's reviews.

For Altaic, in particular, it is possible to reconstruct a whole series of grammatical morphemes that covers not only the derivational sphere but also all the major categories of nominal declension and verbal conjugation [EDAL: 173–229]. Granted, for AV all of these reconstructions are little more than «isolated morphological comparanda» [Vovin 2005: 74], which is a frightening, but essentially meaningless epithet, as these «comparanda» fit within the general scheme of proposed phonological correspondences, have comparable meanings and, most importantly, do not so much serve as proof of the Altaic theory — the major proof lies with the lexics — as manage to give a certain degree of insight into the problem of the genesis and development of the paradigmatic systems of daughter branches of Altaic.

In addition, AV does not see fit to discuss any particular examples of EDAL's morphological comparisons; instead, he concentrates the reader's attention on the suspicious fact that «derivational morphology has a lion's share of the chapter, while inflectional morphology takes less than nine pages. It is no secret that derivational morphology is borrowed much easier than inflectional… out of the eight pages and several lines allotted for inflectional morphology, more than three pages are occupied by numerals and pronouns, which, strictly speaking, belong to lexicon, and not to morphology, unless their paradigms are discussed. However, this is not done» [Vovin 2005: 74].

It may be suggested that if we are to use statistical arguments here, it would be more appropriate to measure the amount of reconstructed morphemes in absolute numbers rather than pages and lines. In these terms, a set of 33 derivational morphemes (many of them homonymous) is suggested for PA on p. 220, although 10 of these are later reprised on p. 226 in the «inflectional» section since they were probably used to form different verbal categories. The «inflectional» section proper includes 20 nominal and verbal morphemes, *not* counting pronouns and numerals. The final count is thus 20 inflectional vs. 23 derivational vs. 10 «border cases».

The reason of the textual disproportion mentioned by AV will be obvious to anyone who is acquainted with EDAL itself: derivative morphemes, being more tightly bound to the root than inflectional ones, are regularly accompanied by a large selection of examples, whereas for most of the inflectional morphemes it suffices to adduce them on their own. As for pronouns, strictly speaking, their paradigms *are* discussed, if only briefly, on p. 225, where the oblique stems (*mi-n-, *ma-n-; *si-n-, *su-n-) are given along with the direct ones (*bĭ, *si) for the 1st and 2nd person pronouns.

In short, while the morphological systems of Turkic, Mongolic, Tungusic, Korean, and Japanese cannot be shown to fully match paradigmatically — just like those of Russian and Hindi — there is still plenty of morphological evidence, both within EDAL and in a great other number of publications on Altaic (including, funny enough, some of AV's own works, like [VOVIN 2001a]).

Turning to Sino-Caucasian, AV's indignation may be understood when he proclaims that «one is only left to wonder why the rich morphology of 'North Caucasian', Yeniseian, and Sino-Tibetan is left out in these attempts to establish 'Sino-Caucasian' macrofamily. After all, 'North Caucasian' and Yeniseian are morphological world-record holders, and Sino-Tibetan... has very complex morphology, too» [VOVIN 2002: 157]. He, however, fails to mention that *complexity* of morphological systems has nothing whatsoever to do with their *origins*. No matter how many different grammatical morphemes and inflection types related languages can hold, if they can be shown to have arisen secondarily within the languages themselves, they are completely irrelevant to the issue of proving genetic relationship.

For instance, many Sino-Tibetan languages do have very complex morphology. The traditional — and probably correct — view on Proto-Sino-Tibetan morphology, however, is that «there was no relational morphology... but there was derivational morphology in the form of prefixes, suffixes, and voicing alternations of the initial consonants» [LAPOLLA 2003: 22 with literature]. In other words, whenever one observes a complex system of nominal declension or verbal conjugation in Sino-Tibetan, they are almost certainly of recent origin, having arisen from secondary grammaticalization of pronouns or postpositions. As for the derivational morphology, some of it is indeed reconstructible for PST, but, in AV's own words, «derivational morphology is borrowed much easier than inflectional», not to mention that, just as in Altaic, all of these morphemes are monophonemic, which, according to AV, «makes the possibility of chance resemblances rise dramatically high» [VOVIN 2005: 74].

Likewise, it is a well-known fact that the Yeniseian verbal conjugation system is one of the world's most complex; that complexity, however, refers to the sphere of *morphosyntax*, i. e. paradigm formation, rather than to the sheer inventory of grammatical morphemes, which, in comparison, is not very large. Moreover, it can easily be demonstrated that these morphemes mostly go back to the corresponding forms (direct and possessive stems) of personal pronouns [G. STAROSTIN 1995] — meaning that it will make little sense to directly compare Proto-Yeniseian verbal paradigms with their Caucasian equivalents, although the pronouns themselves are, of course, quite easily comparable (and form a substantial piece of evidence in favour of Sino-Caucasian).

And yet again, it is not altogether hopeless to look for morphological evidence when discussing Sino-Caucasian. It is hardly justifiable, either, to say

that morphology was «left out» of the comparison by S. A. Starostin. Already the first large collection of SC etymologies, presented in [Starostin 1982] and consisting of 153 parallels between North Caucasian and Yeniseian (and occasionally Sino-Tibetan), includes several grammatical comparisons, such as negative particles, plural suffixes, pronominal stems and one or two derivational suffixes. Still more parallels are included in his comparative dictionary of Sino-Caucasian, now available online at http://starling.rinet.ru. Finally, serious research on Sino-Caucasian morphology has recently been undertaken by John Bengtson (see his extensive article on the subject in this volume).

In the end, the situation closely mirrors the one with Altaic. It makes about as much sense to demand that Proto-Sino-Caucasian be proven on the basis of paradigmatic morphology as to demand that Proto-Indo-European be proven on the basis of historically attested documents written in Proto-Indo-European; but, just as historically attested documents written in ancient Indo-European languages can add to our understanding of Proto-Indo-European, so can the attested paradigmatic morphology of Sino-Caucasian languages add to our understanding of Proto-Sino-Caucasian, and we are working on it. Let us now turn to the second group of AV's arguments.

II. *Phonetic Correspondences: Rigorous Or Realistic?*

AV: «...phonetic correspondences maintained in a given proposal of a genetic relationship must be really recurrent... the majority of phonetic correspondences presented in the introduction [to EDAL — *A. D., G. S.*] have multiple variants or exceptions... proposed phonetic correspondences are frequently violated» [Vovin 2005: 76–77]. More or less the same assumptions are also made in [Vovin 2002] in the «Regularity vs. long-range comparison» section on pp. 161–163.

This second criticism of the Moscow school methodology, on the surface, looks more serious than the first. The question of whether relationship is best demonstrated through morphology or through lexics is exactly that — a *question* (albeit one that, we believe, has been fully answered above); but when it comes to regularity of phonetic correspondences — the pillar of comparative linguistics — there can be no second opinion on the issue: correspondences *must* be regular. That said, before proceeding to anything else, an understanding must be reached on what exactly constitutes regularity in correspondences.

First of all, no one will probably disagree that *regular* correspondences do not necessarily presuppose *simple*, or *one-to-one* correspondences. If two languages each have a set of, say, 50 identical phonemes, it would be unreasonable to say that they will be considered related if and only if each

of the 50 phonemes in language *A* corresponds to its exact correlate in language *B* and to nothing else. On the contrary, languages with rich phonological systems, unless they are very closely related, are usually expected to have relatively complex systems of correspondences; and the further they are related, the more complex they will become.

In the case of Sino-Caucasian, any proposed system of correspondences cannot help but be extremely complex, if only because North Caucasian languages are among the most phonologically complex in the world. No matter which family or families they are related to — and they probably *are* related to someone unless the original speakers were imported from Alpha Centauri — the proposed system will necessarily include numerous cases of many-to-one (or even many-to-many) correspondences, mostly of the merging type, of course, but with possibilities of reflex splitting as well.

Unfortunately, with increasing complexity of postulated correspondences comes the inevitable: a reluctance — or perhaps even an inability — on the part of the potential observer to discriminate between the *complex but regular* and the *irregular per se*. This seems to have been the case in [VOVIN 1997b], where AV, after having tried to identify phonological correspondences between Old Chinese and North Caucasian on the basis of S. STAROSTIN's proposed 13 matches in the SWADESH list [STAROSTIN 1995b], pessimistically concludes: «There is anything but regularity. Of course, one may argue that PNC has a much richer consonant inventory than OC, but this argument still will not explain how PNC **n* manages to correspond to three different OC phonemes without showing any traces of phonological conditioning» [VOVIN 1997b: 311].

The suggested three different OC phonemes are (a) *n*, (b) *nh*, and (c) auslaut *-t*, established on the following examples: (a) NC **čǎnʔV* — OC 新 **sin* 'new'; NC **nǎwši* — OC 二 **nij-s* 'two'; (b) NC **ʕwǎnʕV* — OC 耳 *nhəʔ* 'ear'; NC **świänĭ* — OC 年 **nhīn* 'year'; (c) NC **hwĕʔnV* — OC 血 **s-whī-t* 'blood'. However, it does not even take going beyond a more careful look at the original text of [STAROSTIN 1995b] to notice that (a) the opposition between OC *n* and *nh* is eliminated on the Proto-Sino-Tibetan level (Old Chinese voiced aspirates all have a secondary origin; in any case, this is not a question of Sino-Caucasian but rather Sino-Tibetan comparative phonology), with OC **nhəʔ* going back to PST **nǎ* and **nhīn* going back to **s-nī-ŋ*; and (b) that the correspondence «PNC **n* — OC **t*» simply *does not exist*, because OC *-t* in **swhīt* is actually a suffix; the original PST form looks like **s-ʔʷīj* (cf. Burmese *swijh* 'blood').

The *real* correspondences for NC dental nasals that can, and should, be established on the proposed 13 matches are: (a) «PNC **n*, **nʔ*, **nʕ* — PST **n*»; (b) «PNC **ʔn*, **nh* — PST **j*». Not only are the matches thus perfectly conditioned (depending on the quality and position of the original laryngeal in the cluster), they are even recurrent (within the 13 matches,

PNC plain *n corresponds to PST *n twice), although, of course, in order to demonstrate the regularity in a fuller way, one has to look at additional data. But the situation is actually even more amusing; let us look closer at the list of consonantal correspondences that AV gets between OC and NC:

Old Chinese	Proto-North-Caucasian
*n	*j, *nʔ, *n
*nh	*nʔ, *n
*-t	*n, *ç, *c, *c̦
*j	*nh, *ʁ, *w
*l	*l, *ɫ̱
*ŋ, *m	*m
*s	*h, *ç, *ṣ̌, *ś

AV calls this «anything but regularity». Now that we got the nasal issue out of the way, is there anything *else* that can be called «irregular» on this list? The only other case where one NC consonant corresponds to two OC phonemes is PNC *ç vs. OC *s and *-t, and it seems to have perfect conditioning (*-t in the final position, *s elsewhere). «Anything but simplicity» would be understandable, as we observe the predictable merger of multiple NC consonants into much fewer OC ones. But it is hardly permissible for a professionally trained comparativist to confuse «complexity» with «irregularity».

This one and many more similar passages in [Vovin 1997b] have already been analyzed in detail in [Starostin 2002], so we need not concentrate on them second time around. It is worth noting, though, that in AV's reply most of S. Starostin's explanations of why the proposed correspondences are actually regular remained without an answer — which leads us to believe that either AV has silently accepted them (a dubious idea, since in that case at least an apology would be in order) or, more probably, that he decided to spare the general public the boredom of yet another round of complex arguing around tricky rows of laryngeal clusters, concentrating instead on statements like «I should thank my esteemed opponent for making it much more detailed than I intended and ultimately — even less credible» [Vovin 2002: 161] that are, of course, eminently readable but do not seem to serve any practical purpose.

One particular moment has, however, managed to attract AV's attention, and, in our opinion, deserves a detailed answer since it illustrates fairly well everything that is inherently wrong about his approach to criticism. Drawing the reader's attention to a subset of SC correspondences presented in [Starostin 1996] (a «highly impressive masterpiece», in AV's own words), AV singles out the following pair:

(a) PNC *n : PST *n/m
(b) PNC *nʔ, *nʔ, *ʔn : PST *n/m

In [Starostin 2002] these correspondences, accompanied by ex-amples, are presented once again, this time with an explanatory note:

> PNC *n : PST *n/m (*m* usually after labial)
> PNC *nʔ, *nˀ, *ʔn : PST *n/m (*m* after labial)

Obviously, this is a case of not very good phrasing on Starostin's part, because it remains unclear what is meant under «labial» — vowel or consonant (or both), Sino-Tibetan labial or Sino-Caucasian labial? Howe-ver, bad phrasing and false assumption are two crimes of a very different nature, and in order to ascertain which one Starostin should be accused of, one has to pay attention to supporting data, which is as follows:

(a)   PNC *ʔwēnɨ (~u) 'sound, movement of air' : OC 音 *ʔəm 'sound'
         PNC *mhănV 'warm; weak, loose' : PST *nəm 'soft, weak, fluffy'
         PNC *fãnV 'mountain' : PST *ŋăm 'height, precipitous' (/ *ŋān)

(b)   PNC *dwǒnʔV 'musical instrument, drum' : PST *tŭm 'instrument'
         PNC *ṭwōnʔe 'manger, feeding-trough' : PST *tōm (~ ua-) 'jar, bottle'
         PNC *GwinʔV 'house, farmstead' : PST *Q[i]m 'house'

In four out of six examples, the presented PST (in one case, OC) forms do not contain any labials in the first part of the syllable. This makes it fairly obvious that the specification «after labial» was referring to Sino-Caucasian as a whole (with North Caucasian being the more archaic branch in this case, as in most others) and *not* to Sino-Tibetan. Indeed, in each and every one of these examples the reconstructed North Caucasian form begins either with a labial consonant (*m-*, *f-*) or with a cluster contai-ning labialisation (*ʔw-*, *dw-*, *ṭw-*, *Gw-*). What we have in Sino-Tibetan is a regular transfer of the «+labialisation» feature to the end of the syllable.

AV, however, does not notice that, and prefers to interpret «after labial» as a statement necessarily referring to Sino-Tibetan, which empowers him, af-ter quoting the first three examples, to ask the rhetoric question: «Now, there is something wrong either with the definition of a labial, or with the stated rule; can we find any labials preceding *-m* in PST?» [Vovin 2002: 162]. The undiscriminating reader, who, of course, is under no obligation to conduct a detailed analysis of the complicated data himself, is thus left convinced that either S. Starostin is a patented ignoramus who hardly even knows what a labial phoneme is, or, worse, a high-class charlatan intentionally hiding the inconsistencies and irregularities of his fictional «system» behind a wall of complex symbols and their combinations. In reality, the worst accusation one could fling in his face based on this particular slice of data is that the corre-spondences have not been stated with perfect clarity. Which *is* a just accusa-tion — after all, when the systems are that complex, a crystal clear presenta-tion should be in order — but, still, refers to style rather than substance.

Worse still, AV finishes his discussion of the issue with the following passage: «...let us rewrite them [*correspondences between NC and ST nasal consonants — A. D., G. S.*] in such a way that every single PST nasal sonorant will have its all correspondences in PNC (excluding clusters for simplicity):

> PST *m : PNC *m, *n
> PST *n : PNC *m, *n, *ŋ
> PST *ŋ : PNC *m, *n

I think the reader can see for her/himself how 'regular' are SS's correspondences... apparently, regularity can be sacrificed to save the 'Sino-Caucasian' hypothesis» [Vovin 2002: 163].

Granted, the list does look somewhat uncomfortable. But what is probably the most uncomfortable thing about it is that it is *false*. For starters, provided clusters really *are* excluded, the correspondence «PST *n* : PNC *m*» does not exist. (Even if we *do* include clusters, the entire text of [Starostin 1996] still includes only one such example — PNC *çwämʔV̌* 'eagle' : Proto-Lolo-Burmese *ʒwan* 'hawk, kite', which may simply be an incorrect etymology).

Second, the correspondence «PST *n* : PNC *ŋ*» does not exist either. In fact, it could never exist even in theory, for the simple reason that a velar nasal phoneme *ŋ is not present in PNC at all* (!). It can certainly be understood that, given the complexity of the general picture, AV simply lost his footing at one point or another. However, it is normally expected that straightforward accusations of incompetence be held to a pretty high standard on the part of the prosecutor, which is not the case. Considering the unflinching regularity with which AV blames «Nostraticists» for not paying attention to factual data and even falsifying it, situations like these produce a definitely odd impression, to say the least.

Third, let us now rewrite AV's scheme in a way which includes S. Starostin's general interpretation of the picture in terms of PSC:

Sino-Caucasian	North Caucasian	Sino-Tibetan
*m	*m	*m
*n	*n	*n / *m
*ŋ	*m / *n	*ŋ

All the horrible «irregularities» have been reduced to one explained case of a split reflex in Sino-Tibetan (*-n > *-m under the influence of preceding labials) and one unexplained case of a split reflex in North Caucasian (judging by the examples, this also has to do with the presence of labials in the anlaut syllable, although there are exceptions). The picture is thus nowhere near as frightening as it may seem — although, of course, given a specific style of presentation (and throwing in a couple fictitious correspondences for good measure that no one will notice anyhow), it can easily be *made* to look frightening.

AV's manner of «rewriting» correspondences continues fluently in his review of EDAL. Here, particular attention is paid to the system of Proto-

Altaic vocalism, the reconstruction of which, according to AV, firmly proves that «EDAL's 'Altaic' is a construct of human mind» [VOVIN 2005: 117]. It is perhaps no coincidence that most of AV's arrows are aimed at vocalism and not consonantism, because that particular part of Proto-Altaic segmental inventory has always been notoriously harder to reconstruct than consonantal phonemes. Nevertheless, once again, it seems to us that AV is underestimating the progress made in that sphere.

Thus, on p. 117, while discussing reflexes of PA *-i- and *-i̯o- in the first syllable, he falls back on the mistake of confusing «regularity» with «complexity». The following reflexes in previously discussed etymologies are noted:

PA	PT	PM	PMT	PJ	PK
*i̋li 'work, craft'	*i	*üyi	*i	*i	*i
*khìĺa 'fetters'	*i	*e	*i	*a	*a
*sīĺa 'sharp stick'	*i/*ɨ	*i	*i	*a	*a
*si̯őĺe 'to mock'	*iā	*i	*u	*ə	*ā
*ti̯őĺì 'stone'	*iā	*i	*o	*i	*o
*zi̯oĺa 'to shine'	*ạ	*o	*u	*a	*ă

This is followed by saying: «One can clearly see that in the first group only PMT has regular reflexes... in the second group, none of the languages have regular reflexes». What is probably meant is that none of the languages have *the same* reflexes, i. e. the same protophoneme can have different reflexes in the same language. But merely having different reflexes in the same language — normally called «conditioned splitting» in comparative linguistics — should not be lumped together with having *irregular* reflexes. Let us see what the tables of phonetic correspondences, presented in EDAL on pp. 92–93, tell us about these reflexes:

PA	PT	PM	PMT	PJ	PK
*i-i	i	i	i	i	I
*i-a	ɨ [i]	i	i	a	A
*i̯o-e	ẹ, a	e, ö	U	ə [u]	U [jə]
*i̯o-i	ia, ja	i [e, ö]	U	i	U [ă]
*i̯o-a	ia, ja	a, U	U	a	U [ă]

Based on these correspondences, the following reflexes can really be ascertained as «irregularities»: (a) for *i̋li — PM *üyi; (b) for *khìĺa — PM *e; (c) for *sīĺa — none at all; (d) for *si̯őĺe — PT *iā, PM *i, PK *ā; (e) for *ti̯őĺì — none at all; (f) for *zi̯oĺa — PT *ạ. This constitutes 6 irregular reflexes out of a potential 30.

Furthermore, the chart on pp. 92–93 only lists principal reflexes; some of the observed «irregularities» are quite explicitly mentioned in the more detailed comments section, cf.:

(for PA *CiCa): «Mongolian normally has *i, but a variation i/e before the following -e-» [EDAL: 106] (concerning PM *kelbe- < *khìĺa);

(for PA *Ci̯oCe): «Korean may have a labialized *o/u or a diphthong *jə (/*(j)a)» [EDAL: 126] (concerning PK *hār- < *si̯őĺe);

(for PA *$C_i̯oCa$*): «although Turkic has one *-ia-* case here... the normal reflex appears to be non-diphthongized *a* (or *ą*)» [EDAL: 126] (concerning PT *jạĺ(č)- < *zi̯oĺa*).

The bottomline is as follows: (a) *an absolute majority* of the vocalic correspondences are perfectly regular in that they conform to the rules set out in the phonetic tables in the introduction; (b) most of the irregular cases are at the least documented; (c) the only root that presents systematic problems with vocalism is PA *si̯őĺe* 'to mock', which may indeed show that that particular etymology needs reworking.

AV next takes issue with the very fact that multiple correspondences are being set up for first syllable vowels depending on the vocalism of the second syllable: apparently, not only does that provide EDAL authors with great opportunities to multiply reflexes whenever they deem it necessary, but they do not even find it worth their while to list all of these reflexes in their tables and charts. He writes: «Let us take all the vowel sequences that involve reconstructed PA *i̯o* in the first syllable and let us see how this PA *i̯o* is reflected in daughter languages:

sequences with PA *i̯o*	PMT	PM	PT	PJ	PK
$C_i̯oCa$	*u, *o	*a, *o, *u	*a, *iā, *ạ, *(i)a	*a	*a, *o, *u
$C_i̯oCi$	*ō, *o, *u	*i, *ö	*a, *iā, *(i)a	*i	*o, *ō, *u, *a, *ă
$C_i̯oCe$	*u, *ū, *o, *ō	*ü, *ö, *i, *e	*a, *ạ, *ā, *(i)a, *(i)ā, *ẹ, *e	*u, *ə, *ua	*(j)a, *o, *∅, *ui, *ə, *jə, *jə̄
$C_i̯oCo$	*i, *ī	*o, *u, *ö, *ü	*o, *ō, *u, *ü	*ə, *a	*i, *jə
$C_i̯oCu$	*ia, *i, *(i)a	*i, *e, *u, *o	*o, *u	*u, ua	*jə, *jə̄, *ə̄, *o, *ă, *ăi

I trust that the reader can see for him/herself how 'regular' are the vocalic correspondences» [Vovin 2005: 118].

Sadly, it is indeed possible that the reader, terrified of this seemingly chaotic «rewrite» of AV's, may come to the premature conclusion that the concept of regularity of correspondences for EDAL leaves a lot to be desired. The truly discriminating reader, however, might want to actually doublecheck the data himself, and this is what he may discover:

(1) the long reflexes in Tungusic and Turkic have been given by AV mostly for additional intimidation reasons, as the feature of length in these languages is essentially treated as prosodic in EDAL and has little connection to vowel quality. Thus, an entry like «*u, *ū, *o, *ō*» should really be read just «*u, *o*»;

(2) in Tungusic, the first three sequences do yield either *u* or *o*, with unclear distribution; within the sequence *$C_i̯oCu$*, however, the rule states «normally *ia*, but *i* for short *i̯o* after sibilants» [EDAL: 129], and the distribution works on all the given material;

(3a) in Mongolic the standard reflex for *$C_i̯oCa$* is *a* (11 examples); 3 examples of *o* and 1 example of *u* are not quite understood;

(3b) PA *C̣i̯oCi* > PM *-ö-* after labials (*Altaic* labials!) and *-i-* otherwise (8 completely regular cases);

(4a) in Turkic, the notation *(i)a* signifies that either *ia* or *a* is reconstructible here in the protoform; therefore, whenever the corresponding cell simply has *ia* or *a* per se, the additional *(i)a* has to be omitted; furthermore, for the sequences *C̣i̯oCa* and *C̣i̯oCi* PT *a* and *ia* are in complementary distribution (the latter only in contact with palatal consonants);

(4b) the PT reflex *-ü-* for PA *C̣i̯oCo* is non-existent;

(5) the Japanese reflex *-ua-* for PA *C̣i̯oCe* and *C̣i̯oCu* is non-existent (in but two examples — PJ **kua* < PA **ḱi̯ŏbe*; PJ **suà-rá* < PA **si̯ŏ̀gu* — it is actually the result of contraction of two vowels caused by elision of the inlaut consonant);

(6) the only branch where variation here is really peaking is Korean, whose vocalism is known worse than anything else. Yet here, too, some of the reflexes are quite common, while others quoted by AV are limited to one or two exceptions and could be dialectal or conditioned by specific consonantal contexts that are too rare to be firmly established.

With all that in mind, let us now «rewrite» AV's table in a more civilized manner, marking complementary distribution with a slash whenever possible and leaving unexplained variations separated with commas. In addition, let us accentuate the common (statistically relevant) reflexes in boldface:

sequences with PA **i̯o*	PMT	PM	PT	PJ	PK
C̣i̯oCa	***u**, ***o**	**a*, **o*, **u*	***a** / ***ia**, **ạ*	***a**	***a**, ***o**, **u*
C̣i̯oCi	***u**, ***o**	***i** / ***ö**	***a** / ***ia**	***i**	***o**, ***ă**, **u*, **a*
C̣i̯oCe	***u**, ***o**	***ö**, ***i**, ***e**, **ü*	***a**, ***e**, **ạ*, **ẹ*	***u**, ***ə**	**(j)a*, **o*, **ui*, **ə*, **jə*
C̣i̯oCo	***i**	***o**, ***ö**, **u*, **ü*	***o**, ***u**	***ə**, ***a**	**i*, **jə*
C̣i̯oCu	***ia** / ***i**	***i**, ***e**, ***u**, ***o**	***o**, ***u**	***u**	***ă**, **(j)ə*, **o*, **ăi*

It can be seen that taking into account the irrelevance of vowel length, various complementary distributions, and the statistical factor significantly reduces the chaos that AV is so eager to emphasize.

Of course, even with this «reduction» it is also still evident that a large amount of seemingly irregular fluctuation is in order, with Korean and Mongolic particularly out of focus. We have shown that AV's presentation of the «irregularities» is seriously exaggerated, but the reader is still entitled to the question: «Does it really matter if a postulated proto-phoneme have *five* irregular reflexes in daughter languages or 'merely' *two*?»

To answer that question, and further clarify the concept of «irregularity», another analogy is in order. Let us consider the following chart that maps vocalic correspondences between two well-known and quite obviously related (even on an intuitive level) European languages:

German	English: <u>UK</u> ([EPD₁₁] &al.)/<u>US</u> ([PDAE]; [RHCUD], [WEUD] &c.)/both ([LPD], [EPD₁₅] &al.)
[a]	**æ** (hacken· *hack*, hatte· *had*), **æ:** (satt· *sad*); **eɪ** (machen· *make*, acht· *eight*, halb· *half⸗*pence, *half-*penny), **aɪ** (Nacht· *night*), **e** (Waffe· *weapon*, zwanzig· *twenty*, manch· *many*), **ɒ** (lang· *long*, was· *what*, ab· *of, off*), **ɔ:** (Salz· *salt*, Halm· *haulm*/US_ONLY*halm*, ge⸗ ⸗bracht· *brought*, Wasser· *water*, ab· *off*; warm· *warm*), **ɵʊ** (kalt· *cold*, Kamm· *comb*), **ɑ:/ æ:** (halb· *half*, Lache· *laugh*), **ɑ:** (Halm· UK_ONLY*halm*; Harm· *harm*), **u:** (Gans· *goose*, Wam⸗ me· *womb*), **ʌ** (ander· *other*), **ɜ:** (Farn· *fern*), **i:** (walten· *wield*); **ɛə** (starren· *stare*), **ʊ** (Gans· *goose⸗*berry); morphologically problematic: **ɪ** (-schaf⸗*t⸗*-*ship*); **ʊ** (standte·⸗*stood*); early loans: **ɑ:** (⁺Almose· ⁺*alms*)
[ɑ:]	**eɪ** (Name· *name*, bat· *bade*, aß· *ate*), **i:** (schlafen· *sleep*); **ɵʊ** (Nase· *nose*), **u:** (Span· *spoon*, Zahn·⁽⁼⁾*tooth*), **ɪə** (Bart· *beard*, Jahr· *year*), **ɜ:** (waren· *were*, Jahr· UK_ONLY*year*), **ɛə** (fahren· *fare*, da· *there*, Haar·⁽⁼⁾*hair*, waren· *were*, Jahr· US_ONLY*year*), **ɔə** ([*der*] Jahre· *yore*); **ɑ:/æ:** (Pfad· *path*); **ɑ:** (Vater· *father*; Arsch· *arse*); **e** (aß· *ate*, sagte· *said*, Draht· *thread*, Schaf· *shep⸗*herd); **æ** (haben· *have*, bat· *bade*); **ɔ:** (tragen· *draw*, schmal· *small*, Ha⸗ bich⸗*t· *hawk*; Warze· *wart*), **ɒ** (Schwan· *swan*); morphologically problematic: **ɪ** (tat·⸗ *did*), **ʌ** (ge⸗tan·⸗*done*)
[ɛ]	**e** (setzen· *set*, selb· *self*); **ɜ:** (ernst· *earnest* 'serius', Werk· *work*), **ɜ:/ɜ:ɹ** (Welt· *world*), **ɪ** (sechs· *six*, sengen· *singe*), **i:** (essen· *eat*, Gänse· *geese*, Feld· *field*), **ɑ:/æ:** (letzt· *last*, Es⸗ pe· *asp*), **ɑ:** (Herz· *heart*), **æ** (Decke· *thatch*, Espe· *asp*), **ɒ** (Wespe· *wasp*, Erle· *alder*), **eɪ** (brechen· *break*, weg· *away*), **ɔ:** (Erle· *alder*; Zwerg· *dwarf*), **aɪ** (Recht· *right*); unac⸗ cented (!) **ɪ** (elf· *eleven*;)
[e:]	**i:** (eben· *even*), **ɵʊ** (Schnee· *snow*, Zehe· *tow*); **ɑ:** (Herd· *hearth*), **ɜ:** (Erde· *earth*), **eɪ** (Weg· *way*, ¿zween·⸗*twain*?), **e** (Feder· *feather*, zehn· *ten*), **ɪ** (leben· *live*), **ɔə** (Schwert· *sword*, mehr· *more*), **ɪə** (Meer· *mere* 'lacus'), **ɛə** (zehren· *tear* 'scindo'); morphologically problematic: **u:** (wer·⸗*who*); **aɪ** (jeder· *either*); **i:** (jeder·⸗*either*, ¿zween·⸗be⸗*tween*?)
[ɛ:]	**i:** (⁺Käse· ⁺*cheese*); **ɛə** (Bär· *bear* 'ursus'), **ɔ:** (jähnen· *yawn*), **ɵʊ** (sähen· *sow* 'sero', blüh⸗ en· *blow* 'flo'), **ɪə** (Ähre· *ear* 'spica'), **eɪ** (wägen· *weigh*)
[ɪ] = **[i]**	**ɪ** (sitzen· *sit*), **i:** (Schild· *shield*), **aɪ** (finden· *find*, Ge⸗sicht· *sight*), **ɒ** (nicht· *not*), **ɔ:** (nichts· *naught*, vierzig· *forty*), **ɔə** (vierzehn· *fourteen*, viert⸗*el· *fourth*), **ɜ:** (Birke· *birch*, dritte, dritt⸗*el· *third*), **ɑ:** (Hirsch· *hart*); **ɵʊ** (will· *wo-*n't), **ʌ** (will· US_ONLY*wo-*n't, dialectal *wo-*nna), **ʊ** (will· US_ONLY*wo-*n't), **eɪ** (Ge⸗wicht· *weight*), **e** (Himmel· *heaven*); **æ** (b⸗in·⸗*am*); historical curious: **ʧ** (ich· US_ONLY dialectal *ch*⸗am ['ʧæm] 'I am'; formerly a consonantal correspondence)
[i:]	**i:** (Sieb· *sieve*, Knie· *knee*); **aɪ** (fliegen· *fly*, Zwielicht· *twilight*), **ɔə** (vier· *four*), **ɔ:** (vier· *four⸗*pence, *four-*penny), **e** (sieben· *seven*, hielt· *held*), **u:** (kiesen· *choose*), **ɪə** (Bier· *beer*, hier· *here*), **jɜ:** (hier· *here*), **jɪə** (hier· US_ONLY*here*), **jɛə** (hier· US_ONLY*here*), **ɛə** (hier· US_ONLY*here*); **ji:** (ihr· *ye*); **ɪ** (ihm·⸗h⸗*im*); early loans: **eɪ** (⁽⁺⁾Giebel· ⁺*gable*)
[ɔ]	**ɒ** (Frost· *frost*); **ʌ** (Sonne· *sun*, kommen· *come*), **ʊ** (voll· *full*, wollte· *would*), **i:** (Woche· *week*), **ɵʊ** (Gold· *gold*, offen· *open* adj.), **ɜ:** (Wort· *word*), **ɔ:** (Tochter· *daughter*; Hornisse· *hornet*), **ɑ:** (horchen· *heark*-en), **æ** (soll·⸗*shall*), **ɑ:/æ:** (soll·⸗*sha*-n't); lexical uncertain⸗ ity: **ɪ** (ob·°*if*); early loans: **aɪ** (⁺Roggen· *rye*); complex of issues: **ɜ:** (⁽⁺⁾Sporn·⁽⁻⁾⁽⁺⁾*spur*)
[o:]	**ɵʊ** (Kohle· *coal*), **i:** (Floh· *flea*, Not· *need*); **ʌ** (Monat· *month*, Montag· *monday*, Honig· *honey*, Sohn· *son*, Ofen· *oven*), **ɪə** (Ohr· *ear* 'auris'), **jɜ:** (Ohr· UK_ONLY*ear* 'auris'), **ɛə** (wo· *where*, Ohr· UK_ONLY*ear* 'auris'), **e** (Brot· *bread*, wohl· *well*), **ɒ** (Boden· *bottom*, Trog· *trough*; vor· *fore⸗*head), **ɔ:** (Moos· *moss*, Stroh· *straw*, Trog· *trough*), **aʊ** (Vogel· *fowl*, Trog· *trough*), **aɪ** (hoch, hohe· *high*), **eɪ** (groß· *great*), **ɔə** (vor· be⸗*fore*, *fore-*head), **ʊə** (Moor· ⁽⁺⁾*moor*), Moor· US_ONLY*⁽⁺⁾moor* as *Moore*), **oə** ⟨sic!⟩ (Moor· UK_ONLY⁽⁺⁾*moor*); **u:** (Mon⸗*d· *moon*); complex of issues: **ɜ:** (⁽⁺⁾Sporen·⁽⁼⁾⁽⁺⁾*spurs*)
[ʊ] = **[u]** = **[ʊ]**	**ʌ** (jung· *young*, uns· *us*, Mutter· *mother*); **u:** (Wund· *wound* 'vulnus', durch· *through*), **ʊ** (Busch· *bush*), **aʊ** (Mund· *mouth*, ge⸗funden· *found*), **aʊə** (unser· *our*), **ɒ** (Fuchs· *fox*, Fut⸗ ter· *fodder*), **ɵʊ** (Schulter· *shoulder*), **aɪ** (Furcht· *fright*), **ɜ:** (Wurm· *worm*); **ɔ:** (kurz· *short*); **æ** (und· *and*); problematic morphology: **e** (Brust·⸗*breast*), **ɪ** (kucken·⸗*kick*); **ɑ:** (furzen·⸗*fart*)

[u:]	**u:** (Huf· *hoof*, Jugend· *youth*, tun· *do*); **ʌ** (Blut· *blood*, Bruder· *brother*, tu=t· *do=es*, tust· *dost*, tut· *doth*), **ɵʊ** (Stub· *stove*, tu(e)(·)· *do=n't*), **aʊ** (Kuh· *cow*), **i:** ((be=)suchen· ⁽ᴴ⁾*seek, be=-seech*), **ʊ** (gut· *good*, Busen· *bosom*), **ɔə** (Hure· *whore*); <u>early loans:</u> **[j]u:/ʉ** (⁺Jule· ⁺*yule*)
[œ]	**e** (zwölf· *twelve*); **ʌ** (Knöchel· *knuckle*), **ɜ:** (fördern· *further* v.), **aɪ** (möchte· *might* v.), **ɵʊ** (öffnen·⁼*open* v.); <u>early loans:</u> **eɪ** (schröpfen· ⁺*scrape*)
[ø:]	**i:** (schön· *sheen*, Öster=reich· *easter=*n); **ɪə** (hören· *hear*); **ɪ** (König· *king*); **ɜ:** (Föhre· *fir*); **ɛə** (schwören· *swear*); **ɔə** (röhren· *roar*); **ʌ** (Hönigs=*wald*· *honey*), **[j]u:/ʉ** (Möwe· *mew* 'Larus'); <u>early loans:</u> **u:** (lösen· ⁺*loose*)
[ɣ] = **[ʏ]**	**ɪ** (dünn· *thin*, rücke· *ridge*, fünfte· *fifth*); **ʌ** (schützen· *shut*, dünn· *hull*), **ɜ:** (Gürtel· *girdle*), **aɪ** (fünf· *five*), **aʊ** (tüchtig· *doughty*), **ɒ** (Hül=se· *holl=*y)
[y:]	**i:** (grün· *green*, übel· *evil*), **aɪ** (lügen· *lie* 'mentior'), **eɪ** (betrügen· *betray*); **e** (Brüder· *brethr=*en); **ɵʊ** (blühen· *blow* 'floreo'), **ɔə** (Tür· *door*), **wi:** (süß· *sweet*); <u>phologically problematic:</u> **u:** (kühl·⁼*cool*), **ɔ:** (für·⁼*for*); <u>early loans:</u> **aʊ** (⁺Süd· *south*)
[ae]	**aɪ** (Eis· *ice*), **ɵʊ** (Laib· *loaf*, heilig· *holy*, einlich· *only*, Kleider· *clothes*, kleiden· *clothe*); **e** (Fleisch· *flesh*, einig· *any*, drei· *three=*pence, *three=*penny), **wʌ** (ein· *one*), **u:** (zwei· *two*); **i:** (meinen· *mean*, drei· *three*), **ɒ** (heiß· *hot*, Kleid· *cloth*, heilig· *holi=*day, ein· an=*on*), **ɔ:** (breit· *broad*, Kleid· *cloth*, ein· *a=u*ght), **[j]u:/ʉ** (speien· *spew*), **ju:/jʉ** (Eibe· *yew*), **ɜ:** (dreizehn· *thirteen*, dreißig· *thirty*), **ʌ** (zwei· *two*-pence, *two*-penny, drei· *three*-pence, *three*-penny, Laib· ᵁᴷₒₙₗᵧmy *lo*-rd [mɪˈlʌd]), **ʊ** (weib· *wo*-man, -heit· -*hood*, drei· *three=*pence, *three=*penny), **ɪ** (Zweig· *twig*, weib· *wo*-men, drei· *three=*pence, *three=*penny), **eɪ** (Laib· *la=*dy), **æ** (Laib· *La=*mmas, ein· *an*), **ɔ:** (Laib· (my) *lo*-rd, (mi)*lo*-rd [(mɪ)ˈlɔːd]); <u>mor-phologically problematic:</u> **aɪə/aɪ.ə** (Eisen· *iron*); **ʌ** (nein·°*none*); <u>early loans:</u> **eg** (Ei· ⁺*egg*)
[ao]	**aʊ** (Maus· *mouse*), **aʊə** (Schauer· *shower*), **i:** (Rauch· *reek*, blau· *blea*-berry/*blae=*berry); **ʌ** (Daum=*en*· *thumb*, rauh· *rough*), **e** (Haupt· *head*), **ɪ** (Faust· *fist*), **aɪ** (Auge· *eye*), **[j]u:/ʉ** (Tau· *dew*, hauen· *hew*, blau·⁼*blue* ¿compound?), **ɵʊ** (schauen· *show*), **ɔ:** (Klaue· *claw*, ver=dauen, ⁺tauen· *thaw*), **u:** (Raum· *room*), **eɪ** (grau·⁽⁼⁾*gray/grey*, blau· ⁺*blae(=berry)*)
[ɔø]	**aɪ** (neun· *nine*, Mäuse· *mice*), **aɪə** (Feuer· *fire*), **[j]u:/ʉ** (neu· *new*, treu· *true*), **e** (Freund· *friend*), **eɪ** (Heu· *hay*), **ɪə** (teuer· *dear*), **jɜ:** (teuer· ₒₙₗᵧᵁˢ*dear*), **ɛə** (teuer· ₒₙₗᵧᵁˢ*dear*), **jɛə** (teuer· ₒₙₗᵧᵁˢ*dear*), **ɑ:** (teuer· *dar=*ling), **ɜ:** (teuer· *dear=*th); **ɵʊ** (bläuen· *blow* 'fligo'); **jɔə** (euer· *your*), **jʊə** ⟨sic!⟩ (euer· ₒₙₗᵧᵁᴷ*your*), **jʊə** (euer· *your*); <u>morphologically problematic:</u> **aʊ** (beugen·⁼*bow* 'flecto'); **ju:** (eu=*ch*·⁼*you*, Eu=st·⁼*ewe*/ₒₙₗᵧᵁˢ*yow*), **jɵʊ** (Eu=st·⁼ ₒₙₗᵧᵁˢ*yow*)

With minor corrections, the vowels are given according to the «averaged» transcription in the sources mentioned (**e** = **e**/**ɛ**). The restriction notes «ᵁᴷₒₙₗᵧ», «ᵁˢₒₙₗᵧ» and «<u>dialectal</u>» reflect only these sources as well. The US English sounds are given mainly where they differentiate or attest more forms than UK English does. Biphonemic treatment was chosen for the retroflex vowels, which seems appropriate for the purposes of historical comparison.

Although it is impossible to elaborate here on all the less trivial cognate pairs included in the table, a couple examples are perhaps worth commenting on. It seems that *leben* and *live* can be treated as a purely phonetic correspondence, although the vowel quality in English reflects a complex process of reshaping the respective verb class in OE. *Käse* and *cheese* are borrowed («⁺»), but their phonetic development does not prevent dating the loan as already present in PWG. The case of *zwei* : *two* is somewhat more complex than it might seem, since each form here results from chained paradigmatic levellings; nevertheless, the stem shapes surviving into modern languages (OHG *zwei-*, OE *twā-*) can be compared directly (as reflecting a PWG **twai-*); with *zween* : *twain* the issue is far more complicated.

This list is actually incomplete, based on but a brief analysis of several hundred well-known English-German matches; for almost each of the lines, <u>the list of English correspondences</u> may be increased further[9].

[9] A few of these correspondences represent dialectal irregularities and/or morphological variations («⁼»); however, it is not excluded that, likewise, similar irreg-

However, even this incomplete collection of cognates presents an opportunity for a curious thought experiment.[10] The development of vowels in Germanic languages shows a feature resembling the reconstructed development in Altaic: the reflexation of vowels (in particular the short ones) in stem-initial syllables strongly depends on the vowels of the following syllable(s), which allows for representing the correspondences in a way similar to that used in EDAL. Now, let us imagine a scholar who, having access only to data of the major modern languages, managed to correctly identify vowel quantities in stem-initial syllables of Proto-Germanic (e. g. through adducing Modern Icelandic materials). A detailed and adequate reconstruction of PWG second-syllable vocalism is indeed not feasible based only on extant literary languages. Nevertheless, it does not seem impossible that our imaginary researcher could formulate a hypothesis (a bit too simple rather than really incorrect) that most reflexes of the first-syllable short vowels can be explained as representing only three original vowels (*a*, *i*/*e*, and *u*) whose development is affected by three (groups of) second-syllable vowels: *a*, showing neither narrowing nor fronting effect; *i(/j)*, both narrowing and fronting; and *u*, narrowing but not fronting; with various further influences from the consonant environment. Ideally, one could arrive then at a grouping of the reflexes like the following (ideally, since the chart below actually does account for various data from ancient Germanic languages):

«PWG»	Correspondences (German ↔ English)
a... a or *u*	a ↔ æ (hacken· *hack*, hatte· *had*); a ↔ æː (satt· *sad*); a ↔ eɪ (machen· *make*, acht· *eight*); a ↔ ɒ (lang· *long*, was· *what*); a ↔ ɔ: (Salz· *salt*; warm· *warm*, Wasser· *water*); a ↔ ɵʊ (kalt· *cold*, Kamm· *comb*); a ↔ ɑː/æː (Halbe· *half*, Lache· *laugh*); a ↔ ɑ: (Harm· *harm*); a ↔ ʌ (ander· *other*); a ↔ uː (Gans· *goose*, Wamme· *womb*); a ↔ ɜː (Farn· *fern*); a ↔ iː (walten· *wield*); ɑː ↔ eɪ (Name· *name*); ɑː ↔ ɪə (Bart· *beard*); ɑː ↔ ɛə (fahren· *fare*); ɑː ↔ ɑː/æː (Pfad· *path*); ɑː ↔ ɑː (Vater· *father*; Arsch· *arse*); ɑː ↔ ɛ (gesagt· *said*); ɑː ↔ æ (haben· *have*); ɑː ↔ ɔ: (tragen· *draw*, schmal· *small*, Habicht· *hawk*); ɑː ↔ ɒ (Schwan· *swan*)
a... i	ɛ ↔ e (setzen· *set*); ɛ ↔ ɪ (sengen· *singe*); ɛ ↔ iː (Gänse· *geese*); ɛ ↔ ɑː/æː (letzt· *last*); ɛ ↔ æ (Decke· *thatch*, Espe· *asp*); ɛ ↔ ɒ (Wespe· *wasp*); ɛ ↔ ɔ: (Erle· *alder*); e: ↔ ɪə (Meer· *mere*); ɛː ↔ eɪ (wägen· *weigh*); œ ↔ e (zwölf· *twelve*)
i/e... a	ɛ ↔ e (selb· *self*); ɛ ↔ ɜ: (Werk· *work*); ɛ ↔ ɪ (sechs· *six*); ɛ ↔ iː (essen· *eat*); ɛ ↔ ɑ: (Herz· *heart*); ɛ ↔ eɪ (brechen· *break*, weg· *away*); ɛ ↔ ɔ: (Zwerg· *dwarf*); ɛ ↔ aɪ (recht· *right*); e: ↔ iː (eben· *even*); e: ↔ ɑː (Herd· *hearth*); e: ↔ ɜ: (Erde· *earth*); e: ↔ eɪ (Weg· *way*); e: ↔ e (zehn· *ten*, Feder· *feather*); e: ↔ ɪ (leben· *live*); ɪ ↔ e (Himmel· *heaven*); ɔ ↔ iː (Woche· *week*); ø: ↔ ɛə (schwören· *swear*)
i/e... i or *u*	ɪ ↔ ɪ (sitzen· *sit*); ɪ ↔ iː (Schild· *shield*); ɪ ↔ aɪ (Gesicht· *sight*); ɪ ↔ ɜ: (Birke· *birch*); ɪ ↔ ɑː (Hirsch· *hart*); ɪ ↔ e (Himmel· *heaven*); i: ↔ iː (Sieb· *sieve*); i: ↔ e (sieben· *seven*)
i/e... i, u, or *a*	ɪ ↔ aɪ (finden· *find*); ɣ ↔ aɪ (fünf· *five*); ɪ ↔ ɪ (singen· *sing*)
u... a	ɔ ↔ ɒ (Gott· *God*); ɔ ↔ ʊ (voll· *full*); ɔ ↔ ɵʊ (offen· *open*, Gold· *gold*); ɔ ↔ ɜ: (Wort· *word*); ɔ ↔ ɔ: (Tochter· *daughter*, Hornisse· *hornet*); o: ↔ ɵʊ (Kohle· *coal*); o: ↔ ɒ (Boden· *bottom*, Trog· *trough*); o: ↔ ɔ: (Moos· *moss*); o: ↔ e (wohl· *well*); o: ↔ ɔə (bevor· *before*); o: ↔ aʊ (Vogel· *fowl*);
u... u	u: ↔ uː (Jugend· *youth*); u: ↔ ʊ (Busen· *bosom*)
u... a or *u*	o: ↔ ʌ (Sohn· *son*); ɔ ↔ ʌ (Sonne· *sun*); ʊ ↔ ʌ (jung· *young*); ʊ ↔ uː (Wunde· *wound*); ʊ ↔ aʊ (Mund· *mouth*, gefunden· *found*)

ularities and variations also account for some of the Altaic correspondences. In this respect, the two situations are quite parallel to each other.

[10] The authors are indebted to V. Chernov for providing this idea along with accompanying table.

ʉ... *i*	ʊ ↔ aɪ (Furcht· *fright*); ø: ↔ ɪ (König· *king*); ɣ ↔ ɪ (dünn· *thin*); ɣ ↔ ʌ (schützen· *shut*); ɣ ↔ ɜ: (Gürtel· *girdle*); y: ↔ i: (übel· *evil*); ao ↔ ɪ (Faust· *fist*)
ʉ... *i* or *u*	ʊ ↔ əʊ (Schulter· *shoulder*); ʊ ↔ ɜ: (Wurm· *worm*)

There is little methodological difference between our presenting this list with few comments on the actual data, and AV's «rewrite» of the correspondences from EDAL; therefore, we may as well proclaim all these correspondences «irregular» and express indignation at the fact that someone ever nurtured the (obviously prescientific) theory that languages where every vowel seems to correspond to every other vowel and vice versa could be genetically related.

Being familiar with AV's typical objections, we can predict that he might want to remind us that the history of the languages in question is well known to researchers for a period of at least one millennium, and that most, if not all, of these correspondences reflect regular, positionally conditioned developments of the vocalic phonemes of the protolanguage — a picture that is unavailable for «Altaic». But, in fact, this is only the beginning of discussion and not the end of it, since we are entitled to asking the following additional questions:

(a) Let us suppose that no historical or dialectal information is available for the ancestors of modern day English and German. Would it then be possible to reconstruct the relatively simple and economic vocalic system of Proto-Germanic *and* have a satisfactory explanation for *every* case of split vowel reflexation in its two daughter languages?

(b) Provided the answer is «no» (no Germanologist would probably dare say «yes»), is it possible to *come to terms* with the sometimes unexplainable «irregularity» of reflexation and accept that English and German are related on the basis of *other* data?

(c) Provided the answer is «yes», here is the most heretical question so far: is it possible to view those German-English etymologies that show particularly rare vocalic correspondences (e. g. *Haupt — head, Faust — fist, hoch — high*, etc.) as at least provisionally acceptable, or should they be rejected off the cuff?

Now it is possible that some advocates of the «ultra-mechanistic» approach to phonetic correspondences would say that such etymologies should indeed be abandoned until additional information becomes available (which, in many situations, would mean forever). In our humble opinion, this runs against common sense, and not least because, in our effort to concentrate on «negative» evidence, we would be forgetting all about the «positive» one — namely, impeccable correspondences within the consonantal «skeleton» of the compared forms.

Indeed, if we were to produce a table similar to the one above, but this time featuring only consonantal correspondences between English and German, the results would be far less horrifying. Within the Swadesh wordlist, for instance, there are, at most, two English correspondences to

any given German phoneme (e. g. German *t* — English *d*, *t*), and even for these it is almost always possible to suggest complementary distribution.

All of this is but a simple illustration to a well-known phenomenon in historical typology: vocalic systems are usually much less stable in the world's languages than consonantal ones, easily prone to various types of tricky positional developments depending on multiple factors, and are therefore generally harder to reconstruct than consonantal ones. Again, for Indo-European with its attested ancient languages this is not such a crucial problem (although one that still crops up regularly when it comes to particular etymologies), but when it comes to, for instance, Uralic, the situation is far more drastic: even the smaller subgroups display correspondences that are quite comparable in complexity with the Proto-Altaic system (see, e. g., [Zhivlov 2007] on the most recent attempt at reconstructing the vocalic system of Proto-Ob-Ugrian, a branch that consists of but two languages, yet displays more than fourty distinct series of vocalic correspondences), and as for Proto-Uralic itself, there is still no consensus whatsoever on how its vowels are to be reconstructed (most issues raised by W. Steinitz, E. Itkonen, and B. Collinder still appear unresolved) — and this despite the fact that there are plenty of Uralic etymologies (with utterly «irregular» vocalic correspondences) that most Uralists commonly recognize and accept, instead of throwing them into the dustbin because they cannot provide an adequate interpretation to the correspondences.

Returning to our Germanic analogy, it also helps to remember that the linguistic difference between modern English and German measures, at most, 2,500 years of independent development on the part of each; at the same time, the linguistic difference between, for instance, Middle Korean (the earliest *fully* relevant source for Korean studies) and Proto-Mongolic covers *at least* 5,000 years of such development (actually, closer to 6,000). It is perhaps not so surprising that it is these particular branches that display so many «irregular» splittings.

We can honestly say that, if it were not for the generally more conservative Tungusic and Japanese (and, to a lesser extent, Turkic) systems, the task of reconstructing Proto-Altaic vocalism in *any* form would have remained unsolvable to this day. As it is, we are the first to admit that the presented system is well open to criticism — preferably *constructive* criticism, that is, one that will help explain some of the «irregular» reflexations, just as similar constructive criticism sometimes helps to understand many similar «irregularities» within Indo-European, Uralic, Semitic, and other language families.

For now, however, the pillar of Altaic reconstruction — as well as of the Sino-Caucasian one, where the problem of complex phonetic development of the vocalism is further aggravated by the presence of various types of ablaut (morphological gradation) in both the North Caucasian and the Sino-Tibetan branches — can be identified as recurrent correspondences within the con-

sonant systems of its daughter branches. It is hardly a matter of coincidence that it was the *vowel* system and not the consonant one that AV chose to serve as the main object of his rigorous criticism. In fact, the entire text of his (quite lengthy) analysis features but *three* serious remarks about possible irregularities in Altaic consonantism, which we will gladly discuss in details:

(1) «PT *ẹĺ-* 'to walk, to trot, to amble' is based on OT, Tur., Gag. *eš-*, Yak. *is-*, etc., and Chuv. *iš-* (p. 1134). However, Chuv. *iš-* 'to walk in the snow or in the dirt' cannot be a cognate of other Turkic forms, since something like Chuv. **al-* would be expected.[11] Therefore, Chuv. *iš-*, since it violates correspondences can only be a loan, and for the purposes of PT reconstruction the Chuvash form simply does not exist» [VOVIN 2005: 103]. This is accompanied by a footnote on the same page, which reads: «Dybo refers to a rule that in Chuv. PA **ĺ > l* in syllable final position, but *> š* in intervocalic position (p. 138). Since in this case we have a syllable final position, either is something wrong with the etymology ⟨sic — *A. D., G. S.*⟩, or with the rule, which is given without examples, but with reference to Mudrak's dissertation and another obscure publication by Mudrak that are not accessible outside Moscow».

It is certainly true that [MUDRAK 1989], just like many out-of-print Soviet scientific publications of the times, is not easily accessible for the general public (although AV, for one, would certainly have little trouble accessing it had he really wanted to). The rule, however, was not explained in more details in A. DYBO's introductory part not out of a desire to hoodwink the audience, but merely because it is not a very complex rule, and plenty of examples supporting it are available within the text of the dictionary itself, e. g. PT **ẹlit-* 'to hear' > Chuv. *ilt-*; PT **jāĺ* > Chuv. *śol*; but PT **Koĺ-* 'join, unite' > Chuv. *xoš-*, etc., as well as the current example.

AV's main objection here is the result of a misunderstanding akin to the one with Sino-Caucasian labials described above. Nothing is really wrong here either with the etymology *or* with the rule, because Chuv. *iš-* is, of course, a clear-cut case of a phoneme *not* in a syllable final position, but in an intervocalic one: the root is verbal, and in a majority of cases finds itself before *vowels* of the following grammatical morphemes, being never used on its own (in our notation, this is quite conventionally represented by means of a hyphen); forms where it appears before consonantal morphemes have developed accordingly by analogy. In contrast, forms of *nominal* roots usually follow the basic (nominative) form, where PT **-ĺ* does indeed find itself in syllable-final position. It can clearly be observed that most Chuv. items reflecting PT **ĺ* as *l* are nouns, whereas **š* is the usual reflex in verbal stems.

---

[11] It is unclear where exactly does AV get his information on vocalic correspondences between Turkic languages; we certainly cannot expect anything like Chuvash **al-*, since the regular reflexation of PT **ẹ* > Yakut *i* in this position in Chuvash is also *i*, cf. **jẹl* 'wind' > Chuvash *śil*, Yakut *sillie* 'storm'; **dẹri* 'skin' > Chuvash *tira*, Yakut *tirī*, etc.

As for the hypothesis of Chuv. *iš-* being a loan from a different Turkic source, it runs into unsolvable problems concerning the difference in meanings: [ESTYA I: 316] does present as a working hypothesis that it is a borrowing from Tatar, but the Chuv. meaning ('walk in the snow or in the dirt') is unique among other Turkic occurrences, which suggests a long period of independent development rather than borrowing.

On a final note, *even* if the Chuvash form is indeed borrowed, and all cases of Chuvash *š* < PT **ĺ* secondary (which is dubious), the issue is still Turkic rather than Altaic; even if we remove the Chuvash form from this and other similar etymologies, they will still stand on their own as valid Turkic reflexes within Altaic. (AV's «an absence from Chuvash *may be* indicating that PT **eš-* is a comparatively new word» [VOVIN 2005: 103] makes about as much sense as saying «an absence from Anatolian may indicate that PIE **okʷ-* 'eye' is a post-Anatolian innovation». Then again, of course, it may not).

(2) [On the deduction of PK **hār-* 'to slander' from PA **si̯óĺe*]: «The etymology is very weak, because a correspondence of MK *h-* to OJ *s-* (and, presumably, to **s-* in "Altaic") cannot be supported by a significant number of examples (VOVIN 1993a: 340–341), (VOVIN 2000: 149–150), contrary to EDAL's fantastic claim that PA **s-* > PK **h-* in front of PA **-i̯a-* and **-i̯o-* (p. 50–51) that contradicts any phonological or phonetic reality».

The reader is not left deceived; the quoted passages in AV's own works do indeed convey that idea. Unfortunately, no details are given, which is a pity, because AV's phrasing does not exactly let us know whether he explicitly *rejects* the possibility of such a correspondence or is merely *unsure* of its validity. Considering that he is now sceptical about a genetic relationship between Korean and Japanese [VOVIN 2005: 72], we may surmise the former. However, if the reader actually takes the time to check what is written in [VOVIN 1993a: 340–341], not only will he find that back then, AV was fairly confident that such a correspondence *did* exist, but even that the proposed distribution of reflexation (i. e. conditions under which Altaic **s-* developed into Korean *h-*) was fairly close to the one suggested by EDAL authors: «I come to the conclusion that *s-* and *h-* in Middle Korean as reflexes of PA **s* are in complementary distribution, with *ho-* being a special reflex of PA **si-*»[12].

The three etymologies on the basis of which AV arrives at his conclusion, albeit in a somewhat modified manner, also happen to be included in EDAL. One more etymology, offered by him in [VOVIN 2000] (PK **hòk-* 'small, few' ⬦ PJ **sùkù-* 'few') is also present, with the appropriate reference. Granted, four is hardly a significant number. However, the actual number in

---

[12] The entire passage is a critique of J. WHITMAN's proposal [WHITMAN 1985] for a Proto-Japanese-Korean phoneme **š-* > MK *h-*, OJ *s-*, with some of WHITMAN's supporting etymologies rightfully rejected and others reinterpreted as representing PJK **s-* before **-i-*.

EDAL is not exactly four, but twenty-one, and even if AV does not agree with the particular etymologies, this should at least have been mentioned. As for the «fantastic claim» that PA *sĭa- and *sĭo- can develop into *h-sequences in Korean, it is hardly any more fantastic than AV's original claim that Korean *h-* may stem from PA *si-; in fact, it is *less* fantastic, because a change like *sĭ- > *h-*, with devoicing of the semi-consonantal *yot* after an initial voiceless fricative and subsequent simplification of the cluster, is far more understandable than a «lenition» before one high front vowel and nothing else.

In any case, regardless of the actual quality of the criticized etymology, it is hardly wise on AV's part to reject it out of hand before at least letting us know what he now thinks of those three or four etymologies that he himself used to endorse, as well as others such as PK *hjə̀ 'nit' (cf. PT *sirke id., PM *sirke 'louse', PTM *sire- id., PJ *sìrám(u)i id.), PK *hǎi 'sun; year' (cf. PTM *sigūn 'sun'), PK *hím 'sinew' (cf. PTM *sumu id., PK *hók 'wart' (cf. PM *söɣel id.), etc. Considering that vocalic reconstruction in Altaic is still in progress, other interpretations of this correspondence may be suggested in the future, but there hardly seems to be any rationale for denying its existence.

(3) On the etymology of PTM *ʒola 'stone' < PA *tĭṓl̀ì id.: «the initial consonant does not match with other Altaic languages, but EDAL authors know the way around: a diphthong in PA is created to support the correspondence. Thus, we are facing again the situation when the reconstruction is done 'from above'» [VOVIN 2005: 111].

To be more precise, a diphthong is created not in Altaic, but rather in Turkic, where the reconstruction is *diāĺ based on the Chuvash reflexation *čol* (and the reconstruction of the vocalism is actually accepted by AV, who is merely content to change the protoform to *tiāš on the next page[13]), so the explanation is not nearly as *ad hoc* as AV would have one believe. What is more important, however, is that the correspondence between PT *d- and PTM *ʒ- is recurrent: cf. PT *d(i)ār 'narrow' — PTM *ʒir- 'dense'; PT *dīĺ 'tooth' — PTM *ʒul- 'wedge', etc. In addition, PTM *ʒube- 'two' = PK *tubu id., further confirming the suspicion that in certain cases PA *t- gets palatalized in PTM. Finally, it is not excluded that some of EDAL's etymologies with initial *č- should actually be reconstructed with *t-, e. g. *čĭòlú > PT *dolu 'hail', PTM *ʒal-ka 'fine snow', PJ *tùrárá 'icicle'; *čĭobe 'ten' > PTM *ʒuba-n, PJ *təwə, since the potentially diagnostic Korean forms are missing. In brief, while it is

---

[13] Upon modifying EDAL's PT *d- back to the more traditional *t-, he remarks that this will be «causing a problem with correspondence to initial consonant in Tungusic», but in reality it will not, because a rejection of the reconstructed opposition «voiced : voiceless» in PT will not influence the regularity of the proposed correspondences; it will only make us ignore some important details of the development of PA stops in Turkic. The question, of course, is why exactly should we *want* to ignore them.

possible that the exact reason of palatalization in Tungusic will be formulated differently in the future (depending on further refining of the vocalism), there is little reason to doubt the regularity of the correspondence.

Wrapping up this section, it should be stressed once and once again that a lack of regular phonetic correspondences is the heaviest accusation one can fling at a piece of comparative linguistic research, and one that, if proven right, automatically annuls that research in its entirety. In the light of this, it is quite invigorating for all of us to see how little space AV dedicates, in both of his critical reviews, to this crucial issue, preferring instead to concentrate on other matters (to be taken care of below) — and where he *does* dedicate it, it mostly concerns peculiarities of the vocalic system[14]. In our view, this in itself constitutes significant supportive evidence for the existence of both Sino-Caucasian and Altaic as relevant genetic taxa, represented by numerous reliable etymologies.

It should also be added that one sign of a truly objective review is when the reviewer in question is able to concentrate not only on what he perceives as negative sides of the work, but positive ones as well; if for every case of irregularity within EDAL one finds several more cases of regularity, this, in our opinion, should have been explicitly mentioned. However, this is not done, and the reader may be left thinking that, indeed, «the majority of phonetic correspondences presented in the introduction have multiple variants or exceptions» [Vovin 2005: 77], when in fact (a) this is not really true and (b) even where it is true, the unmotivated «variants» and «exceptions» are almost always limited to a handful of etymologies, the elimination of which would do little harm to the Altaic theory in general[15].

---

[14] Unless it is dedicated to particularly baffling general statements like «In addition, there are gaps. For example, reliable Turkic etymologies with reflexes of PA *m- or *ñ- are few and far between» [Vovin 2005: 77]. Last time we looked, there were 80 Altaic etymologies in *m- with Turkic reflexes (out of 157), each and every one of them featuring a regular development *m- > *b- in that particular branch; and 32 Altaic etymologies in *ñ- (out of 50) with the regular Turkic reflexation *ñ- > *j-. If that may be called «few and far between», AV obviously holds the authors of EDAL to a higher standard than Indo-European. Of course, he may argue that few of these etymologies are «reliable»; but then again, the main idea of his review seems to be that *no* Altaic etymologies are really «reliable», which makes this particular point rather meaningless.

[15] It should probably be noted that, for some of our opponents at least, regularity of correspondences can sometimes be overlooked if a «good chance» arises to present a counterargument against an Altaic etymology. Thus, the inlaut cluster *-lŋ- in the PTM word *xilŋü 'tongue' is reconstructed in EDAL based on a perfectly regular set of correspondences where Manchu and Jurchen -l(V)ŋ- are in contrast with *l*-less nasal clusters in other languages: Manchu *ileŋgu*, Jurchen *hileŋ-ŋu* : Evenki *inńi*, Even *ienŋv*, Negidal *inńi*, etc. The same correspondence (with minor variations) is seen in PTM *palŋa* 'palm (of hand)' (Manchu *falaŋGu* : Evenki *hanŋa*, Negidal *xańŋa*, etc.) and *xul-

Finally, many of AV's statements suggest a firm idea that when a system of phonetic correspondences is proposed for a hypothetical language family, it can only be credible when *all* of the phonetic developments in daughter languages are made clear. Hence the observations about «gaps» in the Altaic system; hence also the phrase «now SS allows even rules themselves to be irregular» [VOVIN 2002: 162], referring to STAROSTIN's inability to explain why a PSC root structure like *C_1VR_2V can, within Sino-Tibetan, develop into either *C_1VR_2 or *C_1R_2V under unknown conditions and poking fun at his «appeal to unknown prosody»[16].

If this is indeed what AV passes for «standard methodology in mainstream historical linguistics», we are afraid that he will not be able to find a lot of support from mainstream linguists who are actually engaged in practical work on reconstructing protolanguages. There is hardly one language family in this world, big or small, ancient or young, where one hundred percent precise, mechanistic rules have been formulated for the development of all of the protolanguage's phonemes. If one or two unresolved problems with one part of the phonemic inventory are enough to invalidate the entire reconstruction, we might as well say goodbye to historical linguistics altogether.

ŋu- 'navel' (Manchu *uleŋgu* : Evenki *uŋurē*, Even *inŋö*, Negidal *uŋńon*, etc.). All three words are comparable with external parallels that have the *same* meaning *and* an inlaut lateral: cf. PM **kele-* 'tongue', **haliga(n)* (< **p-*) 'palm', **köjil-sü* 'navel'. And yet, following the argumentation of G. DOERFER [DOERFER 1995: 257], S. GEORG, in both [GEORG 1999/2000: 163] and [GEORG 2004], insists on reconstructing the PTM form for 'tongue' as **xiń-ŋi*, claiming that the forms in Manchu and Jurchen arose «unter dem starken analogischen Einfluß des gemeintungusischen Verbums **ile-* 'lecken'». Let us even assume that such a development is possible (in his original argumentation, DOERFER adduced the parallel between this case and Latin *lingua* 'tongue' < *dingua* under the influence of *lingere* 'to lick', which does not work because the same language also has *lacrima* 'tear' < *dacruma*, with no analogical influence to speak of) — but where, in this case, is the «stark analogisch Einfluß» necessary for the development of Manchu forms like *uleŋgu* and *falaɢu*? At best, one could try to think of a dissimilation **xińŋi > *xilŋi*, but even that solution does not work very well, since there are plenty of cases where Tungusic nasal clusters correspond to nasal consonants/clusters in Manchu without any traces of *-l-*, e. g. Evenki *anŋanī* 'year' : Manchu *ańa* id. < PTM **ańŋa*. Thus we see a fairly obvious correspondence sacrificed for mere sakes of «hypercautiousness».

[16] The interesting thing is that since then, a basic solution *has* actually been found for this particular problem: according to [STAROSTIN 2005], there is indeed an underlying prosodic factor, namely, tense/lax articulation: namely, NC roots with lax (non-geminated) consonants (a feature that can be defined on the entire wordform, thus making it prosodic in essence) correspond to ST roots with the structure *$C\d{R}V$, whereas NC roots with tense (geminated) consonants correspond to the ST structure *CVR, e. g. PNC **qwīrhV* 'horn' — PST **krua* id., but PNC **ččǎ̆lV* 'branch, tree' — PST **Cal* 'wood', etc. This rule *knows* possible exceptions, but from a statistical point of view, they are irrelevant (and could possibly be discarded in favour of better etymologies).

For instance, according to strictly «mainstream» Dravidian linguistics [Kri-shnamurti 2003: 121–122], no exact rules are known to determine when Proto-Dravidian *c-* becomes lost in South Dravidian languages (e. g. PD *cēr-* `plough' > Tamil *ēr*) and when it is preserved (e. g. PD *cēr-* `join, unite' > Tamil *cēr*). This is, however, tacitly acknowledged as an unexplainable irregular bifurcation of reflexes — implying that the reason behind it may lie, for instance, in dialectal interaction. It is also possible that this opposition may reflect two different phonemes in the protolanguage, although mainstream Dravidologists usually decline that solution. No one has ever thought of using this argument (and quite a few more can be added) to doubt the very existence of Proto-Dravidian. AV will, of course, argue, that the day is saved with morphological and other hard evidence, like he does for Indo-European in [Vovin 2002] — well, likewise, there exists plenty of hard evidence for Altaic and Sino-Caucasian to save the day for them as well. For some mystical reason, though, AV prefers to ignore it, although he does not do the same for Indo-European.

In short, experience shows that a search for *rigorousness* in comparisons must always be tempered with striving for *realism*. The reconstructed phonological systems of Altaic, Sino-Caucasian, or any other macrofamily will *never* look like ideal mechanistic constructs; important and *statistically relevant* parts of them will, but there will always be exceptions as well. In fact, we can quite positively state that any 'Altaic' or 'Sino-Caucasian' phonological system proclaiming to be one hundred percent «productive-predictive» — satisfying the late R. Austerlitz's criterion that is so heartily advocated by AV[17] — will *necessarily* be flawed, as it will be proclaiming impeccability for a distant relationship theory when not even a single close relationship theory is ever impeccable. Try as we might, not even a hundred similar invocations from AV or

---

[17] «Ideally, the relationship between the languages constituting a given language family should be productive-predictive... given a form in language A, we should be able to predict what the corresponding form will be in language B based on the set of proposed regular phonetic correspondences... Needless to say, productive-predictive relationships do not exist in EDAL» [Vovin 2005: 77]. Cf. the following argumentation in [Vovin 1994: 32]: «Some scholars... use productiveness and predictiveness *instead* of the comparative method, and this is a serious methodological flaw, for there are considerable limitations to productiveness and predictiveness in cases of phonetic merger... If we start to predict Sanskrit forms on the basis of Tocharian, we find ourselves in a complete mess. If, however, an anti-Altaicist recognizes the genetic relationship between Sanskrit and Tocharian despite the impossibility of predicting a Sanskrit form on the basis of Tocharian, then what inhibits him from recognizing the same kind of relationship between Japanese and Korean, or Japanese and Tungusic? Does it mean that our methodology has to be different in the case of Altaic? I do not think so: the comparative method is the same for all language families». Needless to say, we fully concur with the author of the quoted passage.

other critics can make us do for Altaic and Sino-Caucasian what has not even been done for Semitic, Dravidian, or Indo-European.

III. *Semantics*: *same meanings or typological reliability*?

After discussing phonetic correspondences, it is only natural to proceed to with accusations put forward by AV concerning the semantic side of the comparisons. These mostly refer to the Altaic theory, but glimpses of the same can occasionally be evidenced for Sino-Caucasian etymologies as well.

AV: «...the semantic side of comparisons, especially of those that the EDAL authors propose for the first time is often vague and even stretched beyond any credibility. Certainly, semantics is the most subjective area, but there should be some reasonable limits» [VOVIN 2005: 81].

The most usual problem here is that different researchers understand different things under «reasonable limits». What lies within such limits for one specialist in one field, may seem utterly incredible to another specialist in a different field, and vice versa. Unfortunately, comparative linguistics still lacks anything even remotely close to a formal model of historical semantics, and there is thus no way to prove that a certain postulated semantic change is possible over a given time period, while another one is not. The only possible solution here is to rely on one's own experience, and postulate only such semantic changes that can be «calibrated» by comparing them with either historically attested changes in languages with a well-known history or (more frequently) changes in morphemes reconstructed for commonly accepted families (such as Indo-European) whose phonetic structures correspond so well to each other that the probability of a chance coincidence is very low.

Likewise, there are no formal algorithms that aid us in reconstructing the original meaning in the protolanguage. At present, semantic reconstruction is at a very preliminary stage[18], and all the meanings given for Proto-Altaic forms are approximate, which is to say that if the adduced meaning reads 'a kind of tree' or 'a kind of bird', it, of course, means, that we were not able to establish precisely *what* kind of tree or bird the word used to mean in Proto-Altaic, rather than that its *actual* meaning was 'a kind of tree' (as in, 'I sat down in the shade of a kind of tree with a kind of bird singing above my head'). Why AV views this solution as criminal is way beyond us, particularly since, as every specialist in comparative linguistics is supposed to know, it is frequently used in etymological dictionaries of well established families (e. g., K. RÉDEI's

---

[18] Relatively formal semantic reconstruction for just one specific Altaic field of terminology (body parts of the arm/shoulder region), although yielding important results, had required over 300 pages worth of work [DYBO 1996], so, obviously, the same procedure could not be performed for the dictionary in its entirety.

classic Uralic dictionary [Rédei 1988] is full of meanings like 'eine Art Fisch', 'eine Art Vogel', but, to the best of our knowledge, that particular sin has never been brought up by anybody as a significant flaw of the dictionary).

AV seems to be unjustifiably overconfident in the powers of semantic reconstruction for «well-established families, like Austronesian, where we have comparatively small but quite reliable lists for common fauna and flora for common subspecies without any attempt to compare a 'crab' to a 'lobster' or a 'clam' and reconstruct Proto-Austronesian 'a kind of sea creature with shell'» [Vovin 2005: 81]. Perhaps not, but how about 'lobster' and 'shrimp' (Proto-Austronesian *qudaŋ, in R. Blust's reconstruction, with the original meaning uncertain)? And it is certainly not true that 'crab' and 'lobster' are never interchangeable in Austronesian: cf., for instance, Dobuan *kalimana* 'crab' vs. Kambera *kalimau* 'lobster' (both from Proto-Austronesian *qali-maŋaw). In addition, there really is a whole lot of common Austronesian lexics for flora, fauna, and other semantic fields where the exact meaning cannot be established due to too much variation in daughter languages, as even the most brief scan of any available comparative materials on this family will reveal; AV's choice of Austronesian as a «semantically rigid» family is completely misleading, since such families, unless they are one or two thousand years old, simply do not exist.

Apart from that, we would be the first to admit that there are plenty of semantic comparisons in EDAL that might seem off-putting even to specialists. Two main arguments must be mentioned in our defense: (a) the «non-trivial» semantic developments are well balanced by quite obvious and unsurprising ones over the course of the dictionary (it is hard to engage in statistics over this matter, but the interested reader may well perform a check on his/her own); (b) in an absolute majority of «non-trivial» cases the authors follow the guidelines laid out above: any semantic changes are acceptable provided they can be confirmed typologically.

For example, on p. 82 of his review AV takes issue with the fact that the authors dared to put words with the respective meanings 'horn' and 'gum' within the same etymology. Unfortunately, he provides the reader with no further information on the context of such a weird comparison. The exact forms selected for comparison on p. 948 of EDAL are: Tungusic *muŋi 'tendon'; Mongolic *möγer-sü 'cartilage, gristle'; Turkic *bujŋuŕ 'horn'; and Korean *mừ'ừm 'gums'. The reconstructed meaning is actually just a listing of the most common meanings in the protolanguages and is given as 'horn; cartilage, tendon'; the original form is reconstructed as *mújŋi.

If the original Proto-Altaic meaning here was really just 'horn' or 'gums', the direction of the semantic development would indeed be hard to establish. However, it would be useful to consider the semantic element

that unites all these forms, and this is essentially 'bodily outgrowth, excrescence (of callous or gristly nature)'. For typological parallels, cf., for example, Latin *gingiva* 'gum' < PIE **geng-* 'lump, swelling' [WP I: 637–638], or rather frequent concurrences of the meanings 'horn' and 'toe- or fingernail' (= 'corneous excrescence') within the same root (e. g. Avestan *srū-* with both meanings). Also, it is interesting to note that in classical Tamil, the same word *eyiṟu* can have the meanings 'gum', 'tooth' and 'tusk (of elephant, wild hog)' (!) [TL: 529] whereas in another Dravidian language, Malto, the word *qaru* can mean either 'gums' or 'cartilage' [MAHA-PATRA 1987: 22]. 'Tooth' and 'tusk', in their turn, are meanings that are closely connected and often interchange with 'horn'.

In short, while the exact pairing 'horn' — 'gum' is not met in etymological literature on relatively young language families, a hypothesis that derives both of them from an original 'callosity, excrescence' is perfectly acceptable as a working one, and is further strengthened by solid phonetic correspondences that do not violate any rules given in the preface to EDAL. It should be admitted, of course, that the way the meaning is given in the dictionary may be somewhat misleading, but the comparison most definitely stands. Certainly the idea that Proto-Altaic speakers had «horns instead of teeth growing from their gums» [VOVIN 2005: 82] is not any more bizarre than the idea, for instance, that Proto-Indoeuropeans had their foreheads located on their breasts (Gr. στέρνον 'breast' is usually associated with Germanic **stirn-ō(n)-* 'forehead' — not directly, of course, but both through a verbal root **ster-* 'to be wide, spread' [WP II: 638]), or that Proto-Fenno-Ugrians had unusually bloody hair (the etymology of FU **puna* 'hair' in [RÉDEI 1988: 402] also happens to include Finnish *puna* 'red, blood', due to a supposed development 'hair, wool' > 'red hair (wool)' > 'red in general'; the development is typologically unique, but from a phonetic standpoint the comparison is impeccable).

There is actually a plethora of «funny» moments like these that can be dug out from respectable dictionaries of well-accepted families, but for some reason AV singles out EDAL as the funniest one, concluding that «any kind of a body part term can be compared in EDAL to any other body part» — a rather sweeping assumption to illustrate it on but one example. Even if one still prefers to discard the etymology, there are dozens of straightforward body part connections between the various branches of Altaic, with no cases of «eye» corresponding to «foot» or «finger» to «liver», as AV's phrase might make one believe.

Upon reading AV's particular criticisms of EDAL's semantic comparisons, however, it eventually becomes obvious that for him, etymologies should be discarded not just every time the suggested semantic changes

are typologically unusual, but, in fact, every other time a semantic change is supposed *at all*. Consider the following cases:

(a) The comparison of PJ **pàsìr-* 'to run' with Tungusic, Turkic, and Korean roots respectively meaning 'walk, hurry', 'walk, trot, amble', and 'tread, trample' is rejected partially due to the fact that «OJ *pasir-* apparently had much broader semantics than modern *pasir-* 'to run': it could also mean 'to move quickly', and 'to fall down (of stones)'» [Vovin 2005: 104]. The implications are not easy to understand. Does that necessarily mean that the original meaning was not 'run', but 'move quickly'? And even if it does (which it does not), what is the unsurmountable obstacle that prevents us from comparing a word with the meaning 'to move quickly' with a word with the meaning 'to walk'?

(b) AV correctly points out the authors' inaccuracy in glossing OJ *isawo* as 'craftsman, diligent person'; the actual meaning, based on dictionary entries as well as textual occurrences, should rather be presented as 'brave, vigorous person' [Vovin 2005: 96–97]. However, considering the semantic variety in the external parallels (not just 'work, craft', but also 'deed, action; diligent, active, firm'), this by no means invalidates the etymology — at worst, it shows that the authors have not managed to lay their semantic argumentation down properly, for which an apology is indeed in order. Yet if the reconstructed meaning 'work' is actually understood in a manner closer to '*activity, occupation*', the Japanese forms fit in perfectly — unless AV wants to insist that a semantic derivation of 'brave, vigorous' from 'active' is impossible.

(c) Similarly, AV protests against bringing in PK **īr* 'work, profession' to fit in the same comparison, since «In MK, as far as I can judge on the basis of its textual usage, the word means 'thing', 'matter', 'affair', 'fact', 'deed' as well as 'work', 'job', but uncontroversial examples for 'work' and 'job' start to appear only in later texts. The meaning of 'thing', 'fact' and 'matter' still exists in the modern language... and I believe that it is the primary meaning of the word, that is also supported by the fact that MK Chinese-Korean dictionaries translated the Chinese character 事 'matter', 'thing' as MK *:il*» [Vovin 2005: 97–98].

It is telling, however, that the original meaning of the Chinese character 事, used to render MK *:il* in dictionaries, was neither 'matter' nor 'thing' but actually 'affair (business), mission, *service*', as well as the predicative meaning 'to serve', as is well established from the earliest texts [Schuessler 1987: 548–549] and, furthermore, confirmed etymologically; the more abstract meanings ('thing', 'fact', etc.) start to develop actively only beginning from the classical period. This is not to say that it is necessarily the original meaning that should have been reflected in Korean translations; but it does mean that the semantic development 'work, service' > 'matter, thing' is fairly normal for the area, and could be the case for Korean as well. As for chronology of usage in Korean, this can hardly be a decisive factor in our rejecting the etymology.

(d) The comparison of TM **sulu* 'rogue; to mock' to various Altaic roots with meanings like 'to talk nonsense' (Mongolic), 'wild, astonished' (Turkic), 'slander' (Japanese, Korean) is again rejected on the basis of «far-fetched and distorted semantics» [VOVIN 2005: 108], simply because the authors of EDAL dared to render the Russian translation of the Ulcha word *sulu* 'шалун' as 'rogue', and the Russian translation of the Orok word *sụlụ-da-* 'дразнить' as 'to mock'. Granted, 'prankish person' and 'to tease' (as suggested by AV) would have been better equivalents. But what does that exactly have to do with the validity of the etymology? 'To play pranks, to tease', 'to mock', and 'to slander' are all meanings that lie within the same semantic field; there is absolutely nothing 'far-fetched' or 'distorted' about the comparison, unless, of course, each and every meaning shift is supposed to be 'far-fetched'.

(e) The derivation of PJ **sàs-*, **sàsì* 'to prick, stab, sharp stick' from PA **sīla* 'tooth, sharp stick' is rejected on the following grounds: «It must be noted that *sasi* 'sharp stick' is a fruit of authors' imagination... it is, of course, *sas-i*, a nominalized verbal form... the basic form is a verbal one: *sas-* 'to pierce, stab, stick (into)'. You can *sas-* in Japanese with a knife, an owl ⟨*sic!* — *A. D., G. S.*⟩, a spear, a sword, a hairpin, or whatever you like — there is no exclusive connection whatsoever to 'sharp sticks' or even less 'teeth', with which you certainly don't pierce in Japanese, you bite, as in many other languages. The further problem for this exciting etymology is that *sas-* in all historical periods of Japanese, including modern language, used to mean quite different things: pulling a boat with a pole, setting a trap, igniting fire, pouring rice wine, pointing with a finger, wearing a sword on one's side, holding an umbrella, etc... the archetype of the meaning seems to be connected with 'pointing' or 'insertion', and not necessarily with 'stabbing' or 'piercing'» [VOVIN 2005: 106–7].

Some of the oddities of this passage may be observed without additional commentary, but several things should be made clear. First, the word *sasi* 'sharp stick', be it a secondary nominalization or a primary nominal stem, really exists in the language (a more precise meaning is 'sharpened bamboo or metal tube'), even if it does not seem to be textually attested in OJ — but «not attested in old texts» is certainly not the equivalent of a «fruit of the authors' imagination». Second, the question of whether we are dealing with one or more roots here is still open to debate: despite having the same accentuation, the verbs are transcribed with at least three or four different Chinese characters, and homonimity, or, to be more exact, contamination of several roots through secondary phonetic merger, cannot be ruled out[19]; particularly suspicious is the connection between 'pointing'

---

[19] Some of the meanings quoted by AV quite probably owe their existence to literal translation from Chinese — e. g., 'pulling a boat with a pole' can hardly be

and 'inserting', relatively rare in the world's languages. Third, even if all of these words *are* the same root, what exactly is the overwhelming evidence that forces us to assume the meaning 'to point' as primary? As long as the latter is safely attested in OJ, the etymology is perfectly acceptable.

(f) Everything, however, pales in comparison to this (concerning the Mongolic reflex for the above-mentioned PA **p'ĕĺo* 'walk, run'): «PM *hülde-* 'to chase' is based on MM *xulde-*, WM *ülde-*, KH *ülde- / öldö-*, and Bur. *ülde-*. WM, KH and Bur. *ülde-* indeed have meanings 'to chase away, to drive out, to expel, to chase'... MM *xulde-* also has a range of meanings: 'to chase', 'to chase away', 'to trail'... We can see again that EDAL authors edit or oversimplify the semantics to serve their purpose: if 'to chase' certainly involves movement, 'to chase away' or 'to drive out' does not necessarily have the same connection, since one can chase away or drive out somebody just verbally» [Vovin 2005: 102–103]. One cannot help but wonder if this is some sort of linguistic joke: even if AV manages to find verifiable examples of semantic derivation «to chase away verbally» > «to chase away by pursuing on foot» (we tried, but failed), what exactly is his way of asserting that the meaning «chase away verbally» in select Mongolic idioms is primary? Yet if he actually thinks it is secondary, how does that work against the etymologization? And isn't this more or less the same as saying, for instance, that one should not compare Skt. *váhati* 'to carry, bear' with Proto-Germanic **wigan-*, because the latter frequently means 'to move', and, of course, one can easily *move* things around without having to carry them at all?

These and several similar cases are, in our humble opinion, not at all how one should treat semantic relations in comparative research. Would we reject the — phonetically ideal — connection between Hittite *tuzzi-* and Lithuanian *tautà*, Proto-Germanic **ϑiud-*, Old Irish *tuath*, etc., simply because the former means 'army, troop' in contrast to 'people, folk' for the latter? Or, supposing the author of the etymology made a mistake and translated the Hittite form as 'people': would we then take advantage of the subjective inaccuracy to eliminate the etymology? And just how many well-accepted etymologies, within Indo-European or other commonly re-cognized families, would we have to dismiss as unscientific?

Finally, one more semantics-related criticism is tightly bound with the «disdain for philology» accusations (see below), but should rather be dealt with in this section. According to AV, the authors of both EDAL and SCH frequently reconstruct meanings that could not exist in the respective protolanguages simply because the realities they represent could not be

---

separated from the fact that much the same meaning is attested for Chinese 刺 *cì* (used to transcribe Japanese *sas-*), where, however, there is absolutely no doubt about the original meaning ('to pierce, stab').

known to the actual speakers. «One should not forget that every language has a history, and that this history is intimately connected with the cultural and sociopolitical history of the people who speak a given language. Authors of EDAL have no respect for and no need of cultural history of languages they compare» [VOVIN 2005: 75]. But, once again, this accusation rests on the twin shoulders of (possibly intentional) misunderstandings and hardly warranted preconceptions. Let us illustrate this through several examples from both [VOVIN 2002] and [VOVIN 2005].

(a) «Starostin 1995a: 189 compares PY *ʔeχV 'iron' with PNC *riẊ(w)e 'copper, gold'... with Sino-Caucasian claimed to be 10,000 years old, what evidence for metallurgy do we have for such a remote past?» [VOVIN 2002: 161]. The comparison is actually much older, being first offered in [STAROSTIN 1982: 214]. However, neither in that source nor in the newer ones is there an explicit or implicit statement that, based on this comparison, we should necessarily reconstruct the concept of metallurgy for the speakers of Proto-Sino-Caucasian.

Let us not forget, like AV does, that it is actually a relatively frequent phenomenon when the same meaning *A* in the protolanguage independently shifts to meaning *B* in two related languages, *particularly* if it is easy to «apply» the old word to newly emerging realities. Thus, Tamil *vari* and Telugu *vrāyu* both have the meaning 'to write', which certainly does not empower us to suggest that speakers of Proto-Dravidian (at least 4000 BC) were able to write. However, they could certainly 'make incisions' or 'carve', which is probably the original meaning (fortunately, it happens to be preserved in many modern languages as well, but we may not always be so fortunate).

How can a word designating some sort of metal appear in a given language once its speakers become acquainted with that metal? Obviously, there are two possible ways: (a) the word may be borrowed from a 'technologically superior' neighbor or (b) it may be newly formed, usually through a metaphoric (e. g. from colour terms) or metonymic (e. g. from names of artefacts) meaning transfer. That Proto-Sino-Caucasian could have a word that independently developed into 'iron' in Yeniseian and 'gold' or 'copper' in North Caucasian *can*, of course, be put under doubt[20]; but a truly serious way to «get rid» of that etymology would be to either demonstrate that it is phonetically unsatisfactory, or that at least one of the compared forms can be etymologized differently *and* better. The first point does not work here (the correspondence of PNC *r- to PY anlaut zero is regular and recurrent, as shown in [STAROSTIN 1982]), and the second one works only if the Yeniseian word is indeed an old loanword

---

[20] Let us also note, while we're at it, that even though the possibility of «proto-Sino-Caucasians» being familiar with professional metallurgy is a remote one, a general knowledge of some kinds of metal and its properties (gold, meteoric iron, etc.) cannot at all be excluded for the period of 10,000 BC (or even earlier).

from an Iranian source (as is cautiously suggested in [Starostin 2005a]). Until the Yeniseian-Iranian connection is made clear, however, the etymology stands and does not violate a single principle of the comparative method.

(b) «Starostin 1995a: 181 compares PY *ʔalVk 'dog sledge' with PNC *hwə:l(V)kwe: 'cart, wheel'. As in many other similar cases, the proto-Yeniseian reconstruction is based on a unique attestation in Ket (but not in other Yeniseian languages). The bisyllabic root in Yeniseian *may* also be suspicious. But above all, it is interesting to know whether SS realizes when carts and wheels were invented» [Vovin 2002: 161].

This particular etymology has since been rejected by S. Starostin himself; it is not present in the latest version of the Sino-Caucasian database, nor can it be found in the glossary in [Starostin 2005a]. However, this certainly had nothing to do with the fact that the Ket form is isolated in Yeniseian (all the «many other similar cases» have a very simple explanation — there is much more lexical material available from Ket than from any other Yeniseian language, all of which have been extinct since the XIXth century), nor with the suspicious disyllabic character of the Ket root (which might easily reflect earlier morphology), *nor* with the root's semantics, which could easily just designate some sort of movable container in Proto-Sino-Caucasian. Instead, it has been rejected because a better etymology for the Ket word has been found: borrowing from Evk. *alaɣ* 'tug (in reindeer harness)' ([Vasilevich 1958: 22]). Coincidentally, the rejection of the Sino-Caucasian etymology has been explicitly stated within the text of EDAL (p. 757, in connection with the etymology of the Evk. word). AV, of course, may simply see his opinion vindicated and leave it at that, but the point is that theoretical speculations of the «isn't it strange that...?» type simply do not constitute a significant scientific argument when not backed up by hard data.

(c) (while discussing the PA etymology *k'ìĺa 'fetters'): «I wonder what would PA speakers use fetters for in the 6th millennium B.C.E., when nomadic pastoralism did not yet exist» [Vovin 2005: 98]. This is a particularly baffling statement. First of all, we are not sure about the sources of AV's knowledge here: it is quite commonly accepted that nomadic pastoralism was concurrent with, or even preceded, the Neolithic Revolution (far earlier than 6,000 BC), and there is no reason whatsoever to picture speakers of Proto-Altaic as a strictly hunting/gathering society. But second and even more important, one does not absolutely need to herd cattle in order to have knowledge, or make use of some sort of restrictive ties, be it 'fetters', 'ropes', 'strings', or anything of the kind.

Winding up this section, it should be mentioned that the ultimate test for determining the degree of genetic relationship — the lexicostatistical one — is actually based on comparisons between words the semantics of which

matches perfectly. Thus, no matter how many dubious semantic comparisons one may encounter in the Altaic or Sino-Caucasian dictionaries (and we strongly assert that the number will not be critical, unless, of course, one prefers to label *every* semantic shift as dubious), their validity is always reinforced by the remaining straightforward comparisons, ones that constitute the «core» of the evidence. Which, logically, brings us to AV's fourth point.

IV. *Convincing evidence: short lists or core etymologies*?

In both [VOVIN 2002] and [VOVIN 2005], AV frequently brings up the idea that, in order to be convincing, genetic relationship between two or more language groups should not only be formally proven through the existence of tight phonetic laws, but also intuitively evident based on a small number of highly impressive lexical (and morphological) items. Cf.:

«...actually one needs very few examples, provided the regular correspondences are really there» [VOVIN 2002: 155];

«...the possibility of a genetic relationship should be easily demonstrated with a small set of paradigmatic morphology and/or cognate vocabulary that clearly reflect the proposed correspondences. Only in this case a proposed language family will be uncontroversial» [VOVIN 2005: 77].

Unfortunately, the last quoted phrase suggests that what is at stake here are really two different issues. In our humble opinion, it is one thing to formulate a scientific hypothesis — in this case, a hypothesis of genetic relationship — and an entirely different thing to procure for that hypothesis the status of «uncontroversial». It is no big secret that scientific phenomena vary greatly in respect to their being intuitively evident. To put it very roughly, it does not take more than one apple to illustrate the law of gravity, but it takes much more than that to arrive at something like, for instance, Maxwell's electromagnetic equations.

The same applies to genetic relationship. Being quickly and easily able to demonstrate that two languages are related depends on a whole number of factors, primary among which, of course, is the time distance that separates the languages, but which also include their degree of structural complexity, typological peculiarities, quantity and quality of available data, and others.

Thus, one might reasonably ask the question: why does AV, in order to support his thesis that «one needs very few examples», resort to demonstrating it on the basis of data from Malayo-Polynesian, even though that particular language family is not his primary area of expertise?

The answer, when you come to think of it, is easy. Obviously, using such families as Turkic, Mongolic, or Germanic (where the relationship is indeed unquestionable, right down to mutual intelligibility among spea-

kers of different, even if not all, languages), from AV's perspective, would not be a very good choice: all of these families are relatively young, in comparison to «Sino-Caucasian» and «Altaic». It is therefore necessary to bring up something older. But here we may encounter an unexpected problem, which AV avoids mentioning: with language families that are more than 2,000 years old, it is actually *not that easy* to come up with short lists of lexical items that will immediately suggest relationship to a critically minded reviewer. And even if one actually manages to come up with such a list of lookalikes, it is practically impossible to ensure that this list will be thoroughly conforming to all the proposed phonetic laws.

Let us illustrate this last thesis with a simple example. Below we quote a set of examples of comparisons between Old Slavonic, Old Indian, Greek, and Latin, taken from an old handbook on general linguistics [REFORMATSKY 1955: 314]. This table of lexical cognates, reproduced here with a few obvious mistakes corrected, is clearly given in the original edition with the aim of convincing the reader — on an intuitive basis, since the regular correspondences are not specified in the text — that all these languages *are* related.

Meaning	Old Ch. Slavonic	Old Indian	Ancient Greek	Latin
mother	мати (матер-)	mātā́ (mātár-)	μᾱ́τηρ	māter
sister	сестра	svásā (svásar-)		sorŏr
daughter	дъшти (дъштер-)	duhitā́ (duhitár-)	θυγάτηρ	
two (*f.*)	дъвѣ	dvé	δύο	duae
hundred	съто	çatáṃ	ἑκατόν	centum
thou	ты	tvám	τύ	tū
I	азъ	ahám	ἐγώ	egŏ
nose	носъ	nā́sā		nāsus; nāris
wolf	влькъ	vŕ̥kaḥ	λύκος	lupus

Based solely on an intuitive basis, few would probably argue that this list conveys any idea other than genetic relationship, and that it is not eminently suitable for the goal it sets out to achieve. However, when it comes to regular correspondences, the only items on the list that demonstrate complete regularity according to general sound laws for Indo-European, turn out to be 'mother' and 'two'! Everything else is most definitely problematic. For instance, Old Slavonic *-e-* does not regularly correspond to Latin *-o-* ('sister'), nor do the declension types of the two; Old Indian *d-* and *-h-* do not respectively correspond to Greek *θ-* and *-γ-* ('daughter'); nor does Old Indian *-h-* to Latin *-g-*, while Slavonic *a-* does not correspond to Greek and Latin *e-* ('I'); Old Slavonic *-ŭ-* is not a regular match for Old Indian *-a-* and Latin *-en-*, nor is the *he-* element in Greek easily explainable ('hundred'); the root vowels and morpholo-

gical paradigms of the words for 'nose' do not match between Slavonic, Old Indian, and Latin (for the latter, *nāris* 'nostril' being just about the better correspondence than *nāsus* 'nose' itself, with its non-rhotacistic -*s*-); we meet with serious morphological problems in the Old Indian and probanly Greek (with an initial *σ*- that also defies regularity) words for 'thou'; and the Greek and Latin words for 'wolf' are completely irregular. Even with the «good» words, it should also be noted that Reformatsky intentionally chooses the rarer (Doric) form *μάτηρ* (and for the second pronoun form, *τύ*) in Greek, whereas the more commonly known — Ionic-Attic — forms are *μήτηρ* and *σύ*.

Let us not forget, either, that what we are dealing with here is the *best* kind of comparative evidence (otherwise it would not have been chosen for a general handbook), and, furthermore, that it is evidence from ancient Indo-European languages, not modern ones — separated by, at most, four thousand years, not six or more.

Granted, over the years Indo-European comparative linguistics has managed to work its way around most (but hardly all) of these issues. But this already has little to do with intuition: instead, such problems are solved by bringing into consideration factors like dialectal developments, tricky and rare positional changes, morphological gradation (ablaut), and even borrowing from related languages (like Latin *lupus*). The relationship still remains very probable even without this further analysis, but trying to make a selection of Indo-European etymologies that would at once satisfy the «easy demonstration» criterion *and* thoroughly match all of the general phonetic laws proposed for the family is an extremely hard task, perhaps plain impossible.

However, that is exactly what AV needs in order to discredit the existing Sino-Caucasian (or Altaic) comparisons. And this is why he has to resort to Malayo-Polynesian. Unlike Indo-European, in fact, unlike *most* of the world's families, Polynesian languages usually share the pleasant feature of having relatively simple phonetic inventories, little, if at all, burdened by consonant clusters, multiple oppositions in manners of articulation, and complex morphonology; and, while it is possible to show that in many cases this has been the result of historical simplification, Polynesian languages overall give the impression of being exceptionally simple in terms of their historical development — in relation, of course, to the other world families rather than on their own.

Even so, it takes time and effort to make a particularly «convincing» selection of test cases for Malay and Hawaiian, like AV does with his pairings of words for 'eye', 'I', 'road', 'deaf', 'die', and 'fish' (p. 155). He then goes on and compares these forms with their Paiwan equivalents (p. 156), claiming that, although we have thus increased the time depth by 2,000 years, this does not affect the regularity of the correspondences. This all looks quite plausible.

But then, still a couple pages later, something very strange happens: AV offers us his list of comparisons between the various Formosan languages and

Malay (the same one that was discussed in the first section of our article), claiming that they do *not* share any common body part terms and implicitly inferring that it is not only impossible to demonstrate that Formosan languages are related on a «small set of cognate vocabulary», but, in fact, hard to demonstrate it on lexical evidence as such — and that morphological evidence necessarily has to be brought in in order to make the relationship palatable!

Apparently it means that for the pairing of, say, Hawaiian and Paiwan AV *can* find this «small set of cognate vocabulary», which makes the relationship obvious, yet for the pairing of, say, Squliq Atayal and Bunun such a set is utterly unavailable: «if one sides with 'basic vocabulary'-based approach, the Austronesian theory goes out of the window» [Vovin 2002: 159]. Let us get this straight: just a few pages before, the plausibility of the Austronesian theory was being illustrated through a small slice of the basic vocabulary, and now it is being stated that basic vocabulary is not a good means to prove genetic relationship, and that, if not for strong morphological ties between Bunun and Atayal, said relationship would still remain questionable.

Such a blatant self-contradiction can seemingly only be resolved in one possible way: we would need to assume that there is no such thing as a universally applicable criterion of establishing genetic relationship. In fact, everything goes (morphology, lexicon, etc.), but only as long as it does not take a lot of written space to lay down the information. If the relationship cannot be shown on a «small set of cognate vocabulary» with impeccable correspondences (preferably fitting on less than half a page), it should be shown on an equally small and impeccable set of morphological markers. If it cannot be shown on that either, it cannot be shown at all and, no matter how huge the evidence, «mainstream linguists» will forever remain unconvinced of the proposal.

It goes without saying that this approach has nothing whatsoever to do with scientific principles. Apart from placing everything into the realm of the subjective and intuitive (just how «small» is the cognate set supposed to be? just how many morphological markers are needed?), it makes a mess out of a clear understanding of genetic relationship — apparently, we can now select whatever is more suitable to us in each particular case. If lexics works, we can go along with lexics; if it does not, we still have morphology. If there is no morphology (isolating languages), we will once again go along with lexics.

In reality, of course, there is nothing controversial about the Austronesian situation as long as the correct interpretations are given, and these are as follows:

(a) all of the Austronesian languages quoted in [Vovin 2002] — Hawaiian, Malay, Paiwan, Atayal, Bunun, etc. — *are* related, and in each and every case it is possible to provide evidence for that relationship based solely on lexical comparison, without having to resort to morphology (see the discussion above);

(b) in the case of *some* Austronesian languages (Hawaiian — Malay) this relationship can be intuitively seen even on a small selection of cases, due to their simple phonetic structures and relative scarceness of the historical processes they have undergone over the past several millennia;

(c) in the case of *other* Austronesian languages (Bunun — Atayal, etc.) this relationship cannot be seen intuitively and has to be properly demonstrated through intermediate reconstructions instead of directly bringing modern languages into the comparison, not to mention small, compact cognate sets from modern languages.

In other words, the same conclusion that we reached while discussing the morphology criterion applies here as well: it is *reasonable* to expect a «short and sweet» demonstration of genetic relationship, but under certain circumstances — such as significant chronological distance combined with multiple changes in formerly complex phonological systems — it should be expected that such a demonstration is *not* possible, and a more detailed presentation of evidence will be in order.

Such is the case with Sino-Caucasian, where, on one hand, we observe a plethora of bisyllabic stems, widely varying in phonemic inventory and frequently loaded with rare consonant clusters (North Caucasian), on the other, an equally abundant number of monosyllabic stems with severe phonotactic restrictions in every position. It would be very pleasing, of course, if we could, for instance, support the correspondence between PNC *ʕwănʕV 'ear' and PST *nă id. by other etymologies that have the same consonantal structure in PNC, like AV does with his double example of Malay *mata* 'eye', *mati* 'die' — Hawaiian *maka*, *make* id. The problem is that there are simply very few items in PNC that have an initial cluster *ʕw- or an inlaut cluster *-nʕ-, and one would necessarily have to bring structural considerations in the picture, e. g., asserting that clusters of the type «resonant + laryngeal» generally behave the same way in Sino-Tibetan, that is, lose the laryngeal element but preserve the resonant — which is confirmed by numerous examples in both [STAROSTIN 1996] and [STAROSTIN 2005a].

One might reasonably object that for Altaic at least such a situation is not the case: the phonological systems of its daughter branches are not too complex, nor has the root structure in any of them undergone such drastic changes as in Sino-Tibetan. A «short and sweet» demonstration of relationship along AV's lines might, therefore, be expectable here. Fortunately for the Altaic theory, though, and quite contrary to AV's claims that «it is impossible to take a small set of words and demonstrate regularity within that set» [VOVIN 2005: 77], such a demonstration *is* possible. Consider, for instance, this small selection of EDAL parallels between Turkic and Japanese — the two most remotely placed branches of the family (so none of the etymologies can be ascribed to mutual contact):

Turkic	Japanese	Meaning
*bir	*pitə	'one'
*buč-uk	*puta-	'half' (T), 'two' (J)
*č(i)ak-	*tak-	'strike fire' (T), 'burn' (J)
*bus-	*pìsə́-ka	'hide' (T), 'hidden' (J)
*dur-	*tàt-	'stand'
*ïr	*útà	'song'
*dāg	*tàkà-	'high' (J), 'mountain' (T, J)
*čug	*tùkà	'bundle'
*kir	*kìtà-na-	'dirt' (T), 'dirty' (J)
*sïk-	*súkú-ma-	'to press'

Granted, the vocalism is not «self-evidently regular»: it takes extra examples to show that such vocalic correspondences as T *u* : J *a* are also recurrent. However, the vocalic argument has already been discussed in detail in the previous sections. As for the consonant correspondences, though, all of them are perfectly regular and recurrent, and in addition they are established on words with perfect or near-perfect semantic matches[21]. The advanced reader will, of course, be able to construct multiple similar lists out of the large material provided by EDAL, be it between Turkic and Japanese or between any other daughter branches. In addition, the reader is also entitled to evaluate such lists on a statistical basis and deciding for him/herself whether their emergence can be due to chance.

Of course, it is necessary to mention that these kinds of short lists can only have sufficient convincing force if several additional conditions are met. Thus, (a) it must be certified that the correspondences between the items in question are indeed due to genetic relationship and not contact (otherwise, for instance, it would take about five minutes to slap together a similar list proving a genetic relationship between Chinese and Japanese); (b) the compared lexemes should belong to the basic lexicon, yet preferably not be onomatopoeic or 'expressive', in order to emphasize the completely arbitrary link between form and meaning; (c) if comparisons are made between intermediate reconstructions, the latter should be arrived at by conventional comparative means and, preferably, not include highly questionable cases. It is all right to have dubious etymologies with a large enough corpus (both EDAL and the Sino-

---

[21] The selected examples do not, of course, mean that Turkic *d* always corresponds to Japanese *t*, etc.; they represent only a very small slice of available comparisons. However, that is not the point in this particular case: the case is to demonstrate regularity on a small set, and this is exactly what is being done. No small set of cognate items can ever ideally represent the complete system. (With the possible exception of Hawaiian and its thirteen phonemes, of course).

Caucasian dictionary have plenty of those), but short demonstrative lists consisting of less than a dozen entries should, of course, be as clear as possible.

Both AV's Malayan-Hawaiian-Paiwan list and our Turkic-Japanese list are impeccable in these three respects (mutual contact is unlikely due to areal reasons, lexical comparisons involve no sound symbolism, and the compared forms are either taken from attested languages or represent fairly uncontroversial reconstructions).

However, a good example of a list that violates *all* three conditions and should therefore be placed under the highest level of suspicion would be the short selection of Laurent SAGART's parallels between Old Chinese and Austronesian that AV quotes as representing, in his opinion, far more convincing evidence for SAGART's «Sino-Austronesian» theory than for STAROSTIN's «Sino-Caucasian». Let us reproduce that list in full (taken from [VOVIN 2002: 165], but originally stemming from several of SAGART's presentations on the issue [SAGART 2001; SAGART 2002]):

Gloss	PAn	Old Chinese
head	quluH	ᵇhluʔ
elbow	sikuH	ᵇt-r-kuʔ
vomit	utaq	ᵃthaʔ
brain	punuq	ᵃnuʔ
female breast	nunuH	ᵇnoʔ
earth	-taq	ᵃthaʔ

It might be suspicious that the above selection is actually made from a list of but 71 lexical comparisons (as opposed to over a thousand available Sino-Caucasian etymologies), so no matter how regular the correspondences may look, it will not be easy to confirm them on a lot of additional material. Both AV and L. SAGART himself, though, may reasonably object that quantity is preferable over quality, and that six completely regular etymologies constitute better proof than six hundred «partially» regular ones. Yet, as it turns out, the quality of even this list is quite severely exaggerated by AV, because:

— in reality, there is no such Old Chinese form as *trkuʔ* 'elbow'; existing evidence from Middle Chinese only allows us to reconstruct OC 肘 *truʔ* > MC *ʈə́w*. The reconstruction **trkuʔ* is justified by Sagart «on account of the graphic similarity of the early graph for this word with *jiǔ* 九 **ᵇkuʔ* 'nine'» [Sagart 1999: 96–97]. Without going into too much detail on Old Chinese prehistory, let us merely note that modifying reconstructions that are based on systematically applied procedures (MC *ʈ* regularly reflects OC **tr-*) on the grounds of isolated and highly subjective «similarities» does not provide us with highly reliable evidence. Thus, criterion (c) (exclusion of seriously questionable cases) finds itself violated;

— including the word for 'female breast' violates criterion (a), since this word is generally known to belong to the 'expressive' part of the lexicon (plenty of the world's languages have words for 'female breast' that contain nasal resonants *n* or *m*); again, it is possible to present it as supporting evidence, but only within a large corpus of evidence;

— most importantly, though, there is still no way to ascertain whether the remaining four items on the list are really indicative of genetic relationship or are the result of borrowing from Austronesian into Sino-Tibetan or vice-versa. Moreover, even if just two of them happen to be borrowed (e. g. 'vomit' and 'earth'), then the similarities between the other two can be ascribed to chance (especially considering that Sagart, in order to make Old Chinese comparable with Austronesian, also has no choice but to postulate a radical reduction of the root structure in the former) — or, perhaps, be indicative of a much deeper relationship than Sino-Caucasian (Sino-Caucasian + Austric?).

It is somewhat strange that AV, so eager to fall back upon the old idea that each and every non-random similarity between various branches of Altaic should be ascribed to language contact in [Vovin 2005], does not even mention the possibility that similarities between Sino-Tibetan and Austronesian should be explained the same way — despite the fact that this particular matter is actually one of the few moments where «long-rangers» and «mainstream conservatives» (if one is inclined towards such a dichotomy) mostly agree [Starostin & Peiros 1984].

Indeed, if we be allowed to dwell on the issue for another moment, AV's sudden and passionate embracing of Sagart's «Sino-Austronesian» in [Vovin 2002] (as opposed to his definite rejection of it in [Vovin 1997b]) is hard to understand both on general principles as well as in respect to the particular reasons that are quoted. With all due respect to Laurent Sagart, a serious and well-qualified scholar in his own rights, we do not think that he himself would acknowledge the following claims:

«The Sino-Austronesian theory... has managed to acquire supporters from a general historical linguistic audience, and this support seems to grow with every day» [Vovin 2002: 165] — since AV does not give any actual names (probably because there are none — we, at least, are unaware of any general historical linguists who claim to accept Sino-Austronesian as a theory proven beyond reasonable doubt), this phrase has a slight whiff of propaganda to it;

«Sagart is a wonderful philologist... everyone who has read his latest book (Sagart 1999), knows that practically all the examples from Old Chinese are accompanied with at least verse and line references to Old Chinese texts, if not with actual examples» [ibid.]. There is no doubt whatsoever that Sagart is quite competent in Old Chinese, but at least one author of the present article, having spent quite a lot of time working with «The

Roots of Old Chinese», can safely vouch that only *several* examples from Old Chinese in that book are accompanied with textual references, and only in those select few cases where SAGART views it as necessary to confirm some of his (occasionally rather far-flung) hypotheses. We will return to this issue later (in the «philology» section); for now, however, it should be stressed that quite a lot of his linguistic information is actually drawn by SAGART from dictionaries (which AV normally tends to despise, see below) — and is not confirmed by textual references for the simple reason that *the words in question are not actually found in literary texts at all*;

«Sagart's theory is based on impeccable correspondences in morphology between Old Chinese and Austronesian» [ibid.] — see above on the morphology issues; here let it, however, be known that the correspondences in question, presented in SAGART's papers, amount to a scattering of isolated morphological markers (no «paradigmatic» morphology!), and even then, much of it rests on his very questionable internal reconstruction of Chinese derivational prefixes and suffixes, which has been criticized by many authors, including, odd enough, a detailed critical review by Ding Bangxin published in the very issue of JCL that contains AV's unbridled praise of «Sino-Austronesian» [TING PANG-HSIN 2002].

Of course, offering a serious critical analysis of SAGART's theory lies beyond the scope of this paper. The above remarks merely show that AV, in contrasting his «six-word argument» for Malayo-Polynesian with a similar one for «Sino-Austronesian», ends up effectively discrediting it as a working tool (not that it ever was much of one from the beginning, since nobody really establishes genetic relationship based on six words — at best, such demonstrations are mildly suitable for general handbooks designed for non-specialists).

This is not to say, of course, that evidence for genetic relationship cannot be «graded» in any way. In any etymological dictionary there is always a «core» part and a «peripheral» part, with the former fulfilling the role of major proof and the latter adding to the general picture, but being dependent on the former: the stronger the «core», the more reliable becomes the «periphery».

No hundred percent formal definition can be given for these terms, and there is always a debatable margin, but it is generally evident that «core» etymologies are the ones that (a) belong to the basic lexicon, (b) are consistently regular in terms of correspondences (which does not necessarily presuppose complete regularity in every segment in every compared language, as we have just ascertained with Indo-European), (c) do not show much semantic drift, and (d) are fairly well represented within the daughter branches/languages. Overall, such «core» etymologies might constitute from 20% to 50% of the entire dictionary, depending on the time depth between compared languages.

One serious methodological error committed by AV (this concerns particularly his Altaic critique, and is actually applicable to a lot of anti-Altaist rhetoric by other authors as well) is that he concentrates too much on criticizing the «periphery» while at the same time neglecting to criticize the «core». EDAL in its published form contains about 2,800 Altaic etymologies, which vary greatly in terms of reliability, yet the basic «skeleton» of Proto-Altaic reconstruction really rests on the shoulders of maybe three or four hundred etymologies, many of them based on earlier work by pioneers of Altaic and some newly added by the authors. It is exactly these three or four hundred that, in our humble opinion, position the Altaic theory beyond the point of reasonable doubt[22] *and* make it possible to accomodate «weaker», «peripheral» evidence as well; and it is exactly the same principle that underlies any work that bills itself as an *etymological dictionary* rather than a *select comparative lexicon*.

It may be reasonable to ask, then, why is it so that the «core» Altaic etymologies are not specifically highlighted or «rated» within the dictionary so it would be easier for the non-specialist to separate them from the peripheral ones. The reasons for this are many. First, this is not an accepted practice in comparative linguistics: nobody «rates» Indo-European or Uralic etymologies, for instance, and since Altaic is just another language family like any other one, it seemed strange to award it this extra favour. Second, reiterating what was already mentioned above, it is not always easy to decide whether a certain etymology is «strong» or «weak». But most importantly, the strength of many etymologies can be seen on the spot even without giving them a special «promotion». Usually, the best ones are those that figure prominently on Swadesh's 100-wordlist, as well as the ones frequently found in earlier literature and also belonging to the basic lexicon. These form the basis for everything else.

It goes without saying that serious criticism of Altaic should first and foremost be directed at «core» evidence. This point is well understood by some anti-Altaicists, such as, for instance, Stefan Georg; in his short, but, in our humble opinion, much more straightforward than AV's, critical review of EDAL [Georg 2004], he tries — once again — to discredit some of the lexicostatistical matches between Mongolic and Tungusic. The discred-

---

[22] It should be stressed that most of these etymologies can already be found in earlier publications on Altaic (although, within EDAL, some of them have undergone modifications), starting with the classic works of N. Poppe, G. Ramstedt, R. A. Miller et al. and ending with [Starostin 1991], which means that EDAL is essentially justified in treating the Altaic languages as *linguae quarum affinitas est demonstrata*, although the final proof of language relationship should always be presented in the form of a large etymological dictionary (see below).

itation is hardly convincing, as shown in S. Starostin's reply [Starostin 2005b], but at least the direction of the critique, aimed at the very foundations of Altaic, is understandable. AV, however, chooses a different route: focusing on the «peripheral» evidence, yet at the same time cleverly disguising his criticism as objectively aimed at invalidating the very basic premises of EDAL's reconstruction.

The «objective» method chosen by AV is as follows: «I will only use those etymologies that have reflexes in all five 'Altaic' branches, since those etymologies provide much stronger support to the Altaic theory... all the etymologies will include one and the same correspondence: in this case PA non-initial *-ĺ-... that is supposed to be one of the pillars of the Altaic hypothesis» [Vovin 2005: 92]. This leaves the reviewer with 8 etymologies to be criticized («being represented in all 5 branches they objectively represent the best proof»); upon all of them being successively «destroyed», it is implied that any other EDAL etymology will also fall apart when analysed in sufficient detail.

However, it should be perfectly clear that AV's initial statement, though seemingly reasonable upon first glance, should be discarded as completely wrong. The strength of the etymology does not necessarily depend on its being represented in all the branches of the family. For instance, an isogloss between Turkic, Tungusic, and Japanese that is based on perfectly regular correspondences and trivial semantic equations and is also well represented within all or most of the subbranches of Turkic and Tungusic will objectively constitute better proof for Altaic than an isogloss between all five branches that would incorporate occasional irregularities, dubious semantic shifts or would be drastically underrepresented in some of the daughter branches[23].

---

[23] Out of pure interest, a similar selection was chosen by us from the material of the fairly non-controversial Dravidian family: namely, all entries which happen to be represented in all major subbranches (with the exception of the isolated Brahui) and have the inlaut retroflex resonant *-ṛ- (which is statistically even more frequent than PA *-ĺ-). We got 10 etymologies (out of more than 2,000); not a single one of them stemmed from the Swadesh wordlist (more than half, in fact, were from the cultural lexicon), and most had serious phonetic irregularities in at least one or two of the compared branches. In addition, Proto-North Dravidian displayed as many as *nine* ⟨sic!⟩ different reflexations of *-ṛ-: *-y-, *-r-, *-s-, *-l-, *-n-, *-c-, *-ṛ-, *-ṭ-, and zero (only -s- occurred twice). It goes without saying that, if no additional material is taken into account, one might easily suggest that North Dravidian languages should be considered unrelated to the rest of the family.

A similar experiment was conducted for Indo-European, where we randomly selected a phoneme of moderate frequency (namely, PIE *g^w) and looked for roots that contain it (in *any* position) and are met in five of the most «important» branches of IE: Indo-Iranian, Greek, Latin, Germanic, and Slavic. The count ended at

As a result, AV's «objective» criterion, which leads him to selecting only 8 etymologies out of 76 containing PA *-*ĺ̥*-, results in only two «core» etymologies being criticized: *sīĺ̥a* 'sharp stick, tooth' and *tị̯ól̥̀ì* 'stone', both of which belong to the basic lexicon — and it is hardly surprising that it is in these particular cases that AV's criticisms seem especially artificial to us (all of them will be discussed below). At the same time, no attention is paid to such obvious matches as PT *jāĺ* 'young' : PM *ǯalaɣu* id.; PT *keĺč* 'belt' : PJ *kǝ́sí* 'waist'; PT *jāĺ* 'age, year' : PM *na-su* id. : PK *náh* 'age'; PT *jāĺ* : PM *nilaɣu* : PTM *ń(i)ali-* : PK *năr* (all meaning 'raw' or 'fresh'); PM *hulaɣan* : PTM *pula-* : PK *pìrk-* (all meaning 'red'); PM *ele-sü* 'sand' : PJ *ísá-* id.; PT *Kɨĺ* 'winter' : PTM *gil-* 'cold' and quite a few others where phonetics, semantics, and representativeness match better than in some of the admittedly weaker cases dismissed by AV. (It should probably also be mentioned that AV confuses «non-initial» and «intervocalic» — there are several more etymologies with inlaut *-*ĺ̥*-* that are not discussed simply because the consonant forms part of a cluster, even though reflexes are found in all daughter branches, e. g. *kèĺčo* 'to scrape, rub', *kʿoĺba* 'couple', *pìĺǯi* 'to become overripe' et al.).

Therefore, the «dismissal by induction» principle advocated in [Vovin 2005] does not really work, because, either intentionally or through a simple methodological failure, the majority of the criticisms are directed at peripheral rather than core evidence. In addition, contrary to the frequently met assertion that «ten bad etymologies are worse evidence than one good one», size *does* matter, if only because, in this particular case, for instance, it is more reasonable to accept *one* explanation that accounts for 76 cases of systematic similarities than to find several dozen different explanations that would interpret each one of these as a case of accidental resemblance (or, at best, borrowing — although borrowing is actually invocated by AV in a surprisingly small percent of cases).

---

six (*g^wou-* 'cow', *nog^w-* 'naked', *g^wer-* 'to gather', *g^weị̯-* 'to live', *g^wer-* 'to swallow', *g^werụ̯-* 'throat, neck') and the absolute majority of the etymologies had various phonetic, morphological and semantic problems in at least several branches.

Thus, Slavic *govędo* 'cattle' cannot be segmented into *gov-ędo* on Slavic soil; Old Indian *gaṇá-* 'flock, troop', sometimes derived from *g^wer-* 'gather', is hard to reconcile phonologically with that root; Greek γυμνός 'naked' features an irregular assimilation; Old Indian *gárgara-* (supposedly < PIE *g^wer(ụ̯)-* 'throat, neck') does not really mean 'gullet', as is supposed in some dictionaries, including [WP], but 'whirlpool', so the semantics is poor, etc. etc.

All of this leads us to conclude that AV has, in effect, discovered a fail-safe method of «invalidating» *any* proposal of genetic relationship — and in a manner that is superficially quite convincing to the general audience, too.

The same applies to «core etymologies»: namely, *reasonability* of the criticism. Thus, S. GEORG's main point in [GEORG 2004] is well understood: it is implied that the authors of EDAL are twisting the intermediate reconstructions in such a way that they look better together (e. g., reconstructing PTM *ńia-sa 'eye' instead of the more traditional *jasa so it could form a match with Mongolic *ni-dü, etc.; for more details on that particular etymology, see below).

However, as long as there is at least *some* basis to such «twisting» — and it is always based on real evidence from languages, no matter how scarce or indirect it may seem — the proper question is: is it true that this kind of methodology would allow us to «prove» *any* relationship? For instance, is it at all possible to take the 100-wordlist for, say, Proto-Germanic, «twist» some of the etymologies based on alternative reconstructions or dialectal irregularities, and through this «twisted» reconstruction demonstrate that Proto-Germanic really has much more in common with, say, Proto-Samoyed or Proto-Omotic, than it has with any other branch of Indo-European?

We, at least, know of no such cases, and strongly suspect that nothing of the sort is even remotely possible. A very clear demonstration can be made on the «Sino-Austronesian» example. Thus, no matter how much L. SAGART has managed to «twist» his reconstruction of Old Chinese (and it is substantially different from that of his predecessors and controversial on many points), he is still able to score but *five* matches between Austronesian and Old Chinese on the «ultra-stable» 35 item section of the SWADESH wordlist as defined by S. YAKHONTOV [SAGART 2002] (two of them *still* with irregularities in correspondences; the sixth comparison, OC 骨 *kut* : PAN *kukut* 'bone' does not count because the basic meaning of the word in PAN is really 'joint') — in comparison to *thirteen* matches between OC and North Caucasian [STAROSTIN 1995], even though the latter consists of no more than three dozen languages, whereas the very hugeness of Austronesian seems to provide far more promising opportunities for comparison, with many different roots often «competing» for occupying the Proto-Austronesian slot for notion so-and-so. Obviously, SAGART can criticize his opponent for tailoring the reconstructions to fit his own tastes, but that is not the issue, since the same criticism can be addressed to SAGART himself. The issue is that STAROSTIN's OC reconstructions (justified on Chinese evidence) fit in with North Caucasian better than SAGART's OC reconstructions (justified on *alternative* Chinese evidence) fit in with Austronesian, and that there really seems to be no way that *any* more or less credible OC reconstruction would fit in with Austronesian.

Likewise, before agreeing with AV that all of EDAL's Altaic comparisons, including «core» ones, should be dismantled based on a miriad of reasons, we would like to see at least one example of a comparison that (a) should be conducted between two clearly non-related (or very distantly related, which, from

a purely practical point of view, is frequently the same thing) language families, such as Germanic and Omotic — or, better still, Germanic and Fenno-Ugric (since these two are known to have been in contact with each other); (b) had the same convincing force as Altaic, that is, was based on more or less regular correspondences (at least for consonants); and (c) yielded as many exact semantic matches on the 100- and 35-wordlists. The compared reconstructions themselves may be as controversial as one likes, provided the controversial solutions are based on at least *some* evidence stemming from within the family.

Until this is done — and none of the known attempts so far have worked, to the best of our knowledge[24] — we will continue to view the Altaic (and Sino-Caucasian) theories as confirmed by evidence beyond a reasonable doubt. The phrase «the very fact that it is possible to compile a dictionary of common Altaic heritage appears to be a proof of the validity of the Altaic theory» in EDAL may have come under deserved criticism from both AV and Stefan Georg [Georg 2004], but we may use this opportunity to reformulate it in a more explicit way: «The very fact that it is possible to present *a realistic comparative-historical model of Proto-Altaic that is based on evidence from daughter languages without contradicting it* appears to be a proof of the validity of the Altaic theory»[25].

V. *Internal vs. external reconstruction.*

We are now coming to a group of AV's arguments that, in comparison to the ones discussed above, are less crucial from a purely theoretical point of view; nevertheless, since it is these arguments that AV spends the most time and space laying out, a detailed response is in order here as well, including analysis of concrete examples.

---

[24] The most interesting of such attempts was conducted by AV himself between Japanese and Uto-Aztecan [Vovin 1999], with AV coming to the conclusion that the lexicostatistical procedure, while working well for Japanese and Turkic, yields catastrophic results in the other case. One could also mention oddities like the late L. Trask's mock-English-Basque comparisons [Trask 1996], which also do not prove anything since they are not based on regular correspondences at all (e. g. within three successively encountered words we find Basque *h* : English *h*, Basque *h* : English *k*, and Basque *k* : English *k*!).

[25] We should emphasize very strongly that a relationship hypothesis *can* be proposed between *any* two or more given languages or language families, no matter how odd it may look on an intuitive level. In order to be corroborated, however, the hypothesis has to be transformed into a working comparative-historical model, i. e. it must be shown that the compared languages may fit within the requirements prescribed by the comparative method, primary among them the presence of systematic correspondences that can be established on a large part of the basic lexicon. These requirements are fulfilled for both Altaic and Sino-Caucasian.

AV: «The internal and comparative reconstructions within an uncontroversial genetic grouping must be thoroughly done before attempting any external comparisons... Under no circumstances can external etymologies take any precedence over internal etymologies based on thorough analysis of a language itself» [VOVIN 2005: 74].

The two parts of this statement look similar, but are in fact widely different. One should always distinguish — and this is a common place in mainstream comparative linguistics — between *comparative* reconstruction, in which a protolanguage is being arrived at through regular correspondences in daughter languages, and *internal* reconstruction, in which an earlier, unattested stage of a given language (or protolanguage) is arrived at through the means of that language alone, usually based on the analysis of its archaic features, systematic irregularities, and derivation.

The first type of reconstruction is normally achieved through a strict, objective procedure that follows the same universally accepted methods regardless of the language family that is being reconstructed. This is, of course, the type of reconstruction that the authors of EDAL/SCH cannot help but acknowledge, and all of the Altaic reconstructions in EDAL are faithfully based on intermediate reconstructions for Proto-Turkic, Proto-Mongolic, etc., with the exception of those cases where the Altaic form has been preserved in but one language (and even then, of course, it has to be compatible through regular correspondences).

Pure *internal* reconstruction, however, is a different matter. In most cases, there are no standard, universal procedures or principles to help the researcher deduce one stage of the language from another. Thus any internal reconstruction is, by definition, subjective, and while that does not invalidate the method per se, it is up to external comparison to offer definitive evidence in favour of or against it.

If we were to take AV's statements at face value, we would, of course, have to analyse the classic example of Russian *zontik* 'umbrella' in a strict, rigorous way based on internal Russian evidence: consisting of the root *zont* (which exists in the language all by itself) and the productive diminutive suffix *-ik*, i. e. as originally 'little umbrella'. Fortunately, we know from well-attested sources that the word was originally borrowed from Dutch *zondek*, lit. 'suncover', and that the form *zont* is the result of a later morphological reinterpretation. But if «under no circumstances can external etymologies take precedence over internal», well, too bad for historical facts; instead, we should probably start hunting for the *true* etymology of the old Russian nominal root *zont*.

There are, of course, numerous (and fairly «mainstream» in origin) cases where internal reconstruction was triggered or made possible only through external comparison. Indo-European comparative linguistics (the

most «mainstream» branch of comparative linguistics) is, in particular, full of such cases, since Proto-Indo-European itself was originally reconstructed based not on internal reconstructions of its daughter branches (Indo-Iranian, Germanic, Slavic, etc.), but on comparison of select archaic languages within these branches. There is no way, for instance, that one might arrive at a proper reconstruction of Proto-Germanic accentuation without bringing in evidence from Greek or Old Indian.

In purely lexical terms the possibilities of successfully etymologizing two genetically different words on the basis of one another would also soar, were it not for external comparisons. We could, for instance, easily derive Skt. *īṣā* ʻpoles of a carriageʼ from the verbal root *īṣ-* ʻto moveʼ (the semantic derivation is quite transparent) — if not for the fact that the former goes back to PIE **(H)oi̯es-* ʻshaftʼ (cf. Hittite *hissa-*, Slav. **oje, -ese*, etc.; WP II: 298), whereas the latter represents PIE **eis-* ʻto move (rapidly), to be agitatedʼ (cf. Avestan *aēš-* ʻto hurryʼ, Latin *īra < *eis-a* ʻangerʼ, etc.; WP I: 299–301). Or, turning to Germanic, we could brilliantly deduce Proto-Germanic **fawjan* ʻto siftʼ (> OHG *fowen*) from **fawa-* ʻfewʼ, through a crystal clear derivation process (ʻto siftʼ = ʻto make fewerʼ) with a perfectly productive suffix; fortunately, it is the external parallels that help us know that the former should rather be traced back to PIE **peu̯(ə)-* ʻto clean, to siftʼ [WP II: 14], while the latter goes back to **pau̯- (*pəu̯-)* ʻsmallʼ [WP II: 75].

Of course, it is only possible to accept that these stems should be separated from each other because it has already been established, on the basis of less ambiguous evidence, that Sanskrit or Proto-Germanic are related to Hittite, Slavic, Latin, etc., which provides us with sort of a *carte blanche* to reject certain arguable «internally based» solutions in favour of «externally based» ones when we feel the latter are more reasonable. In other words, we probably will not include these and other examples in our small stock of «core etymologies»; they will be building up on the already well-grounded theory rather than serving as its foundation.

But the interesting thing is that EDAL actually follows this exact principle. Yes, there are etymologies where external solutions are preferred over internal ones, just as is the case with every other linguistic family, «controversial» or not. However, they are in a minority compared to etymologies where a possible internal solution is simply unknown. Yet let us see what AV has to say about that:

«When quite an obvious case that can be easily explained internally contradicts a proposed external etymology, the former is discarded as being ʻmost probably a folk etymology in the light of external evidenceʼ, a typical phrase that is repeated time and again throughout the dictionary like a magical mantra» [Vovin 2005: 75].

Just to oblige AV, we actually performed a statistical check on how many times the words «folk etymology» (together with derivates like «folk-etymologically», etc.) are met in different sections of the vocabulary. The results are as follows: 18 times in the general Altaic section (out of 2805), 17 times in the Turkic section (out of 2017), 5 times in the Mongolic section (out of 2172), 0 times in the Tungus-Manchu section (out of 2435), 5 times in the Korean section (out of 1204), 14 times in the Japanese section (out of 1705). A few of these occurrences overlap within the same etymologies, and in addition, it is not every time a «folk etymology» is mentioned that it is actually asserted (e. g. for PA *pằjkù 'owl': «cf. some similar bird names: Bur. *bug-bātar* 'owl', if the analysis 'demon-hero' is a folk etymology» [EDAL: 1073] — the main comparison is with PM *beg- 'barred owl', and the etymology is an uncontroversial one regardless of what one thinks of these additional parallels).

Altogether, that makes for, at most, *fifty* etymologies that are partially (only partially!) based on rejection of more traditional internal etymologies, out of an overall number of 2805. To be completely fair, one should probably add all the etymologies which make reference to «secondary contamination» of two different roots in particular branches (although that is a somewhat different issue), which would account for about fifty more «controversial etymologies». Even if we reject every single one of them (and there is really no reason why we should), this would not even in the gentlest way possible shake the foundations of the Altaic theory.

But, just for the sake of it, let us take a look at a couple actual cases where «folk etymology» is involved. Japanese in particular is a major culprit here, which should come as no surprise: with its isolated status and particularly with the radical simplification of its consonant inventory and root structure (as compared to Altaic), resulting in a lot of secondary homonymous lexemes, it is generally much easier to reinterpret old Japanese items on Japanese territory than, say, old Tungusic items on Tungusic turf — and since many such reinterpretations date back to old Japanese traditions, this makes it particularly uncomfortable to accept or even suggest new solutions, no matter how odd the tradition might look like[26].

Thus, the Japanese parallel for PA *kʼi̯ŏbe 'touchwood, tree fungus' (safely reconstructed on the basis of PT *Kiab 'tree fungus; dry grass', PM *köb- 'moss', and PTM *xub(u)te 'touchwood; rotten' [EDAL 802]) is glossed as (Tokyo only) *kino-ko* < PJ *kua. It is explicitly stated that «the word is

---

[26] A particularly radical approach has been advocated by J. JANHUNEN, who goes so far as to propose that *all* Japanese roots must originally stem from monosyllabic morphemes [JANHUNEN 1994]; this fantastic idea, based exclusively on the author's understanding of «internal reconstruction», has been convincingly rebuked in quite a few of AV's own publications, most notably [VOVIN 1994].

usually analysed as tree-child» (*ki-no-ko*), but the latter interpretation is rejected as a folk etymology. Now, first and foremost, this is being done simply because such a derivation is completely unprecedented. Unless we have serious reasons to believe that the Japanese language, unlike any other language on earth, shifted its meanings primarily through poetic metaphor (and *numerous* internal Japanese etymologies are formed that way), we have little reason to assume that the *-ko* element in *kino-ko* (the first two syllables may indeed be intepreted as 'of the tree') has anything to do with PJ **kúa* 'child'. From a typological point of view, the word for 'mushroom' may, of course, have different origins, but in most of the understood cases it is connected with such notions as 'excrescence, knob' or 'sponge, slime'. We have, therefore, all the reasons and more to doubt the 'tree-child' interpretation just as AV doubts the 'horn / gum' connection (the latter, in fact, is much more plausible).

The scenario here is exceedingly simple: PA **kŭŋi* 'child' and **kʼi̯ŏbe* 'fungus' began to sound the same in PJ, due to the merger of **k-* and **kʼ-* and regular elision of the inlaut consonants **-ŋ-* and **-b-*, and *ki-no-* 'of the tree' was added to the latter in order to avoid homonymity (just as it is sometimes added to *me* 'bud, sprout' > *ki-no-me* 'tree-bud').

Outside Japanese territory, «folk-style» internal etymologization is quite popular in Turkic, where it frequently follows the Indoeuropeanist scenario of deriving as many nominal stems as possible from verbal ones, regardless of how arbitrary the connection may be. It is obvious that, for instance, an external etymology that ties PT **bors(m)uk* 'badger' in with PM **borki* 'old badger' and PK **ùsìrk* 'badger' should at the very least be regarded on equal footing with one that tries to derive the PT stem from **bur-sɨ-* 'to stink' [EDAL 374]; personally, we happen to think that, in spite of certain phonetic difficulties (the inlaut *-m-* is hard to account for, but neither is it easily explainable internally), the external etymology is by far preferable, since the semantic connection is more direct (not to mention that a badger is hardly any 'stinkier' than any of its colleagues in the carnivorous department).

Of course, still more of such «folk etymologies» can be found within AV's review itself. Such is, for instance, the derivation of PK **tōrh* 'stone' (MK *:twolh* in AV's notation) from the verbal root *:twol-* 'turn (intr.)', which claims to follow [Martin 1996: 36]. What is not mentioned is that Martin himself discusses this possibility only briefly («it may be that the *h* < *k* was a diminutive suffix in the two Korean words. Or, the word for 'stone' may be a deverbative noun from the intransitive stem **twolo-...*»), and that the derivation has not at all found «general acceptance» — most probably due to the fact that such a semantic development is utterly unprecedented. Not

everything that 'turns' is a 'stone', and in every one of the very few known cases where we have etymologically related words for 'turn' and 'stone', the latter is never the main word for 'stone', but rather a specific term for 'round stone', 'pebble' (cf. Lettish *uõla* id. < PIE *wel- 'to roll'). Of course, it is not impossible to imagine a scenario according to which a word like Lettish *uõla* would eventually become 'stone' in general, but such a scenario would require a lot of historical time, whereas the hypothesis of «Martin-Vovin» would seem to indicate quite a recent development. We will therefore stick to the old etymology that pairs MK *:twolh* 'stone' with PT *diāĺ* 'stone' — based on quite regular correspondences, too.

In other cases, previously suggested internal etymologies seem more reasonable, yet the alternative of an external explanation still remains; the reader may incline towards the former or the latter, but, simply because one is more probable than the other, this does not mean that the external solution should not even be mentioned. So, AV's lengthy and painstaking explanation of why OJ *pani-wa* 'clay figure' can be perfectly well derived from *pani* 'clay' + *wa* 'ring' [Vovin 2005: 120] is rather expendable, since EDAL's authors are familiar with that derivation [EDAL: 1084] and have no objections against it (otherwise those would be stated), merely suggesting that the Japanese form *may* be the result of a semantic reinterpretation of an old Altaic stem, a situation that is by all means not unprecedented among the world's languages. The entire etymology (PA *pélabanV* 'stone or clay figure') certainly does not belong to the «core» layer, and is a tempting, but indeed somewhat dubious addition to the overall corpus, and no general or particular issues of the Altaic theory depend on whether we eventually accept or reject it.

Overall, we might count about a couple dozen such cases at best, and we should probably be grateful to AV for pointing out their weaknesses to EDAL's potential audiences — if not for the fact that these weaknesses are subtly extrapolated over the whole dictionary; according to AV, the etymology for *pélabanV* «represents all problems typical for the EDAL», a statement that we find hard to agree with. Problematic, yes, but hardly typical (see the statistics above).

We now turn to the second part of the issue: the supposedly arbitrary segmentation of attested or reconstructed stems into original «roots» and «suffixes» to facilitate further comparison. This is what AV has to say on the subject:

«It is fallacious to draw morphemic boundaries in a given language and create a non-existing 'morpheme' on the basis that the same 'morpheme' can be found in its distant relative, especially when the relationship is not proven. It is even more fallacious to segment the root *alga* in the language A as *al-ga* and the root *aldu* in the language B as *al-du* without

any internal evidence in A and B for suffixes *-ga* and *-du* respectively, just because A and B happen to share the common part *al»* [Vovin 2005: 74].

Apparently AV is certain that the principles that he exposes here are in perfect accordance with «mainstream comparative linguistics», or else they would not be stated with such assuredness. Unfortunately, they are not, or, at least, not quite. One example will suffice for an introduction. In Proto-Indoeuropean, two basic stems for the meaning 'hand' can be reconstructed: *ĝhes(o)r-* (based on Hittite *kessar*[27], Greek χείρ, Tocharian *tsar*) and *ĝhesto-* (based on Old Indian *hásta-* and some old Latin and Baltic compounds; the reconstructions are given according to [Pokorny 1959/1969: 447]). Within each individual subgroup, there is not a shred of evidence that the stems in question can be segmented, respectively, into *ĝhes-(o)r-* and *ĝhes-to-*, after which the two «roots» can be compared and the «suffixes» just thrown away. Certainly such suffixes, particularly *-to-*, do exist, but they are mostly used to form nominal derivates from verbal roots, and there is no separate verbal root *ĝhes-* either in any of these subgroups or even in Proto-Indoeuropean.

Based on AV's theoretical assumptions and description of the very similar hypothetical *al-ga/al-du* case, and also judging by analogy with those Altaic etymologies that he rejects on similar ground, we should consider it our primary hypothesis that *ĝhes(o)r-* and *ĝhesto-* represent not only two different stems, but probably two different Proto-IndoEuropean *roots*, unsegmentable in any way and having nothing to do with each other. Yet this is not being done. The two stems are placed together in [Pokorny 1959/1969] (under separate entries, but marked by numbers 1 and 2 which indicate an etymological connection) and are constantly treated as representing one etymology without any major headache by specialists in the field. Given that similar examples crop up quite frequently, AV's general statement is just plain wrong.

What is the reason for suggesting a common origin for these two stems? One such reason may be the general theory of root structure in Indo-European: since it is commonly accepted that most of the roots were monosyllabic (of the CV(C)C type), based on various sorts of evidence, the final syllables in these two stems are automatically analysed as old suffixes. However, this reasoning alone would be insufficient: it is well known that *nulla regula sine exceptione*, and it is quite possible that Indo-European might have had a limited number of bisyllabic roots (even though many Indoeuropeanists would be contesting that idea).

---

[27] It should be stressed that the *-ar* in Hittite is not even heteroclitical, as in, e. g., *watar* 'water', g. sg. *weten-as*; cf. the well-attested dative form *ki-is-sa-ri*. So there is really no way the word could be segmented on Hittite soil.

It is more reasonable to say here that, since the two stems are (a) in complementary distribution across the languages, (b) show perfect phonetic correspondences within the CVC part that could be seen as the original root, and (c) form perfect semantic matches, the *probability* of their being related significantly exceeds that of the opposite. In the light of this, it is *permissible* for us to suggest that both *-(o)r-* in the first case and *-to-* in the second one represent some sort of old root extensions with meanings that we are unable to determine.

A different question is: can we make use of the **ĝhes-(o)r-* / **ĝhes-to-* case in order to *prove* Indo-European relationship, i. e. place this etymology on equal footing with, e. g., the examples from [REFORMATSKY] given above? This again reverts us to the problem of «core» and «peripheral» etymologies. Obviously, if *all* the etymologies that we suggest involve this kind of arbitrary segmentation, this would look suspicious. But that is not the case. AV mentions the «skyrocketing number of cases» in which the authors «willingly ignore the structure and/or morphology of a given language» [VOVIN 2005: 74–75], but the question here is not exactly how «skyrocketing» is the number of etymologies that are problematic in that respect, but rather how large is the number of etymologies that are *not*. And, in our opinion, it is sufficiently large to justify the «arbitrarily segmented» cases — even though, as will be shown below, the segmentation is nowhere near as arbitrary as it may seem to AV.

Before proceeding, though, it may be useful to point out that in several cases AV violates his own rules by implying that «arbitrary segmentation» *can* be allowed — but only if it is performed for the sake of providing the word with an internal etymology. Thus, in order to prove the link between OJ *isi* 'stone' and *isô* 'rock' (and thus to remove the former from the Altaic etymology for 'stone'), AV writes: «OJ *isi* 'stone' probably < **isï* < **isu-i* with a suffix **-i*, and OJ *isô* 'rock' < **isu-a*» [VOVIN 2005: 112]. This would imply that OJ not only had a well-known nominalizing suffix *-i* (on which see above), but also a different *-i* that was used to form nouns from nouns with an unclear function, and, in addition to that, it also had a suffix *-a* that had a similar usage. To the best of our knowledge, however, there is no otherwise clear evidence for the existence of such suffixes in OJ (and if there is, why has AV failed to mention it?)[28].

Elsewhere, AV forbids the authors of EDAL to «arbitrarily segment» the PJ verb **sásír-* 'to slander' into **sás-* and **-ir-*, stating that there is no evidence for postulating a suffix **-ir-* for Japanese [VOVIN 2005: 110]. It should be

---

[28] MARTIN [1987: 65] mentions the possibility of a fossilized suffix *-a-* that was used to derive nouns from verbs, e. g. *hor-a* 'cave' < *hor-* 'to dig', but no examples are given where *-a-*-containing forms can be derived from nouns.

noted, though, that since the exact reflexation of the second syllable vocalism in Japanese is unknown, the «arbitrary segmentation» should probably be established as *sə́sí-r-*, and that there is abundant evidence for a verbal suffix *-r-* in the language. Instead of this simple solution, AV suggests an extremely complex scenario according to which the word should be segmented as *sə́-sír-*, with the second morpheme representing «an obsolete suffix». To make this look reasonable, he claims the presence of such a suffix in MJ *nono-sir-* 'to make big noise, to shout' (as opposed to MJ *nono-mek-* 'to make noise talking in a loud voice'), and then «discovers» the same opposition between *sə́-sír-* and MJ *so-mek-*, *sè-mèk-* 'to accuse/reproach holding a grudge'.

Apparently, it does not bother AV that (a) he is establishing the suffix *-sir-* based on but two examples (unless he has in mind some similar ideas that would explain all Japanese words like *hashiru* 'run', *mushiru* 'pluck', etc. as combinations of monosyllabic roots *ha-*, *mu-*, etc., with his «obsolete suffix»); (b) no explanation whatsoever is produced as to the possible meaning or function of the supposed *-sir-*; c) the relation between *sə́sír-* and MJ *sèmèk-* (for the other proposed form, *somek-*, we have been unable to locate a reliable source) is, at best, questionable, since neither the vocalism nor the accentuation match too well[29]. None of this matters — probably because a complex and bewildering internal etymology that nevertheless stays firmly grounded in Japanese soil is *a priori* better than a simple solution that takes us beyond that territory.

We are, of course, not saying here that these internal hypotheses have no right to exist and have always to be rejected in favour of external explications. But neither do we find it acceptable to reject evidence for Altaic every time we cannot explain a supposed morpheme boundary on Japanese, Mongolic, or Tungusic territory. Let us remind the reader for a second that the date of splitting for Proto-Altaic, based on lexicostatistical calculations, is estimated at approximately 5,000 BC. This means that all of its daughter branches had undergone a period of at least four thousand years of independent development (and some, like Korean and Mongolic, close to six) before splitting in their turn, thus allowing us to reconstruct intermediate stages. With agglutination as one of the major productive means of most Altaic languages (and probably characteristic of Proto-Altaic as well), it would be surprising *not* to expect quite a lot of fossilized derivational suffixes — with meanings that are extremely hard to reconstruct, since the suffixes are no longer productive.

In one of his footnotes, AV describes the dangers of assuming too much of this «fossilization» in the following way: «An 'unproductive suffix' of body

---

[29] The MJ form quoted by AV (*semekikemu*) is related to modern Japanese *semegu* 'quarrel, struggle; maltreat, molest', which, in turn, may be related to *semeru* 'attack; blame, reproach' [Martin 1987: 749]. If so, there is no reason at all to suggest the presence of the *-mek-* suffix here.

parts *-du(n)/-dü(n)* in Mongolic fares no better than an 'unproductive prefix' of body parts *g-* in Russian, e. g. one may suggest that the following words all contain this 'prefix': *g-olova* 'head', *g-laz* 'eye', *g-uba* 'lip', *g-rud'* 'chest', *g-orlo* 'throat', *g-lotka* 'gullet', *g-ortan'* 'larynx', *g-olen'* 'shin', *g-landy* 'tonsils', and *g-orb* 'hump'» [VOVIN 2005: 105]. The analogy, however, does not work; in fact, contrary to AV's intentions, it clearly and expressly shows the importance of tempering any internal reconstruction with external comparison.

Namely, if the Russian language had no known related languages — or, at least, its relationship with anything else were to be «controversial» among «mainstream» linguists — we would be perfectly justified in carrying out an internal reconstruction that would *suggest* that the presence of so many body parts beginning with *g-* might indicate its status as that of a fossilized body part prefix. (Some nice internal etymologies could be based on that, too, for instance, coupling *g-orlo* 'throat' with *or-at'* 'to yell' and suchlike). The reason we are one hundred percent certain that in all these cases *g-* not only forms part of the roots in modern as well as Old Russian, but had been part of these roots for thousands of years before, is that we can *compare* these words to their *external cognates* in Slavic, Baltic, and other Indo-European languages and observe that the same consonant, or one of its regular correlates, is present there as well; in addition, it can be demonstrated through this comparison that in different cases Russian *g-* goes back to several different Indo-European consonants (*$*g$-*, *$*g^w$-*), completely destroying any possible grounds for the initial suggestion.

The Mongolic situation is not that much different. Here, upon noticing that several nominal stems all end in *-du(n)/-dü(n)*[30], it also *may* be suggested that this sequence represents an old fossilized suffix, and this time the suggestion is rather confirmed than disproved by external evidence: cf. *$*ni$-dü* 'eye' vs. PTM *$*ńia$-sa* et al., *$*ho$-du* 'star' vs. PJ *$*pəsí$* et al., *$*si$-dü* 'tooth' vs. PT *$*sīĺ$* et al., *$*mo$-du* 'tree' vs. PJ *$*mə́rí$* et al. (all four cases form perfect semantic matches).

AV, of course, uses this opportunity to accuse the authors of circular logic[31]: we are using this evidence to prove the existence of «Altaic», when

---

[30] It is unclear why AV keeps insisting on labeling this morpheme as a «suffix of body parts»; none of the authors of EDAL ever pretended that it is exclusively joined to words denoting body parts, and neither did N. POPPE, who had first drawn attention to it in [POPPE 1973: 233–234].

[31] «Circular logic» appears quite frequently as an argument against Altaistics and long-range comparison in general, usually in the following form: it is inadmissible to use external evidence from family *A* to resolve ambivalent issues in family *B*, if the genetic relationship between *A* and *B* has not been convincingly established on non-ambivalent evidence. In the absolute majority of cases, however, if families *A* and *B* are not related, it is simply *theoretically impossible* to resolve any

in reality it should only be put forth once «Altaic» has already been proven on the basis of better evidence. To this we can only reply by suggesting a little statistical test: if AV takes any four Russian words that begin with *g-* (not necessarily body parts, but preferably belonging to the basic lexicon, considering that all four Mongolic items belong to the 100-wordlist), throws the *g-* away, and then shows us that the rest of the forms correspond to four words with *identical* meaning in *any* other language, we will admit that he has made a strong case[32]. Until that is done, we are completely justified in our segmenting the Mongolic items exactly the way it is done in EDAL and using that as strong (but, of course, far from only) evidence for Altaic[33].

Let us now individually consider some other cases in which the authors of EDAL are accused of «arbitrary segmentation».

(1) PJ **kàtànà* 'sword, knife' is derived in EDAL (p. 531) from PA **gǎř[à]* 'sharp edge'; AV rejects the derivation in favour of an internal etymology that analyzes the word as *kata-* 'one' + *-na* 'blade', reasonably pointing out that the word normally refers to a single-bladed sword (as opposed to OJ *turugî* 'double-edged sword') already in OJ. This etymology, which he calls «quite transparent» in his review, did not, however, seem all *that* transparent to him a decade earlier, when he wrote: «there is not enough evidence for hypothetical PJ **na* 'blade', reconstructed on the basis of *-na* in *katana* 'small sword', 'knife' and *kana* 'plane' by Martin» [Vovin 1993a: 130].

One wonders, of course, what exactly is that extra evidence for PJ **na* that came up over the last ten years and made Martin's etymology so transparent. However, we are not really putting that against AV, since we ourselves believe that such a word is *potentially* reconstructible (it is at least suggested in several dictionaries, e. g. [JDB: 512]). Yet even taking into account that the internal analysis goes way beyond Martin (the same explanation is given in [JDB: 193]), there is still a major problem with this analysis: in the oldest texts

---

issues in *B* based on *A*. In reality, such «circular logic» is not at all different from the «circular logic» used, for instance, in the analysis of numerous ancient scripts, many of which could only have been deciphered based on a prior *assumption* that they were written in a certain language (such as Linear B).

[32] We can even try to disregard such important circumstances as, for instance, the biphonemic status of Mongolic *-du* vs. monophonemic status of Russian *g-*, as well as the fact that it is normal for Mongolic to have adnominal suffixes, but not normal for Russian to have adnominal prefixes.

[33] An alternative hypothesis, currently espoused by O. Mudrak [Mudrak 2007], is that *-d(u)-* is historically not a suffix in Mongolic, but part of the root, and that Mongolic *-d-* is a regular reflexation of PA **-ĺ-* in the intervocalic position. This is questionable (although it helps to avoid any accusations of «arbitrary segmentation»), but, in any case, the etymologies remain valid regardless of whichever solution one prefers.

(as well as later ones) the word is *never* transcribed with the character 片, as is the common case with all similar compounds quoted by AV, such as 片足 *kata-asi* 'one foot', 片側 *kata-gawa* 'one side', etc. Instead, it is always written semantographically as simply 刀 'sword' or 小刀 'small sword', as opposed to *tsurugi*, written either as 劍 '(Chinese) sword' or 大刀 'big sword'.

Given this consistent spelling, it can be concluded that (a) already in the OJ period the word *kàtànà* was not perceived as having anything to do with *kata-* 'one' and (b) the main difference between it and *turugî* was seen as their respective size — even though sharpness of the edges might have been a criterion as well. As for the 'one-blade' etymology, it may have become popular through analogy with the well-attested compound *kata-ha* 'single blade(d sword)', appearing in the language much later and this time faithfully transcribed as 片刃.

To this must be added that we really know little about what kind of weapon was designated in OJ by the term *kàtànà*. Its current and most well-known incarnation — the long curved single-edged samurai sword — is actually quite late (coming into heavy use around the 15th century). Before that, the primary fashion for single-edged swords was the *tachi* (太刀), but even this is generally thought to have been carried over from China not earlier than the 10th century (around the same time that the famous dictionary *Wamyō ruijushō* [倭名類聚抄] was compiled, where the 'one-bladed' definition of *kàtànà* is met for the first time). And yet we already have the term *kàtànà* attested throughout the *Kojiki* (古事記, 8th century). Whatever the ultimate solution may be, the *kàtànà* of these early monuments should hardly be associated with the iconic samurai sword of the 15th century, meaning that AV's confidence here is slightly anachronistic.

In the light of this, we are at peace with the idea that *-nà* in *kàtànà* is not a suffix, but really an old word meaning 'blade', although it had already fallen out of independent usage long before the OJ period (and, therefore, could hardly have been borrowed from Korean — see below). As for what concerns the first part of the compound, an external etymology that links it to PK **kárh* 'sword' (actually following MARTIN [1966: 251], who reconstructs Proto-Korean-Japanese **khal-nal* 'blade of sword') and then further with several semantically more distant parallels in Turkic and TM[34],

---

[34] AV calls the TM parallel **gara* 'bough, stick' «incredible»; it might not be out of place, then, to call to mind some of the well-established parallels of the English word *stick*, e. g. Icelandic *staki* 'pole, spear', or, while we're at it, cf. PIE **deru-* 'tree' > Greek δόρυ 'spear'. The transition from 'sharpened piece of wood' to 'sword' is equally likely, if the semantic emphasis in the original root was on 'sharp', cf. Tamil *kūr-ccu* 'pointed stick' = Kannada *kūr-ahu* 'sword' [DEDR: 173]. Such links

seems to us a much better solution than the one advocated traditionally; and if AV's understanding of «respect for cultural history of languages» (more on that in the «philology» section below) equates it with uncritical acceptance of every instance of lexical analysis found in Japanese dictionaries, we must respectfully disagree.

(2) AV: «PT *ī́lč 'work, deed' is a questionable reconstruction... but even if one agrees with it, PT *ī́lč is going to be a kiss of death to further Altaic comparisons, because it again contains the element *č that cannot be explained away as a suffix in Turkic. Thus, we have to deal with *ī́lč as a root» [Vovin 2005: 96].

Actually, it is not at all hard to explain the final *-č away as a suffix, both here and in other cases. It is quite likely connected to the well-known suffix -š, used to form abstract nouns from verbal stems, cf. in Old Turkic: *toqu-* 'to hit' > *toqu-š* 'blow (n.)', *küsä-* 'to wish' > *küsä-š* 'wish (n.)' [DOT: 663]. This suffix has an equally well-known variant -č in reflexive verbal stems ending with -n, i. e. a nasal resonant, cf. *ökün-* 'to repent' > *ökün-č* 'repentance', *ajman-* 'to be afraid' > *ajman-č* 'fear', etc. [DOT: 650]. The same suffix is listed as the reflexive correlate to -iš in [Clauson 1972: XLIII]. The development *-n-(V)š > *-nč suggests that in Early Turkic and Proto-Turkic this deverbal suffix could have easily been pronounced as -č after resonants, and, therefore, analysis of *ī́lč as *ī́l- 'to work (vb.)' + *-č '*deverbal suffix*' nicely fits in with both internal and external data.

(3) In the same etymology, PTM *(x)ilga- 'diligent, crafty, etc.' is «deconstructed» by AV form by form in the following way:

(3a) for Evenki *ilga-n* 'dodgy, brave, active, diligent': «The root of the Ewenki word is just *ilga-*... but here the segmentation process comes to a halt: there is no Ewk. suffix *-ga-* that would allow us to see Ewk. *ilga-* as *il-ga-*» [Vovin 2005: 95]. But there most definitely *is*: a non-productive, but well discernible PTM suffix *-ga-* with a general causative meaning, and uncontroversial information on it can be found, e. g., in [Kormushin 1978: 67];

(3b) for Even *ilgъrъlbъ-* 'to grow, develop, become strong': «...can be segmented as *ilgʌrʌ-lbʌ-*, with suffix *-lbʌ-* denoting acquisition of a quality in derived verbal stems (Novikova 1980: 30). There is a suffix *-rʌ-* in Ewen,

---

seemed obvious enough to the authors so as not to require additional commentary, but if that is not the case, an apology is in order.

While we are at it, equally incomprehensible is AV's scepticism over the semantic connection 'arrow' — 'sharp stick' (not necessarily just 'skewer') in [Vovin 2005: 107]. Perhaps we can slightly dissipate it by adducing such evidence as PIE *k̂el- [WP I: 431], with reflexes such as Old Indian *śala-* 'staff; dart, spear', *śalya-* 'dart, lance, spear, pike, arrow'; Prussian *kelian* 'spear'; Old Norse *hali-* 'sharp end; tail', Greek κῆλον 'arrow' (and this is only one of the innumerable illustrations of how these meanings can be connected).

but it is used to derive verbs from nouns that denote liquid or dry substaces (Novikova 1980: 19), therefore it is not applicable here, and the process of segmentation stops at *ilgʌrʌ-lbʌ-*» [VOVIN 2005: 96–97].

Again, we must disagree. The Even word is met in two Eastern sub-dialects, where *-rʌ-* can reflect either *-ra-*, *-ri̇-*, or *-ru-*. All of these morphemes are active TM suffixes that derive verbs from nouns [SUNIK 1962: 107–108]; the Even form *ilgỹrỹ-lbỹ-* could easily be derived from a nominal stem *ilgỹ-n*, which, in its turn, is traceable back to *ilga-* — and is actually the same as Evenki *ilga-n*;

(3c) for Manchu *ildamu* 'elegant, refined, agile, etc.': «apparently includes an obsolete suffix *-mu*, which is not mentioned in Manchu grammars, but can be established due to the existence of the verb *ilda-* 'be quick-witted, be agile, be bright'. The latter form cannot be segmented any further, though, since the root *il-* that is badly needed for this etymology to survive simply does not exist in the language and even contradicts the rules of Manchu phonotactics» [VOVIN 2005: 96].

There is indeed a slight contradiction, since *-ld-* normally yields *-nd-* in Manchu, but only slight, since there are at least some well attested cases where *-ld-* is preserved or there is a variation between the two clusters, e. g. *aldasi / andasi* 'halfway' < PTM *alda-*; *aldu-ŋa* 'strange, unusual' < PTM *aldu-* 'news', etc. As for the segmentation problem, the suffix *-da-* is not productive in Manchu, but it is also a well-known PTM suffix that usually forms verbs from nouns (cf. Manchu *ul-de-fun* 'shovel' < PTM *xulē-* 'to dig' [TMS II: 265]), but occasionally also verbs from verbs.

Altogether, then, there is no problem whatsoever of joining all the three forms under one PTM etymology.

Another problem that has to do with the authors of EDAL supposedly inventing fanciful external etymologies, instead of sticking to the tried and true, is the perennial issue of distinguishing borrowings from true cognates. So much has been written on this topic by Altaicists and anti-Altaicists alike that it is futile to even begin to hope to say something radically new on it within the small scope of this article. The preface to EDAL deals in details with two of the most significant parts of the issue: Turco-Mongolic and Mongolo-Tungusic contacts, trying to establish a firm set of criteria (based primarily on different subsets of phonetic correspondences, as is the standard procedure within the comparative method) for distinguishing cognates from loanwords; yet this is only briefly mentioned by AV «en passant» as he then concludes that «apparently, EDAL's authors believe that any given lexical similarity between Korean and Tungusic, Korean and Mongolian, and Korean and Japanese are due to genetic inheritance and not to contact» [VOVIN 2005: 84–85].

Actually, they do not: several cases of such borrowings are explicitly indicated in EDAL, even though such occurrences are indeed met only rarely in comparison to the other two groups of cases. But the principal questions to be asked here are, first, do we actually *have* at least relatively strict criteria to be able to distinguish between, e. g. Japanese-Korean cognates, and Japanese-Korean contacts? And if we do not, what exactly is it that makes the claim of «borrowing» stronger than that of «relationship»? On p. 73 of his review, AV claims that «the burden of proof» is «on those who propose a genetic relationship», and we incline to agree that a claim of relationship is a stronger one than a claim of non-relationship. However, we could never agree that a claim of relationship is in any way more drastic than a claim of language contact, or that the «burden of proof» lying on the shoulders of those who explain all similarities between Japanese and Korean as results of convergence should be any less heavy than the one that lies on the shoulders of «Macro-Altaicists».

Thus, returning to the issue of *kàtànà* above, it is certainly not enough to merely *state* that the two parts of this word are both Korean loanwords (< PK *xata-* 'one' and < MK *nolh* 'blade'). Since one significant distinction of loanwords (which reflect a younger stage of language development) from genetically inherited cognates (reflecting an older stage) is that, roughly speaking, loanwords look *more similar* than cognates, this statement should have been strengthened by demonstrating that

— it is *normal* for a word like EMK *xatun* (河屯, **γâ-don* in MC transcription) to transform into Japanese *kata-*, losing the final nasal; it is also *normal* for EMK *x-* to correspond to Japanese *k-* in all the relevant borrowings (if any more can be found, that is);

— it is *normal* for final liquids in Korean to be dropped in Japanese borrowings, with subsequent delabialization of **o > a*[35];

---

[35] Curiously, exactly the opposite scenario is advocated in [Georg & Vovin 2005: 188], where the authors propose that Korean *mal* (MK *mál*) 'measure for liquids and grains' has been borrowed into Japanese as *masu* 'a measure, container for measuring grain'. Not only does that presuppose that *masu* < earlier **malu* or **malú*, thus corroborating an important pro-Altaic point (because it can hardly be imagined that a resonant, even an auslaut one, would have been borrowed into Japanese as *-s-*), but it directly contradicts the theory that *nolh > nà* (**nasu* would be expected). The authors' additional argument — that Japonic *masu* is attested only in Central Japanese, which makes it suspicious — only makes the proposed scenario even less comprehensible, because it means that the authors attribute the borrowing to a very recent period (already *after* the migration to the Japanese islands and even after the split of Proto-Japonic), and this, in turn, means that there would be no way for a form like *mal* or *mar* to be rendered as *masu* (since the

— finally, since the word *kàtà-* 'one (of two)' is not met independently in Japanese, it is unreasonable to suggest that an independent Old Korean numeral was specially borrowed into another language with the sole intention of forming compounds with the specific semantics 'one (of two) objects'. Such a thing would only be possible if the form were borrowed not by itself, but *already* as part of several compound lexemes, from whence it could «take root» in the language and gain a newly found productivity. However, AV's argument only works for both parts of *kàtànà* independent from one another; to the best of our knowledge, there is no word like **xatu-nʌrh* or whatever it may have looked like attested at any stage in the Korean language, and not a single other Japanese compound beginning with *kàtà-* can be traced back to its Korean predecessor.

Nothing of the sort has been produced by AV next to his statement, nor, as far as we know, anywhere else. In light of this we would prefer to see a genetic connection between at least PJ **na* 'blade' (provided the word really exists) and PK **nǎrh* id., as opposed to, e. g., OJ *tera* 'temple', borrowed from an older form of MK *tjǎr* id. (where the final resonant in Korean is actually *preserved* in Japanese). As for the connection between **na* and **nǎrh*, it may be reflecting a complex correspondence, with the Japanese form actually going back to **ni̯āl(o)-gV* (the actual rule for this consonantal loss in Japanese is explained in [EDAL: 58]).

Let us now examine the other cases discussed in [VOVIN 2005], where EDAL authors are supposed to reject the suggested etymologies in favour of contact explanations, and see whether there are any serious problems with this alternative or if it should be unanimously accepted.

(1) MK *tyél* 'temple', which the authors include on the Korean side in a relatively old Altaic etymology (**dằlà* 'to close, hide'), is rejected by AV on the grounds that it is in reality an old borrowing from MC 刹 *ṭhạt* (we use S. A. STAROSTIN's transcription instead of E. Pulleyblank's as in AV's original); he further claims that «this etymology is universally accepted in the field of Korean historical linguistics» [VOVIN 2005: 95].

Given that in most cases a strong appeal to «universal» or «general acceptance» in AV's reviews is synonymous with «alternate, but not any less controversial theory» (see his evaluation of L. SAGART and U. SEE-FLOTH above), it makes sense to analyze this case a little deeper. The reason that MK *tyél* (*tjǎr* in EDAL's orthography) was included is obvious: it is not a regular Sino-Korean reading of MC *ṭhạt*. In particular, MC *-ạ-* is never rendered in Korean as *-ə-*, the regular reflexation being *-a-*, and this

---

change **ĺ > *s* took place long before that time). In our opinion, this already means taking scepticism beyond anything that can be called «reasonable limits».

time, paying sufficient attention to vocalism *is* necessary, since what we are dealing with is not a six thousand year long period of phonetic change, but quite a rigid system of «recoding» from MC phonetics to Korean that took place over, at most, a few generations. Since the word does not belong to the «Sino-Korean» stratum, it is either an obscure dialectal borrowing or a native Korean word. In both cases, additional evidence is necessary to strengthen the arguments; EDAL provides such evidence for the «native» solution, but AV, for reasons unknown, does not provide it for the «loan-word» solution[36]. Nor does he even mention the obvious problem of treating the word as belonging to the Sino-Korean layer.

(2) PM **čilaɣu* 'stone', according to AV, is not genetically related to PT **diāl* id., but borrowed from the Bulgar branch of Turkic[37] (cf. Chuvash *čol* id.). The claim is categorical («most certainly»), yet strong evidence to support it is lacking, because:

— borrowing of such basic lexical items as 'stone' usually presupposes a situation of extremely heavy language contact[38], meaning that we should be able to easily identify a huge stratum of loanwords from Bulgar

---

[36] Except for an indirect argument that the word *tyél* «is not any kind of 'temple', as it can only mean 'Buddhist temple'… MK *tyél* is used as a translation equivalent of not only character 寺 'Buddhist temple', but also of the character 刹 '[property of a] Buddhist temple'» [Vovin 2005: 95]. If AV thinks that any word meaning 'Buddhist temple' cannot be native, but must always have an external provenance, he might be surprised to learn that the Chinese word 寺 *sì* that he is quoting, although mostly used to denote a Buddhist shrine in MC times, is already attested in old Chinese texts long before the advent of Buddhism, meaning either 'hall, residence' or 'eunuch' (= 'the one who dwells in a special residence'; the second meaning is actually attested earlier than the first, but is obviously derived from it; see [Karlgren 1964a: 253], [Schuessler 1987: 583]).

[37] Throughout his review, AV consistently misspells «Bulgar» as «Bulghar», although the second spelling is not academic and, in fact, misleading (the word is always pronounced with a velar stop, not fricative).

[38] But this is different from stating that the possibility of borrowing in such cases is excluded, which is what AV would have the reader think of EDAL's authors who «apparently believe that words such as 'stone' cannot be borrowed [Starostin 1991: 38]». Again, though, the reference to Starostin is completely misguided. The corresponding passage on p. 38 of his book reads: «Poppe considers the TM forms for 'stone' to be borrowed from Mongolian, which is hardly possible» — because, of course, PTM has initial **ǯ-* where PM has initial **č-*, and certainly not due to general considerations of what can and what cannot be borrowed. We would be grateful to AV if he could explicitly provide a reference to any published work that belongs to Starostin — or, for that matter, any of EDAL's authors — where it is stated that word so-and-so cannot ever be borrowed from one language into another. Until then, his tentative decoding of their «apparent beliefs» will have to remain under doubt.

into PM, belonging to a variety of semantic fields. Such an identification has not, however, been made; unlike the case for later contacts between non-Bulgar Turkic and Mongolic, the hypothesis of Bulgar-Mongolic inter-action is based on too little evidence[39];

— even supposing the hypothesis *is* correct and the PM word is related to Chuvash *čol* as a loanword, why does the PM word has a structure that is so seriously deviant from the Bulgar prototype? Neither the front vowel in PM, nor the suffixal *-γu* are seen in Bulgar. Considering that the supposed Bulgar-Mongolic contacts must have taken place no earlier than the beginning of the 1st millennium A.D. (after the initial split of Proto-Turkic and «Bulgarization» of the inherited Turkic lexicon) and no later than the end of that millennium (before the disintegration of PM; probably even earlier, since there is no attested evidence for Bulgar presence in Eastern Asia), this gives PM a very small amount of time to make **čilaγu* out of **čol* or **čul*;

— trying to correlate Bulgar palatalization of dentals with the corresponding PM palatalization just does not work, because all reliable instances of PM **č-* < PA **t-/*t´-* would then have to be explained as Bulgarisms — even when the Bulgar forms are unattested, e. g. PM **čig* 'separately' = PT **tēk* 'odd; only, solitary'; PM **čilga-* 'to splash, overflow' = PT **d(i)āĺ(ɨ)-* id.

---

[39] The idea of Bulgar-Mongolian contacts is occasionally advocated by S. GEORG, e. g., in [GEORG 1999/2000: 171], where he, for instance, asserts that PM **širüken* is borrowed from Bulgar **jirük* (< PT **jürek*). The proposed scenario is as follows: (a) PT **jürek* > early (unattested) Bulgar **jirük* through semi-regular assimilation of the vowel to the initial consonant; this form has been borrowed into PM; (b) early Bulgar **jirük* > Chuvash *čərə* instead of the expected **śərə* due to analogy with *čərə* < **tirig* 'alive'.

The unavoidable problem with this scenario is that, although we agree that irregular *č-* in Chuvash is best explained by analogy with *čərə* (the same solution is adopted in EDAL), this is in all probability a very late development. First, there is no strong evidence for a PT **jü-* > **ji-* development in Bulgar. GEORG postulates it on the basis of correspondences between the Anatri and Viryal dialects: PT **ü* > Anatri *ə*, Viryal *ö*, but in the 'heart' etymon both dialects show *ə* < PT **i*. Cf., however, PT **jüŕük* 'finger-ring' > Anatri *śərə*, Viryal *śörö*, with no trace of such an assimilation. Therefore, «Proto-Chuvash *-i-*» in 'heart' could only be motivated by the aforementioned analogy with *čərə* 'alive' < **tirig*, an analogy that could obviously take place only *after* the phonetic change **t-* > *č-* (it is hard to see how a form like **jürek* could be influenced by **tirig*, but not difficult to see how **śare/*śöre* could be influenced by **čərə*). But it is well known that the change **t-* > *č-* did not take place in Bulgar until at least the late 13th century [RÓNA-TAS 1986; ERDAL 1993: 126–127], at which time we have the first instances of variation between *t-* and *č-* in epigraphic material (cf. Old Bulgar *tekü* 'hen' > Chuvash *čъxъ*, etc.). And, of course, by the 14th century it is already much too late to talk about any potential «Bulgarisms» in Proto-Mongolic.

(3) PK **ak-su* 'heavy rain' is removed from PA **agà* 'rain, air' on the following basis: «PK **ak-su* is reconstructed solely on the basis of SSK *ak-swu* and *ek-swu* (no attestations in MK or dialects), which are both transparent loans from Chinese: *ak* 'bad' (惡) + *swu* 'water' (水) and *ek* 'oppressive' (抑) + *swu* 'water' (水)» [Vovin 2005: 121].

Unfortunately, no reference is given, so it is hard to establish the origin of the proposed etymology. Nevertheless, unless AV's idea is that both forms were concocted already on Korean territory from two Sino-Korean morphemes (not very likely), we would have to presume that both of these compounds can easily be found in Middle Chinese, thus corroborating the «loanword» solution. And the first of these, 惡水 (MC **ʔâk-świ*), really exists — but its meaning, however, is usually defined as 'dirty (stagnant) water' [HDC VII: 554], with no direct connections to 'rain', let alone 'heavy rain'. The second one, to use AV's own terminology, is a 'ghost word': no such compound is ever attested in Chinese texts or modern spoken language (although we do sometimes have it as a verb+object idiom: MC *ʔek świ* 'to suppress the water(s)', said, e. g., of Yu, subduer of the Great Flood). Even if we suppose that such a compound *may* have existed, though (but why should we?), it is hard to believe that it was used to designate any kind of falling water, since in Chinese the word 水 *shuǐ* (MC *świ*), throughout all the attested epochs, has always designated *flowing* rather than *falling* water. If the Korean words meant anything like 'flood, inundation', the borrowing suggestion would be more plausible. As it is, there is nothing «transparent» about these «loans», and external explanation is at the least equally reasonable, if not more so.

And, building up on that, we would like to remark that loanwords, in order to qualify as such, really *must* be transparent. The majority of postulated contacts between Altaic-speaking peoples belong to fairly recent chronological epochs (at least, when compared to the proposed date of separation for Proto-Altaic), meaning that in each of these cases a clear, concise, and uncontroversial scenario should be offered, along the lines of those that *have* been offered for, e. g., Slavic-Germanic or Fenno-Ugric/ Indo-Iranian contacts. Any kind of lexical resemblances that do not clearly fit within those scenarios — violating phonetic correspondences, differing in meaning or morphological structure — *may* be indicative of older genetic relations rather than contacts. If anything, the «burden of proof» is, in fact, heavier in the case of «loanword solution», because smaller time depths allow for less variability between the compared items. The Altaicist is nowhere near as «responsible» for explaining the final -*γu* in Mongolic **čila-γu* (an adnominal suffix that may have appeared in PM anytime over a period of six thousand years) as the «Mongolo-Bulgarist» who has to explain how it got there in Mongolian

(over a period of six *hundred* years) if it was not there in Bulgar. This is not to say that the Altaicist should not *try* to explain — but the «Mongolo-Bulgar-ist» *has* to explain, and these are two different things.

We will conclude this part of the discussion with the following state-ment: *very few instances of internal reconstruction can be accepted as «proven» when not validated by external comparison*. This is more or less a commonplace with short-range comparison, but is regularly forgotten when the external re-lationship is not intuitively obvious. Substituting internal reconstruction — which is very frequently ambivalent, unsystematic (or «pseudo-system-atic», striving for building a mechanistic model instead of a realistic lan-guage), and uncontrolled by independent evidence — for systematic ex-ternal comparison, and then denying linguists the right to perform such comparison because it «contradicts the solid achievements of internal recon-struction» is essentially just trading in science for scientism. And if this is the main, or one of the main, reasons why some of our opponents call the Altaic hypothesis «premature» and a «pragmatically poor foundation on which to build a sustained research program» [UNGER 1990: 479], we must respectfully disagree. In our view, it is much more premature to make a claim that the word for 'stone' is a Bulgar borrowing into Mongolian.

VI. *Philology or fantasy*?

The next argument, or group of arguments, consistently occupies the central spot of all of AV's critical evaluations of the Moscow school — even though, in our humble opinion, it is also the most vague and incon-sequential of all of them. We mean his constant references to the ignorance of «philology» on the part of EDAL/SCH authors, and to how their reliance on dictionaries rather than textual evidence produces an increasingly negative effect on the results of their work.

Cf.: «The fact that any dictionary is not a source for a historical lin-guist was already well proven in the 19th century, which have ⟨sic! — A. D., G. S.⟩ seen the birth of philology as a scholarly discipline. It is certainly true that consulting the actual texts rather than the dictionaries puts much more burden on a researcher than just consulting a dictionary, because it requires something more than an ability to read a vocabulary entry: namely, the sufficient knowledge of a given language to read and understand the texts, as well as of its philology and culture. This pains-taking process of checking texts and cultural data rather than dictionaries may seem to be not worth the candle for some (but not all) long-rangers» [VOVIN 2002: 160].

«Historical linguistics is intimately connected with philology. It means that overall wordlists and dictionaries are not the source for serious work in historical linguistics. It is fallacious to put several dictionaries on one's desk and just pick and choose data according to one's whim, without verifying the reliability of the data through texts that are the most reliable sources for a historical linguist» [Vovin 2005: 76].

On the surface, both these passages look quite convincing; however, it only takes a minute of trying to fit them within any practically-oriented scenario to notice some of the ensuing absurdities.

First, it is unrealistic to expect that any single researcher working on an etymological dictionary of a widespread *language family* will be able to meticulously verify his data on all the members of that family through textual evidence, let alone present all the results of such a verification within the dictionary. Not even mentioning the problem that for some of these languages textual evidence is not even *available* (this is particularly actual for Sino-Caucasian languages, where wordlists are frequently the only source of data), this would make technically impossible the creation of any sort of etymological dictionary that operates on data from more than ten or twenty languages.

Neither A. Walde and J. Pokorny, nor K. Rédei, nor T. Burrow & M. B. Emeneau (creators of some of the most well-known and respected etymological dictionaries), saw it fit to confirm *all*, or even *most*, or, in fact, *any* lexical entries that are given in their dictionaries on the basis of such evidence. AV is, of course, entitled to state that none of these dictionaries constitute «serious work in historical linguistics», but it is not exactly clear how that statement would tie in with some of the general agreements in «mainstream linguistics».

Certainly all of these scholars had professional knowledge of many, or even most, languages that constituted their respective families, but the plain fact is that, in this kind of work, textual information is almost never included, and the absolute majority of the data references usually address the reader to... *dictionaries* of the languages — the worst possible sources for a historical linguist! Even worse, one might occasionally fall upon a *factual mistake* in any of the dictionaries — and make the logical conclusion that J. Pokorny had never consulted the *Ilias*, or that T. Burrow, in his overoptimistic reliance on «wordlist linguistics», had not even once opened the *Pattupattu*; certainly not a single line in the Dravidian etymological dictionary will directly suggest the opposite to the reader.

The main reason why all of these so-called «historical linguists» so stubbornly keep referring to dictionaries is, of course, easy to understand. Theoretically, a good dictionary of the language is supposed to provide

the researcher with the exact kind of information that is necessary *and* sufficient for etymological purposes: namely, an exhaustive list of adequately transcribed lexemes, accompanied by a trustworthy rendering of their meaning in the language of the translation (if the dictionary in question is bilingual) or by an equally trustworthy intepretation of their meaning in the dictionary's own language. An *excellent* dictionary of the language will supplement this information with examples of textual usage or, at least, a list of textual sources for as many given lexemes as possible.

Thus, the reason why we can lend our trust to Pokorny, Rédei, etc., is that their work rests on the premise that they are mainly using *good* to *excellent* dictionaries — a premise that may not always be true, but generally works. Certainly the same goes for Altaic and Sino-Caucasian. If the authors of EDAL/SCH are at any time misled by erroneous data they pick out from language dictionaries, the *primary* fault lies with these dictionaries and their authors, and not with those researchers who are using the data[40]. It sometimes seems as if AV is demanding that anyone who ever wishes to indulge in macro-comparative linguistics must be at the very least an *Übermensch*, able to read, translate, and analyze any imaginable literature in any language that is relevant to the studied macrofamily. Unfortunately, it is hard for us to agree with him on this point. If, being an acknowledged specialist in Korean, he is willing to use his philological knowledge in order to point out factual errors in EDAL, this is a noble and helpful task. If, however, he is willing to use it in order to convince the general audience that EDAL authors are not *Übermenschen*, he needn't have bothered, since no one is.

Throughout both of his reviews, AV constantly appeals to a good grasp of «philology» as the one thing that separates qualified scholars from ignorant EDAL/SCH authors. Perhaps he will find it hard to believe, but all of us are actually quite respectful of classic philology and careful study of documented evidence. What we are highly suspicious of, though, is using the grand word «philology» to mask the uncritical acceptance of highly subjective and hypothetical speculations that look promising mainly because they demonstrate the author's deep knowledge of the subject, but not because they follow some kind of strict methodology or are based on sufficient evidence. We will try to illustrate these suspicions with two examples: one drawn from the work of L. Sagart, whom AV calls a «wonderful philologist» in [Vovin 2002], one from the work of AV himself.

---

[40] Unless, of course, we deal with errors committed by EDAL/SCH authors themselves during the transfer of data from primary sources, on which see below.

In his 1999 monograph, L. Sagart — frequently on the basis of «philological» evidence — contested numerous Sino-Tibetan etymologies put forward by I. Peiros and S. Starostin, in an attempt to demonstrate that Chinese and Tibeto-Burman languages are not genetically related (he has, however, since amended that view). One such argument concerns one of the Chinese words for 'ram', OC 豮 *b(h)ər (*b(h)ən) > Modern Chinese *fen₂*, which the authors compared with select Tibeto-Burman parallels and traced back to PST *bhVr 'goat, sheep' [Peiros & Starostin 1996, № 52].

On this Sagart writes: «However, as pointed out by Karlgren (gloss 746), this *fen₂* 豮 is homophonous and etymologically identical with *fen₂* 墳 'big-horned' (*Shi Jing*, Ode 233); rams and he-goats have large horns, while she-goats have smaller horns and ewes no horns at all. Thus semantic evolution from 'big-horned' to 'male' is straightforward. The Chinese term *fen₂* 墳 'big-horned' is itself a semantic specialization of *ben₁* 賁 'great, big'. The Tibeto-Burman forms adduced by Peiros and Starostin are better regarded as borrowings from Chinese than as true cognates, in view of their clearly secondary semantics» [Sagart 1999: 195]. In a footnote below, the full quotation from Ode 233 is also adduced: 牂羊墳首 *zang₁ yang₂ fen₂ shou₃* «the ewes have big (horned) heads», and an explanation of the line: «the Mao commentary observes that it is meaningless to speak of big-horned ewes — this is the very point made by this ode».

The argument can be viewed as philologically impeccable: in this case, it even includes both the verse and line reference *and* the actual text example, so highly priced by AV (see above). And yet it is deeply flawed, since significant parts of the evidence are not made available to the reader. For one thing, while *fen₂* 豮 (*b(h)ər in Starostin's reconstruction, *bɨr or *bɨn in Sagart's) is indeed homophonous with *fen₂* 墳, their etymological identity is but a hypothesis, based on an old gloss in the «Ěryǎ» (爾雅) dictionary where it is said that 羊, 牡豮牝牂 «of sheep, males are *fén*, females are *zāng*». Karlgren [1964b: 165]: «Later comm. have added that the *fen* in the Erya gloss is the same word, and that in this ode 'big head' means 'horned head' in contrast to the 'small head' of the hornless animal».

So the *real* evidence based on which we claim etymological identity for these two words is not really *textual*, but *lexicographical* — moreover, rooted in Hàn-era (漢, $\frac{206\text{ BC} - 8\text{ AD (W.)}}{25 - 220\text{ AD (E.)}}$) dictionaries where semantic interpretations are frequently confused with phonetic glosses, and explanations of obscure words found in the classics are frequently adduced on an *ad hoc* basis.

Second, Sagart gives the basic meaning for *fen₂* 墳 as 'big-horned'; however, Ode 233 includes the only passage in the *Shī Jīng* (詩經) that could justify this interpretation. In another context (Ode 10), it means 'raised bank of a river'; and Karlgren himself acknowledges that it is also obviously the same

word as transcribed by the character *fen₂* 頒 in Ode 221, where the line is 有頒 其首 *you₃ fen₂ qi₂ shou₃* «big (lit. having bigness) are their heads». Now it would be all right if that line were also referring to cattle; unfortunately, it is referring to *fish*, and, since it is hard to imagine anyone calling a bunch of fish 'big-horned', the natural implication is that in both cases the word should simply be translated as 'big' or 'swollen' (the latter is even preferable since it helps to bring 'raised bank of a river' < 'swelling' inside the same semantic cluster).

In conclusion, the line in Ode 233 should either be translated as «the ewes have swollen heads» (if one wishes to follow ZHŪ XĪ's [朱熹] interpretation: = «swollen heads, but lean bodies», i. e. suffering from hunger), or «the ewes have rams' heads», i. e. «things are not what they should be»; the second version is closer to MÁO's (毛) interpretation, and it would suggest that in this line 墳 = 羒 'ram' (on the other hand, the weakness of this variant is that it ignores the «big-head» parallelism with Ode 221). As it is, SAGART has presented only *part* of the argument, and in addition, he has not mentioned that in some later texts and dictionaries (e. g. «Guǎngyùn», 廣韻) the word 羒 is defined as 'white ram', i. e. denoting a breed of the animal. Its comparison with Tibeto-Burman entries on a genetic level is therefore perfectly justified[41], and while «philology» may somewhat complicate the case here, it gives us no major reasons to reject the etymology.

The second example, significantly more relevant to the particular issues discussed in this article, concerns AV's «philological» analysis of the Korean word *meli* 'head' in conjunction with some possible Altaic parallels [VOVIN 2000: 143–145]. In this work — written in the twilight of the author's «pro-Altaic» period — AV argues for the importance of Early Middle Korean data (unfortunately, only available in a limited quantity and in Chinese transcription) for the reconstruction of Proto-Korean and, subsequently, Altaic.

One of the examples is the Korean word for 'head' (MK *mèlí*; *mə̀rí* in EDAL's transcription), which in EMK (*Kyerim yusa*[42], №161) is found glossed as 麻帝 (MC **mạ-tìej* in STAROSTIN's transcription). Drawing upon

---

[41] Besides, if the Tibeto-Burman words were really borrowings from Chinese, as SAGART claims, why would they have meanings like 'goat' and 'male mountain goat' and not 'ram' or 'white ram'?

[42] The name of this source (鷄林類事, 계림유사) can depend on the adopted transcription standard; AV uses the Yale-based spelling *Kyeylim*, although the McCUNE standard *Kyerim* is generally more preferable, as in [SASSE 1976] et al. (for some odd reason, in his bibliography AV actually changes *Kyerim yusa* in the title of W. SASSE's monograph to *Kyeylim yusa*!). In EDAL the spelling *Kirim* is used, which is inconsistent with the general Korean spelling in the dictionary (*Kjerim* should have been a better choice — AV's critical remark is quite on target here); the source will be transcribed as *Kyerim* in this article, following the McCUNE–REISCHAUER romanization. The latter is also used for rendering Korean proper names, but the language data are given in EDAL's own transcription.

this transcription, AV reconstructs EMK *matay or *matæ; proclaims MK *màrí* to be the result of a later lenition **-t-* > **-r-*; and renounces the old Altaic etymology, traceable back to G. Ramstedt, that pairs the Korean word with Turkic *bálč 'head' (EDAL also adds here PM **malǯa-* 'bald' and PTM *meli-mu 'neck joint, back part of neck, etc. < PA *mḗlǯu 'head').

The refusal of EDAL's authors to recognize this reconstruction and, subsequently, abandon the Altaic etymology is apparently viewed by AV as little other than subjective stubborness on their part («I suspect that the reason for rejection of the *Kyeylim* data is very simple: it reveals certain facts in the history of the Korean language that destroy important 'Altaic' etymologies» [Vovin 2005: 89]). However, the exact reason why AV's *matay is to be questioned *was* actually given in EDAL, and the fact that AV failed to mention this in his review is, in our humble opinion, quite telling. We quote it here: «It is most probable that MC *-t-* was used here just to transcribe Korean *-r-* (since Middle Chinese, as well known, lacked *r-*)» [EDAL: 911].

Granted, this alternate explanation was given briefly and is not without problems of its own; within *Kyerim yusa*, we frequently find instances where MK *-ri-* (*-li-*) is transcribed by MC syllables with *-l-*, e. g. gloss №85: 弼陀里 *bit-thâ-lɨ 'dove' (MK *pìtùrí*). But neither is AV's version. The suggestion that Proto-Korean **-t-* > MK *-l-* is in direct contradiction with the simple fact that MK has lots of instances of intervocalic *-t-*, whose regular development in Modern Korean is that of palatalization before front vowels, e. g. MK *mōtír-* 'bad' > *mōǯil-*, *ìpàtí* 'feast' > *ibaǯi*, etc. Therefore, if the Proto-Korean word for 'head' were really something like *mati (AV's reconstruction *matay is not convincing because MC 帝 *tìej* quite clearly indicates a front vowel; the final *-j* is there simply because MC lacked a separate syllable *ti*), we would normally expect it to stay that way in MK and then develop into something like *maǯi* in the modern language.

Another issue is that MC *-d-* and *-l-* are not the only two correspondences to MK *-l-* (*-r-*). For instance, the MK word for 'cloud', *kúrùm*, is transcribed in *Kyerim yusa* as MC 屈林 *khüt-lim (№4). Perhaps this might motivate AV to reconstruct something like Proto-Korean *kutrum (which will not only signify a major breakthrough in reconstructing consonant clusters in Proto-Korean, but also successfully help to eliminate the unnecessary connection between this word and PA *kʽòlmV 'shadow, cloud'); we, however, would prefer viewing this as simply a different way of conveying Korean *-r-*, further illustrating the inconsistency of Chinese transcription.

Thus, based on *one* instance of unsystematic transcription, AV confidently postulates a significant irregularity in the language, and based on that irregularity, rejects a perfectly regular correspondence for Altaic[43].

The most interesting part, however, comes later, and this is where the real «philological» evidence really steps in. Having rejected the more traditional Altaic etymology, AV replaces it with an interesting new Koreo-Japonic one, comparing his provisional **matay* with no less than the name of the fabulous serpent *Ya-mata-*(*no woroti*), prominently featured in Japanese mythology as represented in the *Kojiki* and *Nihon Shoki* (日本書紀) chronicles. Now the name *Ya-mata* is traditionally analyzed as 'eight-forked', and accordingly spelled 八俣 in the first of the mentioned texts and 八岐 in the second one. At the same time, it is quite evident from both of these texts that the same serpent was represented as having eight heads and eight tails (see [VOVIN 2000: 144] for textual details).

AV: «There are eight heads and tails but only seven forks between them. The only way out of this confusion is to assume that *-mata* in *Ya-mata* really means 'head' and not 'fork'... It is also likely that by the early eighth century the real meaning of the word **-mata* 'head' in *Ya-mata* was already forgotten, and only the context of the myth in combination with simple arithmetic allows us to reconstruct it» [VOVIN 2000: 144–145].

AV's application of arithmetic is commendable. What he seems to be forgetting, though, is that, when used in compounds, the word *-mata* 'fork' usually denotes not the space between the «prongs» of the alleged 'fork', but the very «prongs» themselves, e. g. 二又 *futa-mata* 'bifurcation, parting of the ways' (lit. 'two-fork', not 'one-fork'); 三又 *mitsu-mata* 'three-pronged fork' (lit. 'three-fork', not 'two-fork'). The same usage of the word 'fork' and its derivatives is observed in many languages around the world, starting from European ones (the word *bi-furcated* implies the presence of only *one*

---

[43] Even if the EMK word *were* indeed pronounced with an inlaut dental, this would still not be a sufficient incentive to reject the Altaic parallel. No other words on the *Kyerim* list are spelled with the character 帝, and the few instances where Korean words with the syllable *ti* or *ti* are attested on the list are always represented by a MC voiceless consonant, e. g. *ti-* 'to fall' = MC 恥 *ṭhí* (gloss №8; note the lack of the expected 帝). This would indicate **madi* rather than **mati*, and since the PA word in EDAL is reconstructed with the cluster **-ĺǯ-* (based on PT **baĺč* and PM **malǯa-, *melǯe-*), this could simply be a specific development of the cluster in Proto-Korean (**-ĺǯ- > *-d-*). However, this is all a matter of vague speculation, since there is simply too few evidence in favour of any of these hypotheses. No systematic conclusion can be drawn based on one instance of a resonant being represented by a dental in Chinese transcription, which, as AV himself remarks on a more suitable occasion, «is not the IPA» [VOVIN 2005: 113].

forking) and ending with Chinese (三叉 *sān-chā* 'trifurcation'), which may have influenced Japanese usage (but may also have not).

All of this means that there is no inconsistency whatsoever in between calling the serpent 'eight-forked' and then explicitly stating that it had 'eight heads', and thus, no serious grounds for surmising an earlier *mata* 'head' that would later, based on «folk-etymological analysis» (and we all know AV's usual feelings about this), converge with *mata* 'fork'. In addition, there are reasons to doubt that a word as important as 'head', even having fallen out of active usage, would leave no trace whatsoever in the language apart from being buried in the name of a mythical character.

Both of the discussed examples, in our opinion, illustrate a sad, but imminent truth: «philological» evidence is usually ambivalent and inconclusive when it is based on isolated quotations rather than systematic analysis. The indisputable fact that L. Sagart and A. Vovin really know their *Shī Jīng* and *Kojiki*, per se, contributes nothing to a better understanding of the prehistory of the Chinese or Japanese language. Their hypotheses *may* be considered and evaluated, but it is only through a systematic and non-contradictory explanation of *all*, rather than *select*, available facts, that these hypotheses can be considered proven «beyond a reasonable doubt» and put ahead of all others in terms of historical probability.

Thus, the aforementioned Chinese and Korean forms both fit within the systematic set of rules established for the Sino-Tibetan and Altaic families. Just because some inconclusive evidence has been found, based on philological analysis, that they *may* be removed from that system, does not automatically mean that they *have* to be removed from it. The great benefit of something forming part of a system is that such a placement makes less room for subjectivity; the well-known principle that «every word has its own history» works well when that history is well attested, but not so well when that history is invented. And this is what both of the quoted examples are about — *inventing* rather than *uncovering* the respective history of both words.

Now let us turn to some actual examples where, according to AV, ignorance of «philology» leads the authors of EDAL/SCH to formulating unscientific conclusions.

(1) AV: «Starostin 1995a: 205–06 compares PY **baʔŋ* 'ground' with ST **bhu:m* 'grave', citing among several languages Tib. *ābum* 'grave' (more up-to-date transliteration is *'bum*). However, Tib. *'bum* does not mean 'grave', it means '100,000'. 'Grave' in Tib. is *'bum-pa*. Well, *-pa* is probably a suffix, but the elementary accuracy requires a researcher to cite the words as they are. The main problem is, of course, that Tib. *'bum-pa*, glossed even in Jäschke's dictionary as 'grave, *sepulchre*', is a kind of aristocratic burial, when a grave

structure is build *over* a dead body or a relic. No digging into the ground is involved» [VOVIN 2002: 160].

AV is certainly right that *'bum* should be correctly glossed as *'bum-pa*. His identification of the main problem, however, is exaggerated. Grave structures built over dead bodies frequently find their prototypes in raised *burial mounds*, which, of course, are made of *earth*, and a semantic development 'bank of earth, hill, mound' > 'burial mound' > 'sepulchre' is perfectly reasonable. The weirdest moment about this is, however, that AV does not adduce the remaining Sino-Tibetan part of the etymology, which, besides Tib. *'bum-(pa)*, also includes such words as Burmese *pum* 'heap; grave, mound', Kachin *bum²* 'mountain', Lushai *phūm* 'to bury, inter', and OC 堋 **pǝŋ-s* 'to bury'. All of these words exhibit regular correspondences and add significant evidence to the postulated development of the meaning in Tibetan; and, to the best of our knowledge, a better etymology for the Tibetan word has not ever been offered by anybody.

In addition, it is also peculiar that AV, fresh from reproaching S. STAROSTIN for ignoring *textual* evidence in favour of inherently flawed language dictionaries, draws *his* own information about the usage of the word *'bum-pa* from not one, but *two* different dictionaries of Tibetan ([JÄSCHKE 1987] and [ZHĀNG 1993]). If dictionaries are not a serious source for historical linguists, how can he be so certain of what *'bum-pa really* means?

(2) AV: «Starostin 1995a: 206 repeats Karlgren's mistake in interpreting OC 芾 **pǝt* as 'knee-cap' (Karlgren 1950: 131) (although he does mention that it is a kind of apron). But since he needs 'knee-cap' for his comparison with his PY **bat-* 'knee', the actual cultural data are sacrificed and the word is misglossed as 'knee-cap' to make the comparison with PY look better. However, as it is quite apparent from the Shijing II: IV.5... OC **pǝt* is a ceremonial apron worn at court... An apron covers knees in no more specific way than pants or skirt do» [VOVIN 2002: 161].

At least in this passage AV provides the «verse and line» reference (as well as the entire quote from *Shī Jīng*, which we are not reproducing here because of its length). However, there is so much that is either wrong or understated about this argument that it is not even clear where to start.

First, the reference to [KARLGREN 1950: 131] is misguided; the actual translation is not 'knee-*caps*', but 'knee-*covers*', and the same word is also used to interpret OC **pǝt* in his *Grammata Serica Recensa* [KARLGREN 1964a: 137]. Granted, one could suggest that there is not much difference between the former and the latter, but elementary accuracy etc. etc.

In reality, however, the difference is more important than it might seem to AV. KARLGREN's translation is not *ad hoc*, but is instead based on the definition and description of this particular piece of clothing in various sources,

dating back to Hàn times. Cf. one such description in Zhèng Xuán's (鄭玄) commentary on one of the Odes: 芾, 大古蔽膝之象也 «The *fú* is a sort of *knee-covering* in high antiquity». Not only that, but even compounds with explicit reference to 'knees' were also used to denote apron-like clothing: Zhèng Xuán's 蔽膝 *bì-xī* is used in that exact meaning in multiple contexts at least since the Hàn period [HDC IX: 542]. All of this means that for Chinese people of that time, aprons of the 芾 type covered knees in a *much* more specific way than pants or skirts — that is, they *stopped at the knees*, and, therefore, derivation of OC **pət* 'apron = knee-cover' from 'knee' is more than likely.

Furthermore, once again, AV hides the rest of the etymology from the unsuspecting reader. PY **bat* 'knee' is compared not only with OC **pət*, but also Tib. *pus-mo* 'knee', Kachin *lə-phut²* 'knee', Proto-Kuki-Chin **pus* 'knee', etc. etc. — even if we prefer to remove the OC entry, the comparison still remains perfectly valid. There is no particular and urgent «need» to provide an OC parallel for the Yeniseian word, since its presence or absence does not affect the etymology; the OC word is inserted simply because its semantics, if, of course, Hàn-time interpreters are at all to be trusted, offers a reasonable connection with the Tibeto-Burman entries. No evidence whatsoever can be found in OC literary monuments that would somehow contradict this interpretation.

(3) On the potential derivation of PJ **dàsìrà* 'shrine, enclosure for worship of deities' from PA **dăla* 'to close, hide': «There is a traditional etymology for OJ *yasirö*, taking it to be a compound of *ya* 'house, dwelling' + *sirö* 'substitute', since deities in Shintô tradition are considered to come to dwell in shrines. However, this is likely to be a later interpretation, since the earliest textual evidence indicates two important things: a) that OJ *yasirö* refers to sacred place or ground itself where a deity is venerated, and not to the building, and b) that *yasirö* can mean 'deity' as well» [Vovin 2005: 94]. The first point is then illustrated with a quotation from the *Man'yōshū* VIII: 1517 («MÎWA-nö YASIRÖ N-Ö YAMA 'mountain that is the sacred place of Miwa'»), although AV himself admits in a footnote that there are several other interpretations of the character 祝 that is supposed to represent the word *yasirö* in this passage.

At least it is reassuring to see AV take arms against the «folk etymology» of *yasirö* < *ya-sirö*. Regardless, however, of whether the *Man'yōshū* quotation really includes this word or not, this is again a typical example of how it is possible to misuse the «philological» method. Of course, there is nothing wrong about wishing for more precision in denoting the semantic scope of the analyzed lexeme, and, by all means, AV is qualified to exercise that precision. The etymologization of *yasirö* < **dăla*, however, does not have to depend upon whether we have given an exhaustive list of all the meanings of the word or not. Instead, it is built on the following facts:

(a) the word *yasirö* is most frequently used to denote a special place that has been reserved for worship;

(b) special places reserved for worship are frequently known to be separated from their surroundings (enclosed), be it in a *literal* sense (walled or fenced) or in a *figurative* one (delimited by spiritual presence)[44].

The fact that, on Japanese soil at least, there are *exceptions* to (b), does not weaken the etymology. To say that the original meaning of *yasirö* should not have anything to do with «enclosing» is like saying that Russian *город* 'town, city' cannot be compared with Lithuanian *gar̃das* 'fence', because, of course, nowadays it is a pretty rare sight when a town or a city is surrounded by fences. As for the observation that *yasirö* can also mean 'deity' itself, well, Old Indian *gō-*, besides meaning 'cow', can also mean 'milk', 'flesh', 'skin', 'mother', 'star', 'sun', and 'moon'. It is useful to know all these if one plans to extensively pursue Vedic philology, but it is hardly reasonable to use this polysemy as a pretext to remove the word from the Indo-European etymology of *$g^w\bar{o}u$-* 'cow'.

(4) On the derivation of PTM *sila(-bun)* 'skewer' from PA *sī̆la* 'sharp stick, tooth': «...followed by a note: 'Most languages also reflect the verb *sila-* 'to roast on a spit, to put on a spit'. I am afraid that the meaning 'to put on a skewer' is again invented by the authors, and the verb *sila-* ~ *silo-* (to cite it more exactly) means just 'roast on a skewer' [TMS II: 82]. I can understand that EDAL's authors do not know that PMT suffix *-pūn* makes the deverbal nouns designating tools, but only a religious belief in Altaic can obscure their vision to such a degree that they fail to realize that PMT noun *sila-pūn* 'skewer' is morphologically complex, and therefore, must be an obvious secondary derivation from *sila-* 'to roast on a skewer' that is morphologically simple. Now, a quick excursus into roasting ways of Tungus peoples will demonstrate that roasting on a skewer was traditionally the only common Tungusic way to roast things that involved any kind of an instrument, and that it is known to the majority of Tungusic tribes... Roasting on charcoals seems to be attested only for Udege. Thus, *sila-* means just 'to roast', it is a reference to the cooking method itself, but not to the way it was done. Thus, PMT *sila-pūn* is just a 'tool for roasting', nothing more» [VOVIN 2005: 104–105].

Let us start from the beginning. First, the meaning 'to put on a skewer' is certainly *not* invented by the authors. Provided [TMS II: 82] is to be believed (it is, after all, a *dictionary*), this meaning is attested in at

---

[44] The authors of EDAL do not presuppose that the expression 'enclosure for worship of deities' must necessarily mean some kind of *building*. Perhaps this is not entirely clear from EDAL itself, but the etymology does include the reference to [MILLER 1971: 86], where this connection was suggested for the first time, and in his work, MILLER glosses the word in more details: 'enclosure for worship of native deities; later, buildings erected within such enclosures'.

least several dialects of Evenki and Even languages. Granted, it is less frequent than 'to roast on a skewer', but a reviewer who constantly accuses his oponents of not paying attention to details should probably be setting an example of higher standards himself.

Second, leaving the ethnographic references to cultural ways of the Tungusic peoples aside for a second[45], we should note that PTM actually has *two* main verbal stems for which the meaning 'roast' can be reconstructed; besides *sila-*, there is also *dalga-* [TMS I: 193–194]. The basic semantic oppositions in attested languages are as follows (only data from those languages where both stems are attested are given):

Evenki: *sila-* 'roast on a spit' / *dalga-* 'roast (on a spit or pan)', 'bake (on coals)'. Even: *hil-* 'put on a spit' / *dalgat-* 'roast on a spit'. Negidal: *silat-* 'roast on a spit' / *dalga-* 'roast fish on a spit'. Orok: *silotči-* 'roast on a spit' / *dalda-, daldatči-* 'roast in a pan, to grill'. Nanai: *silo-* 'roast on a spit' / *dalGa-* 'roast (gen.), bake'[46].

While some uncertainties as to the exact details of semantic reconstruction may exist here, several facts are crystal clear: (a) PTM had *more* than one verb to denote the idea of 'roasting'; (b) out of these two verbs, *sila-* is *always* associated with a spit, while *dalga-* only occasionally. The logical conclusion is that it is unreasonable to reconstruct PTM *sila-* 'to roast', as it is quite specifically 'roast *on a spit*', and occasional attestations of the meaning 'put on a spit', as well as the fact that the noun 'spit' is always formed from *sila-* and never from *dalga-*, only confirm that conclusion further. Yet the existence of *dalga-* is not even mentioned in AV's argument, which is curious for somebody who so frequently accuses his opponents of «sweeping the data under the rug».

If, therefore, PTM *sila-* is 'roast on a spit', such a verb could quite easily be formed from a noun meaning 'spit' (in fact, there is hardly any other possible way of derivation), after which the original noun became lost and a new deverbative (*sila-pūn*) was formed. We can understand that AV may find such a scenario hard to believe, but perhaps his trust can be gained if an extremely similar typological parallel is adduced: cf. Dutch *roosten* 'roast', derived from Middle Dutch *roost* 'grill, roaster'; the Middle Dutch *roost* is no longer used in the modern language, but instead, we have *rooster* 'grill, roaster', derived from

---

[45] AV's statement is accompanied by an impressive number of references, but, although in many of the works cited it is indeed stated that roasting on a spit is the primary way of roasting anything among members of the described population, very few of them explicitly state that other ways of roasting are/were unknown to them. And AV is most definitely mistaken that «roasting on charcoals seems to be attested only for Udege»: cf. [Turayev 2001: 67], where it is quite clearly stated that one traditional way of preparing wild fowl among the Oroch people was to smear it in clay and roast on coals.

[46] Two languages — Solon (*dalga-*) and Udege (*daga-*) display the meaning 'to burn down (tr.)', but this is a secondary development ('roast' > 'overroast' > 'burn down').

the verb *roosten* [DE VRIES 1992: 590]. And all this considering Dutch did not even have a time period of six thousand years to undergo all these developments. Similar cases can be adduced for other language families as well.

This example will conclude the «philology» discussion. We consider AV's theoretical grounding of this argument to be oversimplified, poorly stated, and potentially offensive not just to the EDAL/SCH authors, but to an overwhelming majority of historical linguists working on large scale problems, and his practical examples subjective, occasionally misinformed and inconsistent with his own theoretical remarks. That said, let us move on to the next group of arguments.

VII. *Ignorance or irrelevance*?

One important factor contributing to the overall low quality of EDAL and other Altaic publications of the Moscow school, according to AV, is that the authors of EDAL are almost completely self-sufficient and routinely ignore important research on the same issues by «non-Moscow» specialists in the field[47].

Cf.: «The authors of EDAL often make pronouncements and statements with a voice of authority that are outdated or mistaken. Quite frequently they also present a one-sided judgement not supported by any argumentation on an issue that is known to have a history of a heated debate in the past and for which no universally accepted solution is known. Sometimes they just ignore what their opponents have been repeatedly saying. Overall, their knowledge of the current state of the art in the field and/or of the most pertinent sources is sadly lacking» [VOVIN 2005: 85].

AV is, of course, entitled to his opinion about how much the authors of EDAL know and how much they do not; as for us, we will avoid any direct evaluation of the amount of our knowledge because it is hard to measure such things on an objective basis. Two things, however, have to be made clear, in conjunction with this issue: what exactly *is* EDAL, or, at least, what it proclaims to be, and what it is *not*.

What EDAL actually *is* is an up-to-date compendium of information on Altaic etymology as seen through the phonological and lexical reconstruction of its authors. This means that major efforts were made in order to compile all, or, at least, an absolute majority of all the Altaic etymologies offered over the past century; add in extra material; and systematize all of it through a unification of phonological correspondences. Accordingly, all previous etymologization (with the exception of those comparisons that do

---

[47] This line of argumentation is not, however, applied by AV towards the Sino-Caucasian research of S. A. STAROSTIN — most probably because AV himself is not an acknowledged specialist on either Sino-Tibetan or North Caucasian and therefore considers himself unqualified to make such a statement.

not fit in and had to be rejected) has been documented, and all the relevant sources indicated within the etymologies and included in the bibliography.

What EDAL, however, is *not* is «A Handbook on the History of the Altaic Debate». This does not mean that such a book would not be useful. On the contrary, a detailed, objective, and all-encompassing presentation on the evolution of Altaic studies would be extremely welcome these days (although the reader will certainly understand if we do not recommend that it be written by AV). But the goals and the scope of such a volume are completely different from those of EDAL.

This is why, despite the abundance of published (and unpublished) literature on Altaic and its daughter branches, it was decided that the main emphasis in the preparation of EDAL would be placed only on those works that have *major* relevance for the version of the Altaic reconstruction that is advocated by EDAL's authors. With all due respect to authors of works that were omitted from the bibliography or quoted only briefly without any prolonged arguments, this was the only possible way so as *not* to let EDAL be transformed into the aforementioned «Handbook».

It should also be said that, in his effort to demonstrate how ignorance of their opponent's views constantly leads EDAL's authors to catastrophic conclusions, AV is taking up an oddly «non-discriminating» stance. Reading through his review gives the feeling that he is presently willing to embrace *any* argument, «traditional» or «modern», «mainstream» or «controversial», presented by *any* scholar, as long as it stands in the way of EDAL's Altaic theory. Sometimes he goes as far as to present two self exclusive arguments without a clear indication of which one it is that he himself prefers[48] (but, apparently, *any* of them will do as long as they help avoid the Altaic curse). Furthermore, in this quest to put as much distance between himself and «Altaic» as possible, he even adopts a series of arguments that are only *superficially* «anti-Altaic», but not *substantially* so.

Thus, he now (for the first time in his scholarly career) prefers to reconstruct the traditional Turcological **š* (and, probably, **z*, although this is not mentioned explicitly) in the place of the «pro-Altaic» reconstructions **ĺ* and **ŕ* [Vovin 2005: 94], despite the fact that even Gerhard Doerfer, to whose memory the article is dedicated, was inclined to viewing the Chuvash representation of these phonemes (*l, r*) as archaic rather than innovative [Doerfer 1984], and was absolutely right in saying that one's choice of the variants to

---

[48] Returning to the already discussed problem of PM **čila-ɣu* 'stone' as a potential borrowing from «Bulgar», it must be noted that the same passage, with a reference to G. Doerfer, also states: «First, the reconstruction is likely to be **čila-xun < *čila-pun*, with a suffix **-pun* that is used to derive nouns from verbs. True enough, the corresponding verbal stem is not attested» [Vovin 2005: 111]. We are confused: is this, after all, a nominal borrowing with the later addition of a suffix *or* a verbal derivative from a lost stem?

be reconstructed is *irrelevant* to the Altaic question (one of his articles even bears the name «Zetacism/Sigmatism Plays No Role» [DOERFER 1988]) — because comparison and reconstruction, as is well known in historical linguistics, depends upon phonological correspondences and not phonetic similarity. And yet AV now insists on reconstructing, e. g., PT *jaĺ- as *jaš-, because, of course, any representation that will make PT look less like PA is *a priori* better than any other. What he does not seem to realize, or, at least, does not mention, is that even if we go through the entire dictionary and change every instance of PT *ĺ to *š, this will have exactly *zero* influence on the validity of the proposed correspondences and etymologies.

Nevertheless, just to ensure the reader that the authors *are* familiar with many of the arguments they have «ignored» within EDAL, below we will try to answer — briefly, but concisely — most of the specific allegations made by AV in the «ignorance of the current state of the art» section of his review (pp. 85–92).

(1) On the reconstruction of initial voiced stops for PT: «Dybo in a sketch of Turkic comparative and historical phonology makes no reference to works by DOERFER (1969, 1971) that represent a thorough refutation of Illich-Svitych's proposal regarding the reconstruction of voiced initial *d- and *g- for Proto-Turkic. As Doerfer pointed out, the initial voiced *d-* and *g-* in Oghuz languages cannot be old, because they are found in recent loanwords from Persian that have original voiceless initial consonants, e. g. Persian *tōbra* 'bag' (< Indo-Aryan *tōbra*) > Tur. *dorbacik*; Persian *taġār* 'vessel' > Tur. *daġarcik* [DOERFER 1971: 276]... Until the authors of EDAL manage to disprove DOERFER's argument, I, like most Turcologists nowadays, will side with DOERFER and reconstruct only voiceless *t- and *k- for PT» [VOVIN 2005: 85].

The main reason why the authors of EDAL have, up to this moment, been refraining from «disproving DOERFER's argument», is that, in our opinion, it is an argument based solely on inadequate knowledge of the material — not Turkic, of course, where G. DOERFER has always been an acknowledged specialist, but comparative Indo-Iranian. However, since AV now presents it in such an ultimative manner, a discussion is probably in order. Let us look a little more closely at both examples.

First of all, AV's «Indo-Aryan» is misguided; obviously, what he meant was «Indo-Iranian». Second, there is no general consensus about the direction of the borrowing; several acknowledged Iranologists insist that it is the *Persian* word *tōbra* that has been borrowed from a Turkic source, cf., e. g., [TSABOLOV 2001: II, 214]. Third, in an earlier publication [DOERFER 1965: 594], G. DOERFER happens to be less categorical about any historical scenarios, merely suggesting a connection between the Turkic form and the Indo-Iranian ones.

Fourth, there is no serious evidence that «Indo-Aryan» *tōbra* really represents a form that can be reconstructed for either Proto-Indo-Iranian,

Proto-Indo-Aryan, or Proto-Iranic. The assumption really rests on the corresponding entry in R. Turner's dictionary [Turner 1966: № 5972], where we see a reconstructed «Indo-Aryan» form *tōba* and «Iranic» **tūbraka*. It should be heavily stressed, though, that Turner's dictionary is *not* a reconstruction of Proto-Indo-Aryan, but rather just a compendium of various lexical layers in Indo-Aryan languages that includes inherited lexicon along with numerous borrowings into later stages of Indo-Aryan languages. Without a doubt, what we are dealing with here is a case of the latter, because Indo-Aryan **tōba* and Indo-Iranic **tūbraka* cannot be easily reconciled with each other and are certainly *not* inherited from Indo-European.

Moreover, Indo-Aryan **tōba* essentially rests upon the shoulders of Sanskrit *ṭōpara-* 'small bag' — which neither corresponds well phonetically to the rest of the Indo-Aryan forms, nor is, in fact, an old word, being attested only once in a late (14th–15th century!) monument; and Indo-Iranic **tūbraka* is equally unattested in any of the old Iranian languages, so in reality what we are dealing with is just Persian *tōbra*, later on borrowed into both Iranian (Bakhtiyari *turba*, etc.) and Indian (Panjabi, Hindi *tobrā*) sources. Amazingly, even though Turner quite clearly states that the Panjabi and Hindi forms are borrowed from Persian, Doerfer [1971] surmises just the opposite (Persian *tōbra* < Indian).

Doerfer's main argument in favour of the Turkic form being borrowed from Persian is that in the latter it is attested starting from the 10th century (Gārdēzī *tōbra*), whereas Turkic only starts featuring it in 14th century monuments (Codex Cumanicus, etc.). However, this is not really an argument; there is no such word in Middle Persian or any of the anciently attested Iranic idioms. On the other hand, traces of the Turkic form can be found even in Hungarian, where *turba* 'sack' is hardly < Osmanic Turkish, since the vocalism indicates a much earlier time of borrowing (most probably, PT **tōpra/*tōrpa* > Old Hungarian *tūrba* > mod. *turba*). If the Hungarian borrowing is from Danubian Bulgar, the word is unquestionably of Proto-Turkic origin; otherwise, it is at least Common Oghuz-Karluk-Kypchak. There are also possible Altaic parallels, but we will not dwell on that for now; even the already adduced evidence is enough to show that the probability of borrowing from Persian is extremely low.

However, none of this really matters in the light of the fact that there is simply *no voicing in Oghuz* in this particular case! «Modern Osman *dorbacık*», mentioned by Doerfer in [1971], simply does not exist (but *torba* does); nor are there any voiced consonants in the reflexation of this root in any other Oghuz languages. One form that *may* exist is Old Anatolian Turkish *dopraj̆uq* 'ein kleiner Sack', quoted in [Doerfer 1965: 593] from [Vámbéry 1901], but Turkish manuscripts of that particular epoch are well-known to frequently confuse the Arabic letters for *d-* (د) and *ṭ-* (ط), used in words with back

vocalism (see the published text of *Kalila-ve-Dimna* in [ZAJĄCZKOWSKI 1934] for this kind of orthographic confusion), and the form cannot be considered diagnostic, at least, not until strict orthographic rules for the manuscripts in question have been established. No such rules are formulated by DOERFER either in [DOERFER 1971] or in [DOERFER 1969], where Old Turkish graphic forms with *d-* are analyzed based not on statistical analysis of actual manuscripts, but rather on their being mentioned in a *historical dictionary* of Turkish [TARAMA SÖZLÜĞÜ] (and we all know AV's attitude towards dictionaries).

The second example is even more instructive. It is a Common Turkic root for 'sack', found in Turkish *dağar*, Gagauz *daar*, Azerbaijani *daɣar*, Tuvan *t'aar*, Khakas *taar*. The root has been borrowed both into Persian [*taġar* 'sack (as a measure)'] and into Mongolian (*taɣar* 'sack'); it also has a decent Altaic etymology (cf. Mongolic **togu-ɣan* 'cauldron', Evenki *taɣa* 'a k. of basket'). In [DOERFER 1965: 512–519] it is suggested that the word *may* have been borrowed from Turkic into Persian, although he doubts that the word is really Common Turkic and thinks that it may have been borrowed into Turkic as well from an unknown source. But in [DOERFER 1971] the suggestion is suddenly reversed — even though there is no Iranian etymology whatsoever for Persian *taġar*, and, again, the word lacks any attestation in old Iranic sources. DOERFER's old «non-argument» is repeated here again: the word is attested in Persian since 1035, and in Turkic — since 1072. Another argument is that there is no Turkic root like **ta-* or **tag-*, from which **tagar/*dagar* could be derived; this is even less convincing, given that no internal etymology can be offered for the Iranian root either. So, what we have is a detailed etymological treatment of the word in [DOERFER 1965] with one conclusion, and a very brief and unargumented remark in [DOERFER 1971], with a radically opposite conclusion. With all due respect to the change in G. DOERFER's opinion, we prefer to side with his former interpretation.

In [DOERFER 1969], these «loanwords from Persian» are only augmented by one more example: Turkic *deng* 'equal', compared with Persian *täng*. Again, however, the Persian form does not have any solid Iranian, let alone Indo-European, etymology (and this is also acknowledged by DOERFER himself, who suggests that the Persian form is an earlier borrowing from a Chagatai source). In addition, the Persian form really means 'pack, sack (of sugar)', and could hardly have served as a borrowing source for a common Oghuz root represented by forms like Turkish *denk* 'equal, pair; balance', Gagauz *denk* 'equal', Turkmen *deñ* 'equal, same', etc. A somewhat better etymology would try to derive them from Persian *däng* 'half a load (of horse); balance', *dängädäng* 'equal' (lit. `*däng* vs. *däng*'), but, first of all, this would already strengthen the case for primary voicing in Oghuz rather than weaken it (since these Persian forms begin with a voiced consonant), and, second, even for these forms it is more reasonable to suggest a Turkish origin rather than vice versa. Note that

the Oghuz forms also have a regular parallel in Chuvash (*tan* 'equal'), unambiguously suggesting a PT origin. Beyond that, the PT form *deng is most likely a Chinese borrowing (cf. OC 等 *tə̄ŋʔ* 'equal; rank'), with a PT voiced consonant regularly reflecting an OC unaspirated stop.

Recent research on the voiced/voiceless opposition in Turkic has helped to discover even more regularities in the evolution of this opposition within Turkic — cf., e. g., [Dybo 2005] — and thus corroborate V. M. Illich-Svitych's original theory that tried to establish certain patterns in the seemingly chaotic distribution of attested variants. This does not mean, of course, that we are not able to make good use of a lot of G. Doerfer's research on Turkic historical phonology either. On the contrary — EDAL, as well as numerous other publications by EDAL's authors, draws a lot of important conclusions from a careful analysis of the linguistic material adduced in Doerfer's works. The same, unfortunately, cannot be said of our opponent, who prefers to focus on Doerfer's theoretical discussions and conclusions rather than the actual etymologies.

(2) In the brief section on problems with EDAL's internal reconstruction of Tungus-Manchu, AV writes: «Dybo and Starostin either ignore (or are unaware of) all the research that was done on the reconstruction of PTM between Cincius's seminal monograph and their own publications… Doerfer cogently demonstrated that an additional vowel *ö has to be reconstructed in PMT [Doerfer 1978], and this reconstruction creates havoc in some of EDAL's vowel 'correspondences'» [Vovin 2005: 87].

It is rather hard to take these statements at face value. First and foremost, the «honour» of reconstructing an additional *ö for PTM does not belong to Doerfer; this vowel was already introduced by J. Benzing [1955], and this is acknowledged by Doerfer himself on the second page of his article. Second, the reconstruction of *ö did not at all depend on the recognition of any new types of hitherto undetected vocalic correspondences; rather, it was an isomorphic reinterpretation of the system as originally proposed by Cincius [1949] (in the latter the vowel that Benzing interprets as *ö* usually corresponds to Cincius' *ụ̈*).

Third, and most important, the system as proposed by both Cincius and Benzing was *thoroughly analyzed* in [Starostin 1991], where some of the qualitative distinctions introduced were reinterpreted based on the proposition that PTM did not have vowel harmony, and, therefore, that the series of correspondences for the first syllable depended on the quality of the vowel in the second one. Very roughly speaking, then, Benzing's and Doerfer's *ö in PTM corresponds to the following PTM structures as proposed in [Starostin 1991] and followed in EDAL:

Doerfer 1978	EDAL	Examples
$ö_1$ - $ö_2$	u - u	'fat': D. *börgö-, EDAL *burgu-
		'squirrel': D. *hölökī, EDAL *xulukī
		'six': D. *ńöŋön, EDAL *ńuŋun
$ö_1$ - $ä_2$	u - e	'mountain': D. *hörä, EDAL *xurē
		'dig': D. *hölä, EDAL *xulē
		'wound': D. *pöjä, EDAL *puje
$ö_1$ - $ü_2$	u - $ü$	'bladder': D. *höžükü, EDAL *xužük
		'power': D. *kösǖn, EDAL *kusǖ-n
$ö_1$ - i_2	u - i	'cradle': D. *döri, EDAL *duri
		'eagle': D. *gösi, EDAL *gusi
		'belt': D. *(h)ösī, EDAL *usī

The isomorphism between the two systems is not one hundred percent rigorous, and one may argue about the respective flaws and advantages of the «+ vowel harmony» vs. «– vowel harmony» interpretations (a detailed discussion would take up at least a couple dozen pages), but one thing is for certain: it is essentially a matter of *phonetic interpretation* of the system rather than one of disagreeing over the actual correspondences, and regardless of whether one prefers the treatment of TM vocalism in EDAL or in Doerfer's article[49], it is impossible to foresee any kind of havoc in EDAL's correspondences; were we to adopt Doerfer's position, all we would have to do is change some of the symbols. The inevitable conclusion is that either AV is intentionally distorting the facts, or that he has simply confused the notions of phonetics and phonology. Symbolic notation can sometimes be treacherous.

(3) On pp. 87–89, S. A. Starostin is accused of intentionally ignoring a large number of sources on Korean dialects (including archaic ones such as Chejudo and Yukch'in), as well as pre-Hangŭl sources like the already mentioned *Kyerim* wordlist along with even older material (the *hyangga* of the 7th–11th centuries). The dialects are said to «preserve certain features that are crucial for the reconstruction of Korean», while the *hyangga* information «on the earlier history of Korean phonology and morphology is quite revealing».

Unfortunately, these statements are somewhat weakened by the fact that, throughout the rest of the review, AV fails to build up upon them and demonstrate where exactly EDAL treatment of Korean material leads to objectively incorrect results, either for Korean itself or for external etymologization of the items. The only serious example is the Korean word

---

[49] For additional insights into specific problems concerning interpretation and reconstruction of the PTM vocalic system, see [Dybo 1991] and [Dybo 2000] (it is unclear whether AV is aware of the existence of these works).

for 'head', which was already discussed above — not only is this one of the instances where pre-Hangŭl material is actually *discussed* in EDAL (thus contradicting AV's allegations), but, as has been shown, AV's reconstruction of inlaut *-t- in this word is inconclusive.

As for dialectal information, the only time AV demonstrates how it should be used is on p. 116, where EDAL's PK form *săr- 'to burn' is changed to *solok- based on the Northern Phyenganto form *salkwu-*. Not only is this reconstruction *also* inconclusive (the same dialect also has a *k*-less form *salwu-* and additional information is necessary to prove that the -*k*- is to be extrapolated onto a higher level), but even if we do adopt it, the comparison with Altaic * zi̯oĺa 'shine' still remains valid, because the last consonant of the «root» may reflect the old Altaic verbal suffix *-k´- [EDAL: 216–218]. The fact that PK *solok- «cannot be divided any further» on Korean territory is completely irrelevant (see section V).

This certainly does not mean that Korean dialectal data, or the data on pre-Hangŭl Korean, cannot be of *any* use to Altaic comparative studies and should always be ignored. But the best use one *can* make of such data is when it is treated systematically, and not just based on isolated quotations of peculiar forms and subjective hypotheses. A systematic reconstruction of Proto-Korean, taking into account all (and not just select) relevant data from archaic dialects, treating the pre-Hangŭl data in a non-speculative manner, and represented by a large comparative-etymological dictionary, is, unfortunately, still lacking (it was hoped at one time that AV himself would take up this burden, but in the end, this never happened). Until then, the main source for Korean data — being at the same time phonetically archaic, systematically transcribed, and extremely well represented — is 15th century (and later) Middle Korean. If AV intends to challenge this statement, he is welcome to try.

(4) Further down, Starostin is taken to task for adopting the phonemic status of intervocalic *-b- and *-d- in Proto-Korean. Starostin: «we accept here the basic reconstruction proposed in [Ramsey 1986a], rather than the poorly grounded theory of intervocalic voicing *-VCV- > *-VZV- put forward in [Martin 1996]» [EDAL: 164]. AV: «The reader should be aware that in fact 'poorly grounded theory' is the one that is currently accepted by a majority of Korean historical linguists both in Korea and abroad, and that it was not invented by Martin. However, in his seminal 1996 book, Martin provided inescapable evidence for the lenition theory which is now swept away by Starostin in one phrase, and, as always, without any argumentation» [Vovin 2005: 89].

The reference to the «majority of Korean historical linguists» is, as usual, given without any actual statistics — not that it really matters, because AV's

ever-increasing tendency to apply the «voting principle» to scientific matters goes against not only much of his own research[50], but against common sense as well. Putting that aside, though, it is not difficult to see why STAROSTIN allows himself to refrain from an extended discussion on the issue: the solution advocated in EDAL does not belong to any of the authors of EDAL, but to a distinguished Korean historical linguist (Robert RAMSEY). The truth is that there are simply two concurrent hypotheses: the voiced intervocalic stops one by RAMSEY and the lenition one by MARTIN. Out of the two, RAMSEY's agrees better with comparative Altaic data and is therefore accepted in EDAL.

Interestingly enough, here is what MARTIN himself has to say on the position of his opponent (RAMSEY): «In my opinion there is no need for these extra consonants in proto-Korean when all the facts are taken into account, but it remains an interesting idea and is particularly attractive to those seeking to tie Korean verb stems to the proto-Altaic reconstructions proposed by Poppe and others, though Ramsey wisely sticks to internal evidence... he has sought to eliminate distinctive accent from proto-Korean and he postulates the additional set of obstruents to do so» [MARTIN 1996: 2]. So this is really what it all boils down to: ascribing primary phonological relevance to *segmental* or *suprasegmental* features in proto-Korean. And since the first solution finds additional support in external evidence, STAROSTIN can hardly be blamed for taking RAMSEY's side in the dispute. Again, as in (3) above, we are dealing with an either accidental or intentional confusion of the problems of reconstruction of different oppositions with those of phonetic interpretation.

(5) To conclude with Korean, here is another interesting passage: «PK *$\bar{\imath}r$ 'work, profession' based on MK $\bar{\imath}r$ is a live witness to a repeatedly manifested ignorance of the most basic facts about the Korean language history. It is very well known, contrary to EDAL's unfounded statement: 'In MK orthography, length was marked by two dots...' (p. 165), that these two dots do not represent the length, but indicate the RISING pitch, thus MK :*il*. The reference to that can be found in the most elementary books on the history of Korean... which, of course, has not been consulted. Furthermore, there is overwhelming evidence... that monosyllabic roots with RISING pitch have a disyllabic origin (Ramsey 1991: 227), (Martin 1996: 2–5), etc. Thus, the PK form of the word in question is *ilu LH» [VOVIN 2005: 97].

---

[50] For instance, as far as we know, AV has not yet gone on record as to renounce his phonetic interpretation of the MK grapheme △ as a palatal nasal *ñ* [VOVIN 1993], over which he happens to be in agreement not with the «majority of Korean historical linguists» (including *both* MARTIN and RAMSEY, who interpret it as the voiced fricative *z*), but rather with the authors of EDAL. Thus, the «voting argument» is seen to be applied inconsistently.

No further references to Starostin's text are given, and the connotations are not difficult to interpret: (a) Starostin knows nothing about historical Korean suprasegmental phonology; (b) Starostin knows nothing about the existing theories that try to interpret historical Korean suprasegmental phonology in *segmental* terms. Just for the sake of the argument, though, let us pay a little more attention to the unfortunate p. 165 of EDAL:

«In MK orthography, length was marked by two dots and thus *perceived as a prosodic feature of a syllable* [italics are ours — *A. D., G. S.*]»; compare this with: «No vowel length was transcribed as such, but *from philological sources we know that syllables with rising pitch were pronounced long* [italics are ours — *A. D., G. S.*]» [Ramsey 1986b: 2].

And, slightly above, we find:

«All Mkor. vowels could be long or short, and *it was convincingly demonstrated by Ramsey 1978* that the long vowels should have originally resulted from contractions and *a reduction of the vowel of the next syllable*. In many individual cases, however, this is not quite clear, so we preserve the feature of length for 'Proto-Korean' — *although it certainly is not of Altaic origin* [italics are ours — *A. D., G. S.*]».

No further comments are probably necessary on this issue, so let us move on to the following one.

(6) Starostin's reconstruction of a four-vowel system for PJ (*a, *i, *u, *ə) is criticized as follows: «It has been already pointed out by Hattori Shirô more than twenty years ago that many Ryukyuan dialects preserve PJ *e and *o that have been raised to *i and *u, merging with original *i and *u in other Ryukyuan dialects and in most of the Japanese dialects [Hattori 1978–79], with the exception of EOJ and the Hachijō dialect that also preserve the distinction at least partially» [Vovin 2005: 91].

Again, one cannot help but wonder at this further inconsistency, because in this passage AV is certainly not advocating the «mainstream» view and does not appeal to the authority of «most Japanologists» as he sees fit to do in the Turkic or Korean cases described above. Starostin's reconstruction of the PJ vocalism (which was, by the way, published *before* Hattori's; see [Starostin 1975]), in this particular case, happens to coincide almost completely with the most authoritative version so far — Samuel Martin's [Martin 1987: 67] (the only major difference being phonetical, with Martin's *o corresponding to Starostin's *ə). Martin's arguments concerning Hattori's system and why the evidence to accept it is unsufficient can also be found in the same publication [Martin 1987: 68].

Leaving, however, the subjective question of «authority» behind, the primary reason why the six-vowel reconstruction for PJ is ignored in EDAL is that, as of yet, no two scholars have agreed on the same scenario for its

development both inside and outside Ryukyuan, and practically every-one — HATTORI, UNGER, THORPE, SERAFIM, MIYAKE, etc. — reconstructs *e* and *o* in a different subset of lexical entries, based on a different set of criteria drawn from different kinds of evidence. In addition, the same evidence is occasionally explained by other scholars (e. g., WHITMAN) as secondary and innovative. A good overview of the various conflicting theories can be found in [MIYAKE 2003] (unfortunately, the article could not have been used in the preparation of EDAL for chronological reasons).

Thus, when AV claims that the first syllable vowel in PJ *(d)ísì* 'stone' «is *e-*, that can be safely reconstructed on the basis of the Ryukyuan evidence» [VOVIN 2005: 112], he is exaggerating the degree of safety: for instance, in [THORPE 1983: 335] the Ryukyuan reconstruction is *isi*, based on more than twenty dialectal forms, all of which have *i-* in the anlaut position. At the very least, given the uncertainty of the situation, AV might have given the evidence that allows us to see the *e-* in Ryukyuan. The fact that he did not, along with the enumeration of the forms in [THORPE 1983], leads us to suspect that such evidence does not exist.

AV's brief demonstration that PJ *ki-* > Shuri *ci-*, but PJ *ke-* > Shuri *ki-*, given in a footnote without any references [VOVIN 2005: 91], is equally unconvincing, because the exact conditions of Shuri palatalization are so far unclear, and it is known to occur before other vowels as well (cf. *ci-iru* 'yellow' < PJ *ku-* or *kə-* [MARTIN 1987: 449]). In addition, in an absolute majority of the forms reconstructed with *ke-* in [THORPE 1983] (where the palatalization in Shuri and other dialects does not take place) it can clearly be seen that Ryukyu *-e-* is secondary: cf. PR *ke* 'tree' < PJ *kəi* (not *ke), PR *kebusi* 'smoke' < PJ *kai(N)-puri*, PR *ke* 'hair' < PJ *kai*, etc.

About the only possible exception is the word for 'wound', quoted by AV, which is reconstructed by THORPE as *kezu* but corresponds to OJ *kjizu* < PJ *kensu* or *kinsu*; however, the lack of palatalized variants for this word among the Ryukyuan dialects may simply be explained by the fact that they have been replaced by borrowings from standard Japanese (and this is partially acknowledged by THORPE: «the Amami-Okinawa area has been inundated with borrowed forms similar to *kizu* (OJ, standard Japanese)» [THORPE 1983: 351]).

In short, given that (a) the general situation is hugely controversial and (b) piling up problematic vocalic reconstructions for PJ on top of the already problematic vocalic reconstruction for PA would not be reasonable, in the matter of Japanese vocalism it was decided to go along with the *mainstream* reconstruction of the system. It is not excluded, of course, that eventually some clarification will be achieved, and that it will be possible

to incorporate a revised system of PJ vocalism within the Altaic system; but for the moment, we do not see any sufficient reason to revise it.

At least one scholar who might see fit to agree with us on the issue is AV himself, cf.: «While Martin reconstructs four basic vowels (*a, *i, *u, *ə), following the lead of Susumu Ōno, Leon Serafim proposes to reconstruct six, adding *e and *o to the system. In an earlier proposal, Serafim considered the reconstruction of ten, if not more, vowels. Although his later proposal is much more persuasive than the earlier, the extremely rare occurrence of *e and *o in initial position leaves room for doubting the validity of their reconstruction. *For the time being, it is safer to adhere to the traditional Ōno-Martin reconstruction* [italics are ours — *A. D., G. S.*]» [VOVIN 2001b: 94]. Compare this with [VOVIN 2005: 91]: «PJ does not have four vowels as alleged by Starostin, but six: *a, *u, *o, *i, *e, *ə, and that, as the reader will see below, is going to influence correspondences».

Granted, in between the years 2001 and 2005 we see AV's stance on the issue somewhat gravitating from skepticism to approval, as additional argumentation in favour of the six-vowel reconstruction has been introduced by J. WHITMAN, M. MIYAKE, and others ([VOVIN 2003: 425]). Nevertheless, the main flaw of the idea — namely, that there is too little general consensus about which exact items are to be reconstructed with *e and *o — has not in the least dissipated, nor is it entirely transparent how the various sources of evidence (many of them extremely ambiguous) can, at the present moment, be satisfactorily reconciled with each other; no such reconciliation has been presented so far by any of the supporters of this idea. Overall, the whole issue, at least, the way it is presented in AV's review, seems vastly exaggerated, and raised with the foremost aim of finding another pretext to accuse EDAL authors of ignorance[51].

It can thus be seen that, for the most part, the literature that is «ignored» in EDAL falls into the «questionable» category, that is, presents the scholarly world with hypotheses based on ambivalent evidence that can really be subject to several different interpretations. In all such cases, the standard methodology of EDAL is simple: if a hypothesis based on such ambivalent internal evidence finds additional confirmation *outside* the gi-

---

[51] A very similar situation can be observed with AV's position on whether Yonaguni *d-* is an innovation (< PJ *y-*) or an archaism. In [VOVIN 2001b], AV seems to incline more towards the latter solution, arguing with WHITMAN's hypothesis that PJ *y-* > Yonaguni *d-* on the basis of various historical evidence. In [VOVIN 2005], AV already calls the reconstruction of PJ *d-* on the basis of the Yonaguni reflex «highly controversial». Needless to say, no explanation for this sudden change of attitude is provided; in fact, it is not even explicitly *mentioned* that there *was* a change of attitude, and this will certainly confuse anyone who might want to read [VOVIN 2001b] and [VOVIN 2005] back-to-back.

ven branch of Altaic, it is given the green light. If it does not, we reserve the right to ignore it. Thus, the Oghuz distinction between initial voiced and voiceless stops seems to be confirmed as archaic through regular correspondences with PM and PTM. Likewise, the correspondence «Yonaguni *d-* : all other Japanese dialects *y-*» is interpreted as representing PJ *d-* because the same solution is *also* advocated by external evidence[52]. On the other hand, the controversial increase of PJ vocalism based on Ryukyuan evidence, that *may* also be interpreted as secondary, corresponds to nothing outside Japanese and is therefore allowed to be neglected.

AV, of course, may and will accuse us of circular logic, but this would be an incorrect understanding of the principle, since what we are saying is not, e. g., «Japanese is Altaic, therefore Japanese has to be reconstructed in the most 'Altaicized' way possible», but «Japanese is Altaic, *because* it is possible to reconstruct Japanese in an 'Altaicized' way (e. g. compatible with Turkic, Mongolic, etc. on many issues), but, for instance, impossible to reconstruct it in a 'North Caucasian' or 'Austronesian' manner». This has nothing to do with circular logic; rather, it has to do with testing of various scientific hypotheses.

To conclude this section, it must be stated, of course, that *some* publications that might contain crucial information might have been missed by the authors; the Altaic field boasts an immense number of studies that reflect several extremely old traditions, and occasional gaps in knowledge are unavoidable. But within AV's lists of recommendations we find very few such publications — most of them in the Korean section, next to none in the others. Overall, we would think that a statement like «their knowledge of the current state of the art in the field and/or of the most pertinent sources is sadly lacking» would need to be validated by better factological evidence than by what is presented on pp. 85–92.

This is not even mentioning that in some respects at least this knowledge, in our humble opinion, slightly exceeds that of our esteemed opponent, and this can be easily illustrated on two examples.

(a) AV: «In the section on Altaic numerals, EDAL cites Old Bulg[har] *tvi-rem* 'second' (p. 223). The word *tvirem* is attested in the famous 'List of Danube Bulgar Princes', but the problem is that we simply do not know what it means, as there are many interpretations of the Bulgar words on

---

[52] But certainly not *required* by it, since reconstructing this phoneme as PJ *y-* will not affect the correspondences with Altaic. Actually, even though AV's review might make it seem that the authors are unfamiliar with the recent argumentation *contra* PJ *d-*, this is not the case; the possibility of reconstructing *d-* instead of *y-*, based on the very same reasons that AV is quoting (a 15th century Korean transcription of the name of the Yonaguni island, 與那國[島]), was explicitly stated by S. A. Starostin already in [Starostin 1991: 135], although he did not think the evidence sufficiently conclusive.

this list, and most of them are controversial. Maybe EDAL's authors managed to achieve a major breakthrough in deciphering this list, but without revealing to the rest of the scholarly world the results of their study with all pertinent argumentation, it makes assigning to this word the meaning 'second'... look like elementary cheating» [VOVIN 2005: 85–86].

Incidentally, one of EDAL's authors — O. MUDRAK — *has* achieved such a breakthrough, or, at least, has a perfectly well-working theory on how the list should be deciphered. Granted, the results of his research, for technical reasons, remained unpublished until 2005[53], i. e. two years after the publication of EDAL, but the work itself was completed much earlier, even though AV cannot be blamed for not being familiar with it. The study itself quite convincingly proves, along with other things, that each of the lines in the «List» documents the rulership period of the corresponding prince, and that the last word in these lines is usually related to the corresponding ordinal number in Chuvash, cf. such unmistakable examples as **тоүтомъ** = Chuvash *tъᵂvatъm* 'fourth', **вечемь** = Chuvash *viśəm* 'third', **алтомъ** = Chuvash *oltъm* 'sixth', **читемь** = Chuvash *śičəm* 'seventh'. Herein also **твиремь** = Chuvash *təbərəm* 'second' (AV's argument that the Chuvash word does not really mean 'second', but 'second out of two', 'following', is completely irrelevant to the issue, since it has no influence on Altaic etymologization).

As for the word **диломъ**, found in conjunction with **твиремь**, AV follows the hypothesis that derives it from PT **jïlan < *dïlan* 'snake', stating that this «may pose another significant problem to EDAL's reconstruction of PT *d- and *g- on the basis of Oghuz evidence, since it demonstrates... that Common Turkic *j- <*d-*... under this scenario, one is faced with two *d- in PT, which is, of course, impossible» [VOVIN 2005: 86].

Well, first of all, it is, of course, quite possible, because in most cases Common Turkic *y- does* stem from an earlier *d-, only not Proto-Turkic, but Proto-Altaic; at the same time, PT *d- < PA *t- (see the corresponding charts in EDAL). We could therefore represent it the following way: (a) PA *t- > Turkic *t-, PA *d- > Turkic *d-; (b) Turkic *t- > Proto-Bulgar *t-, Proto-Non-Bulgar *d-; (c) Turkic *d- > Proto-Bulgar *d- (> Chuvash *ś*-), Proto-Non-Bulgar *j-.

The only dubious thing about this scheme is that, in order to become Chuvash *ś-*, Proto-Bulgar *d-* would probably still need to go through the *j-* stage, which would make us position *two* independent developments of *d- > *j-* in both branches of Turkic. This is not very probable — and that is why there is serious reason to doubt that Old Bulgar *d-* is in any way connected with PT *j-. In fact, according to O. MUDRAK's reading, the word **диломъ** does not mean 'snake' at all, but rather 'tiger', and is to be compared with PT *ta-

---

[53] It is also available online at http://starling.rinet.ru.

lɨm 'predator'; the actual word for 'snake' is represented by **верени** = Osm. *ev-ren* 'large snake, dragon', Kypchak *ewre:n* 'adder', etc. [CLAUSON 1972: 14].

Thus, even though the public explanation did arrive a little late (for which we owe the scholarly world an apology), at the very least we can assure AV that no «elementary cheating» whatsoever is involved.

(b) The other example concerns the much-debated PA etymology for 'eye', namely, the PTM reconstruction **ńia-sa*[54]. AV, following G. DOERFER and S. GEORG, insists that the actual form should be **ja-sa*, which would then preclude it from comparison with PM **ni-dü*, PK **nún*, etc., since the nasal consonant in Tungusic is essential here. Within TM, the word is indeed mostly met with an initial *j-* or zero (Manchu *jasa*, Evenki *ãsa*, Orok *isal*, etc.), but two forms are found with an initial nasal that may well be archaic: Nanai *nasal* (dialectal forms also include *ńisal* and *ŋasar*) and Jurchen *ŋia-ĉi*. The reconstruction **ńia-sa* follows this evidence; the lack of the nasal in the majority of the forms is explained as follows: «Initial **ń-* may develop into *j-* between front vowels and **-ia-*, although exact rules are not yet quite clear, because of a great deal of confusion between **n-* and **ń-* (sometimes also **ŋ-*) in this position» [EDAL: 159].

In his response, AV first discards the Nanai form, arguing that the authors of EDAL «stubbornly cling to Starostin's old idea» «in spite of the cogent demonstration by Doerfer that it is in fact an innovation [Doerfer 1995: 252–253]. Doerfer's arguments are quietly passed over, as usual...»

---

[54] AV claims that the Japanese form in the etymology, **mài(N)* 'eye', «cannot be compared with other 'Altaic' forms» as well, referring to a discussion in [VOVIN 2001: 107–108]. That discussion can be briefly summarized as follows: (a) the etymology depends on whether the reconstruction is **mài* or **màiN*, since it is only the presence of the final *-N* (= *-ŋ*) that is able to explain the development PA **niā-(ŋ)-* > [with assimilation] **ŋiā-ŋ-* > PJ **mai-N*; without *-N*, we would expect the initial consonant to remain **n-* in PJ; (b) the form **màiN* is reconstructed on the basis of Hateruma *mìń*, but it has been shown that in some Hateruma lexemes the final *-ń* is secondary; (c) therefore, we cannot reconstruct PJ **màiN* and the form cannot be compared with Altaic.

To this we may reply as follows: (a) if Hateruma *-N* is *sometimes* secondary, it does not automatically mean that it is *always* secondary; (b) MARTIN's arguments for a secondary velar nasal in [MARTIN 1987: 74–75], to which AV draws our attention, work well when arguing for truncation of longer words in Hateruma (e. g. *páN* 'foot, leg' < **pànkì*), but nowhere near as well when arguing for secondary epenthesis; the argument of OJ *potokey* 'Buddha' being Hateruma *putugiN* is also lost because no one knows what the element *-key* (*-giN*) in that word really is; (c) if the reconstruction can reasonably be *either* **mài* or **màiN* (and this is indicated in EDAL's notation: **mài(N)*), we are perfectly entitled to select the second variant for the Altaic comparison, following the standard comparative procedure.

[Vovin 2005: 83]. Well, there is a good reason to «quietly pass over» Doerfer's arguments in that article: namely, that there are *none*. The Nanai form is commented upon in the following manner: «a disturbance emanated from Central Nanai dialects». No explanation *at all* is provided for that disturbance, apart from a vague reference to «taboo phenomena» and a recommendation for linguists to «leave aside that which cannot be clarified (as, e. g. the Nanai forms with *n-, ŋ-* in the terms for 'eye')».

So this is the «cogent demonstration» AV is referring to: the appearance of *n-* in Nanai cannot be explained and should be left alone. To be fair, in a more recent article by S. Georg Nanai *nasal* is tentatively explained as being motivated by analogy with PTM **ńamū-(kta)* 'to weep; tears' [Georg 1999/2000: 151–152]. But PTM **ńamū-* itself has quite a transparent structure: it is **ń(i)a-* 'eye' + **mū-* 'water' [Starostin 1991: 30][55], and provides further evidence that the nasal in 'eye' is archaic. Whether the *n-* in Nanai *nasal* was always present in that form or was «reinstated» there based on an analogy with the same root in a different environment, then, is not very relevant to the issue.

The second piece of independent evidence for a nasal in PTM comes from Jurchen, where the word is found as *ŋia-ĉi*. This comparison was not included in [Starostin 1991] and is therefore not discussed in [Doerfer 1995]. S. Georg, in his recent reply to [Starostin 2005], remarks just briefly that «no such word exists» [Georg 2005: 455], but AV explains that in more details. According to him, the authors of EDAL are misled by the Chinese transcription of Jurchen *ja-ĉi* as 牙師, which they read **ŋạ-ṣi* (AV gives the first syllable as *ŋia:*, following E. Pulleyblank's transcription) as if it were Middle Chinese. In reality, though, the form has to be read according to the Early Mandarin standard: **ja-ṣi*, since the dictionary in which it is attested (*Huá-yí yì-yǔ*, 華夷譯語) dates back to Early Mandarin — and for that period, the reading of 牙 as **ŋạ* is anachronistic, especially since the same character is also used to transcribe Jurchen words with obvious etymological *j-*, e. g. 牙力 *ja-li* 'meat, muscle'. «Thus, Jurchen offers no evidence for any initial nasals in PMT. Quite on the contrary, it offers evidence against it,

---

[55] Georg's objection that the PTM word for 'water' has to be reconstructed as **mȫ* (based on Even *mȫ*) and is therefore unrelated to **ńamū-* [Georg 1999/2000: 151] is completely irrelevant here: even if we *do* accept the idea of PTM **ö* > Even *ö* and change the reconstruction from **mū* to **mȫ*, the Even form for 'tear' is *ińŋụmtụ̈*, and the secondary vowel quality here does not allow one to see based on this form whether the «original» vowel was **u* or **ö*. The exact same compound is found in Mongolic (PM **ni-l-mu-sun*), Korean (MK *nún-mɨ́r*) and Japanese (**nà-mì-(n)ta*), all meaning 'tear' = 'eye-water' and traceable back to the same PA compound. Of course, for Georg and AV all of this may be just «chance comparanda», but then anything can be «chance comparanda».

confirming that the PMT form is just **iā* 'eye'. Incidentally, Jur. *jali* 牙力 'meat' appears as such and not as *ŋiali* in another EDAL entry on p. 985. If the EDAL authors possess some esoteric knowledge that allow them to treat the same Chinese transcription in different ways, it would be extremely helpful if they could share it with the rest of us» [VOVIN 2005: 84].

By all means, the EDAL authors are delighted to share, except the knowledge is not nearly as «esoteric» as AV suggests it to be. First and foremost, it is obvious that AV has very little idea as to where the EDAL authors get their Jurchen readings from. In this passage, as well as elsewhere in the book, he relies exclusively on [KIYOSE 1977] and [KANE 1989]; however, both of these sources, detailed and valuable as they are, represent attempts at reconstructing Jurchen phonetics and phonology of the *Míng* (明) period as transcribed by the *Huá-yí yì-yǔ*, i. e. not earlier than the 15th century. At the same time, the Jurchen script, loosely based on Chinese characters but properly tailored to the needs of Jurchen phonology, dates back to the early *Jīn* (金) period, around the first decades of the 12th century [KIYOSE 1977: 22] — and the fact that Jin-era Jurchen and Míng-era Jurchen could have certain phonetic differences is well recognized by specialists. One of these differences, already suggested by L. LIGETI [1953] and cautiously accepted by G. KIYOSE, is as follows: «The velar nasal consonant **ŋ*, besides *ŋ* as an allophone of *n*, may have existed in the Jin period, and then merged with *g* in Ming Jurchen» [KIYOSE 1977: 59].

Next, before making up hasty assumptions about the authors' sources for Jurchen phonetics, it might have been useful to consider this little passage in the preface to EDAL, on p. 238, where it is quite explicitly stated that «for Jurchen we use the phonetic transcription proposed in [MUDRAK 1985, 1988]». While these two works may not be easily accessible to the general reader, this does not mean that they do not exist, and an objective reviewer, before pronouncing judgement on conclusions based on them, should have at least informed the reader of this fact (not to mention that it is very unlikely that AV, being an expert on extinct Tungus-Manchu languages, is unaware of their existence — they could have hardly slipped through his fingers while he was still residing in the former Soviet Union and occasionally interacting with the Moscow school).

Both of these works are dedicated to providing a coherent phonetic decipherment of the Jurchen script, based not only on later period Chinese transcriptions, but also on the internal logic of the script itself, as well as suggestions from external Tungusic data. In particular, they confirm LIGETI's old idea that a separate phoneme *ŋ* existed in Jīn-era Jurchen, since a whole series of syllabic characters can be identified as beginning with a consonant different from *n*, *y*, *g*, or any other, with external parallels strongly advocating for a velar nasal reading.

Thus, the phonetic value of the character 兌 is reconstructed by Kiyo-se as **ga*; however, the two lexical items in which it is met go back to PTM roots with **ŋ-*: **ga-la* 'hand' (504)[56] < PTM **ŋāla* (cf. Evk. *ŋāle*, etc.), **am-ga* 'mouth' (494) < PTM **amŋa* (cf. Evk. *amŋa*, etc.). On the other hand, a «regular» syllable **ga* also exists in Jurchen, transcribed as 半 and regularly corresponding to items with a velar stop in PTM, e. g. **ga-gwa-i* (399) 'to take' < PTM **ga-* (cf. Evk. *ga-*, etc.). Based on this distinction, Ligeti reasonably reconstructs the first syllable as Jīn-era **ŋa* > Míng-era **ga*.

Similar oppositions can easily be established for other syllables as well; Ligeti [1953] makes an equally good case for **ŋe* and **ŋi*, while Mud-rak [1988] also argues for **ŋu*, **ŋo*, and others. Specific cases can be debated, but overall, there is little doubt that a distinct series of syllables with *ŋ* was specifically devised for the Jurchen script.

Turning now to the word for 'eye', it is transcribed syllabographically as 光巴 (496). The second character here is reconstructed by Kiyose as **ši* [Kiyose 1977: 91] and transcribed as *ĉi* in [Mudrak 1988] and EDAL. The first character is transliterated by Kiyose as **ja* (**ya*) [Kiyose 1977: 68], based, obviously, on both the later Chinese transcription 牙師 **ja-ṣi* and comparative data from Manchu (*jasa* 'eye').

The important thing to notice, however, is that this is *not the same character* that is used to transcribe the already mentioned word for 'meat, muscle': Jurchen 炎休 *ja-li* (511). *Huá-yí yì-yǔ* era Chinese transcription, of course, does not distinguish between the two, rendering the initial syllable as 牙 *ja* in both cases, but the Jurchen script makes it clear that there *was* a difference in the pronunciation as late as the Jin epoch at least. Neither of the two characters is logographic (or the words would not be transcribed by two symbols each)[57], and there is no reason whatsoever to assume that the Jurchens needed two different characters to render the same syllable (*ja*).

Therefore, in [Mudrak 1988] the two characters are differentiated by assigning the phonetic value *ja* to 炎, and the phonetic value *ŋia* to 光. The variant with an initial velar nasal is selected for the following reasons: (a) it should be graphically distinct from both *ja* and *nia*, for which we

---

[56] Numbered references for Jurchen items are given according to [Grube 1896] (the same numeration is reproduced in [Kiyose 1977]).

[57] Kiyose [1977: 85] seems to be suggesting a logographic nature for 炎, glossing it as «**ya* (牙) 'flesh'», whereas 光 [1977: 68] is glossed as «**ya* (牙) phonogram». However, such a distinction is arbitrary: both words are transcribed by *two* characters (*ja-li*, *ja-ĉi*), which unambiguously suggests a syllabic reading for both. Perhaps Kiyose is misled by the fact that the graphic combination 炎休 is encountered twice in the Berlin National Library Text, and both times in words with similar meanings: *ja-li* 'muscle' (511) and *ja-li* 'meat' (521). But, of course, this is simply the same Jurchen word in both cases.

have their own characters; (b) it has to be compatible with both the Chinese transcription (牙 *ja*) and the external parallel in Manchu (*jasa*); (c) it fits well within the reconstructed Jin-era Manchu phonology where we postulate the presence of an initial *ŋ-*, and, therefore, can reasonably expect a syllable like *ŋia-*; (d) the reconstruction with a velar nasal is made more probable due to graphic similarity with 元 *ma* and other similar-sounding syllables (cf. the use of a similar principle in pairs like 犀 *hi* — 早 *fi* [KIYOSE 1977: 85], [MUDRAK 1988]).

These two examples, in our opinion, are enough to demonstrate that sometimes the «ignorance» accusation can be succesfully reversed. The necessary implication, though, is that it is perhaps best to simply make use of comparative, historical, and philological work done by specialists in the field, regardless of their national background and university affiliation, and criticize that work based on constructive suggestions and alternatives, rather than spur-of-the-moment considerations and a «holier than thou» attitude.

VIII. *Mistakes: crucial or insignificant*?

The final group of evidence-based arguments consists of numerous accusations of *mistakes* and *carelessness* on the part of EDAL's authors that are said to become obvious upon detailed analysis and, apparently, render a large part of the work useless (SCH, apart from the already discussed «philological» cases, is not accused of the same flaw, so only EDAL will be discussed in this section). The argument as such is not stated in a concise way in [VOVIN 2005], but is scattered all over the place as authors are accused of mixing correctly quoted forms together with «lexical ghosts» [VOVIN 2005: 77], relying on outdated dictionaries, and mislabelling attestations. Additional examples can be found in [VOVIN 2001], and the same problem is dealt with at length in a different review of EDAL that judges it from a pro-Altaicist perspective [MILLER 2003/4].

It is our honest duty to inform the reader that, indeed, EDAL is not free of mistakes, moreover, ones that happen in all the mentioned categories. Below we will list all the cases mentioned in [VOVIN 2005] where our opponent is right in pointing out factual errors, at the same time noting to what extent the admission and correction of these errors will influence the Altaic etymologies in question.

(1) Written Mongolian *gelgüri-* 'to stroll' is, as indicated, a misprint carried over from [TMS I: 178] [VOVIN 2005: 77–78]. However, since EDAL also includes the correct form *geldüri-* (as well as an array of forms in related languages), the exclusion of *gelgüri-* from the etymology will have no negative bearing on even the Mongolic comparison, let alone Altaic.

(2) MJ *situ 'damp place, dampness' is, indeed, attested slightly later than the end of the MJ period, and its severely limited attestation (as quite convincingly shown by AV), along with other problems, should probably lead to its removal from PA **si̯ằŕi* 'earth, sand, marsh' (unfortunately, since the comparison was advocated by the late S. A. Starostin, we are unable to learn if he had any additional arguments for this case[58]; AV's «complete hoax, consciously concocted by EDAL's authors to advance the Altaic hypothesis», will just have to remain on AV's own conscience). Overall, this will not affect the PA etymology too much, since its reflexes are abundantly attested in all the other daughter families [EDAL: 1269].

(3) Khaladj *šašqa-* 'to slander' < PT **siāĺ-* < PA **si̯óĺe* 'mock, slander' is not *really* a 'ghost', because the form does exist — except that it is not Khaladj, but Shor (the form accidentally got into the wrong database field during the preparation of the dictionary). In Shor, however, it is attested with the meaning 'to slander' ([SRRSD: 67]; cf. also «matyrisch *šaška* 'verleumden'» in [Räsänen 1969]), which cannot, therefore, be ignored when comparing the Turkic word with its potential Altaic cognates.

(4) Written Mongolian *bilaɣu* 'carp' *is* a ghost, to be more precise, an accidental leftover that should have remained in the reconstruction field (PM *bilaɣu) but not present in the actual 'Written Mongolian' entry[59].

(5) Middle Korean *sằr-* 'burn, make a fire' does repeat a dictionary mistake and should be *sắr-* with high pitch instead of low. This will influence the tonal reconstruction in PA, but not lead to the exclusion of the Korean form from the etymology.

Overall, then, we may count five factual errors indicated by AV over the course of the entire review, and our sincere gratitude goes to him for pointing them out. To this, we might add about a dozen cases where the

---

[58] One possible consideration is that, in addition to the poorly attested *situ / shitsu*, in Modern Japanese there also exists a verbal stem *shitoru* 'to be damp', which is curious, since verbal stems like these are not usually formed from words of Chinese origin.

[59] Actually, this is one of the rare cases where we tend to accept AV's objections about the daughter-branch etymology in general; namely, that Khalkha *buluu cagaan* 'carp' and Buriat *bulūsxai* 'taimen, ide, roach' (the reason why Khalkha *buluucgai zagas* 'ide' was ignored is that it is quite possibly a Buriat loanword in Khalkha) constitute too flimsy evidence for PM *bilaɣu 'a k. of fish' < PA *pắli (since the internal etymologization from *buluu* 'thick end, lump; plump' works relatively well). On the other hand, the same entry in EDAL offers an *alternative* set of Mongolic cognates: Buriat *baĺuɣan* < *baliugan (Okinsk dialect), *baĺus* (Darkhat dialect) 'a k. of carp (Phoxinus laevis, гольян) or small grayling (хариус)' [Anikin 2000: 115]. The two forms presuppose a PM *baliɣu, which is further confirmed by Khalkha *balius* 'grayling' [BAMRS I: 222], and all of them correspond well to both PT *bālɨk 'fish' and Manchu *falu* 'a k. of bream'.

«mistake» means attributing a slightly incorrect or incomplete English meaning to a select form or protoform (e. g. we agree with AV that Japanese *sukum-* has to be glossed as 'to cower' rather than 'to become numb' — but this is actually *better* for the etymology, since the corresponding words in other daughter branches usually have meanings like 'to shake, tremble, quiver (from cold or fear)'); not a single one of these cases we deem sufficient to abandon the suggested etymologies. We also acknowledge and apologize for occasional misprints and bibliographic inconsistencies indicated by AV, like the incorrect inclusion of [POPPE 1930] into the bibliography as «Poppe 1933–1937»[60].

Furthermore, we are not deceiving ourselves or the reader that these are the *only* mistakes in EDAL. A few more instances can be found in [MILLER 2003/4][61], and others will certainly become evident in the future, for which the authors assume full responsibility. It is hardly necessary to remark that a large work of EDAL's scope, for various technical reasons, is practically *bound* to contain a certain amount of errors, but we sincerely hope subsequent revisions and re-editions of the work will diminish the overall number.

Two things, however, need to be stressed:

(1) *Factual errors* should never be confused with alternative explanations. Unfortunately, this kind of confusion seems to be heavily endorsed in AV's review, rich on categorical statements like «nonsense», «cannot be a cognate», «contradicts phonological reality», etc., etc. Regardless of whether all these statements are justified, most of the time they deal with AV's currently proposed historical scenarios (some supported by the «mainstream», others hardly so), and little else.

Oftentimes the confusion can only be seen as intentional. Whenever the authors of EDAL suggest a folk-etymological reinterpretation of an Altaic root on Mongolic, Japanese, Tungusic, etc., territory, they are immediately branded as ignorant of some grammatical feature or cultural reality: «we do not need any external parallels to rewrite Mongolian grammar», «I can understand that EDAL's authors do not know that PMT suffix *-pūn*

---

[60] Although this is hardly any more criminal than, for instance, misspelling the surname of the notorious Orientalist S. A. KOZIN as *Kozincev* [VOVIN 2005: 76].

[61] A particularly unfortunate mistake indicated in that review is connected with the reconstruction of PA *balu* 'sable' > MJ *furuki* id., upon which it is said: «R. A. Miller's hypothesis that *furuki* was borrowed from Mongolian seems rather far-fetched» [EDAL: 326–327]. The hypothesis, was of course, that the origin of the form can be found in the Tungus-Manchu word for 'fox', not the Mongolic word for 'sable' [MILLER 1996: 199–200]. We still find MILLER's hypothesis somewhat far-fetched, but certainly more reasonable than the mistaken Mongolic borrowing, and, of course, apologize for the misquotation.

makes deverbal nouns», «all this information can be safely secured from the most elementary books on Japanese archaeology», etc. etc. Obviously, this should be interpreted as indications of factual mistakes committed by the authors due to their unfamiliarity with the material. But in most cases, the authors *do* indicate their familiarity — although briefly — before suggesting that a different interpretation *may* be in order.

It is precisely this failure to distinguish between *mistakes* and *alternate explanations* that seriously undermines much of the work conducted by anti-Altaicists (and «splitters» in general, as has already been indicated by V. A. Dybo several decades ago: «In order to be considered scientifically correct, criticism has to at least make a rigorous distinction between perfectly established facts, on one side, and hypotheses (i. e. opinions and guesses about these facts, supported by more or less reasonable argumentation on the part of the researcher), on the other» [V. Dybo 1984: 7; translation is ours — *A. D., G. S.*]).

E. g., when S. Georg, in an example already discussed above, writes «Mong. **kelen* ≠ Tung. **xiń-ŋi*» [Georg 1999/2000: 163], this is presented as a fact («es scheint deutlich zu werden»), when in reality this line should read: «there is a *probability* that the Mongolic and Tungusic forms for 'tongue' are unrelated because the Tung. form *can* be reconstructed without an inlaut lateral resonant». But it also can be reconstructed *with* such a resonant, according to a different, but also quite probable (according to our opinion — *much more* probable) historical scenario — and if such a reconstruction fits in with the proposed external relationship, this should, quite necessarily, count as a piece of evidence — one that can be tested, but cannot be discarded.

(2) When it comes to *real* factual mistakes similar to the ones listed above, it turns out that in most cases they do not threaten the etymology at all. Sometimes — in a very limited number of cases — a *part* of the etymology may be removed based on corrected forms or attestations. There has not yet been a single case spotted by EDAL authors where a detected factual mistake on their part would have to lead to a complete abandoning of the supposed etymology.

At the same time, neither can we say that our esteemed opponent is completely guilt-free in that respect. We have already spotted inconsistencies in his analysis of the phonetic correspondences suggested by authors of EDAL and SCH (addition of «extra» correspondences and even phonemes), incorrect transfer of meanings from primary sources («knee-cap» instead of «knee-cover», etc.), false accusations of invented meanings («put on a spit» vs. «roast on a spit»), exaggerated appeals to «universal acceptance», and ignorance of progress achieved in various subbranches of Altaistics by the Moscow school. In addition, we are frequently left bewildered by radical shifts of position from AV's earlier publications that are not accompanied by detailed explanations. Finally, on a sidenote, the text

of [VOVIN 2005] was not checked carefully by an English-language editor, and some of its orthographic peculiarities might lead to positively odd associations (e. g. «you can *sas-* [pierce] in Japanese with an owl» [VOVIN 2005: 106] — the Japanese people certainly have their share of cultural features that look odd from a Western point of view, but something makes us seriously doubt that piercing things with *owls* is one of these).

As a last example of how, based on a mix of subjective opinions, inadequate data handling, and inconsistent application of the same principles, it is possible to turn a constructive piece of work into a chaotic mess, we would like to present a detailed analysis of the following lengthy footnote in [VOVIN 2005: 78], dealing with Tungusic and Korean reflexations of PA **gĕle* 'to come, to go' (preliminary apologies to AV for briefly recurring to «E-mail style»).

AV: «Established regular correspondences within uncontroversial families mean nothing when the need arises for an 'Altaic' etymology. Thus, aforementioned PMT **gel-* is reconstructed on the basis of Ewk. *gel-*, cited above, and Orok *gilin-* 'to start moving [трогаться с места]' (left unglossed in EDAL's English gloss, p. 538) with reference to [TMS 1975: 150]. However, Ewk. *-e-* does not correspond to Orok *-i-*».

The correspondence Evk. *-e-* : Orok *-i-* is not mentioned in the table on p. 162, but it most certainly exists and is well-documented on at least a few cases both in TMS and EDAL, cf. PTM **lebē-* 'to be stuck (in a swamp)' > Evk. *lewē-*, Orok *lemu-*, *liwe-* [TMS I: 514]; PTM **perke-* 'to bind, tie round' > Evk. *herke-*, Orok *pitu-* 'man's girdle' [TMS II: 369–370] (Orok *-t-* is a regular reflexation of the PTM cluster **-rk-*). It is one of the many cases of nontrivial vocalic development in TM bisyllabic stems, and, even though the exact conditions are hard to determine, there is no reason whatsoever to think that Evk. *lewē-* and Orok *liwe-* have nothing to do with each other. So, is this a mistake on the part of EDAL's authors or just a case of AV's insufficient knowledge of comparative Tungusic material?

AV: «Second, Orok *gilin-* does not mean 'to start moving' in general, but to 'take off' (about a string/caravan of reindeers attached to a sledge). This can be easily learned even from existing Orok dictionaries».

Actually, this can be easily learned even from the corresponding entry in TMS, and the authors are well aware of this fact. While it may have been more reasonable to gloss the Orok meaning in a more exhaustive manner, the important thing here is that such a glossing would only *help* the etymology: Evk. *gel-* 'to get hardly (with difficulty) on one's way' and Orok *gilin-* 'to take off (of a caravan of reindeers)' both share the semantics of 'moving with effort', which makes it even more obvious that the two forms belong together.

AV: «Third, no internal Orok evidence is provided for the suffix *-in-* in EDAL. However, in spite of the fact that Ikegami does not list this suffix in

his article on Orok verb stem-formative suffixes, there is a possibility that it is an obsolete suffix...».

There most certainly is, although it has absolutely nothing to do with AV's upcoming argumentation for it. One has only to consider such Orok verbal stems as *pelin-* 'to hurry' [TMS II: 364] (contrary to AV, most definitely related to **peli-* 'to walk'), or *ūnžin-* 'to shamanize' (cf. Nanai *un-*, *un-ži-*, Oroch *uńi-* id.) [TMS II: 277]. The suffix *-(i)n-* was most probably used to form nouns from verbs, which could then be themselves used as verbal stems (a fairly common process in the world's languages; cf. the fact that the stem *pelin-/helin-* is usually nominal — 'hurry *(n.)*' — in the other TM languages). In any case, EDAL is not really a handbook on internal reconstruction on TM morphology, just as Walde & Pokorny's dictionary of Indo-European is not a handbook on fossilized Greek suffixes.

AV: «...but the news are bad for the 'go/come' etymology: there is Orok *gilbə* 'reindeer that goes attached to the sledge where people are riding' [Ikegami 1997: 70]. Thus, Orok *gilin-* obviously has nothing to do with 'coming' or 'going', as it is a derivation from a name of a very specific type of domesticated reindeer».

On the contrary, the news are very good for the 'go/come' etymology. Forgetting his own principles, AV demonstrates a clean-cut case of 'doing wordlist linguistics', quoting a word from a dictionary without looking at it in a broader context. In addition, while there may be a vague problem about segmenting *gilin-* as *gili-n-* or *gil-in-*, there is a much more concrete problem about segmenting *gilbə* 'reindeer' into *gil-bə*, since there is no evidence whatsoever for a nominal suffix *-ba* (*-bə*) in Orok. Perhaps AV implies that *gilbə* is an unsegmentable root, but in that case, how can we derive *gil-in-* from *gilbə*? Finally, if AV had bothered to check the comparative evidence, he would immediately discover that Orok *gilbə* is nothing but the nominal correlate of the well-attested PTM verbal stem **gilbe(n)-* 'to tie together (a caravan of reindeer)' [TMS I: 150], cf. Evk. *gilbe-* id., Even *gilbъn-* id., Orok *gilben-* id. The nominal stem *gilbə* in Orok actually has a *plural* meaning ('*цуг оленей* = a caravan of deer'), but perhaps AV was misled by the uncomfortable fact that in English, the plural of (*rein*)*deer* is also (*rein*)*deer*. Thus, Orok *gilbə* 'a caravan of reindeer' and *gilin-* 'to start moving (of a caravan of reindeer)' have nothing to do with each other, and this is not even mentioning that AV's derivation of *gilin-* from *gilbə* runs into unresolvable semantic problems: to the best of our knowledge, it is very atypical for a word with the meaning 'animal *X*' to serve as the basis for a derivative word that means '(to start) moving (of animal *X*)'.

AV: «Finally, MK *ká-* 'to go' belongs to the accentual class 3 with highly irregular accentuation [Ramsey 1991: 232], the only reason it is 'reconstructed' in EDAL as PK **ká-* with high pitch seems to be that in its

dictionary form *ká-tá* it indeed has high pitch. But reconstructing high pitch even for MK, let alone PK, is to beg the issue».

The reference to Ramsey is correct; however, if AV had actually followed the reference a little further, he would probably see that the main reason it is 'reconstructed' in EDAL as PK *ká-* with high pitch is that it is also 'reconstructed' as PK *ká-* with high pitch by none other than Ramsey himself. Cf.: «In Middle Korean, the accentual behaviour of Class 3 and Class 4 stems was extremely complex... although it is not clear how these irregularities in pitch arose historically, *the stems appear to have been uniformly high pitched at the Proto-Korean stage of the language* [italics are ours — *A. D., G. S.*]. Evidence for this surmise comes from compounding phenomena» [Ramsey 1991: 234]. This is followed by detailed argumentation in favour of the high-pitch reconstruction. Now it is possible that for AV, this argumentation may not seem convincing, but it is nevertheless bizarre that Ramsey's presentation is cut in two and only the first statement is quoted. Isn't this exactly what AV himself calls «sweeping data under the rug»?

AV: «Finally, there is some oblique evidence that it should be reconstructed with a final consonant, but not the *-l*, as EDAL suggests without providing any evidence, but with *-n*».

First, there is a misunderstanding here, rather typical for AV. EDAL does *not* suggest that the Middle Korean, or even Proto-Korean, form should be reconstructed with *-l*, *-n*, or any other consonant. The reconstruction in the Korean part of the entry is presented as *ká-* with a vocalic auslaut. It is for *Altaic* that a final *-l(e)* is reconstructed, and a discussion on potential reconstruction of resonant-ending stems in PA, with provided evidence, can be found in several places in EDAL (most notably in the preface, on p. 23).

It is AV, on the contrary, who does not provide any evidence for the reconstruction of PK *kán-*, making us suspect that the aforementioned evidence is so «oblique» as to be dangerously close to «non-existent» (as is frequently the case when his remarks are not accompanied by references, cf. above the discussion of PJ «*esi*» stone and the «borrowing» of Korean *ak-swu*, *ek-swu* from Chinese). Perhaps he thinks that *all* CV-type verbal stems in Korean go back to *CVn-*, based on their morphological behaviour, but that is really a non-argument.

AV: «Thus we can safely put at least the Korean and Tungusic parts of the etymology to rest».

We can interpret this as a confession that the *other* parts are fully satisfactory — i. e., that PT *gel-* 'to come', PM *gel-* 'to walk slowly' (with the exclusion of the non-existing form *gelgüri-*), and PJ *kə̀-* 'to come' constitute a perfectly valid comparison. Considering that the anti-Tungusic and anti-Korean arguments are wiredrawn and do not so much under-

mine the etymology as confirm it, the entire discussion only strengthens the Altaic theory and the methodology that underlies it.

Nevertheless, we are not willing to hold any of these details against our opponent as evidence of his «incompetence» and, consequently, the «supremacy» of the Altaic and Sino-Caucasian hypotheses. They simply confirm that *errare humanum est*, and we think that, in order to produce an objective evaluation of one's views on a particular problem, we must learn to look above the populistic principle of «if a work has several mistakes in it, it is worth nothing».

IX. *Religious belief or scientific research*?

After (sometimes before) presenting his argumentation on the many methodological flaws of the Moscow school, AV allows himself to ponder upon the possible reasons that lie behind this surprisingly careless attitude towards language material on the part of people who are supposed to be professional scholars in the field. His conclusions are strikingly similar in both [Vovin 2002] and [Vovin 2005]:

«The only tangible explanation… is that the Altaic hypothesis at least in its Moscow version became a set of beliefs highly reminiscent of a religion. However, religion and science cannot coexist, because the first is based on faith, while the second seeks the explanation of the facts that are in need of explanation. EDAL does not explain these facts, it simply creates the 'evidence' for the pre-existing belief that is, of course, non-evidence» [Vovin 2005: 123].

«The critical scholarship seems to be completely replaced by a religious belief in macro-families. No wonder that this attitude led to the situation when 'sino-Caucasian' is supported by virtually no one outside of Moscow Nostratic school and several Proto-Worlders, who compare anything with anything with a complete disregard to regularity of correspondences» [Vovin 2002: 167].

AV's *penchant* for branding his opponents as «religious believers» is well-known, although several years ago it used to be more frequently applied to *anti-* rather than pro-Altaicists (e. g., on Karl Krippes' condescending review of [Starostin 1991]: «Krippes, like the majority of his fellow anti-Altaicists, gives his reader religious sermons, not supported by any evidence or documentation» [Vovin 1995]). In the light of this it is hard to understand whether, over the last decade, AV has really made the major progression from a «religious belief» in Altaic to «science» or has merely switched from one church to another. However, loud words aside, let us evaluate AV's conclusion independent of its pragmatic context.

First and foremost, not even the staunchest anti-Altaist or anti-Sino-Caucasianist will ever deny that there *are* intuitively obvious similarities

between the compared languages that *may* be interpreted as evidence in favour of genetic relationship. Most of the discussions are usually centered not around that particular question, but rather around two ensuing ones: (a) are these similarities *sufficient* to consider such genetic relationship demonstrated beyond a resonable doubt; and (b) are there any *alternate* solutions other than genetic relationship to explain these similarities, and if there are, which ones should be preferred?

The scientific — or, according to AV, religious — methodology of the Moscow school, having gradually been worked out over the past half century, would probably answer these questions in the following way:

(a) *No*, not sufficient. The similarities, as prescribed by the comparative method, must be systematic and recurrent (which does not presume «one hundred percent mechanistically recurrent»), and, moreover, they must necessarily permeate all the levels and spheres of the compared languages — phonology, grammar, and vocabulary.

This is a very important point: the languages cannot be *partially* related. Thus, if language *A* shares a lot of nature terms with language *B*, but at the same time has no common body parts, or vice versa[62], *and* it is impossible to demonstrate that, for instance, language *A* has borrowed all of its words for body parts from language *C*, such a situation is highly suspicious. This is why significant evidence for genetic relationship *must* be presented in the form of a bulky etymological dictionary. A list of one or two hundred cognates is substantial, but if these cognates represent but a few percent of the languages' respective vocabularies, the imminent question is «what about the *rest* of them?», and a realistic theory of genetic relationship must always be prepared to answer that question.

It is exactly that quest for *realism* and *completeness* that lies beyond the «uncontrolled urge» of the authors to «create» numerous Altaic and Sino-Caucasian parallels. And this is why we find it unacceptable to agree with slogans like «less is more» or «quality over quantity», so popular with opponents of long-range comparison. While there definitely has to be a certain «core» of «quality etymologies» (see above), neither the Altaic nor the Sino-Caucasian nor, in fact, any other hypothesis of relationship, long- or short-range, can be considered satisfactory if not backed up by a *large* amount of cognate forms.

(b) Yes, there *may* be alternate explanations to genetic relationship, and the deeper we find the distances between compared languages, the

---

[62] This does not mean that all the «common body parts» between languages *A*, *B* and *C* must necessarily have identical meanings (see the discussion above on body parts in Formosan languages, or in English and Irish); but a significant part of them in each language has to be phonetically and semantically compatible with equally significant parts in the other two.

more such alternatives will be coming to mind. One scientific principle that we share, however, is that of Ockham's razor: «entities should not be multiplied beyond necessity». In practical terms, this would mean that if we take five words across three languages displaying the same recurrent phonetic correspondences, and then someone comes along and says that

— word №1 *can* be alternatively reconstructed in language *A* with a different initial consonant (but can also be reconstructed as presented);

— word №2 should be rejected due to sound symbolism (even though it displays a fully regular set of correspondences);

— word №3 has a stretched semantics in language *B* — not entirely incompatible with languages *A* and *C*, but hard to believe;

— word №4 *may* be a borrowing from *A* to *B* and from *B* to *C* (but we do not have rigorous proof for that);

— word №5, left alone, may just be an instance of chance resemblance (but not necessarily so);

— we have a perfectly valid right to claim that explaining all five cases based on *one* solution — genetic relationship — is preferable from a strictly scientific point of view[63].

The only possible reason why this solution might *not* be preferable would be an intentional *aversion* to postulating genetic relationship of *any* kind[64]. For instance, if it were a well-known fact that languages are *very rarely* genetically related to each other, and that linguistic diversity on Earth is essentially due to other factors (such as convergence of independently arising idioms), such an aversion would be understandable. But we all

---

[63] A good example of this procedure can be found in [Georg 1999/2000], in which the author rejects the match between PT **el* 'hand' and PTM **ŋāla* id., stating that the correspondence «PT **∅-* — PTM **ŋ-*», proposed in [Starostin 1991], is too dubious to make the etymology look convincing. He then briefly discusses some of the supporting evidence, such as PT **it/it* 'dog' — PTM **ŋinda* id.; PT **u-* 'sleep' — PTM **ŋuja-* id., finding various *potential* flaws with each of the examples (unclear vocalic correspondences, morphological segmentation, etc.); yet absolutely nothing is made of the fact that what we are dealing with here is a *recurrent* pattern established on *perfect semantic* matches within a *non-arbitrarily* chosen 100-wordlist of basic items. One could try to demonstrate statistically that this cannot be due to chance, but the simplest thing to do would be to try and find any other non-Altaic language with which PTM **ŋ-* could give at least three similar basic lexicon parallels of the same quality — which is, of course, not done by Georg because it cannot be done.

[64] We might also reject it, of course, in case a competing hypothesis presents even *better* evidence; cf. the Sino-Austronesian vs. Sino-Caucasian case discussed above, where the Austronesian parallels for Chinese are better explained as traces of contacts (or, perhaps, an even higher level relationship than Sino-Caucasian).

know that in reality, it is the reverse situation that is true: an overwhelming majority of the world's languages are known to be related to at least *some* other languages, and there is little reason to think that the situation was significantly different six, ten, or fifteen thousand years ago.

And this is what really constitutes the most vulnerable moment in AV's argumentation (or, rather, argumentation which AV has borrowed from his former opponents, without giving them too much credit). He insists that, in order to demonstrate the validity of Altaic or Sino-Caucasian, we should remove *all* the evidence that can be explained alternatively. In [VOVIN 2005], based on eight etymologies, he proceeds to demonstrate how, in his opinion, this should be done — and in the process ends up removing all the evidence whatsoever, proudly informing the reader that all of it «is just *fata morgana* based on an *ad hoc* morphological analysis, invented semantics, inadequate knowledge of the languages involved and their philology, irregular phonetic correspondences found inside and outside a given branch, and severely limited attestations that have a different explanation» (p. 117).

AV probably means that what he is doing here is *science*, as opposed to his opponents' religious activities. Yet we have serious reasons to doubt it. As we have amply demonstrated above, the majority of AV's rejections and accusations in [VOVIN 2005] lack absolute proof: at best, a few of them represent alternate explanations with widely varying proportions of «convincing force»; at worst (when no explanations are presented, only criticism), as forced arguments that make no influence whatsoever on the original probability of the hypothesis.

In addition, only two etymologies on the list have a connection with the core of the basic lexicon («sharp stick, tooth» and «stone»), and it is hardly a coincidence that it is precisely in these two cases that not a single «argument» put forward by AV in favour of their rejection can be considered convincing[65],

---

[65] One part of the PA etymology *sĪĺa* 'sharp stick/tooth' that has not yet been discussed is Turkic *sīĺ/*sĭĺ; the manner of its rejection in [VOVIN 2005: 106] is, in our opinion, a prime example of how a «hypercritical» approach can border on the absurd. Cf.: «it is remarkable that the meaning 'tooth' is represented exclusively by Chuvash *šăl*» — nothing remarkable about that, since Chuvash forms one of the two primary branches of Turkic. «Besides, even Common Turkic *šiš/šiš* means 'skewer' rather than a 'sharp stick'» — no, it does not, because the actual meanings range from 'skewer' to 'fork' to 'spike' [CLAUSON 1972: 856–7]. «A solution presented by DOERFER that Chuvash *šăl* 'tooth' is related to Common Turkic *tiš* 'tooth' and is actually the development from the latter through palatalization and subsequent assimilation *tĭš > *čiš > *šiš > šăl* is much more elegant and simple solution, that does not require unusual semantic shift within Turkic»: it is not a more elegant and simple solution, because it is based on at least three assumptions — (a) that we know for certain that the PT consonant was really *š (considering that in the later stages of his research, DOERFER himself preferred

being based on incorrect semantic reconstruction (PTM **sila-* 'roast' instead of 'roast on a spit'); inability to recognize quite obvious patterns of ancient fossilized morphology (PM **si-dü(n)* 'tooth') — and, at the same time, postulating quite obscure patterns of *recent* fossilized morphology (PJ **ísì* 'stone' < **isu-i*?); overlooking regular phonetic correspondences, both internal (PT **sīĺ* 'tooth, sharp stick') and external (PTM **ʒola* 'stone'); excessive semantic demands that contradict semantic typology (PJ **sàs-* 'to prick, stab'; PK **sár* 'arrow') — and, at the same time, dangerously *lax* semantic demands that *also* contradict semantic typology (PK **tōrh* 'stone' < 'to roll'?); unrealistic and contradictory borrowing scenarios (PM **čilaγu* 'stone' < Bulgar **čUl*?); and reconstructions based on non-existing evidence (PJ **esi* 'stone' instead of **d(isi)*).

AV's seemingly thorough conviction that many of these explanations are inherently better than the ones provided by EDAL (as seen in his incessant usage of words and expressions like «transparent», «no chance to compare it with 'Altaic'», «falls apart», «chance comparanda», «does not withstand the scrutiny», etc. etc.) does not help matters much, because this can only mean that he, too, now views the idea of genetic relationship between Altaic languages as requiring much more rigorous «proof» than *any* other explanation — without explaining *why*.

One of AV's pessimistic conclusions states that «EDAL will surely become another part of Nostratic Holy Scriptures for true believers» [Vovin 2005: 123][66]. In the light of this, AV will probably be surprised to hear the

---

the «sigmatism» solution for Turkic historical phonology, this very fact would have rendered his earlier etymology unacceptable); (b) that we can suggest irregular assimilation for «Proto-Chuvash»; (c) that the «lambdacism» took place *after* the assimilation. There is also nothing unusual whatsoever about the semantic connection between 'sharp stick' and 'tooth' (cf., among innumerable examples, Greek γόμφος 'peg' : Old Indian *jámbha-* 'tooth'), so rejecting the phonetically ideal match between Chuvash *šăl* and Common Turkic *šiš* 'sharp stick' can only be justified by the usual «counter-Altaic drive».

[66] And this idea, apparently, builds up on AV's general atittude towards Nostratic as expressed in [Vovin 2002: 164]: «I am a mild supporter of the Nostratic theory; some parts of it are probably destined to survive, although probably not the whole theory as it is currently enshrined in Moscow». The authors of this article, all of them representatives of the Moscow school, would be extremely curious to learn AV's thoughts on how the Nostratic theory is «enshrined», considering that both of them have somewhat differing opinions on various sides of Nostratic, and that those sides of it over which they *are* in agreement are frequently very different from the original postulates by V. M. Illich-Svitych, as well. Unfortunately, the citation is now five years old, and since it is now impossible for AV to continue to be a «mild supporter of the Nostratic theory» (because the same argumentation that AV uses in [Vovin 2005] to dismantle «Altaic» can just as easily be applied to Nostratic — see, e. g., [Campbell 1998]), our curiosity will probably have to remain unsatisfied.

following: there is not a single etymology in EDAL, not even among the crucially important basic lexicon, that its authors would be willing to «defend to the death» regardless of the circumstances. On the contrary, we are almost sure that many of them are bound to undergo significant revision (some already have). However, rejection or acceptance of etymologies certainly cannot depend on simply being *told* by AV, or anyone else, to reject it because it does not «withhold scrutiny» (whatever that may mean). Instead, we try, where possible, to follow a slightly more objective criterion, namely:

— an etymology formerly suggested for a certain word or intermediate reconstruction may, and will, be rejected if a *better* one is found. To make this last point clearer, it can be said that «etymology X is *better* than etymology Y» if (a) X requires making *fewer* unsystematic hypothetical assumptions than Y; (b) the hypothetical assumptions underlying X are typologically more probable than the ones underlying Y.

Thus, comparing two different etymologies for PM **čilaɣu* 'stone' — (A) genetic relationship with PT **diāĺ* id. or (B) borrowing from Bulgar *čol* id. — we will discover the following:

Etymology (A), in addition to *systematic* evidence such as the regularity of consonantal *and* vocalic correspondences, makes *one* unsystematic hypothetical assumption: namely, that after the disintegration of Altaic (sometime during a period of about six thousand years) the word has received an additional suffix *-ɣu* on Mongolic territory.

Etymology (B), on the other hand, rests on completely unsystematic evidence. No stable «Bulgar layer» in Mongolian has been identified, let alone regular correspondences between loanwords (which makes the vocalic issue particularly dubious), and the assumption that Mongolian has added an unproductive suffix *-ɣu* over a period of several hundred years is notably weaker than the one proposed above. Thus, unless much stronger evidence is presented for the «loan» etymology, we will not be rejecting (A) anytime soon.

On the other hand, application of the same principle does lead us to reject, for instance, the inclusion of Mongolian *bejle* 'prince of the 3rd rank' as the appropriate reflexation of PA **bóju* 'esteem' ([EDAL 369–370]; the original comparison, in the form **biōji*, was suggested in [STAROSTIN 1997: 329]). AV identifies the Mongolian word as a borrowing from Manchu *beile* id. [VOVIN 2001b: 111–112], and this alternate etymology is *better* than the formerly suggested one because it requires fewer assumptions (e. g. in the spheres of morphological segmentation and semantics).

What, then, if a *better* etymology cannot be found, but the existing one still displays problems? Above we have shown that many of these problems as presented by AV are, in fact, fictitious; but issues of irregular correspondences and hypothetical morphological segmentation still remain, and it would

be futile to assert that EDAL is impeccable in those respects. Our position here is simple: as long as there is an established «core» part of the evidence to form the foundation of the hypothesis, «problematic» etymologies should be admitted inside the corpus, provided (a) they are not *too* problematic, i. e. follow the basic consonantal correspondences and are semantically compatible; (b) some reasonable assumption can be made to explain away the «problems»; (c) we are ready to let go of the comparison if a better solution can be found.

And, before AV and the other anti-Altaicists start using this as an excuse to issue the usual «Lax standards!» battle cry, we would just like to politely remark that these are the *exact same* principles as applied to *all* traditional fields of historical-comparative linguistics, beginning with Indo-European. The same methodology is applied to Altaic, Nostratic, Sino-Caucasian, and whatever other families, macro- or micro-, «traditionally accepted» or «controversial», there are. There is the «core» material, and then there is everything else, some of which fits in perfectly, while some is *very likely* to fit in, but depends on an unprovable, albeit data-motivated hypothesis in order to do so[67]. For Altaic and Sino-Caucasian, it is an accomplished task. For Sino-Austronesian, hardly so. For Germanic-Omotic, completely impossible.

In brief, these are the «religious beliefs» that constitute the basis of the comparative research as carried out by the Moscow school. Whether this has anything to do with science depends, of course, on one's understanding of science. AV's current understanding of it, from what we can tell, follows the principle that *a hypothesis cannot be accepted until all the alternate ones have been eliminated*. So, suppose we have shown that there are more arguments in favour of a genetic relationship between Chuvash *čul* and Mongolian *čilaɣu* than there are in favour of their relationship through 'horizontal' transmission — but for our esteemed opponents, this is *still* not enough, since we somehow have to demonstrate that there are *no* arguments in favour of the second hypothesis. But even if we manage to do that (perhaps by writing a lengthy monograph explaining all the implausibilities of the Bulgar-Mongolian contact scenario), there will still be no way to prevent our opponents from saying «well, perhaps this is not really a borrowing, but then it's just a chance resemblance».

Yet this is not the way science really works, since in true science, *all hypotheses and theories are in principle subject to disproof*, meaning that it is

[67] The degree of this motivation varies widely depending on particular cases, but it is very hard to estimate the various probabilities on a strictly formal basis (e. g., if it is suggested that a form found in one language out of ten is archaic, the probability that this hypothesis is correct is certainly not 0.1, but depends on the classification, the established correspondences, the phonetic nature of the phoneme(s) under discussion, etc.). However, what really matters most of all is whether such a probability *exists* in the first place.

never really possible to eliminate *all* the alternatives; it is merely possible to show that some alternatives are more probable than others. Appeals to «rigorousness», «scepticism», and «critical scholarship» are an essential part of the scientific method, but all of them can only be taken so far before one turns out to be left with no scientific theories at all other than the classic *scio me nihil scire*. The *truly* important thing in empirical science is to know where exactly one should draw the line.

Thus, the relationship between Germanic and Slavic is more probable than one between Germanic and Omotic, but it is certainly not «rigorously proved», as there is always a slight — minimal, in fact, but nevertheless *existing* — probability that all the resemblances between Germanic and Slavic are due to chance and/or convergence. The main reason why scientists accept German-Slavic relationship is simply because, at the present time, it is *the most economic and the least contradictory way* to explain a large number of systematic (and an almost equally large number of *un*systematic) similarities between these two language families.

Likewise, there is much more potential evidence for Sino-Caucasian than there is for Sino-Austronesian, and much more potential evidence for Altaic than there is for, e. g., Japanese-Austronesian, Turkic-Sumerian or any other theory of relationship dealing with one or more Altaic branches. In addition, «Proto-Altaic» is a reasonably economic and realistic way to explain multiple similarities between Turkic, Mongolic, Tungusic, Korean, and Japanese languages without contradicting any of our *knowledge* (as opposed to *implications*) about these languages and the peoples that speak or spoke them. The same goes for «Proto-Sino-Caucasian».

Of course, it is not exactly easy for the general reader to see that evidence within AV's review of EDAL, considering how carefully AV avoids discussing «core» etymologies and how intricately «problems» that are not really problems, but questionable alternatives, are mixed together with pointing out a small number of real errors. With Sino-Caucasian, it's actually worse, because [VOVIN 2002], instead of a systematic evaluation of the material, focuses almost entirely on general questions of theory or methodology (all of which have been touched upon in the present article); what the reader really gets is a series of subjective thoughts on why STAROSTIN's Sino-Caucasian is *wrong*, without being able to understand what it actually *looks like* (and without being shown even one relevant factual mistake in STAROSTIN's materials).

It does not help matters much that AV's writing style over the past few years has turned from moderately sarcastic to downright offensive. At times it seems as if the major goal of [VOVIN 2005] is not so much to present a critical overview of the Altaic dictionary as it is to convince the reader that its authors are little more than frauds. Time and time again, over the entire fifty-

page length of the «review», the authors of EDAL are accused of ignorance («the authors do not know the Japanese language and are doing wordlist linguistics», «repeatedly manifested ignorance of the most basic facts about the Korean language history»), intentional carelessness («have no respect for and no need of cultural history of languages they compare»), and cheating («have attempted a willful misrepresentation of data on more than one occasion»).

Passages like the following one are quite typical: «Starostin managed to report an important discovery... namely, that OJ is the Latin of Japonic since all modern varieties of Japonic including Ryukyuan can be derived from OJ. If this were true, Starostin would certainly deserve the Order of the Rising Sun from the Japanese government for his outstanding contribution to Japanese historical linguistics. Unfortunately, it is not, since the Ryukyuan languages preserved certain features not found in OJ» [Vovin 2005: 91]. Even if this assessment of Starostin's short intro on the Japanese section of EDAL *were* true, one could still question the admissibility of this writing style in an academic publication — especially when coming from AV, who was still well away from his illustrious academic career when the late Starostin was already doing comparative research on Japanese, *including* Ryukyuan [Starostin 1972, 1975]. However, even the assessment itself happens to be contrived: the reference to derivation in the original text only concerned the *phonology* of the languages, and besides, it was stated quite clearly that «however, some phonetic features of the Ryukyu dialects... may be actually archaic and preserve the situation preceding that of Old Japanese» [EDAL: 166]. The odd thing is that AV actually *quotes* this in his own review, but then apparently forgets about it so that Starostin, once again, is made to look like a pretentious ignoramus.

In fact, one can almost sense the temperature go up along the way. Thus, on p. 73 we read: «The existence of the 'North Caucasian Etymological Dictionary' has so far failed to persuade specialists and general linguists alike in the existence and validity of 'North Caucasian'[68]». On pp. 122–123, however, we already get a much stronger statement: «EDAL... is, therefore, another etymological dictionary of a non-existing family, which

---

[68] This — as well as a similar passage in [Vovin 2002] — is accompanied with a lone reference to [Schulze 1997], a mildly critical review of [Nikolayev & Starostin 1994] whose author does indeed remain sceptical about the existence of North Caucasian and has multiple objections over specific issues, but nevertheless recognizes the dictionary for what it objectively is: an important step forward in understanding the prehistory of and relations between Nakh-Daghestanian and Abkhaz-Adyghe languages. At the very least, it is indeed a *review*, as its primary goal is to honestly inform the reader about the details of Nikolayev & Starostin's reconstruction, rather than to make him/her believe that Nikolayev & Starostin obviously do not have even the most basic understanding of the languages they are studying and are just doing «wordlist linguistics».

should come as no surprise, since one of its authors has already co-autho-red an etymological dictionary of another non-existing family: North Cau-casian». Not even Brill Publishers manage to escape AV's wrath: «Brill... has a long and distinguished tradition of producing excellent books in various fields of the humanities. However, there are some warning signs that this good reputation may not last forever» [VOVIN 2005: 122].

The situation becomes even less comprehensible once it is realized that all of this blusterous indignation stems from someone who, less than a dec-ade ago, was a convinced «pro-Altaicist»; conducted active research on evid-ence in favour of genetic relationship between the daughter branches of Al-taic; praised [STAROSTIN 1991] as a significant work on Altaistics, albeit with understandable reservations [VOVIN 1995]; and, oddest thing of all, provided reasonable counterarguments to the exact same anti-Altaicist accusations that he is now himself reiterating all over [VOVIN 2005] (see, in particular, [VOVIN 1994] and [VOVIN 1999], where issues like internal reconstruction, phonetic correspondences, and productive-predictiveness are tackled much in the same way as we tackle them in the present article). It is, of course, one thing to eventually become persuaded by your opponent's arguments if you cannot reasonably counterbalance them; but it is an entirely different thing if you *do* find the reasonable answer, state it publicly, and *then,* all of a sudden, switch to the view of the opponent, forgetting everything about your former self.

AV, of course, has a brief explanation for that: «I believe that my over-all knowledge of Altaic philology has slowly improved over the years, and with that came a better understanding of various linguistic facts and data» [VOVIN 2005: 72]. Yet a meticulous comparison of AV's publications from around the late Eighties to the present day does not reveal any major breakthroughs in the amount of AV's professional knowledge, which ever since the beginning of his career has steadily remained at a high level; and it is not to be doubted that he was well acquainted with the absolute ma-jority of «anti-Altaic» arguments long before 2005.

It can also be doubted that a simple statement like «I feel no shame for changing my views, as I believe that a scholar must be open to change, and should feel free to turn back on his past mistakes» [VOVIN 2002: 168] is, from an ethical point of view, fully enough to justify AV's line of conduct — partic-ularly in the light of the fact that, ever since his «transformation», he has for the most part been feeling free to turn back on the past «mistakes» of *others* rather than his own ones, of which, judging by the list of his publications, there must have been many. Considering that AV, as we have shown above, now holds (in the field of comparative linguistics, at least) the undisputed record for both the sheer number of switched opinions and the speed with which these shifts have taken place, we think that he at least owes the scientific world a detailed explanation before embarking on any «external» critical activities.

All of this leads us to a sad, but, in our humble opinion, inevitable conclusion: [Vovin 2005] should not be qualified as a true *review* of EDAL. Instead, it is a meticulously crafted propaganda piece, bent on misrepresenting EDAL as a work of unscientific fantasy and its authors as unprofessional frauds — coming as no surprise, since its author has already engaged in a similar, slightly shorter and milder, but even less argumented enterprise concerning Sino-Caucasian [Vovin 2002].

It is not within our competence to make assumptions about *why*, in both of these cases, professional criticism has been replaced by propaganda. However, given the notable frequency of the already mentioned appeals to «mainstream linguistics», «general acceptance», etc., as a mild guess, we might cautiously suggest that AV has simply gotten tired of representing the currently unpopular «minority» of pro-Altaic/pro-long range comparison linguists, and, these days, feels more comfortable on the «safer» grounds of the anti-Altaic / anti-long range «majority». Of course, we would be glad to have ourselves proven wrong; but the general structure and outline of AV's «criticism» in both of the discussed articles cannot convince us that the primary driving force behind them had anything to do with stimulating objective, scientific discussion.

In addition to these general problems, [Vovin 2005] also sets a new low standard for academic debate in the field of Altaistics, which is well-known, of course, for its amount of vitriolics on both sides; however, not even the unquestionable anti-Altaicist champion of the sarcastic approach, the late Gerhard Doerfer, has ever stooped to directly accusing his opponents of forgery and basic ignorance of «the most elementary facts» of the languages under study so many times over the course of one article. Even some of his most scathing criticisms, e. g. [Doerfer 1995], read like the acme of politeness in comparison with much of AV's writing — most probably, motivated by overtly populistic purposes, since generalized accusations frequently produce more effect on the unprepared reader than systematic discussions based on wholesale analysis of linguistic data.

The conclusion is particularly disturbing in that none of the authors of the present article have any intention to doubt AV's own professionalism and deep knowledge of the material (at least, within his self-professed main areas of study — Japanese, Korean, and Manchu; his grasp of Turkic, Mongolic, and Tungusic data *may*, however, be put under doubt, let alone any expertise on daughter branches of Sino-Caucasian). A serious review on his part, then, one that would honestly evaluate the positive sides of EDAL on par with the negative ones, and concentrate more on listing the factual mistakes than on using all sorts of pretexts to «destroy» EDAL etymologies *en large*, would have been extremely welcome. Unfortunately, we were not presented with one[69].

[69] From AV, that is; cf., however, [Miller 2003/2004], which we also disagree with on some points, but consider well-balanced and, overall, more informative

That said, we still think that the truly discriminating reader will not remain satisfied with merely reading all sorts of reviews or responses to those reviews, but will want to draw the final conclusions on the basis of the data itself, becoming acquainted with the primary sources — EDAL for Altaic[70], [STAROSTIN 2005a] for Sino-Caucasian. Not even the most detailed review can cover all of the material presented in these sources, and even the most objective review is liable to leaving a lot of the best data unmentioned. Nor do we hold any illusions that the present article might somehow «convert» the sceptically-minded part of the audience. However, it is not really supposed to do that on its own; instead, we merely suggest using it as a brief methodological guide as the dedicated reader makes his/her way through the jungle of Altaic or Sino-Caucasian etymologies. If, within some of these pages, he/she manages to find at least a few answers to those questions that inevitably arise upon the perusal of real etymological data, we may consider our main (and, in fact, only) goal to have been fulfilled.

As for AV himself, our response will probably be of little use to him, given how actively he now distances himself from ignorant frauds doing «wordlist linguistics». His prediction that «journalists and Proto-Worlders» who «will hail EDAL as a 'great achievement'» will «all drown in the River of Time, since they do not know the Ford» [VOVIN 2005: 123] is easily applicable to the authors of EDAL as well (and is probably *meant* to be applied to them). The only thing that remains for us is to offer him our heartiest condolences for having drowned fifteen years of his own scholarly career in the same River, and hope that the rest of it will be spent in a suitably more timeless manner.

Abbreviations

EMK	Early Middle Korean		PAn	Proto-Austronesian
EOJ	Early Old Japanese		(PI)E	(Proto-)Indo-European
Evk.	Evenki		(P)J	(Proto-)Japanese
MC	Middle Chinese		PK	Proto-Korean
MJ	Middle Japanese		PM	Proto-Mongolic

than [VOVIN 2005] (of course, R. A. MILLER is one of the most well-known proponents of the Altaic theory, but it must be kept in mind that his views are seriously different from those of the authors of EDAL on many points).

[70] AV's mention of the set price for EDAL calls it «exorbitantly expensive» [VOVIN 2005: 72], in which he may be right — unfortunately, EDAL authors themselves bear no responsibility for this expensiveness, nor are they making a single cent on the publication's sales. The good news for those who cannot afford the three volumes is that EDAL is entirely based on the set of comparative Altaic databases that is currently in free public access at http://starling.rinet.ru — offered both in on-line browse/query and downloadable archive form.

MK	Middle Korean	PNC	Proto-North Caucasian
OC	Old Chinese	PST	Proto-Sino-Tibetan
OE	Old English	(P)T	(Proto-)Turkic
OHG	Old High German	(P)TM	(Proto-)Tungus-Manchu
OJ	Old Japanese	PWG	Proto-West Germanic
Osm.	Osman Turkish	PY	Proto-Yeniseian
PA	Proto-Altaic	Tib.	Tibetan

Literature

ANIKIN 2000   A. E. Аникин. *Этимологический словарь русских диалектов Сибири [An Etymological Dictionary of Russian Dialects in Siberia]*. Moscow–Novosibirsk.

BAMRS   *Большой Академический Монгольско-Русский Словарь [The Big Academic Mongolian-Russian Dictionary]*, vls. 1–3. Moscow: «Academia», 2001.

BAXTER & MANASTER RAMER 2000   William H. BAXTER, Alexis MANASTER RAMER. *Beyond Lumping and Splitting: Probabilistic Issues in Historical Linguistics*. ∥ *Time Depth in Historical Linguistics*, ed. by Colin RENFREW, April MCMAHON & Larry TRASK. McDonald Institute for Archaeological Research, Cambridge, pp. 167–188.

BENZING 1955   J. BENZING. *Die tungusischen Sprachen. Versuch einer vergleichenden Grammatik*. Wiesbaden: Steiner.

CAMPBELL 1998   Lyle CAMPBELL. *Nostratic: A Personal Assessment*. ∥ In: *Nostratic: Sifting the Evidence*. Ed. by Joseph C. SALMONS and Brian D. JOSEPH. John BENJAMINS, Amsterdam/Philadelphia, pp. 107–152.

CINCIUS 1949   В. И. Цинциус. *Сравнительная фонетика тунгусо-маньчжурских языков [Comparative Phonology of Tungus-Manchu Languages]*. Leningrad: Учпедгиз [State Publishers for Student's and Pedagogical Literature].

CLAUSON 1972   Gerard CLAUSON. *An Etymological Dictionary of Pre-Thirteenth Century Turkish*. Oxford: Clarendon Press.

DE VRIES 1992   Jan DE VRIES. *Nederlands Etymologisch Woordenboek*. Brill, Leiden — New York — Köln.

DEDR   Thomas BURROW, Murray B. EMENEAU. *A Dravidian Etymological Dictionary. Second Edition*. Oxford Publishers, 1984.

DĪNG BĀNGXĪN 1978   [DĪNG Bāngxīn] 丁邦新. *[Gŭ Bēinányŭdě nǐcè]* 古卑南語的擬測 *[A Reconstruction of Proto-Beinan]*. ∥ [Shǐyǔ

suǒ jíkān] 史語所集刊 [Bulletin of the Institute of History and Philology (BIHP)], 49.3, pp. 321–392.

DOERFER 1965   Gerhard DOERFER. *Türkische und mongolische Elemente im Neupersischen*. B. II. Wiesbaden: Franz Steiner Verlag.

DOERFER 1969   Gerhard DOERFER. *Ein altosmanisches Lautgesetz im Kurdischen*. ∥ *Wiener Zeitschrift für die Kunde des Morgenlandes*, № 62, pp. 250–263.

DOERFER 1971   Gerhard DOERFER. *Khalaj Materials*. Uralic and Altaic Series, vol. 115. Bloomington & the Hague: Indiana University and Mouton & Co.

DOERFER 1978   Gerhard DOERFER. *Urtungusisch *ö.* ∥ In: DOERFER, G. and WEIERS, M. (eds.). Tungusica I. Beitrage zur nordasiatischen Kulturgeschichte. Wiesbaden: Otto Harrassowitz, pp. 66–116.

DOERFER 1995   Gerhard DOERFER. *The Recent Development of Nostratism.* ∥ Indogermanische Forschungen, B. 100, pp. 252–267.

DOT   *Древнетюркский словарь [Dictionary of Old Turkic].* Ed. by NADELIAEV Vladimir M., TENISHEV Edgem R., SCHERBAK Aleksei M. Leningrad: Nauka.

V. DYBO 1984   В. А. Дыбо. *От редактора. Предисловие к: В. М. Иллич-Свитыч. Опыт сравнения ностратических языков, т. 3.* [Editor's Foreword to: V. M. ILLICH-SVITYCH. A Tentative Comparison of Nostratic Languages, v. 3. Moscow: Nauka.

DYBO 1991   А. В. Дыбо. *Семантическая реконструкция в алтайской этимологии. Диссертация на соискание ученой степени докт. фил. наук. [Semantic Reconstruction in Altaic Etymology. Ph.D. Thesis].* Moscow, Institute of Oriental Studies.

DYBO 1996   А. В. Дыбо. *Семантическая реконструкция в алтайской этимологии. Соматические термины (плечевой пояс) [Semantic Reconstruction in Altaic Etymology. Body Part Terms (Shoulder Area)].* Moscow: Языки Русской Культуры [Languages of Russian Culture], 1996.

DYBO 2000   А. В. Дыбо. *Ассимиляция и аккомодация гласных в сингармонистических потомках десингармонистического языка: рефлексация одной вокалической структуры в тунгусо-маньчжурских [Assimilation and Accomodation of Vowels in Vowel-Harmonizing Descendants of a Language without Vowel Harmony: Re-*

flexation of a Vocalic Structure in Tungus-Manchu Languages]. ∥ In: Евразийское пространство: звук и слово [Eurasian Area: Sound and Word], Moscow: Композитор [Componist].

DYBO 2005   А. В. Дыбо. *Дентальные взрывные в пратюркском [Dental Plosives in Proto-Turkic]*. ∥ In: Аспекты компаративистики I [Aspects of Comparative Linguistics I]. Moscow, RSUH Publishers, pp. 49–82.

EDAL   A. V. Dybo, O. A. Mudrak, S. A. Starostin. *An Etymological Dictionary of Altaic Languages*. HdO vol. 8-1, 8-2, 8-3. Leiden–Boston: Brill. 2003.

EPD₁₁   Daniel Jones. *Everyman's English Pronouncing Dictionary*. 11th Edition. ∥ A volume in Everyman's Reference Library. London: J. M. Dent & Sons, Ltd. – New York: E. P. Dutton & Co., Inc. Reprinted (with minor corrections and short supplement) 1958.

EPD₁₅   Daniel Jones. *English Pronouncing Dictionary*. 15th Edition. / Ed. by Peter Roach & James Hartman. Cambridge – New York – Melbourne: Cambridge University Press. 1997.

ERDAL 1993   Marcel Erdal. *Die Sprache der wolgabulgarischen Inschriften*. Wiesbaden: Harassowitz.

GEORG 1999/2000   Stefan Georg. *Haupt und Glieder der Altaischen Hypothese: die Körperteilbezeichnungen im Türkischen, Mongolischen und Tungusischen*. ∥ Ural-Altaische Jahrbücher, Neue Folge B, 16, pp. 143–182.

GEORG 2004   Stefan Georg. *Review of Etymological Dictionary of Altaic Languages*. ∥ Diachronica 21: 2, pp. 445–450.

GEORG 2005   Stefan Georg. *Reply to Starostin*. ∥ Diachronica 22: 2, pp. 455–457.

GEORG & VOVIN 2005   Stefan Georg & Alexander Vovin. *Review of Joseph Greenberg's «Eurasiatic: Indo-European and Its Closest Relatives, v. 2»*. ∥ Diachronica 22: 1, pp. 184–190.

GRUBE 1896   Wilhelm Grube. *Die Sprache und Schrift der Jučen*. Leipzig: Harrassowitz.

HATTORI 1978–79   [Shirō Hattori] 服部四郎. *[Nihon sogo-ni tsuite]* 日本祖語について *[On the Proto-Japanese Language]*. ∥ In: [Gengo-no 1–22 rensai] 言語の 1–22 連載.

HDC   *[Hànyǔ dà cídiǎn]* 漢語大辭典. 13 vols. [Shànghǎi] 上海: [Císhū chūbǎnshè] 辭書出版社 [Lexicographical Publishing House], 1986–1993.

HOCK & JOSEPH 1996	Hans H. HOCK, Brian D. JOSEPH. *Language History, Language Change, and Language Relationship: An Introduction to Historical and Comparative Linguistics.* Berlin & New York: Mouton de Gruyter.
IKEGAMI 1997	[Jirō IKEGAMI] 池上二良. *[Uiruta go jiten]* ウイルタ 語辞典 *[Dictionary of the Orok Language].* [Sapporo] 札幌: [Hokkaidō Daigaku Tosho kankōkai] 北海道大学図書刊行会 [Hokkaido University Press].
JANHUNEN 1994	Juha JANHUNEN. *Additional Notes on Japanese and Altaic.* ∥ Journal de la Société Finno-Ougrienne, № 85, pp. 236–240.
JÄSCHKE 1987	Heinrich A. JÄSCHKE. *A Tibetan-English Dictionary.* Delhi: Motilal Banarsidass.
JDB	*[Jidai betsu kokugo dai jiten: Jōdai hen]* 時代別国語大辞典: 上代編 *[Great Dictionary of the Japanese Language by Era: Ancient Volume].* [Tōkyō] 東京: [Sanseidō] 三省堂, 1967.
KANE 1989	Daniel KANE. *The Sino-Jurchen Vocabulary of the Bureau of Interpreters.* Uralic and Altaic Series, vol. 153. Bloomington: Indiana University.
KARLGREN 1950	Bernhard KARLGREN. The Book of Odes. *Chinese Text, Transcription and Tramslation.* Museum of Far Eastern Antiquities, Stockholm.
KARLGREN 1964a	Bernhard KARLGREN. *Grammata Serica Recensa.* Museum of Far Eastern Antiquities, Stockholm.
KARLGREN 1964b	Bernhard KARLGREN. *Glosses on the Book of Odes.* Museum of Far Eastern Antiquities, Stockholm.
KIYOSE 1977	Gisaburo N. KIYOSE [清瀬 義三郎 則府]. *A Study of the Jurchen Language and Script.* [Kyōto] 京都: [Hōritsu Bunka-sha 法律文化社 [Horitsu Bunka basic books].
KORMUSHIN 1978	И. В. КОРМУШИН. *Каузативные формы глагола в алтайских языках [Causative Verbal Forms in Altaic Languages].* ∥ In: Очерки сравнительной морфологии алтайских языков [Essays on Comparative morphology of the Altaic Languages]. Leningrad: Nauka.
KRAUSS 1970	Michael E. KRAUSS. *Eyak Dictionary.* College, AK: University of Alaska.
KRAUSS 1973	Michael E. KRAUSS. *Na-Dene.* ∥ Linguistics in North America, ed. by Thomas A. Sebeok. The Hague: Mouton, pp. 903–978.
KRISHNAMURTI 2003	Bh. KRISHNAMURTI. *The Dravidian Languages.* Cambridge University Press, 2003.

Lᴀᴘᴏʟʟᴀ 2003	*Overview of Sino-Tibetan Morphosyntax.* ⫽ In: The Sino-Tibetan Languages. Ed. by Graham Tʜᴜʀɢᴏᴏᴅ and Randy J. Lᴀᴘᴏʟʟᴀ. Routledge, London and New York, pp. 22–42.
Lɪ 1981	Paul Jen-kuei Lɪ. *Reconstruction of Proto-Atayalic Phonology.* ⫽ Bulletin of the Institute of History and Philology 52.2, pp. 235–301.
Lɪɢᴇᴛɪ 1953	Lajos Lɪɢᴇᴛɪ. *Note préliminaire sur le déchiffrement des 'petits caractères' joutchen.* ⫽ Acta Orientalia Hungarica XII: 2, pp. 211–228.
LPD	John C. Wᴇʟʟs. *Longman Pronouncing Dictionary.* [2ⁿᵈ edition.] Harlow: Pearson Education Limited. 2000.
Mᴀʜᴀᴘᴀᴛʀᴀ 1987	B. P. Mᴀʜᴀᴘᴀᴛʀᴀ. *Malto-Hindi-English Dictionary.* Central Institute of Indian Languages, 1987.
Mᴀɴᴀsᴛᴇʀ Rᴀᴍᴇʀ, Vᴏᴠɪɴ & Sɪᴅᴡᴇʟʟ 1998	Alexis Manaster Rᴀᴍᴇʀ, Alexander Vᴏᴠɪɴ & Paul Sɪᴅᴡᴇʟʟ. *On Body Part Terms as Evidence in Favor of the Altaic Hypothesis.* ⫽ Ural-Altaische Jahrbücher, Neue Folge B, pp. 116–138.
Mᴀʀᴛɪɴ 1966	Samuel E. Mᴀʀᴛɪɴ. *Lexical Evidence Relating Korean to Japanese.* ⫽ Language, vol. 42, 2, pp. 185–251.
Mᴀʀᴛɪɴ 1987	Samuel E. Mᴀʀᴛɪɴ. *The Japanese Language Through Time.* New Haven & London: Yale University Press.
Mᴀʀᴛɪɴ 1996	Samuel E. Mᴀʀᴛɪɴ. *Consonant Lenition in Korean and the Macro-Altaic Question.* Honolulu: Center for Korean Studies, University of Hawaii.
McAʟᴘɪɴ 1981	David W. McAʟᴘɪɴ. *Proto-Elamo-Dravidian: the Evidence and Its Implications.* The American Philosophical Society, Philadelphia.
Mɪʟʟᴇʀ 1971	Roy Andrew Mɪʟʟᴇʀ. *Japanese and the Other Altaic Languages.* Chicago and London: The University of Chicago Press.
Mɪʟʟᴇʀ 1996	Roy Andrew Mɪʟʟᴇʀ. *Languages and History: Japanese, Korean, and Altaic.* Oslo: The Institute for Comparative Research in Human Culture.
Mɪʟʟᴇʀ 2003/2004	Roy Andrew Mɪʟʟᴇʀ. *Review of Etymological Dictionary of the Altaic Languages.* ⫽ Ural-Altaische Jahrbücher, 18.
Mɪʏᴀᴋᴇ 2003	Marc H. Mɪʏᴀᴋᴇ. *Philological Evidence for *e and *o in Pre-Old Japanese.* ⫽ Diachronica XX.1, pp. 81–137.
Mᴜᴅʀᴀᴋ 1984	О. А. Мᴜᴅʀᴀᴋ. *К вопросу о внешних связях эскимосских языков* [*On the External Relations of Eskimo Lan-*

	guages]. ‖ In: Лингвистическая реконструкция и древнейшая история Востока [*Linguistic Reconstruction and Prehistory of the East*]. Part 1. М., ИВ АН СССР [Insitute of Oriental Studies].
MUDRAK 1985	О. А. МУДРАК. *К вопросу о чжурчжэньской фонетике* [*On Jurchen phonetics*]. ‖ In: Языки Азии и Африки (Фонетика. Лексикология. Грамматика) [*Languages of Asia and Africa (Phonetics. Lexicology. Grammar)*]. Moscow: Nauka, pp. 131–140.
MUDRAK 1988	О. А. МУДРАК. *Знаки чжурчжэньского письма* [*The Characters of the Jurchen Script*]. ‖ In: Синхрония и диахрония в лингвистических исследованиях, ч. I [*Synchrony and Diachrony in Linguistic Research*]. Moscow: Nauka, pp. 187–210.
MUDRAK 2005	О. А. МУДРАК. *Заметки о языке и культуре дунайских булгар* [*Notes on the Language and Culture of Danubian Bulgars*]. ‖ In: Аспекты компаративистики I [*Aspects of Comparative Linguistics I*]. Moscow: RSUH publishers, 83–106.
MUDRAK 2007	О. А. МУДРАК. *Развитие сочетаний с *ł в праалтайском языке* [*The Development of Clusters with *ł in Proto-Altaic*]. ‖ In: Аспекты компаративистики II [*Aspects of Comparative Linguistics II*], Moscow, RSUH Publishers, pp. 337–352.
MUDRAK 2008	О. А. МУДРАК. *Kamchukchee/Eskimo Glottochronology and Some Altaic-Eskimo Etymologies Found on the Swadesh list*. In this volume, pp. 297–336.
NAISH & STORY 1963	Constance M. NAISH & Gillian L. STORY. *English-Tlingit Dictionary: Nouns*. Summer Institute of Linguistics: Fairbanks, Alaska.
NAISH & STORY 1973	Constance M. NAISH & Gillian L. STORY. *Tlingit Verb Dictionary* (P. 1: *English-Tlingit*; P. 2: *Tlingit-English*). Summer Institute of Linguistics: College, AK: University of Alaska & Alaska Native Language Center.
NIKOLAYEV & STAROSTIN 1994	S. L. NIKOLAYEV, S. A. STAROSTIN. *A North Caucasian Etymological Dictionary*. Moscow: Asterisk.
NOVIKOVA 1980	К. А. НОВИКОВА. *Очерки диалектов эвенского языка* [*Sketches on Even Dialects*]. V. II. Leningrad: Nauka.
PDAE	John Samuel KENYON and Thomas Albert KNOTT. *A Pronouncing Dictionary of American English*. Springfield, Mass.: Merriam-Webster Inc., Publishers. 1953.

PEIROS & STAROSTIN 1996   Ilia PEIROS, Sergei STAROSTIN. *A Comparative Vocabulary of Five Sino-Tibetan Languages.* Fascicles 1–6. University of Melbourne.

PINNOW 1966   Heinz-Jürgen PINNOW. *Grundzüge einer historischen Lautlehre des Tlingit: Ein Versuch.* Wiesbaden: Harrassowitz.

POKORNY 1959/1969   *Indogermanisches etymologisches Wörterbuch.* 2 Bände. Bern und München (Bern/Stutgart 1989, 2. Aufl.; Tübingen/Basel 1994, 3. Aufl.): Francke Verlag.

POPPE 1973   *Über die Bildungssuffixe der mongolischen Bezeichnungen der Körperteile.* ‖ Ural-Altaische Jahrbücher, B. 45, pp. 223–243.

RAMSEY 1978   S. Robert RAMSEY. *Accent and Morphology in Korean Dialects: A Descriptive and Historical Study.* Seoul: Tower Press.

RAMSEY 1986a   S. Robert RAMSEY. *The Inflecting Stems of Proto-Korean.* Language Research, 22: 2, pp. 183–194.

RAMSEY 1986b   S. Robert RAMSEY. *Proto-Korean and the Origin of Korean Accent.* Paper presented to the Asian Linguistics Colloquium, University of Washington, Seattle, November 21, 1986.

RAMSEY 1991   S. Robert RAMSEY. *Proto-Korean and the Origin of Korean Accent.* ‖ In: Studies in the Historical Phonology of Asian Languages. Ed. by William G. BOLTZ and Michael C. SHAPIRO. Amsterdam/Philadelphia: John Benjamins, pp. 215–238.

RÄSÄNEN 1969   M. RÄSÄNEN. *Versuch eines etymologischen Wörterbuchs der Türksprachen.* Helsinki: Suomalais-Ugrilainen Seura.

RHCUD   *Random House Compact Unabridged Dictionary.* Special Second Edition. New York: Random House. 1987.

RÉDEI 1988   Károly RÉDEI. *Uralisches etymologisches Wörterbuch.* Budapest: Akadémiai Kiadó.

REFORMATSKY 1955   А. А. РЕФОРМАТСКИЙ. *Введение в языкознание* [*Introduction to Linguistics*]. Moscow: Учпедгиз (= Государственное учебно-педагогическое издательство [State Pedagogical Publishing House]).

RÓNA-TAS 1986   András RÓNA-TAS. *A Volga Bulgarian Inscription from 1307.* ‖ *Acta Orientalia Hungarica* XXX 2, pp. 153–186.

SAGART 1999   Laurent SAGART. *The Roots of Old Chinese.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

SAGART 2001	Laurent SAGART. *The Evidence for Sino-Austronesian*. Paper presented for the workshop «*Perspectives on the Phylogeny for East Asian Languages*», August 2–31, 2001, Périgueux, France.
SAGART 2002	Laurent SAGART. *Sino-Tibeto-Austronesian: An Updated and Improved Argument*. Paper presented for ICAL9, Canberra, January 8–11, 2002.
SAPIR 1915	Edward SAPIR. *The Na-Dene languages: A Preliminary Report*. ∥ American Anthropologist 17, pp. 534–558.
SASSE 1976	Werner SASSE. *Das Glossar Koryŏ-pangŏn im Kyerim-yusa*. Wiesbaden: Otto Harrassowitz.
SCHUESSLER 1987	Axel SCHUESSLER. *A Dictionary of Early Zhou Chinese*. Honolulu: University of Hawaii Press.
SCHULZE 1997	Wolfgang SCHULZE. *Review of 'A North Caucasian Etymological Dictionary' by Sergei A. Starostin and Sergei L. Nikolayev*. ∥ Diachronica XIV.1, pp. 149–161.
SEEFLOTH 2000	Uwe SEEFLOTH. *Die Entstehung polypersonaler Paradigmen im Uralo-Sibirischen*. ∥ Zentralasiatische Studien 30, pp. 163–91.
SRRSD	*Шорско-русский, русско-шорский словарь [Shor-Russian, Russian-Shor Dictionary]*. Kemerovo, 1993.
STAROSTIN 1972	С. А. Старостин. *К проблеме реконструкции праяпонской фонологической системы [On the Problem of Reconstruction of the Proto-Japanese Phonological System]*. ∥ In: Конференция по сравнительно-исторической грамматике индоевропейских языков (12–14 декабря). Предварительные материалы [Conference on the Comparative Historical Grammar of the Indo-European Languages (12–14 December). Preliminary materials]. Moscow: Nauka, pp. 72–75.
STAROSTIN 1975	С. А. Старостин. *К вопросу о реконструкции праяпонской фонологической системы [Towards the Reconstruction of the Proto-Japanese Phonological System]*. ∥ In: Очерки по фонологии восточных языков [Essays on Phonology of Eastern Languages]. Moscow: Nauka, pp. 271–280.
STAROSTIN 1982	С. А. Старостин. *Праенисейская реконструкция и внешние связи енисейских языков [Proto-Yeniseian Reconstruction and the External Relations of Yeniseian Languages]*. ∥ In: Кетский сборник [Studia Ketica], v. 3. Leningrad, Nauka, pp. 144–237.

Starostin 1989   S. A. Starostin. *Nostratic and Sino-Caucasian*. ∥ *Explorations in Language Macrofamilies*. Bochum: Studienverlag Dr. Norbert Brockmeyer, pp. 42–66.

Starostin 1991   С. А. Старостин. *Алтайская проблема и происхождение японского языка [The Altaic Problem and the Origins of the Japanese Language]*. Moscow: Nauka.

Starostin 1995a   С. А. Старостин. *Сравнительный словарь енисейских языков [Comparative Dictionary of the Yeniseian Languages]*. ∥ Кетский сборник [Studia Ketica], v. 4. Moscow: Языки Русской Культуры [Languages of Russian Culture], pp. 176–315.

Starostin 1995b   S. A. Starostin. *Old Chinese Vocabulary: A Historical Perspective*. ∥ In: The Ancestry of the Chinese Language. Ed. by William S.-Y. Wang. Journal of Chinese Linguistics Monograph Series 8, pp. 225–248.

Starostin 1996   S. A. Starostin. *Word-Final Resonants in Sino-Caucasian*. ∥ Journal of Chinese Linguistics, vol. 24.2, pp. 281–311.

Starostin 1997   S. A. Starostin. *On the «Consonant Splits» in Japanese*. ∥ In: Indo-European, Nostratic, and Beyond. Festschrift for Vitalij V. Shevoroshkin. Ed. by Irén Hegedűs, Peter A. Michalove, and Alexis Manaster Ramer. Washington, Institute for the Study of Man, pp. 326–341.

Starostin 2000a   S. A. Starostin. *Comparative-Historical Linguistics and Lexicostatistics*. ∥ Time Depth in Historical Linguistics, ed. by Colin Renfrew, April McMahon & Larry Trask. McDonald Institute for Archaeological Research, Cambridge, pp. 223–259.

Starostin 2000b   S. A. Starostin. *Indo-European Among Other Language Families: Problems of Dating, Contacts and Genetic Relationships*. Ms. ∥ Reprinted in: С. А. Старостин. Труды по языкознанию [Works on Linguistics]. Moscow: Языки Славянских Культур [Languages of Slavic Cultures], 2007, pp. 806–820.

Starostin 2002   S. A. Starostin. *A Response to Alexander Vovin's Criticism of the Sino-Caucasian Theory*. ∥ Journal of Chinese Linguistics, v. 30, № 1, pp. 142–153.

Starostin 2005a   S. A. Starostin. *Sino-Caucasian*. ∥ Ms., available online at: http://starling.rinet.ru/Texts/scc.pdf.

STAROSTIN 2005b	S. A. STAROSTIN. *Response to Stefan GEORG's Review of the Etymological Dictionary of the Altaic Languages*. ∥ Diachronica 22: 2, pp. 451–454.
G. STAROSTIN 1995	Г. С. Старостин. *Морфология коттского глагола и реконструкция праенисейской глагольной системы* [*Morphology of the Kott Verb and Reconstruction of the Proto-Yeniseian Verbal System*]. ∥ In: *Кетский сборник* [*Studia Ketica*], v. 4: *Лингвистика* [*Linguistics*]. Moscow: Школа «Языки Русской Культуры» [«Languages of Russian Culture»], pp. 122–175.
G. STAROSTIN 2002	G. S. STAROSTIN. *On the Genetic Affiliation of the Elamite Language*. ∥ *Mother Tongue*, v. VII, pp. 147–170.
STAROSTIN & PEIROS 1984	S. STAROSTIN, I. PEIROS. *Sino-Tibetan and Austro-Thai*. ∥ Computational Analyses of Asian and African Languages, 22, pp. 123–128.
SUNIK 1962	О. П. Суник. *Глагол в тунгусо-маньчжурских языках* [*The Verb in Tungus-Manchu Languages*]. Leningrad: Nauka.
TARAMA SÖZLÜĞÜ	*XIII. Yüzyıldan beri Türkiye Türkçesiyle yazılmış kitaplardan toplanan tanıklarıyla tarama sözlüğü, II*, Ankara. 1965.
THORPE 1983	Maner Lawton THORPE. *Ryûkyûan Language History*. Ph. D. dissertation, University of Southern California.
TING PANG-HSIN 2002	TING Pang-Hsin [丁邦新]. *Morphology in Archaic Chinese: A Review of The Roots of Old Chinese by Laurent Sagart*. ∥ Journal of Chinese Linguistics, v. 30, № 1, pp. 190–209.
TL	*Tamil Lexicon*. Published under the authority of the University of Madras. 6 vols., 1924–1939.
TMS	*Сравнительный словарь тунгусо-маньчжурских языков. Под редакцией В. Цинциус [A Comparative Dictionary of Tungus-Manchu Languages, ed. by V. Tsintsius*]. Leningrad, Nauka, 1975–77.
TRASK 1996	L. TRASK. *Response to Starostin*. ∥ Mother Tongue II, pp. 111–118.
TSABOLOV 2001	Р. Л. Цаболов. *Этимологический словарь курдского языка* [*An Etymological Dictionary of the Kurdish Language*]. Moscow, Восточная литература [Oriental Literature].
TSUCHIDA 1976	[Shigeru TSUCHIDA] 土田滋. *Reconstruction of Proto-Tsouic Phonology*. [Tōkyō] 東京: [Ajia-Afurika Gen-

	go Bunka Kenkyūjō] アジア・アフリカ言語文化研究所 [Institute for the Study of Languages and Cultures of Asia and Africa (ILCAA)].
Tsuchida 1981	[Shigeru Tsuchida] 土田滋. *Puyuma-English Index (Austronesian Languages in Taiwan).* ∥ [Tōkyō Daigaku gengogaku ronshū] 東京大学言語学論集 [Tokyo University Linguistics Papers (TULIP)] 4, pp. 10–63.
Turayev 2001	*История и культура орочей. Историко-этнографические очерки [The History and Culture of the Oroch. Historical-Ethnographical Essays].* Ed. by V. A. Turayev. St.-Petersburg: «Nauka».
Turner 1966	Ralph L. Turner. *A Comparative Dictionary of the Indo-Aryan Languages.* London: Oxford University Press, 1962–1966.
Unger 1990	J. Marshall Unger. *Summary Report of the Altaic Panel.* ∥ In: Linguistic Change and Reconstruction Methodology. Ed. by Philip Baldi. Mouton de Gruyter, Berlin-New York, pp. 479–482.
Vámbéry 1901	H. Vámbéry. *Alt-osmanische Sprachstudien.* Leiden: Brill.
Vasilevich 1958	Г. М. Василевич. *Эвенкийско-русский словарь [Evenki-Russian Dictionary].* Москва: Государственное издательство иностранных и национальных словарей [State Publishing House of Foreign and National Dictionaries].
Vovin 1993a	A. Vovin. *Long Vowels in Proto-Japanese.* ∥ Journal of East Asian Linguistics 2, pp. 125–134.
Vovin 1993b	A. Vovin. *About the Phonetic Value of the Middle Korean Grapheme △* ∥ Bulletin of the School of Oriental and African Studies, LVI, 2, pp. 247–259.
Vovin 1994	A. Vovin. *Genetic Affiliation of Japanese and Methodology of Linguistic Comparison.* ∥ Journal de la Société Finno-Ougrienne, № 85, pp. 241–256.
Vovin 1995	A. Vovin. *Reply to Karl Krippes Concerning His Review of Sergei Starostin's «Altaiskaia problema i proiskhozhdenie iaponskogo iazyka».* ∥ Dhumbadji!, Vol. 2, 2.
Vovin 1997a	A. Vovin. *On the Origin of Register in Japanese.* ∥ Japanese/Korean Linguistics, v. 6, pp. 113–133.
Vovin 1997b	A. Vovin. *The Comparative Method and Ventures beyond Sino-Tibetan.* ∥ Journal of Chinese Linguistics, 25/2, pp. 308–36.

VOVIN 1999 | A. Vovin. *Some Notes on Linguistic Comparison*. ∥ Historical Linguistics and Lexicostatistics, ed. by Vitaly Shevoroshkin and Paul Sidwell, Melbourne: AHL, pp. 67–94.

VOVIN 2000 | A. Vovin. *Pre-Hankul Materials, Koreo-Japonic, and Altaic*. ∥ Korean Studies, Vol. 24, pp. 142–155.

VOVIN 2001a | A. Vovin. *Japanese, Korean, and Tungusic: Evidence for Genetic Relationship from Verbal Morphology*. ∥ Altaic Affinities (Proceedings of the 40th Meeting of PIAC, Provo, Utah 1997), ed. by David B. Honey and David C. Wright, Indiana University, Research Institute for Inner Asian Studies, pp. 183–202.

VOVIN 2001b | A. Vovin. *North East Asian Historical-Comparative Linguistics on the Threshold of the Third Millennium.* ∥ Diachronica XVIII.1, pp. 93–137.

VOVIN 2002 | A. Vovin. *Building a 'Bum-pa for Sino-Caucasian: A Reply to Sergei Starostin's Reply.* ∥ Journal of Chinese Linguistics, v. 30, № 1, pp. 154–171.

VOVIN 2003 | A. Vovin [アレキサンダー・ボビン]. *Introduction.* ∥ In: *[Nihon go keitōron-no genzai]* 日本語系統論の現在 *[Perspectives on the Origins of the Japanese Language]*. / アレキサンダー・ボビン, 長田俊樹共編 [Edited by Osada Toshiki and Alexander Vovin]. [Nichibunken sōsho] 日文研叢書 [Nichibunken Japanese Studies Series], № 31. [Kyōto] 京都: [Kokusai Nihon bunka kenkyu series] 国際日本文化研究センター [International Research Center for Japanese Studies]; pp. 421–428.

VOVIN 2005 | A. Vovin. *The End of the Altaic Controversy.* ∥ Central Asiatic Journal 49.1, pp. 71–132.

WEUD | *Webster's Encyclopedic Unabridged Dictionary of the English Language.* New York: Portland House. 1989.

WP | Alois Walde. *Vergleichendes Wörterbuch der indogermanischen. Herausgegeben und bearbeitet von Julius Pokorny.* Berlin und Leipzig: Walter de Gruyter & Co., 1925.

WHITMAN 1985 | John B. Whitman. *The Phonological Basis for the Comparison of Japanese and Korean.* Doctoral dissertation, Harvard University.

ZAJĄCZKOWSKI 1934 | A. Zajączkowski. *Studia nad językiem staro-osmańskim. I. Wybrane ustępy z anatolijsko-tureckiego przekładu Kalili i Dimny.* Kraków.

ZHĀNG 1993   གང་དབྱི་སུན་གྱིས་ཚོམ་སྒྲིག་བྱས་པ། [Krang-dByi-Sun-gyis̱ ṟtsom s̱grig byas̱-pa] ~ [Zhāng Yísūn (ed.)] 張怡蓀主編. བོད་རྒྱ་ཚིག་མཛོད་ཆེན་མོ།།། [Bod ṟGya tshig-mdzod̲ čhen-mo] ~ [Zàng Hàn dà cídiǎn] 藏漢大辭典 [A Big Tibetan-Chinese Dictionary]. པེ་ཅིང་། [Pe-Čiṅg] ~ [Běijīng] 北京: མི་རིགས་དཔེ་སྐྲུན་ཁང་། [Mi-rigs d̲Pe-s̱krun Khaṅg] ~ [Mínzú chūbǎnshè] 民族出版社 [National Minorities Publishing House].

ZHIVLOV 2007   М. Живлов. *К вопросу о реконструкции обско-угорского вокализма [Towards Reconstructing the Proto-Ob-Ugrian vocalism]*. ∥ In: Аспекты компаративистики II [Aspects of Comparative Linguistics II], Moscow, RSUH Publishers, pp. 281–309.

Статья представляет собой подробный критический разбор двух работ известного лингвиста А. Вовина, в которых ставится под сомнение научность работы Московской школы компаративистики и в первую очередь — научная ценность сино-кавказской гипотезы С. А. Старостина и версии алтайской реконструкции А. В. Дыбо, О. А. Мудрака и С. А. Старостина.

В работе подробному анализу подвергаются почти все аргументы А. Вовина, в том числе такие важные темы, как примат морфологии над лексикой при доказательстве языкового родства, проблемы степени регулярности фонетических соответствий, семантического сходства, учета филологических данных, отделения генетически обусловленных сходств от результатов межъязыковых контактов и др. Авторы приходят к выводу, что претензии, предъявленные к методологии Московской школы, в целом являются необоснованными, а ряд конкретных ошибок, допущенных ими при отборе и анализе материала, не может существенно снижать общую значимость проделанной работы.