George Starostin
Russian State University for the Humanities

Language Classification: History and Method.
By Lyle Campbell and William J. Poser.
Cambridge University Press, 2008. IX + 536 pp.

Before discussing the merits and shortcomings of the monograph under review, it should be stressed that its title may be somewhat misleading for the general reader, not well-versed in linguistic intricacies. First of all, in the context of the publication, "language classification" is mostly limited to the issue of *genetic* classification — i. e., classifying languages depending on their being or not being descended from a common ancestor, as opposed to, e. g., typological classification, which groups languages together depending on degrees of similarity as perceived in some of their elements, regardless of the origins of these similarities (inheritance, borrowing, or independent development).

Second, for the purposes of this particular volume, "classifying" languages is primarily understood as "demonstrating genetic relationship" between two or more languages or language groups (we may call this *external classification*), rather than estimating the degree of relationship between languages in an already well-established language group (*internal classification*). This is an important point, because in comparative linguistics, issues of internal classification are just as frequent and hotly debated (sometimes even more so) as those of external classification. The former, however, do not form the major object of Campbell & Poser's monograph, which is essentially preoccupied with the question of "yes or no?" rather than with issues of degree; thus, in section 5.2, dedicated to the Hittite language, the authors only deal with how Hittite was shown to belong to the Indo-European family, and not with the relative position of Hittite within Indo-European — a problem which is still occasionally debated even as of today.

The basic goal of Campbell & Poser's book, then, is to provide a reliable answer to the very first question on its very first page: "how are languages shown to be related to one another?" A preliminary reply is found on p. 4: "throughout the history of linguistics the criteria... included evidence from three sources: **basic vocabulary**, **grammatical evidence** (especially morphological), and **sound correspondences**". The rest of the book can be roughly divided into three parts: (a) supporting the historic veracity of this claim (chapters 2 to 6, which serve as a basic overview of the history

of comparative linguistics); (b) demonstrating the intrinsic correctness of this claim by showing how practically all other types of criteria can be misleading in establishing genetic relationship (chapters 7, 8 and 10); (c) assessing several proposed, but controversial hypotheses of genetic relationship on the basis of the established criteria (chapter 9, as well as chapter 12 that deals with the issue of 'Proto-World').

Obviously, the question of criteria used to establish genetic relationship should, and does, lie at the very heart of comparative linguistics as a science. Given that Campbell & Poser's monograph is essentially oriented at a scholarly reception rather than the general public at large (I would probably define its primary target audience as linguists with but a passing interest in historical linguistics, as well as non-linguists with a serious interest in human prehistory, e. g. professional archaeologists or geneticists), one may ask the question: "Why do the very foundations of this branch of science still have to be discussed and defended more than a hundred years after its inception?" (provided we take a very restricted view of historical linguistics as a science and count its birth from the emergence of the Neogrammarian model in the 1870s, rather than from the work of William Jones in the late 18th century).

As a comparative linguist myself, I would probably give two interconnected reasons: (a) a steady decline in popularity of comparative-historical linguistics as a whole, starting approximately from the beginning of the 20th century and resulting in a serious lack of knowledge about it even among linguists, let alone specialists in other, only indirectly related, disciplines; (b) the emergence, against this background, of a veritable swarm of "theories" and "hypotheses" about the history of language as a whole or of individual languages and language groups, that have little, if anything, to do with science, yet, quite frequently, find unexpected popularity among non-specialists and amateurs — if properly advertised and brightly presented. These points are not explicitly stated in the book, but I have no doubt that both of these opinions, especially the second one, are shared by both of its authors. The necessity of having a reasonable, formalized, well-defined set of tools that would permit

the researcher and the student alike to separate true historical linguistics from the realms of fantasy is equally well realized by all of us.

The devil, as usual, is in the details. The authors expose their principal target of critique already on that very same first page of the introduction: these are the so-called "long-rangers", proposing "distant linguistic kinship such as Amerind, Nostratic, Eurasiatic, and Proto-World". A lengthy, 10-page list of "hypothesized distant genetic relationships" is adduced at the end of the book, with the authors themselves admitting that these hypotheses "are not all of equal quality"; indeed, they range from solid theories deserving scholarly attention to joke-level comparisons, and it would be useless (and even cruel!) to expect the authors to present us with detailed evaluations of all of them.

Nevertheless, the main culprits, as perceived by the authors, are quickly and easily identifiable, as they are mentioned far more frequently than anyone else in the critical sections. First and foremost, this is the late J. Greenberg (principal mastermind behind the Amerind, Eurasiatic, and Indo-Pacific hypotheses) and some of his followers who continue to operate based on his method of "multilateral", or "mass" comparison, such as M. Ruhlen. Second, this is the Soviet/Russian tradition of macrocomparative research as originated by the late V. M. Illich-Svitych, author of the "Nostratic" hypothesis, and continued by many of his disciples and followers. (It should be noted, though, that the amounts of critique directed at these two "schools" of comparative thought are quite disproportionate: scourging of the "Greenberg-Ruhlen direction" occupies at least twice, if not thrice, the space devoted to scourging of the Russian side.) Distant language relationship theories that are not directly connected to the "Greenberg legacy" or to the "Illich-Svitych legacy" occupy very limited space and are taken mostly from a purely historical perspective (e. g., a brief account of the Ramstedt/Poppe tradition of Altaic linguistics on p. 235–241).

Chapters 2–6 (pp. 13–161) of the book, as I already mentioned, formally represent a brief overview of the main successes of comparative linguistics. They are more often than not well-written, informative, and even entertaining, with a lot of research on the early prehistory of the science and details that are certainly not common knowledge even among those specializing in it. However, most of this overview really serves one purpose: to be able, on the basis of it, to answer in the negative to the following question: "Has anyone ever succeeded in proving genetic relationship on the basis of something other than grammatical evidence and sound correspondences?" (I take "proving genetic relationship" here to mean "presenting evidence in favor of a theory of genetic relationship that would make said theory accepted by the scientific mainstream").

It is not difficult to guess that the authors, at the end of this section, remain convinced that the answer is, indeed, a steady "no". And, from a formal point of view, they are correct: I do not know any such theories myself. Even such proposed language families as Khoisan, Nilo-Saharan, and Australian, which used to find a lot of support from specialists, as correctly indicated to the authors, are now put more and more into doubt by mainstream linguists for the very same reasons — seeming scarcity of evidence from grammar and comparative phonology.

But obviously, "majority votes" and historical analogies alone cannot by themselves disprove the validity of either "mass comparison" or other ways of assessing linguistic evidence in a historical context. Cases when "the majority" has, in the long run, been proven wrong, abound in the history of science; likewise, the very fact that a certain task has, in the past, always been performed according to a set standard of rules, does not necessarily mean that there cannot possibly exist a different standard by means of which it can be performed with equal (or even grander!) success. This issue — that "traditional" comparative methodology is right not because it is "traditional" but because it is, well, *right* — is addressed in Chapter 7. This is where the real problems start.

On the surface, many, if not most, of the arguments presented in this chapter look perfectly reasonable. It essentially functions as a set of "warnings" and "filters" constructed by the authors in order to separate convincing evidence for language relationship from unconvincing evidence or "non-evidence". Numerous obstacles are listed that hamper, or should be considered as hampering, the work of the "long-ranger". First and foremost among them is the issue of borrowing and language contact, coming in the form of lexical resemblances, sometimes even within the basic vocabulary (7.2.1), subsystems of phonetic correspondences (7.3), and grammatical features (7.4); additional problems include lax semantic demands (7.6), onomatopoeia (7.7), sound symbolism (7.8), possibility of chance similarities (7.11), and even direct mistakes in analyzing material — erroneous morphological analysis (7.14), neglect of language history (7.15) and spurious or invented forms (7.16).

The authors list multiple examples of cases where these factors either have or could have contributed to reaching unwarranted, premature, or even downright

wrong conclusions on the part of the researcher (with the lion's share of such examples culled from works by J. Greenberg and M. Ruhlen). Unfortunately, almost all of the arguments they make in this section suffer from one fatal flaw: they represent a broad, over-generalized approach to the problem that refuses to deal with it on a more detailed level, or at least specify the degree to which this particular problem may render useless a certain piece of work. In addition, many of these problems are characteristic not specifically of work on long-distance relationship, but of all kinds of comparative research.

For example, in p. 7.15.2 ("Neglect of known history"), it is correctly stated: "Another related problem is that of isolated forms which appear similar to forms from other languages with which they are compared, but when the known history is brought into the picture, the similarity is shown to be fortuitous". One example from Greenberg's work on Amerind is quoted, with a few others later adduced in Chapter 9. Should we infer that ignorance of language history is an almost obligatory ingredient in all, or most, works by "long-rangers"? All the authors can say about this is that such mistakes are "not uncommon in proposals of distant language relationship" (p. 209) — a phrase that is essentially meaningless, since it is unclear which particular works on distant language relationships have been shown to be useless or even anti-scientific on the basis of this criterion. (It should also be kept in mind that for some researchers, "neglect of known history" of languages sometimes equals "neglect of what I, or mainstream specialists, *presuppose* about history", and it is not uncommon to confuse the two.)

Another type of "over-generalization" of problems consists of the authors' frequent calls to drop from consideration not merely particular *pieces* of potential evidence that raise doubt for a variety of reasons, but whole *blocks* of evidence. Let us take, for instance, section 7.7 ("Onomatopoeia"), the main point of which is formulated as follows: "Onomatopoetic forms may be similar because the words in different languages have independently approximated the sounds in nature, and such cases must be eliminated from proposals of distant genetic relationship, since the similarity may be explained by onomatopoetic mimicry rather than inheritance from a common ancestor" (p. 196).

This statement only looks reasonable before one gives it a second thought — and comes to the inescapable conclusion that it could be *completely* right if and only if certain words and notions in the world's languages turned out to be fully exempt from the arbitrariness principle of the linguistic sign. One example will suffice here (technically speaking, it refers to sec-

tion 7.9, "Nursery forms", rather than 7.7, but, since both deal with violations of the arbitrariness principle, this is not significant; similar examples can be easily drawn from "proper" onomatopoeic lexicon as well). It is well-known that the word for 'mother' all over the world tends to be represented by roots containing a labial nasal consonant (usually of the *ma*-type), from Indo-European **mātēr* to Chinese *mā* and Bantu **-máá*. Since this is by far the most widespread "world" root for mother, and since it is sound-symbolic, an argument that brings together *ma*-type words from different language families will not be seen as important for the task of proving their relationship. But what about the *negative* argument — languages that do not have **ma* for mother, yet agree in having a different root for the same notion, even if it, too, is sound-symbolic?

Such is, for instance, the case with Altaic, where the main word for 'mother' is currently reconstructed as **ĕpʻa*, reflected as Japanese *haha*, Mongolian *ebei*, and Turkic *apa* — all specifically indicating the *female* rather than the *male* parent. Certainly, one could object that the **pa*-type form represents the "nursery lexicon" as well, and, furthermore, cases of *pa*-type words meaning 'mother' rather than the more common meaning 'father' can also be found in non-Altaic languages (e. g. in Telugu the word *appa* refers to both 'father' and 'mother'). Presenting this argument as "crucial" for the Altaic hypothesis would be erroneous. But in the context of *other* pieces of pro-Altaic evidence, it is much more reasonable and economic to explain this situation as the result of a one-time meaning shift from 'father' to 'mother' in Proto-Altaic — rather than the curiously independent "nursery regenesis" of **pa* 'mother' rather than **ma* 'mother' in Turkic, Mongolic, and Japanese.

In other words, instead of simplifying the task and bluntly rejecting everything that even vaguely looks like "onomatopoeia", "nursery words", "sound symbolism" — and the scope of these groups, if one wishes hard enough, can be extended to contain half or more of any individual language's vocabulary — it is essential to approach the issue as a complex one, in which it is important to distinguish between "telling" types of possible cognations and irrelevant ones. Not a hint at such distinctions can be culled from Campbell & Poser's monograph.

In section 7.2 ("Lexical comparisons"), the authors, agreeing with the "long-rangers" usual claim that "basic vocabulary is more resistant to borrowing", warn against the universality of this rule, explaining that "some things in 'basic vocabulary' seem quite subject to borrowing or lexical replacement" — a claim which, to the best of my knowledge, not a single

"long-ranger" has ever argued with. There have certainly been cases in long-range comparison when borrowings in the basic lexicon have been mistaken for signs of relationship (e. g., the situation in Tai-Kadai languages, where multiple borrowings from Chinese, including a significant chunk of the basic lexicon, have for a certain period of time contributed to the perseverance of the erroneous "Sino-Tai" theory). But at the same time, historical-comparative linguistics is old enough to have procured quite a few useful methods to filter out these cases — including statistical tests; procedures of establishing special "sub-systems" of phonetic correspondences for borrowed lexical strata as opposed to "genuine" systems of correspondences (the issue is briefly mentioned by the authors on p. 174); and considerations of geographic proximity (e. g., it would be at the very least odd to insist that Proto-Turkic *bir 'one' has been borrowed from Proto-Japanese *pitə id. or vice versa).

It would, therefore, be refreshing to see, in section 7.2, a straightforward list of conditions under which similarities and correspondences observed in the basic lexicon can be deemed relevant and significant in a proposal of language relationship, and, likewise, a list of types of situations in which such similarities and correspondences rather hint at a situation of secondary contact. No such conditions are given; instead, the reader merely gets a "warning sign" — basic lexicon, too, can be borrowed (hardly a surprise for anyone with even a minimum amount of expertise in the field).

Actually, the very definition of "basic lexicon" is terminologically used by the authors of the book in a different way from, for instance, representatives of the Moscow school of comparative linguistics, and, I dare say, numerous other researchers as well. For the latter, in its most specialized usage, the "basic lexicon" means little more than the highly "compressed" 100-word version of the original list proposed by Morris Swadesh, which, it seems, has for the most part stood the test of time as truly representing some of the most stable items of the lexicon of any particular language; furthermore, even within the 100-wordlist attempts have been made, first by Aharon Dolgopolsky, then by Sergei Yakhontov, and finally, on a more strictly statistical basis by Sergei Starostin, to separate its elements into generally "more stable" and "less stable" parts.

For Campbell & Poser, however, "basic lexicon" is something much more amorphous — clinging to the mostly "intuitive" understanding of the term (p. 166: "basic vocabulary has been understood intuitively to contain terms for common body parts, close kin, fre-

quently encountered aspects of the natural world, and low numbers"), they proceed to make good use of it in passages like the following one: "For example, Udmurt (Votyak, a Finno-Ugric language) borrowed many items of basic vocabulary from Tatar (Turkic), terms for 'mother,' 'father,' 'grandmother,' 'grandfather,' 'brother/sister,' 'elder brother,' 'elder sister,' 'uncle,' 'strong,' 'healthy,' 'deaf,' 'blind,' 'sick,' 'illness,' 'love,' 'land,' 'people,' 'person,' 'cool,' etc." (p. 174, with references). Of all these words, only one — 'person' — forms part of the 100-wordlist, around 90% of which otherwise has reliable Fenno-Ugric parallels. Obviously, the more uncertain and imprecise one makes the limits of "basic lexicon", the easier it becomes to underestimate its importance. I do not mean to say that the authors have invented a particular usage for the term "basic lexicon" — many people use it in different ways — but here they have intentionally chosen the broader one that least suits the purposes of scientific long-range comparison.

An equally muddled approach is seen in section 7.19, dedicated to the issue of pronominal evidence for genetic relationship hypotheses. Taking to task the alleged claim that "pronouns are rarely borrowed", the authors proceed to demonstrate, on pp. 213–214, that "this common perception is nevertheless a misconception", by citing examples and references that show how personal pronouns can indeed be borrowed from one language into another. And yet, this particular demonstration does not in the least shatter the idea that "pronouns are rarely borrowed" — on the contrary, it *proves* it. My impression is that the authors have taken great pains to accumulate a near-complete list of exceptions — which still covers but an absolute minimum of the world's languages. It would have been understandable if their point were to disprove the statement that "pronouns are *never* borrowed", but, to the best of my knowledge, no "long-ranger" or "short-ranger" has ever stated anything like that.

Furthermore, even the examples of borrowed pronouns that *are* given constitute a hodge-podge of entirely different situations. English *they* is, indeed, a Scandinavian borrowing, but this is a 3rd person pronoun, not a 1st or 2nd one, and 3rd person pronouns are usually omitted from the discussion on stability (cf. the Shevoroshkin quote that the authors cite above which only mentions forms like 'I', 'me', 'thou', 'thee'), since they are, indeed, less stable overall (and, for that reason, omitted from the Swadesh 100-wordlist). Indonesian *saya* 'I' is, indeed, a borrowing, but in the language it coexists peacefully, as a polite form, with the inherited Austronesian *aku*; situations like these are plentiful, especially in East Asia, and

whenever they arise, it is usually easy to separate the new "polite" form from the old inherited one.

It is hardly a coincidence that the majority of examples come from various case studies of "Amerind" languages, where issues of borrowing are frequently merely suggested rather than stated with certainty, due to insufficient historical treatment of available data and lack of proper reconstruction. E. g. it is said that "Thomason and Everett... argue that Pirahã has borrowed a majority of its pronouns", but, while such an argument is indeed presented, the paper in question very clearly states that "we can't prove that the pronouns in question are innovative in Pirahã; and we have no evidence (yet) of other borrowings in Pirahã from Tupí-Guaraní" [Thomason & Everett 2001].

Still later, they write: "There are also a number of documented cases of borrowed pronouns in Native American languages. For example, Miskito borrowed its independent personal pronouns from Northern Sumu in relatively recent times: Miskito *yaŋ* 'I' (cf. Sumu *yaŋ*), *man* 'you' (cf. Sumu *man*) (Hale 1997: 154)". The actual paper by Hale, however, does not state this as a fact, like Campbell & Poser do, but rather as a cautious hypothesis — "the entire set of Miskitu independent pronouns could have been borrowed from Sumu" [Hale 1997: 154] — speculatively derived from the likely suggestion that the *third* person pronoun (Miskitu *witin*) is probably borrowed from Northern Sumu (Twahka dialect) *witin*. Even then, it must be kept in mind that Miskitu and Sumu are closely related languages within the Misumalpan family (something the authors do not explicitly tell their readers), and at the very best we could only speak of the Miskitu pronominal forms being *influenced* by Sumu forms rather than borrowed from Sumu.

In one case at least, the situation has advanced to the level of topsy-turvy. Quoting a survey work on Papuan languages, the authors write: "Warembori (an isolate in Papua New Guinea) has borrowed its pronouns: "the first- and second-person singular are transparent Austronesian loans, with even the inclusive-exclusive distinction being taken over [Foley 2000: 392]". Given the extreme rareness of the situation, I was tempted to verify it; surprisingly, the complete quotation turned out to be as follows: "The isolate Warembori (Donohue 1999) exhibits extensive borrowing from Austronesian in basic vocabulary such as kin terms, body parts, and even pronouns. *All except* (italics are mine — *G. S.*) the first- and second-person singular are transparent Austronesian loans, with even the inclusive-exclusive distinction being taken over". No traces of this "all except" can be found in Campbell & Poser's quotation, which is a pity — given that, even under considerable lexical pressure from nearby Austronesian languages, Warembori still retains the original forms for 'I' and 'thou', this is an excellent argument for these forms as representing the most stable and reliable, albeit small, layer of the lexicon.

The second part of Campbell & Poser's argument — taking to task the statement that "shared pronoun patterns defy chance" — completely escapes this particular reviewer's understanding. The central point in this part of the discussion is formulated as follows: "Patterned grammatical material can constitute strong arguments for genetic relationship if nothing else accounts for the shared patterns, but that is often not the case in pronominal paradigms. It is well established also that pronominal systems seem to be subject to analogical reformations, and are also dominated by tendencies towards iconic symbolization, as other deictic markers are" (p. 214). It is not explained why, and whether indeed, "that is often not the case", with not a single actual example to the contrary offered by the authors. It is equally unclear how the first sentence in this quotation ties in with the second one. For instance, the 1st p. sg. pronoun in Old Turkic is *ben*, in Evenki is *bi*; the 2nd p. sg. pronoun in Old Turkic is *sen*, in Evenki is *si*. Where is the "analogical reformation" or the "iconic symbolization" responsible for this pattern match? If my understanding of the historical concept of "analogy" is clear enough, "analogy" works towards making objects more similar to one another, not vice versa; and, although cases of pronominal *endings* being reshaped due to analogy are not unknown, I have yet to hear of a single case where a 1st or 2nd pronoun *stem* has contributed towards reshaping the other member of the opposition.

Likewise, although vague references to "iconic symbolization" of pronouns do crop up every now and then in literature (the authors provide several references), no respected specialist so far has managed to present a convincing argument on why personal pronouns in the world's languages should, due to some non-arbitrary process, look the way they look. At best, we usually agree that pronouns tend to be short and not to incorporate particularly strongly marked phonemes and sound clusters, which is understandable given their exceptionally high frequency. But this is the threshold at which reasonable arguments end — yes, the 1st person pronoun does feature the phoneme *m* in many corners of the world, but why does it have to be *m* and not *n*, or *k*, or *t* (all of these variants are also encountered with different levels of frequency), remains unaccounted for, certainly not within the

framework of any "phonosymbolic" theory. Much less does "phonosymbolism" account for shared patterns, such as M/T, M/S for Nostratic, Z/W for Sino-Caucasian, K/M for Austric, and N/M for Amerind (or, at least, substantial parts of Amerind, see below).

In short, the entire argument on pp. 214–215, in my opinion, only goes to show that the "pronominal paradigm" argument in favor of long-distance genetic relationship remains stronger than ever, since there is very little to offer in opposition to it, apart from vague theoretical speculations on "analogy" and "phono-symbolism" that have precious little basis in fact. This does not mean that issues of genetic relationship can be considered settled based on pronominal evidence *alone*. For one thing, as the authors correctly point out, there is always the negative argument: pronominal systems can be reshaped, losing the original similarity. Chances for the reverse — acquiring accidental similarity to an originally non-related system — also exist, but are much smaller, especially if we are dealing with several identical cases at once; an M/T-type similar paradigm can be attributed to chance if it is met in one language family in Eurasia and one other language family in America, but, obviously, the situation is different when said paradigm is observed to characterize *many* language families in Eurasia and few, if any, language families in America. This, however, merely implies that *lack* of pronominal evidence should not immediately be taken to signify lack of relationship — whereas existence of paradigmatic/pattern-type pronominal evidence immediately warrants at least further serious testing of the hypothesis. It goes without saying that if pronominal evidence constitutes the *only* attractive evidence for language relationship, this is not good for the hypothesis. But cases where successful hypotheses have been developed based on pronominal evidence alone are unknown to me.

The authors fare somewhat better in their subsection (7.19.1–3) on the proposed "Amerind" pronominal pattern N/M, mixing the usual theoretical arguments with more convincing statements like "Whatever the explanation for the frequency of 'first-person' *n*, and for the recurrence of 'second-person' *m*, it will not do to look only at American languages which contain them, ignoring the many American tongues which lack them and the numerous non-American languages where they are attested" (p. 222). This is a fair warning against over-generalization of an issue, and one that warrants a closer look at those "parts" of "Amerind" that are indeed characterized by the N/M pattern as *opposed* to other parts of it that exhibit different characteristics. However, this is an entirely dif-

ferent problem from the ones discussed above: its investigation may, in the end, lead us to exclude some branches from "Amerind", but it is not at all certain to undermine the importance of pronominal paradigms for taxonomy purposes.

One section at least of Chapter 7 (7.2.2, "Glottochronology") clearly shows that the authors have a rather poor understanding of the state-of-the-art on this topic in comparative linguistics. Glottochronology was invented as a method already more than half a century ago, and has since gone through a large number of refinements and modifications; multiple variants of it have been tested, some more successfully, some less, yet not a single word on the history of the method can be found in the discussion on it (less than one page long) in Campbell & Poser's book.

Perhaps the biggest mistake the authors — along with some of their colleagues in other publications — make is to assume that glottochronology *par excellence* is an *independent* method, *opposed* to "standard" historical-comparative linguistics rather than complementing it. Only such an assumption can account for their description of the method: "...glottochronology does not find or test family relationships, but rather just assumes them... In the application of glottochronology the basic vocabulary of any two (or more) languages, related or not, is compared and similar words on the core vocabulary list are checked off" (p. 167).

It is true that *some* lexicostatistical calculations have been performed that way, but neither the father of the method, Morris Swadesh, nor the scholar who has arguably contributed the most to its further development, Sergei Starostin, would have ever agreed with such a definition of it. In Swadesh's earlier — and arguably best — works, glottochronology was primarily a method for internal classification and dating the divergence of languages whose relationship *had already been established* through conventional means of historical-comparative linguistics. (His later works sometimes contain violations of this principle, which should not disqualify the principle itself). For Starostin, who was the first to apply glottochronology to hypothetical macrofamilies, glottochronology could serve as a definitive indicator of relationship — but only if compared words are deemed cognate based on *previously suggested regular correspondences* rather than mere similarity; in other words, glottochronology merely goes to show that the proposed system of correspondences is not bogus, since it works on the compared languages' basic lexicon. Campbell & Poser are correct in saying that glottochronology "assumes" rather than finds language relationships; but, for some reason, they refrain from saying that correctly applied

glottochronology does not "assume" relationship out of nothing, but involves a lot of hard work based on traditional methodology; and glottochronology certainly *does* "test" relationship, checking whether or not this methodology has been applied in the right way.

More "semi-truth" emerges when the authors correctly state that "all its (*glottochronology's* — G. S.) basic assumptions have been challenged" — but fail to mention that some of the 'challenges' have led to constructive emendations of the method which improve its performance on virtually all testable cases [Starostin 1999]. They also bring up the issue of how it is impossible to use glottochronology to establish or confirm deep level relationship due to large lexical losses in modern languages — but, again, fail to mention that some proponents of glottochronology have offered to substitute reconstructed proto-languages for modern languages, showing that the proportion of shared lexics on the 100-wordlist drastically increases that way, overcoming the time barrier (e. g. Proto-Germanic obviously shares more common items on the list with Proto-Slavic than modern Russian with modern English; likewise, the percentage between Proto-Indo-European and Proto-Uralic is notably higher than between any living Indo-European and Uralic language, etc.). It is not that these counter-arguments are, in turn, unassailable; it is simply that they are ignored, and glottochronology is treated exactly the same way as if nothing new whatsoever in this field had been offered since the works of Swadesh in the 1950s — a position that looks undeservedly condescending in the light of numerous works on the issue.

Thus, for some reason, [Starostin 1999], a work with a detailed exposition of the new approach towards glottochronology, is never mentioned in this section, although it is present among the references. (Granted, it would be somewhat odd, in the light of the main points of section 7.2.2, to discuss a work where it is directly stated that "…a thorough comparative historical analysis of the language data should precede any lexicostatistical or etymostatistical study, which will in any case be complementary to rather than a substitute for, comparative work" [Starostin 1999: 47]). Nor do the authors make any references to the papers in [Renfrew, McMahon, Trask 2000], a volume that shows glottochronology, in a wide variety of forms, is still very much alive.

After all the *caveats* in Chapter 7, the authors then proceed from theory to practice, dedicating Chapter 9 to their "Assessment of proposed distant genetic relationships". Basically, it consists of their selecting several of the more popular proposals on the issue and,

one by one, discarding them through a meticulous application of all said *caveats*. Not all hypotheses are given equal attention, and many are left without attention at all (e. g. Dene-Sino-Caucasian), but I do not see this as a major problem, since, given the set of criteria, it is more than likely that not a single "macro-relationship" hypothesis will ever be able to escape the elaborate set of filters as positioned by the authors.

A detailed "assessment of the assessment" will inevitably turn this review into a small monograph by itself. Therefore I will refrain from commenting on those of the authors' criticisms that do not directly refer to my own area of interests or methodological preferences; namely, all of the criticisms directed at hypotheses that have to do with Greenberg and "mass comparison", such as Amerind and Indo-Pacific. As a representative of the Moscow school of comparative linguistics, I hold no definite opinion on these suggested groupings; I concur with the authors in their conclusions that they have not been successfully demonstrated (although sometimes, perhaps, for different reasons), which should not, nevertheless, prevent the true researcher from *ignoring* the evidence presented by Greenberg and others — rather than abandoned, it should be gradually brought more into accordance with the classic comparative method, a procedure that may, in the end, either verify or decidedly overturn Greenberg's conclusions. Ridiculing "mass comparison" is an easy and grateful task, but acknowledging the method as an important first step in establishing genetic relationship would, I believe, be much more in line with the traditional understanding of science.

I will concentrate in more detail on the authors' "dissection" of Altaic and, most significantly, Nostratic — the cornerstone of Russian macro-comparative linguistics, both chronologically, since Il-lich-Svitych's Nostratic dictionary happened to be the first ever example of an etymological dictionary of a linguistic macrofamily, and methodologically, since Nostratic was the first macrofamily claimed to have been established the exact same way as the more "traditional" and "younger" language families.

The authors' "assessment" of Altaic (pp. 235–241) is hardly an independent evaluation of the theory, but rather an attempt of a brief summary of the main-stream consensus on this macrofamily. Around a third of this little chapter simply narrates the history of the Altaic hypothesis, albeit with significant gaps — thus, the recently published three-volume *Etymological Dictionary of the Altaic Languages* [EDAL] does not even deserve a mention (the work is found among the references in the end of the book but is not in any way linked to the text). Another third discusses the "typo-

logical" argument in favour of Altaic (i. e. that common features, such as vowel harmony, agglutination, etc., can be taken as indicative of genetic relationship among the various subbranches of Altaic) — as if, at the present stage of discussion, this argument were still relevant: yet hardly any serious work by any Altaicist in the past fifty years has been known to place typological arguments above other, more important ones.

To be fair, the authors recognize the primary problems for Altaic in other spheres, most importantly "extensive lexical borrowing across inner Asia" (p. 236). Since [EDAL] is not referred to, the authors apparently ignore the solutions presented therein to distinguish borrowed layers of the lexicon in Turkic, Mongolic, and Tungusic languages from inherited ones — although, coincidentally, these solutions look not unlike the ones used to distinguish borrowed and inherited lexicon in Indo-European languages, as described by the authors on p. 174 (namely, identifying "sub-systems" of phonetic correspondences that work on the cultural rather than basic lexicon and thus indicate later contact). The issue of borrowing in Altaic is, of course, far from being completely resolved even to the satisfaction of pro-Altaicists, much less anti-Altaicists, but the position taken on it by the authors is unquestionably biased towards the latter.

Let us now see what the authors have to say about Nostratic. First of all, it turns out that they have very little *new* to say about it: the entire section dedicated to the subject (9.4) is essentially a reprint, with minor corrections and major omissions, of [Campbell 1998]. This is, of course, not a criticism in itself; but it means that most of the argumentation flaws detected in that earlier paper have safely made it, with no amendments, into this larger monograph. The basic principle is the same: go through the existing Nostratic etymologies in Illich-Svitych's dictionary, with an emphasis on Indo-European and Uralic material (partly because for many specialists the "Indo-Uralic" part of the hypothesis has always seemed the most attractive one, partly because L. Campbell's knowledge of Uralic data exceeds that of the data from other branches of Nostratic), and, one by one, "strip away" dubious pieces of evidence that do not agree with the rigid system of criteria presented in Chapter 7. As shown by the examples on pp. 252–263, virtually all the evidence in favour of an inherited connection between Indo-European and Uralic displays some sort of problem — and, for those reasons, the hypothesis is labeled "unconvincing".

Major problems that the authors find with Nostratic in general, and "Indo-Uralic" in particular, are clustered in seven groups that they expose early on for the reader's convenience. These are (a) "descriptive" forms; (b) questionable cognates; (c) sets with only two families represented; (d) non-corresponding sound correspondences; (e) short forms; (f) semantically non-equivalent forms; (g) diffused forms. I will briefly consider all these problems together with some of the illustrative examples on pp. 252–263.

(a) **"Descriptive" forms**. As has already been mentioned above, "descriptive" or "expressive" forms should not be excluded from any hypothesis of relationship, provided they fit within the proposed system of correspondences. Besides, "descriptive" turns out to be a fairly vague notion. Out of Illich-Svitych's original 378 etymologies, Campbell labels 42 as "descriptive" — including such words as *$\dot{k}ol\Lambda$ 'round', *$\Omega E\dot{k}u$ 'water' and *-$\dot{k}a$ 'diminutive suffix'. What makes *-$\dot{k}a$ 'diminutive suffix' in any way more "descriptive" than, for instance, the diminutive suffix *-$j\Lambda$ (№ 151, not listed by the authors) is a question that is probably bound to remain unanswered.

The "descriptiveness" of *$\Omega E\dot{k}u$ 'water' has escaped me for a long time until I reached the following explanation much later in the book, on p. 379, where it is discussed in the context of the "global etymology" *$aq'wa$ 'water': "The similarity of sound suggests to many the imitation of the sound of swallowing water, a nursery form, or of the gurgling of running water". Apparently, I am not one of the mysterious "many", so, in regard to the Nostratic forms only, I feel entitled to ask — if *$\Omega E\dot{k}u$ represents the "descriptive" layer in so-called Nostratic, meaning that the word, due to onomatopoeic reasons, could have independently arisen as both Afro-Asiatic *$\Omega q(w)$ and Indo-European *$\hbar ek^w$- (the two major parts of the equation in Illich-Svitych's comparison), how is it that in both of these families we only find it as 'water' (rarely as 'rain' and 'drink'), but never as 'gurgling of running water' or 'sound of swallowing water'? In other words, is it not a clear-cut case of letting our imagination run a bit too wild? After all, with just a small extra touch of permissiveness, we could easily label half of the words in the world's languages as "descriptive" or "sound-symbolic". In short, this argument should be completely discarded, and the reasonable bits of it incorporated into "non-corresponding sound correspondences".

(b) **Questionable cognates**. Questionable cognates are, in fact, characteristic of every single hypothesis of language relationship, long- or short-range. The issue of whether they should or should not figure in early

stage comparisons is a debatable one. However, as long as they are clearly perceived as questionable, do not constitute the bulk of basic lexicon comparisons, and are accompanied by a reasonable explanation of why they are questionable rather than impossible, I do not see a problem with this. It should also be remembered that even within questionable etymologies, there are frequently parts that are less questionable and those that are more so, so a single question mark next to a certain form should not immediately be taken as a sign to discard the entire etymology once and for all.

(c) **Sets with only two families represented**. This short section contains a gross misrepresentation of Illich-Svitych's work. The authors write: "One of Illich-Svitych's criteria was that only cognate sets with representatives from at least three of the six families proposed as members of Nostratic would be considered as supportive of the hypothesis". This is followed with a supposedly illustrative quotation from an English translation of one of Illich-Svitych's articles. The authors then go on to state that "134 sets from the 378 involve forms from only two families. That is, 35 percent of the forms are questionable on IS's own grounds".

In reality, Illich-Svitych never suggested anything of the kind. The text that the authors are referring to is subtitled "A Probabilistic Evaluation of the Similarities in Question", and specifically proposes that we limit ourselves to cognates represented in at least three families *for the purposes of this probabilistic evaluation* — but certainly not when compiling an etymological dictionary of Nostratic. Furthermore, it looks like the authors did not bother to consult the Russian original of the publication, which (unlike the brief English translation in [Shevoroshkin 1989]) is followed by 32 pages of comparative tables [IS 1971: 6–37] that include 149 lexical and grammatical morphemes, only 2 of which accidentally are represented by reflexes found in only 2 subbranches of Nostratic, and more than half are represented in more than 3. Of course, in Illich-Svitych's own view, these tables must have constituted the strongest evidence for Nostratic. But in no way does this surmise that parallels found only in two branches are not supportive of the hypothesis — just the same way as it would be wrong, for instance, to call Indo-Greek or Slavic-Germanic isoglosses "unsupportive" of Indo-European.

(d) **Non-corresponding sound correspondences**. This can be an occasional issue even in properly conducted macro-comparative research — although the

same could be said of just about any language family (it would probably not be a stretch to say that at least half of all accepted Indo-European etymologies suffer from "non-corresponding sound correspondences" in at least one branch, and that's putting it rather mildly). A small percentage of forms that deviate from regular correspondences should probably be permissible in any such hypotheses, provided there are reasonable explanations for these deviations and also that there be no other problems with the forms.

Nevertheless, violation of sound correspondences is always painful, and it is therefore quite reassuring that all of the examples provided by Campbell on p. 248 as indicative of the problem are, in fact, examples of "non-non-corresponding sound correspondences".

Thus, three of them deal with violations of stop correspondences in Indo-European: Nostratic *$b\Lambda n\d{t}\Lambda$ 'to tie, bind' yields Indo-European *$bhendh$- instead of the expected *$bhent$-, Nostratic *$bi\check{c}\Lambda$ 'to break' yields Indo-European *$peis$- instead of *$bheis$-, and Nostratic *$b\ddot{u}\Ka$ 'to bend' yields Indo-European *$bheug$- / *$bheugh$- instead of *$bheuk$-. However, any seasoned Indo-Europeanist would have immediately understood what is so wrong about the expected forms *$bhent$- and *$bheuk$-: they simply could not have existed in Proto-Indo-European, as they defy the laws of Indo-European phonotactics, in which root sequences of "voiced aspirated — voiceless" were strictly prohibited. "Violations" of sound correspondences are thus in perfect accordance here (and in quite a few other etymologies) with the inner laws of Indo-European. (Somewhat harder to explain is Indo-European *$peis$-, where one has to assume a different development in the period preceding loss of affricate: * $bi\check{c}\Lambda \rightarrow$ *$bhei\check{c}$- $\rightarrow$ *$pei\check{c}$- $\rightarrow$ *$peis$-).

Coincidentally, all of these cases are specifically commented upon in IllichSvitych's own dictionary, rather than Kaiser & Shevoroshkin's set of translated reconstructions, which seems to have been the authors' main (if not only) source of knowledge on Nostratic etymology. Had it been otherwise, the reader would perhaps have been saved from being forced to decipher the meaning of the bizarre Afroasiatic phoneme $t^1(t^2)$ (*sic!*) in the following phrase: "in **$b\Lambda n\d{t}\Lambda$ 'to tie, bind', with Afroasiatic bn (sic!; should really be $bn\d{t}$ — *G. S.*) and Indo-European **$bhendh$ 'tie'... the Afroasiatic reflex of Nostratic **$\d{t}$ should be $t^1(t^2)$..." (p. 248).

As it turns out, the graphic sequence $t^1(t)^2$ is indeed present in the translation of V. Dybo's comparative phonetic tables [Dybo 1989: 114] which the authors had access to. Unfortunately, the figure 1 in this edi-

tion is simply a misprint for a *dot* representing glottalic articulation — and the figure 2 refers to a footnote (!) on the same page which explains the reasons for variability between glottalic *ṭ* and simple *t* in observed Afroasiatic reflexes of Nostratic roots with **$*ṭ$.

By all means, it is lamentable that correspondence tables should come with misprints, but it is even more lamentable that the original — misprint-free — correspondence table in [IS I: 147] has not been consulted, from which it would have been obvious that Afroasiatic *bnṭ* is a perfectly regular and expected reflexation of Nostratic *$bʌnṭʌ$. As for the Afroasiatic phoneme $t^1(t^2)$, it makes its appearance *twice* on one page (Illich-Svitych is further castigated for comparing Dravidian *kuḍḍ-* 'small' with Afroasiatic *q(w)ṭ*) — making me wonder how it was possible not to check the correspondence tables more closely, given the obvious oddity of the notation.

Further down the line it is written: "a brief look at Nostratic forms beginning in **$*p$ reveals that both the Indo-European and the Kartvelian forms arbitrarily begin with either *$*p$ or *$*b$, but this is not regular sound change and is not sanctioned by the standard comparative method". This does not look good for Nostraticists, but it looks even worse for at least some of the more conservative Turkologists and Dravidologists, who have for ages battled with "sporadic voicing" of initial voiceless stops in numerous languages and in numerous stems in the concerned families, without, however, daring to discard the corresponding etymologies; apparently, all of them have been in the wrong, since "this is not sanctioned by the standard comparative method". In reality, the problem is limited to a tiny handful of etymologies with initial *$*p$- and correspondences in Indo-European and/or Kartvelian, where the correspondences are violated maybe two or three times (sometimes representing variation *within* daughter languages, e. g. Nostratic *$palq̇ʌ$ 'foot' → Kartvelian *perq-*/*berq-*). This is hardly a serious problem.

Closing out this list of "non-corresponding correspondences" are two further observations: (a) "For Nostratic **$*ń$ the Indo-European box lists both *y* and *n-*, but an examination of the forms beginning with **$*ń$ shows that it is arbitrary when the postulated Indo-European cognate has *$*y$ and when *$*n$" and (b) "The **$*d$

of (174) should be reflected by Uralic *t* instead of the δ that occurs in the Uralic form listed".

For (a) the true situation is such that Indo-European cognates have *$*y$ in 3 cases and *$*n$ in one case; it may be that this case (Nostratic *$ńida$ 'to tie, bind' → Indo-European *nedh-* id.) shows a sporadic irregularity, or it

may be an incorrect etymology, but most importantly, in [IS II] it is explicitly stated that all examples of Indo-European *$*y$ occur before mid and back vowels, whereas *$*n$ may have been the regular reflexation before front ones. So, the "arbitrariness" of the issue is either false, or seriously exaggerated.

As for (b), this is a straightforward error on behalf of the critic. Nostratic *küda* 'male relation' yields Uralic *küδü* 'wife's husband, husband's or wife's brother' on a perfectly regular basis, because intervocalic *-t-* is regularly reflected by Uralic *-δ-*, not *-t-*, and this time not even a misprint in the phonetic tables in [Dybo 1989] can save things, because there is none.

So much for violation of sound correspondences: I have specifically bothered to comment on all of the authors' examples here because this is the most serious accusation one can present against a hypothesis of relationship, and, as can be seen quite clearly, all of the criticisms without exception fall into three categories: (a) statistically insignificant quibbles, (b) misunderstandings, (c) errors (on behalf of the critics, not Illich-Svitych). It does not help matters, again, that the Nostratic etymological dictionary itself was never even once consulted, with Kaiser and Shevoroshkin's brief summaries of Illich-Svitych's work substituting for the real thing. Does one criticize Indo-European etymology by looking through the brief *List of Indo-European Roots* in *The American Heritage Dictionary of The English Language*?

The sober truth, of course, is that Illich-Svitych, a professional comparative linguist who was raised firmly within the rigid Neogrammarian tradition of Indo-European, simply could not imagine an approach to language comparison that would disregard or neglect the establishment of regular phonetic correspondences. Some of these correspondences may be questionable in that they are not represented by a statistically significant number of examples (these are the areas of Nostratic etymology that require further scrutiny), but virtually none of them are violated during the presentation of material, and those few that are are always commented upon. Any criticism of the "classic" Nostratic model from this angle is inarguably bound to fail.

(e) **Short forms**. It is true — and inevitable — that some of the compared forms are monoconsonantal, including grammatical markers, pronouns, etc. So? On p. 252, some of the most "basic" comparisons are discarded, one by one, by the authors in more or less the following way:

" 'I': Uralic *mi...* 'we': Uralic *mä...* These forms for 'first person' are short..."

" 'thou': Uralic *ti*... This, too, is a short form".

" 'who, what': Uralic **ke-*... 'this is a short form'"...

...and so on, continued on pp. 254–255 with a dozen more examples of "short" forms, put under doubt for no other reason than being "short". Short they may be, of course, but the important thing is not that each single one of them is short, but that every additional link between them and their Indo-European equivalents (or equivalents from other branches of Nostratic) progressively decrease the possibility of all this being due to simple chance. Taking this evidence one stem after another as if they did not constitute part of a single system simply will not do. With equal success we could have taken the classic comparison of the Indo-European verbal paradigm and "destroyed" it in the following way: "Sanskrit *-mi* = Greek *-mi* — this is a monoconsonantal ending, possibly due to chance, involves a nasal; Sanskrit *-si* = Greek *-si* — monoconsonantal ending, possibly due to chance; Sanskrit *-ti* = Greek *-ti* — short form, possibly due to chance; accidentally similar forms can be found in other languages", etc.

Overall, this is essentially a non-argument. Not all of these "short form" comparisons are of equal quality, but as long as the correspondences work, none of them should be discarded from consideration.

(f) **Semantically non-equivalent forms**. The fact that, out of 378 original etymologies, the authors "count 55 forms (i. e. 16 percent) which involve comparisons of forms in the different languages that are fairly distinct semantically", in my opinion, is by itself almost enough to vindicate the Nostratic theory. I seriously wonder if the number of entries that are fairly distinct semantically in Pokorny's Indo-European dictionary can be ground to a halt at 16 percent. Examples given by the authors further show that they have a truly draconic understanding of "semantic distinction", if they feel uneasy about equivalencies like 'day' and 'bright, light' or 'hardened crust', 'crust', 'scab'. The only example in the given group that might cause eyebrows to be raised is 'lip/mushroom' (Nostratic **kanpʌ* 'soft outgrowth') — but all doubts will be dissipated when one considers the Indo-European part of the etymology, represented by the root **gemb-* that, in Slavic languages, yields **gǫba* 'mushroom' *and* 'lip' (in Ukrainian, for instance, even today the word *guba* has both meanings) [IS I: 291–2].

Again, the authors are taking the easy way out: instead of assessing the semantic comparisons offered by Illich-Svitych (as well as other long-rangers) from the point of view of semantic typology, they introduce a rough binary opposition of "identical / non-identical

meaning". Earlier, in p. 7.6 ("Semantic constraints"), they have already warned against excessive semantic permissiveness, correctly observing that "the greater the semantic latitude permitted in compared forms, the easier it is to find phonetic similarity, albeit fortuitous similarity, between compared forms". Yet it has not been mentioned explicitly that comparative linguistics cannot be done properly without allowing for semantic shift — a process that no language with even a very short history is free of — and that, in assessing the strengths and weaknesses of the evidence, we cannot simply lump every pair of compared items into one of two categories — "same meaning" and "different meaning". For Campbell & Poser, there seems to be no difference between a comparison like "day : light", involving a simple and very common semantic shift, "lip : mushroom", involving a rare, but typologically observed semantic shift, and (to invent an example on the spot) "rhinoceros : tablecloth", involving a virtually impossible semantic shift. All of these are simply "semantically non-equivalent".

(g) **Diffused forms**. This is the final argument: everything that cannot be explained away by the preceding factors has to be attributed to "borrowing" or, when direct and simple scenarios of borrowing are hard to construe, "areal diffusion". The usual formula here is that parallel so-and-so between Indo-European and Uralic has been "identified" as a "loan" or "probable loan" (p. 249), although the manner of this "identification" remains obscure, because most quoted sources (A. Joki, K. Rédei and others) do not so much "identify" anything as simply suggest that, since the items in question look similar, they are probably loanwords. Of course, the burden of proof, as usual, lies here on those that suppose genetic relationship: *they* are supposed to prove their point, and are looked at with suspicion until they have, so to speak, jumped through all the hoops, whereas those in favour of an "areal" explanation only have to say "possible borrowing" in order to achieve credibility.

One joint example of two items, I believe, will suffice to demonstrate why this position should be unacceptable. For the comparison between Uralic **wete* and Indo-European **wed-* 'water', the authors themselves admit that "it is one of the more attractive cases for the hypothesis" (p. 254), not forgetting, however, to add that "some identify this as a loanword". Likewise, for the comparison of Uralic **nimi* 'name' and Indo-European **(h)nom-* id. they also say that "this set... is frequently identified as a loanword" (p. 253). One of the quoted sources for both cases is [Rédei 1988], a detailed account of known borrowings from Indo-

European languages into Uralic at several chronological stages of development of both families.

Since both of the Uralic forms clearly represent Proto-Uralic, and, likewise, both Indo-European forms represent Proto-Indo-European (even Proto-Indo-Hittite, if one agrees with the special status of Anatolian languages within Indo-European), it is clear that these "borrowings" have to be attributed to the oldest layer of borrowings from one family into another (Rédei presumes the direction to have been from Indo-European to Uralic and not vice versa, but this is not really relevant for my purpose here). How many are there? Rédei acknowledges seven [Rédei 1988: 651–654]. Seven easily "identified" old borrowings from one proto-language into another that include words for 'name' and 'water'.

Certainly, cases where the word for 'name' has been borrowed are known; likewise, for the word 'water'; likewise, cases where two or more random items on the 100-wordlist have been borrowed from a single source. But in most, probably even all, such cases borrowing of such basic items has only become possible due to a concentrated "bombardment" of the recipient language by lexical items from the donor language — "bombardment" which, obviously, begins with a large number of technical and cultural terms. A situation under which two ancient languages "meet", exchange terms for 'name', 'water', 'give', and 'sinew', and then part company borders on the ridiculous, and at least requires extra proof.

Rédei's own assessment of the situation is as follows: "Die Zahl dieser Wörter ist so klein — insgesamt sieben, — dass sie eben aus diesem Grunde nicht ernstlich als Beweise für die indouralische Verwandtschaft in Frage kommen können" [Rédei 1988: 647]. But, once one considers the issue more thoroughly, it is *exactly* the fact that there are only seven such words that begs for a genetic relationship rather than contact explanation. If Indo-European and Uralic are related within the larger Nostratic phylum, with Proto-Nostratic tentatively projected for some 12,000 years BP, we should not expect a large number of forms, even in the proto-languages, that would be very close both phonetically *and* semantically. We do, however, have a large set of Indo-Uralic comparisons that have non-identical meanings or non-identical phonetic shapes (which, however, still show correspondences) — hard to explain as borrowings, but easy to explain as reflecting original relationship.

On the other hand, if what we are dealing with is a situation of intense linguistic contact between the two languages, we should be able to witness much more than seven parallels; a language that borrows 'name'

and 'water' from its neighbour, as evidenced by all reliably attested historical precedents, should have much the same relations with him as Japanese and Chinese, or English and French, or Brahui and Hindi. A stark contrast is seen here with the attested, and generally undisputed by the supporters of the Nostratic hypothesis, contact lexicon between Fenno-Ugric and Indo-Iranian languages, which includes but one item from the 100-wordlist (*śorwa 'horn') amid numerous instances of cultural lexicon (high order numerals like 'hundred' and 'thousand'; 'honey', 'to milk', 'pig', 'calf'); all of these suggest a very sensible contact scenario, within which even the word for 'horn' fits perfectly (borrowed along with other cattle-related terms). No such easy scenario can be constructed for the earlier "contacts" between Proto-Uralic and Proto-Indo-European, and, I dare say, no such scenario *need* be constructed.

Given that no researcher seems to place under heavy doubt the similarity between the Indo-European and the Uralic words for 'name' and 'water', and given that the attribution of this similarity to borrowing goes directly against all reliable evidence we have of the nature of the borrowing process, these two comparisons alone would have been enough to justify (at least) Indo-Uralic as a serious proposal, worthy of further investigation as a basic working hypothesis. The fact that this is not happening cannot be attributed to anything but bias.

When it comes to discussion of material, etymology after etymology, the presentation is rife with errors, misprints, and misrepresentations of the original text, similar to the ones already mentioned above.

For instance, on p. 255 the Uralic collective suffix -*la*, tentatively traced back to a similar-sounding Nostratic morpheme, is put into doubt: "here again one suspects that IS's reconstruction has been too heavily influenced by the Finnish forms, since Finnish -*la* is a derivational suffix meaning in some forms 'diminutive,' but mostly with the meaning 'place of,' presumably the source of IS's gloss of a collective locative". But nowhere in his work does IS suggest a "collective locative", and, in fact, the corresponding dictionary entry does not even have any Finnish forms (!); the tentative parallel includes Mari -*la* and Selkup -*la* [IS II: 14].

There is no such Dravidian form as *kw-a 'stone' in the dictionary (and never could be, since roots with such structure are unknown in Dravidian), the real form is Kartvelian (p. 258).

On the same page, Illich-Svitych is taken to task for comparing Indo-European *bher- 'storm' and Uralic *purki- 'snow flurry', with the latter form "unjustifia-

bly segmented to leave out the *ki* portion". In [IS I: 189], however, clear and numerous evidence is presented for a basic Uralic form *$purʌ$- 'to snow; snow-bank', from which *$purki$- is easily derived (and even if it is not, the root *$purʌ$- alone can suffice for the comparison).

On p. 259, Uralic *$kōja$ 'fat' is denied cognacy with Indo-European *$gweih^w$- 'to live' — for semantic and phonetic reasons. The semantic connection between 'fat' and 'live' is well-known in semantic typology (without having to go far, consider Russian *жир* 'fat', usually derived from *жить* 'to live'); as for phonetics, for some reason, the authors expect Indo-European to have palatal *k^y- here — "before the front vowel" "by IS's correspondence sets", but the vowel is not a front one, it is a back labial one, as seen in Uralic. This is a misunderstanding of Illich-Svitych's rule, according to which Nostratic *KO-, *KU- → Indo-European *K^we-. The correspondences are perfectly regular.

On p. 260, the authors make use of another misprint in Kaiser & Shevoroshkin's materials, disqualifying the Nostratic root *$küni$ 'wife, woman' because its Turkic counterpart has been erroneously glossed as Uralic. But the Turkic / IndoEuropean comparison is faultless, and there is also a solid Afroasiatic parallel that is not mentioned.

On p. 261, the parallel between Uralic *$kara$ 'thorn, conifer' and roots such as Indo-European *$gher$- 'thorn, branch', Dravidian *$kar(a)$- 'thorn', etc., is put in doubt because the authors, taking the Uralic root to represent Rédei's *$kär_3$ 'willow species', doubt the semantics. But the actual comparison, as clearly seen in [IS: 226], does not involve this root, but rather Finnish *kara* 'thorn, wooden nail', *karahka, karas* 'young fir-tree', Nenets *xārv* 'larch', etc. No semantic problem whatsoever, unless we are prohibited from comparing 'thorn' and 'fir-tree'.

On the same page, we read: (for example № 42) "K&S (Kaiser & Shevoroshkin — G. S.) give no Uralic form and have only two families represented, Indo-European *ken- 'be born,' 'young', and Dravidian *kan- 'give birth.' K&S discuss problems in vowel correspondences in this set. IS (211) **$K'anV$ 'to give birth' has three representatives, but he indicates that the first consonant and all the vowels are questionable". However, K&S do not discuss any vowel correspondence problems in the set; on the contrary, they use it to illustrate the regularity of correspondences. Also, from examples like these it becomes apparent that the authors misunderstand the meaning of capital letters in Illich-Svitych's Nostratic reconstructions: they seem to think that capital letters indicate irregularities in correspondences, but in reality they usually indicate

cases when several variants of the reconstruction are possible because the form is not found in important "diagnostic" languages. Thus, Illich-Svitych's *$Ḵ$- is supposed to mean "either *$ḳ$- or *$q̇$- since the root is not found in Kartvelian"; as representing a correspondence between Afroasiatic *qn- 'to bear', Indo-European *ken- 'to be born; young' and Dravidian *kan- 'to bear', it is absolutely not questionable. Nor is the vowel.

These examples can be multiplied, but overall, I believe, this should give a general picture of how trustworthy the "assessment" of Nostratic in the monograph under review really is. To be fair, the authors do quote a few really weak sets, but their elimination from the material would not seriously reduce the evidence. Over-generalized and oversimplified approaches to methodological issues discussed in Chapter 7; transparent bias in favour of a non-genetic solution even when the latter is more economic and reasonable; and an odd disdain for primary sources of material — all of this contributes to the predictable conclusion: "we do not accept the Nostratic hypothesis... we seriously doubt that further research will

result in any significant support for this hypothesized macro-family" (p. 264).

Concluding this section, I can only say that I find Chapter 9 of the book to be its weakest part after Chapter 7, particularly in its criticisms of macro-hypotheses in Eurasia. With Amerind, as has already been mentioned, the situation is different: here the main target is Greenberg, and since it is generally much easier to criticize "mass comparison"-based theories, and also since both authors of the book are acknowledged Americanists and obviously feel safer in these waters, this section is arguably written in a more reasonable manner. Amerind, unlike Nostratic, is a problematic grouping, primarily because very little historical work (in comparison to languages of Eurasia) has been done on these languages, and the basis for argument is much wider. Still, as long as one does not wish to insist that Amerind has been "proven" by Greenberg beyond a reasonable doubt, there is equally no sense to consider Amerind "dead", as some Americanists informally do, nor are there any reasons to ignore Greenberg's data, provided it has been properly purged from errors.

Speaking of errors, throughout the whole section on "Amerind" (as well as other sections) little or no distinction is made by the authors between errors that invalidate Greenberg's points and errors that are insignificant in comparison. On pp. 270–271 Greenberg is taken to task for mistakenly labeling languages based on the locations where they are, or were, spo-

ken, instead of their true names (e. g. 'Papantla' instead of 'Totonac', etc.). This is not good, of course, but unless the quoted forms themselves are incorrectly quoted, or unless the idioms in question are mistakenly assigned to the wrong language groupings, how is this relevant to the actual comparisons? Or consider this: "Under the set labeled 'kill' Greenberg listed Choctaw *ile* 'do,' together with Hitchiti *ili* 'kill' (both Muskogean languages), but the 'do' of the Choctaw gloss is a scribal error (cf. Proto-Muskogean **illi* 'kill'); Kimball believes the source of the erroneous 'do' is a misreading of the abbreviation for "ditto" used by Greenberg" (p. 274). If this is indeed so, it is a funny misprint, harmful for those that will want to use *Language In The Americas* as a primary source of information (which — and here I fully concur with all of Greenberg's critics — should not be done), but certainly harmless for the Amerind hypothesis. Not so with multiple mistakes in stem segmentation, indicated by Campbell & Poser; but since detailed statistics on the proportion of "mistakes that invalidate the comparison" to "mistakes that are not crucial to the comparison" (or even "mistakes the correction of which make the comparison better") are missing, it is impossible for me to reach a definite conclusion on whether Greenberg is "essentially right" or "essentially wrong" on the matter, and so will it probably be for everyone else with an unbiased approach to it.

One point that seems to constantly escape the detractors of Greenberg and his methodology is that there is only one possible way to make "Amerind", "Indo-Pacific", "Nilo-Saharan" and other macro-hypotheses founded on "multilateral comparison" make a steady retreat from the sphere of both scientific and popular discourse, never to return again: that is, to present better alternatives to Greenberg's classification. It does not suffice to demonstrate, no matter how neatly this is done, that Japanese fits Greenberg's criteria for "Amerind" just as nicely as, say, Quechua (pp. 276–279), since, regardless of this demonstration, the evidence linking Japanese to Altaic rather than Amerind is stronger both in terms of quantity and quality; in fact, it is exactly our *a priori* conviction that Japanese is definitely *not* Amerind (a conviction based on analysis of evidence, of course, not just "common sense") that makes the authors' test on the "Amerindness" of Japanese so believable. And this conviction is in no small part due to the fact that the history of the Japanese language, as well as that of the other Altaic languages, is much better studied than the historical relations between various subgroups of "Amerind".

In other words, instead of wasting endless amounts of time on "deconstructing" the research of Greenberg

— indeed, what could be easier and safer than picking on someone else's mistakes? — it would have been far more productive (although, of course, much more time-consuming) to concentrate more Americanists' efforts on the proper historical-comparative treatment of available Native American linguistic data; the amount of available reconstructions for the area is, even today, absolutely minimal compared to Eurasia (or, arguably, even Africa), and, until more proto-languages for relatively small "subgroups of Amerind" are successfully brought back to life, and both their internal and external connections are reestablished based on the comparative method, there will be no alternative to "Amerind", and not even a hundred guides on methodology will make the hypothesis obsolete. The same goes for "Indo-Pacific" and the rest.

Chapters 10 and 11 drift away from lexically-based long-range hypotheses and, instead, dedicate space to criticisms of several theories on linguistic prehistory that build on typological or areal data, such as J. Nichols' attempts to go beyond the comparative method by using methods that correlate language structure with geographic zones, or R. Dixon's theory of "punctuated equilibrium". For the most part, I concur with the authors' views on these theories, as well as their conclusion: "The approaches discussed in this chapter... are flawed, both in conception and in execution. They afford no new insights which are reliable. They... divert efforts away from more productive lines of investigation. Indeed, there is still much work to be done and much to be learned from the application of

the traditional techniques, especially the comparative method" (p. 329). Given that, as has just been mentioned above, so many language families around the world still lack a proper comparative treatment — for no reasons other than purely technical: a lack of interest and manpower — it makes sense to ask just how much we are really entitled to going "beyond" the comparative method when we are so very far from reaching its borders.

The last substantial chapter again returns us to issues of multilateral comparison, with a detailed critique of "Proto-World" as primarily espoused by M. Ruhlen and J. Bengtson in their works, mostly dating to the 1990s. There is little need to comment on it heavily, since the authors' argumentation about why this hypothesis cannot be convincing is essentially the same as in their earlier treatment of Greenberg's lesser-scale theories, and my earlier comments thus apply to the discussion in Chapter 12 as well.

I will, therefore, limit myself to discussing just one of the "global etymologies" picked upon by the

authors, in order to illustrate both what, to me, seems unjust about their criticisms and what I find reasonable. On pp. 368–370 they discuss the "Proto-World" word **kuna* 'woman', found in [Bengtson & Ruhlen 1994: 306], choosing it as a representative of the "strong" cases of the hypothesis, with the following observations and conclusions:

(a) Bengtson & Ruhlen are accused of "ignoring vowels entirely": "the target in general is a CVC(V) form where differences in the vowels among the languages compared are ignored". Furthermore, they also violate their own consonantal requirements: "For the 'K' velar-like sound, any of the following fits: *k, k', g, q, x, h, w, b, ž, ʔ, č*. For the final consonant, '*N*', any of the following fits: *n, r, m, ã, w, ʔ*, and Ø, among others. Even matches to '*KV*' alone seem acceptable. How difficult could it be to find words matching this broad phonetic target by accident?"

It has always been a great puzzle to me why, in so many works critical of long-range research, the critics cannot resist the temptation to overstate their case. While I cannot pretend to being a great fan of the multilateral comparison method, it is crystal clear to me, from looking at page 306 in [Bengtson & Ruhlen 1994], that the situation with this comparison is much more complicated than the way Campbell & Poser make it look. For starters, wherever possible, Bengtson & Ruhlen adduce proto-language reconstructions, and it is clear (although, unfortunately, not stated explicitly in the preface to the etymologies) that, when a reconstruction is available, it is the reconstructed form that they pay primary attention to and not its descendants.

Thus, to the general reader, not well-versed in the reconstructions of various proto-languages, Campbell & Poser's indication that Bengtson & Ruhlen's **kuna* 'woman' involves compared forms with initial *ž-* or *b-* may seem like the death sentence — it is not enough that they can take any *KVNV*-like form, they can take any *žVNV*-like form as well, or any *bVNV*-like form! But this is wrong. The only such forms found in the etymology are Old Church Slavic *žena* and Old Irish *ben* — both of which are well-known to be regular descendants of Proto-Indo-European **gʷen*, which, indeed, is one of the main compared entries in this global etymology. The forms have been adduced *only* because Bengtson & Ruhlen — and, I presume, Campbell & Poser as well — have no doubts that the initial consonants in them go back to a velar. Note that when Bengtson & Ruhlen start quoting "Amerind" data — for which only a tiny handful of intermediate reconstructions are available — they never even once deviate from initial velar or uvular consonants, being

quite cautious in areas which lack sufficient exploration.

Similarly, the statement that "differences in the vowels are ignored" is also, for the most part, untrue. The etymology brings together such forms as Proto-Afro-Asiatic **k(w)n ~ *knw*, Proto-Indo-European **gʷen*, Proto-Turkic **küni*, Proto-Caucasian **q(w)änV* (Proto-Dagestan **qonV*): all of these reconstructed forms show either a labial vowel or a labiovelar consonant (which is highly likely to have developed out of a labial vowel, especially in families like Afro-Asiatic and Indo-European, where vocalism has, for the most part, assumed an auxiliary morphological function). Among the various "Amerind" forms given without reconstructions, approximately 2/3 confirm to the same pattern, as well as 3 out of 4 tentative reflexations in Australian. Only Proto-Eskimo-Aleut **ʔaʁ(i)na* and two small groups of Indo-Pacific and Austroasiatic forms do not follow this tendency (note, though, that the latter two are honestly marked with a question sign).

(b) Campbell & Poser also write: "As for the glosses accepted which allow a form of this vague phonological shape to be selected as a match, all of the following are encountered among the forms listed in support of the 'woman' global etymology: 'wife,' 'woman,' 'lady,' 'mother,' 'female' (of any species), 'spirit of dead woman,' 'girl,' 'daughter,' 'maiden,' 'daughter-in-law,' 'small girl,' 'young woman,' 'old woman,' etc."

Leaving alone the issue that all of these semantic shifts are well-attested and completely unsurprising among the world's languages, we will agree with the authors that this variation really constitutes a problem. Nevertheless, the true scope of it remains unclear from the way they present this list of meanings. Close analysis of the etymology in question shows that (a) all of the Eurasian reconstructions share the meaning 'woman' or 'wife' (the semantic equivalence of 'woman' and 'wife' in Eurasia is a very common thing, with the two meanings more often represented by one common stem than two different ones); (b) in "Amerind", 'woman' is easily the most widespread meaning (18 glosses), followed by 'mother' (10 glosses), and 'girl' (9 glosses); other meanings are much more rare ('daughter-in-law' is met only one time, 'young woman' — twice, etc.).

Given these details, instead of simply turning down all of the etymology, it may be prudent to dissect it into a "tighter" and a "laxer" part. The "tighter" part would involve entries from Afro-Asiatic, Turkic, Indo-European, and Caucasian, all of which are represented by reconstructions, share the meaning 'wife/woman'

with no significant deviations, and the phonetic shape *KUN- / *Kʷ'VN-*. The "laxer" — and, therefore, more dubious — part of it would constitute Austroasiatic (only one form with a somewhat different phonetic shape), Indo-Pacific (poor representation, phonetic deviation, meaning 'mother' in half of the given forms), Eskimo-Aleut (different phonetic shape), and Australian (poor representation, widely different meanings, including 'spirit of dead woman'). Somewhere in between lies the "Amerind" data — too difficult to assess because too few reconstructions are available; nevertheless, parts of it, especially the South American forms such as found in Macro-Ge, for instance, produce a much better impression (tighter phonetics and semantics) than their North American "counterparts".

If we discard the "laxer" part of the equation, we will be left without a "global" root: its Australian and Indo-Pacific connections will be gone, its American "presence" cut seriously short, and as for African parallels, it did not have any to begin with. The rest of it, however, bound far more tightly in every respect, will have to be taken more seriously by specialists. In particular, limiting ourselves to "tighter" parts of the data makes Campbell & Poser's joke comparisons with Spanish forms like *conyuge* 'wife, spouse', *cuñada* 'sister-in-law', *china* 'girl, young woman', etc., useless, as they would obviously have to be counted as belonging to the "laxer" part — less reliable phonetics, more widely divergent meanings, and, above all, no reconstruction.

The bottomline here is that there might be plenty of wheat hidden in the chaff of "global etymologies", if one is willing to take up the task of separating the two. But once again, none of this is evident from Campbell & Poser's assessment. Being too intent on dismantling the notion that lexical evidence can tell us something about "Proto-World", they fail to notice that there are quite a few other interesting things that it may be able to tell us. I, like most members of the Moscow school of comparative linguistics, prefer to retain an agnostic position on the issue of "Proto-World", believing that only time — and a lot of hard work — will be able to tell us whether it existed or not, and if it did, whether it can be inferred from existing data. But "Proto-World" is one thing, and accumulating lexical evidence about potential large macrofamily groupings — such as, e. g., 'Borean', the current "working codename" for a supposed super-superfamily uniting the four large macrofamilies of Eurasia (see [Gell-Mann, Peiros, Starostin 2009] for more on this issue) — based on realistic phonetic and semantic criteria with respect to known and recon-

structed history of the families in question is another. Unfortunately, Campbell & Poser refuse to distinguish between the "permissive" and the "restrictive" approaches to such comparisons, thus discouraging the potential scholar to engage in any of these activities regardless of their nature.

Bringing this (already overlong) review to its end, I must say that *Language Classification: History & Method* more or less justifies the *History* part of its title, yet could do a lot better with the *Method* part. Positive aspects of the book include the narrative of Chapters 2–6 (stripped from its teleology), recounting the history of historical linguistics; the successful defense of the comparative method from newer (or older) concurrent theories in Chapters 10–11; and the debunking of various misconceptions on the historical development of language in Chapters 8, 12 and a few other places.

All of these things, however, are only tangentially related to the main purpose of the book: an announcement, for both the scholarly world and the general public, that all research — without a single exception — on long-distance relationship among the world's languages carried out so far has been equally or almost equally useless. (I say "without a single exception", because the authors' "optimistic note" on "successful cases of distant genetic proposals, cases which were once controversial, but which have come to be established to the satisfaction essentially of all" on pp. 400–402 is essentially a joke: none of the families listed there, from Sino-Tibetan to Uralic to Uto-Aztecan, etc., approach the time depth usually attributed to even Altaic, let alone Nostratic, Amerind, etc.). This is achieved by first setting up an elaborate system of filters (Chapter 7), much more rigid than the ones usually set up for shorter-range hypotheses, and then testing it on a few known hypotheses — with all the cards marked in advance.

In regard to the reader who has no special interest in historical linguistics, that purpose may be achieved. However, those for whom historical linguistics is a profession will be certain to notice the many flaws of Chapters 7 and 8, such as over-generalization of issues, a preference for binary answers to much more complicated questions, incorrect understanding of proposed methods (e. g. in the section on glottochronology), inability or unwillingness to distinguish between stronger and weaker parts of a given hypothesis, and sometimes even condescending ignorance of source material, leading to mistakes in data presentation.

It is (arguably) clear to most specialists working in long-range comparison that Nostratic (or Amerind, or "Proto-World") cannot be demonstrated the exact

same way as people have demonstrated Indo-European (or Dravidian, or Semitic); evidence for these families is more scarce, and the relationship is anything but intuitively obvious. More scarce, that is, but still plentiful — so much so that the natural question to ask is "what is the necessary minimum to demonstrate it?" Had Campbell & Poser tried to come up with a detailed answer to this question, the work under review might have been infinitely more valuable. Yet not only do they not attempt to answer it, they do not even ask it. Instead, the prospective researcher, according to their guidelines, is basically stuck with no options — either the hypothesis is made to look as strong as Indo-European (in actual reality — *stronger* than Indo-European, with even the slightest breaches of regularity or potentially "sound-symbolic" parallels frowned upon), or it will be labeled "unconvincing" and discarded. In other words — a classic case of "damned if you do, damned if you don't", with no intermediate options; hardly the healthiest of possible attitudes towards a branch of linguistic science.

By now, it should have already become clear that neither "mass comparison" *à la* Greenberg, nor the more conservative approach of the "Nostratic school", advocating for careful use of the comparative method on larger time depths, can be eradicated by concentrating exclusively on the weak spots of these theories; and it is no longer reasonable to indiscriminately label the ranks of their supporters as far-reaching romantics, opposed to the sober realism and rigor of the true professionals in the field. Such a dividing line may have existed a century ago, but clinging to it today is a hopeless anachronism. Just as the more serious "long-rangers" always base their work on historical studies of "short-rangers", approaching them critically, but with respect, so is it high time that the "short-rangers" started paying more attention to what goes on in the long-range field as well, and approaching its theories from a less biased standpoint than the usual "this must be wrong". Unfortunately, *Language Classification: History & Method* postpones this objective evaluation, since one of its purposes is to discourage scholars from any such attempts. This is regrettable.

References

Bengtson & Ruhlen 1994 — John D. BENGTSON & Merritt RUHLEN. Global etymologies // *On the origin of languages: studies in linguistic taxonomy*, ed. Merritt Ruhlen. Stanford: Stanford University Press, pp. 277–336.

Campbell 1998 — Lyle CAMPBELL. Nostratic: a personal assessment // *Nostratic: sifting the evidence*, ed. Brian Joseph and Joe Salmons. Amsterdam: John Benjamins, pp. 107–52.

Dybo 1989 — Vladimir A. DYBO. Comparative-phonetic tables // [Shevoroshkin 1989], pp. 114–21.

EDAL — Sergei A. STAROSTIN, Anna V. DYBO & Oleg A. MUDRAK. *Etymological dictionary of the Altaic languages.* Leiden: Brill, 2003.

Foley 2000 — William A. FOLEY. The languages of New Guinea // *Annual Review of Anthropology* 29.357–404.

Gell-Mann, Peiros, Starostin 2009 — Murray GELL-MANN, Ilia PEIROS, George STAROSTIN. Distant Language Relationship: The Current Perspective // *Journal of Language Relationship*, v. 1, pp. 13–30.

Hale 1997 — Kenneth HALE. Book review article: Campbell, Lyle (1997). *American Indian Languages: the Historical Linguistics of Native America* // *Mother Tongue* 3, pp. 145–58.

IS — В. М. ИЛЛИЧ-СВИТЫЧ. *Опыт сравнения ностратических языков (семито-хамитский, картвельский, индоевропейский, уральский, дравидийский, алтайский): Введение. Сравнительный словарь.* М.: "Наука", Главная редакция восточной литературы [An attempt at a Comparative Dictionary of the Nostratic languages (Semito-Hamitic, Kartvelian, Indo-European, Uralic, Dravidian, Altaic). Moscow: "Nauka" publishers]. V. 1: 1971; v. 2: 1976; v. 3: 1984.

Kaiser & Shevoroshkin — Mark KAISER & Vitaly SHEVOROSHKIN. Nostratic // *Annual Review of Anthropology* 17.309–30, 1988.

Rédei 1988 — Károly RÉDEI. Die ältesten indogermanischen Lehnwörter der uralischen Sprachen // *The Uralic languages: description, history, and foreign influences*, ed. Denis Sinor. Leiden: Brill, pp. 638–64.

Renfrew, McMahon, Trask 2000 — Colin RENFREW, April MCMAHON, & Larry TRASK (eds.). *Time depth in historical linguistics.* Cambridge: McDonald Institute for Archaeological Research.

Shevoroshkin 1989 — *Explorations in language macrofamilies: materials from the First International Interdisciplinary Symposium on Language and Prehistory*, ed. Vitaly SHEVOROSHKIN. Bochum: Brockmeyer.

Starostin 1999 — Sergei A. S<small>TAROSTIN</small>. Comparative-historical linguistics and lexicostatistics // *Historical linguistics and lexicostatistics*, ed. Vitaly Shevoroshkin and Paul J. Sidwell (AHL Studies in the Science and History of Language 3.) Canberra: Association for the History of Language, pp. 3–50.

Thomason & Everett 2001 — Sarah G. T<small>HOMASON</small> & Daniel L. E<small>VERETT</small>. Pronoun borrowing // *Berkeley Linguistics Society* 27.301–15.