

# Japanese, the Altaic Theory, and the Limits of Language Classification

著者	GEORG Stefan
journal or publication title	Perspectives on the Origins of the Japanese Language
volume	31
page range	429-449
year	2003-12-26
その他のタイトル	日本語・アルタイ語族説・言語分類の限界
URL	<a href="http://doi.org/10.15055/00005290">http://doi.org/10.15055/00005290</a>

# Japanese, the Altaic Theory, and the Limits of Language Classification

Stefan GEORG

*University of Bonn*

**Keywords** : Altaic, language classification, comparative method, Nostratic

The question of the origin of the Japanese language is a question of language classification. In the following, I will try to address some of the more fundamental theoretical questions surrounding the problem of “language classification” — what is it, why is it important, how can it be done, and are there any limits beyond which even the most sharp-witted linguistic approach may not go?

Furthermore, since the name of the putative linguistic stock called “Altaic” is probably the one most often mentioned when the question of the origins of Japanese is brought up, I will give an overview — a somewhat *personal* overview to be sure — of the “Altaic theory”, its history, its problems, and, last but not least, its prospects as I see them.

I am not a Japanologist, so I am not sure whether my largely continental view of the Altaic problem will be of any interest for the audience this volume may attract. Yet, I am somewhat confident that the answer may be yes, since the question whether Japanese (and/or Korean) are members of any language family called “Altaic” presupposes the question whether such an entity may be safely assumed in the first place. Further, neither Japanese nor Korean were instrumental in the initial formulations which developed into what was to become “The Altaic Theory”.

The “initial” Altaic languages were Turkic and Mongolian. Tungusic got added to the Altaic family somewhat later (when Tungusic linguistic data other than Manchu became available). In its infancy, the “Altaic” theory was also riddled with a difficult and obscure mixing of (let’s call them *proper*) Altaic languages and Uralic languages.

Whoever tries to find the *beginnings* of the Altaic theory, or at least the earliest scholarly attestations, of the view that some (if not most) languages of Inner Asia (East of Indo-European and North of Sino-Tibetan) may have “sprung from some common source”, might easily run into some kind of frustration. Reading the publications of the most noted European orientalists of the late 18th and early 19th centuries (Klaproth, Castrén, Schleicher, I. J. Schmidt, to name but a few) reveals not much more than the somewhat eerie impression that this idea somehow “was there”, without anyone who can be credited with the fame of a true “founder” of this theory, like William Jones and Franz Bopp for Indo-European or Sajnovics and Gyarmathi for Finno-Ugric.

Instead of such a founder, we find, somewhat surprisingly, that the *critique* of

Altaic seems to be older than any coherent statement of the theory itself; in the late 18th century Peter Simon Pallas has already remarked on the opinion that Turkic and Mongolian may be related:

Die nicht sparsamen Worte, welche die tatarische Sprache mit der mongolischen gemein hat, und deren sich viele in der türkischen Sprache nicht antreffen lassen, können theils einer uralten Nachbarschaft und Gemeinsamkeit beyder Nationen, die wohl niemand leugnen wird, zugeschrieben werden, theils sind es die Spuren, welche die herrschende (sic) Mongolen bey den unterjochten Tataren hinterlassen musten (Pallas 1776, 2-3).

‘The sizable number of words, which the Tatar (= Chaghatay) language has in common with the Mongolian language, many of which are not found in Turkish (=Ottoman), may indeed be due to a century-long symbiosis of both nations - which nobody will deny - or they may be traces left by the ruling Mongols on the (languages of) the subjugated Tatars’

Further, in 1820 the French orientalist Abel-Rémusat polemicizes against any genetic ties between these languages (he adds not only Manchu but also Tibetan to the already attacked, but still non-existent, theory). The question is, what are all these scholars fighting against at a time where virtually nobody actually proposed anything in the way of an “Altaic theory”, let alone one backed by linguistic data?

If we dig a bit further, another surprise is awaiting us: the core idea of Altaic, insofar as the relationship of Turkic and Mongolian is concerned, had not been formulated by European scholars, but it came right from the heart of the Altaic world, from the Middle Turkic (Chaghatay) historian, Abu’l-Ghazi Bahadur Khan, who already in 1659 declared the Turks and the Mongols as close cousins, who have indeed “sprung from some common source”. As it happens, his work, the *Shajara-i Tarakime* “The genealogy of the Turks”, was one of the earliest Turkic historical works, and one which found a wide circulation in the West. (By the end of the 18th century, there were numerous German, French, Latin, Russian, and English translations in print, so this idea - the historical unity of Turks and Mongols — was, so to speak, “in the world”, finding only followers who did not think it necessary to pronounce this fact, and consequently only the critics found any reason to raise their voices.)

The first European proponent of an early version of the Altaic theory was Wilhelm Schott, who started in 1836 to publish numerous papers. In these studies he tried — for the first time — to assemble linguistic underpinnings of the Altaic theory, which had not yet found its name. After examining the overall typological makeup of these languages (Turkic, Mongolian, and Tungus/Manchu), he writes “Essay on the Tatar Languages”

(1836) in which he states:

Eine so innige Geistes-Verwandtschaft ist gewöhnlich mit leiblicher Verwandtschaft gepaart, d.h. alle durch ein so enges Band der Analogie verknüpfte (n) Sprachen pflegen mehr oder weniger divergirende Aeste eines gemeinsamen Stammes zu seyn.

‘Such an intimate mental relationship is usually accompanied by physical relationship as well, this means that languages tied together by such intimate analogy are generally also the branches of a single tree.’

Today, no serious linguist could subscribe to this statement; we have learned that there is no such thing as invariant typological traits of languages which are never subject to change. Moreover, we have learned to use the instrument of areal linguistics to find out that prolonged language contact can, and does, bring about the degree of typological similarities that Schott was only able to explain by common descent.

But in his later writings (his last paper on the subject is published in 1887), Schott makes an effort to find material commonalities of these languages as well. He in fact discovers many, some of which still stand up against scrutiny. However, what is most disturbing about Schott’s work, is that he never took any notice of the tremendous factual and methodological progress Indo-European linguistics underwent during his active years, even though one of his immediate colleagues in Berlin university was none other than Franz Bopp, the founder of Indo-European linguistics. In his writings, sound laws, let alone the neogrammarian concept of their exceptionlessness, simply do not occur: mere similarities rule supreme. Thus, Schott, though his works still are worth the trouble of being read, still belongs to the *prehistory* of Altaic linguistics, rather than to its actual *history*.

The latter surely begins with Schott’s contemporary, the Finnish traveller, linguist and, I dare say, genius, Mathias Alexander Castrén. To be sure, Castrén did not actually publish much on strictly *comparative* Altaic linguistics — his life simply was not long enough for that. Instead, he concentrated on the description of the living languages of Siberia. His dissertation, defended in 1850, was devoted to the systematic comparison of the personal pronouns and, as far as the languages have them, personal verb endings of a language group, which he for the first time dubbed *Altaic* in his diaries from 1844.

But “his” Altaic was both more and less than “our” Altaic. Japanese and Korean were lacking — doubtless because few, if any, scholars in Europe had more than a faint idea of both languages at that time. Also, it did include the languages which we now call Uralic: Finno-Ugric and Samoyedic. Of course, the similar typological makeup of both Uralic and “Altaic” fooled Castrén as well — to a degree. It took almost a century to free the study of Inner and Northern Asian languages from this “lumping” terminology: “Altaic” was used synonymously with “Ural-Altaic” well into the 20th century. The

famous Slavicist and Indo-Europeanist August Schleicher is widely recognized as the first scholar to take William Jones' "common source" seriously and to try to *reconstruct* it in broad outline, and who introduced the familiar asterisk, with which we still mark reconstructed, hypothesized proto-forms. He reserves "Uralic" for the family which still goes by that name and he alone uses "Altaic" exclusively for Turkic/Mongolian and Tungus, in a small booklet published in 1850 and written — I hope I will be forgiven that I take some pride in pointing to that fact — while he was still a young lecturer at my home university in Bonn.

But it was still a long way until the relationship of these languages was scrutinized with the methodological apparatus of the neogrammarians. Uralic linguistics was definitely quicker on this path — not a mean feat, if we come to think of the fact that Uralicists worked, and still work, mostly with recent linguistic data; there are no really old documents available in the languages of their interest. Just like Altaic linguistics it started from the "primordial soup" of an indiscriminate "Ural-Altaic", and *again*, when we try to find out who was the first scholar to declare Uralic and Altaic as unrelated, or at least too distantly related to allow a comprehensive and meaningful reconstruction of a common protolanguage, we are hard pressed to find one. The reason for that is simply stated: Ural-Altaic simply slipped out of fashion at the same pace with which *Uralic* comparative linguistics developed into a true science; Uralic specialists formulated strict sound laws for Finno-Ugric (and later for Finno-Ugric and Samoyed), reconstructed the proto-language(s), subclassified the family, everything in accordance with the proper methodology of comparative linguistics.

While many of the protagonists did pay at least lip-service to the "greater" family, which would have included our "Altaic", they simply did not find Mongolian or Turkic data useful enough for the elucidation of the prehistory of, say, Finnish or Hungarian. Ural-Altaic, thus, simply faded away. However, it did not disappear from encyclopedias (to the present day, one has to add), but it simply ceased to attract Uralic specialists (the former "Altaicists") any longer. Too unclear were the mutual relations of these languages, too vague the notion that they, too, are really *demonstrably* related: *demonstrably* in terms of sound laws and other strict methodological requirements of the art.

At the end of the 19th century, the Hungarian scholar Zoltán Gombocz and the Finnish Mongolist G. J. Ramstedt (who had served as Finnish ambassador in Tokyo before the Second World War) were among the few who tried to change this state-of-affairs by formulating, for the first time in sufficient detail, regular phonological correspondences between Turkic, Mongolian, and Tungusic, which were designed to elevate the vague theory to a new, scientific level. The critics came back from their hibernation (I mention only the decidedly anti-Altaic article by Németh 1912), but on the whole, Ramstedt's time can be described as the first real spring of comparative Altaic. He was followed by Nikolaus Poppe, who published a comprehensive "Comparative Phonology of Altaic" in 1960, where he managed to summarize the Altaic knowledge as it stood in the middle of the 20th century.

When I use the term “knowledge” here, this raises the question whether the relationship of these languages has been proven. For Poppe this was indeed the case, but the critics of Altaic spoke up again, louder than ever before. It is now, with the advent of Gerard Clauson, Gerhard Doerfer, Aleksandr Shcherbak, András Róna-Tas and others on the scene, that a whole new school seems to emerge, often referred to by its opponents as “Anti-Altaistics”. Ever since, heavy polemics have been the hallmark of the “Altaic debate”, which once got to be called the “Civil War of Comparative Linguistics”.

It was the middle of the 20th century, and the relatedness of Indo-European had been demonstrated one and a half centuries before; Uralic, Afro-Asiatic, Austronesian, even Austroasiatic and Niger-Congo were well established families, but the question whether the well-known, well-described, and long attested languages like Turkish, Mongolian, or Manchu were part of any linguistic family, let alone which one, still saw the experts in two camps who fought each other rather fiercely — and they still do forty years later. This immediately brings us to the question of why this was and still is the case: what went wrong in Altaic linguistics?

In the following, I will try to give some answers to this question. First, I will present some of the traditional anti-Altaic arguments from the post-Second World War phase of the debate. After that I will step back in history a little to introduce the approach which currently enjoys the most attention from scholars worldwide: an approach in which Japanese data play a prominent role. I will, of course, formulate some problems I have with this attempt to solve the perennial question of Altaic.

But before I do so, I must say some more principal words. It will become clear that I side more with the critics of Altaic than with its proponents. However, there are different degrees, so to speak, of skepticism: the most radical possible view on Altaic, a view which would claim that the languages in question have *nothing whatsoever* to do with each other, that *any* commonality found between them is just a *fata morgana* and therefore insignificant, is rarely, if at all, encountered today. On the contrary, most critics of Altaic maintain that these languages *do* share a great deal, not only of *similar* elements, but also of *truly shared* — i.e. historically identical lexical and morphological elements, but that these — at least their majority — is better accounted for by assuming large-scale mutual borrowing than common genetic descent. And that, therefore, the main defect of earlier attempts to do Altaic comparative linguistics is a less-than-optimal thoroughness in the attempt to differentiate secondary borrowings from true cognates. And I do mean “from true cognates” since *some* of those scholars who have to live with the label “Anti-Altaicist” *do* indeed accept the possibility that such cognates could be found, but that, before this may be done, a host of borrowings — actually several layers of them — has to be removed before this last layer may be considered as isolated. My teacher András Róna-Tas has always espoused this position — though his skepticism prevailed — while Gerhard Doerfer, for example, is one of a few scholars who explicitly deny the existence of any genetic tie between these languages. (Juha Janhunen is another one, though he recently proposed a fresh investigation of a possible exclusive Mongolian-Tungusic link:

Khinganic.) So there are quite different positions labeled as “Anti- Altaic”, much as there are different “Altaic theories” as well.

But, turning back to the Ramstedt-Poppe variant of comparative Altaic — an impressive intellectual edifice indeed — I mention, briefly, some of the major points of criticism raised against it, in order to show what the debate was about before the 1990s.

First, in terms of the phonological system of correspondences, it cannot be overlooked that it contains *gaps*, much unlike the Indo-European or Finno-Ugric systems. In other words, while in Poppe’s system there are, for example, numerous lexical correspondences involving initial *d-*, *y-*, or *j-* in Mongolian (which are matched by Turkic initial *y-*), there are *no good* examples involving Mongolian initial *m-*, *g-*, or *n-* (e.g., I cannot see that a comparison of a Turkic word for ‘hip’, *bɪkɪn* with a Mongolian word for ‘meat, flesh’, *miqan*, is really a *good* etymology). It’s not that there are *few* such comparisons, there are actually *none* (a correspondence with Zero would of course count as a positive example). The lack of such comparisons looks even more suspicious when we remember that it is exactly these phonemes which are lacking in initial position in Turkic in the first place. Such a situation is highly unusual in the case of true relationship, where ideally every phoneme of every involved proto-language will have a definable match in the other ones, demonstrable by a good number of good etymologies. Yet, this is not the case between Turkic and Mongolian; and such a situation *is*, of course, rather typical for borrowing situations, where a donor language simply could not pass elements to a recipient language containing phonemes it did not possess in the first place.

Another problem with the Ramstedt/Poppe system pertains to vowel quantity: proto-Turkic was, without a doubt, a language with distinctive vowel length, and so was proto-Tungusic. However, proto-Mongolian shows no trace of such an opposition. Now, consider the following etymological sets :

Old Turkic	Mongolian	Tungus		English
<i>āgɪl</i>	<i>ayil</i>	Solon	<i>ayl</i>	“nomadic camp”
<i>āb</i>	<i>aba</i>	Solon	<i>awa</i>	“hunt”
<i>kōk</i>	<i>kōke</i>	Ewenki	<i>kuku</i>	“blue”
<i>kādɪn</i>	<i>qadum</i>	Ewenki	<i>kadum</i>	“father-in-law”
<i>bōz</i>	<i>boro</i>	Ewenki	<i>boro</i>	“grey”

We see that Tungusic shows a short vowel, as opposed to a Turkic long vowel, but Tungusic is a language which *does* know distinctive vowel quantity, so what is going on here?

While it may be possible to rescue the situation with complicated and *ad hoc* sound laws involving accents or even tonal contours, a quite simple explanation is that these words — and of course scores more — were simply borrowed from Turkic into Mongolian, where any distinctive vowel length was cancelled — and from there passed on into Tungusic, where they had to end up with a short vowel. This scenario also reflects the geographical facts of early Inner Asia, so I see nothing wrong with it.

Another, and not just “another” but probably *the* most controversially discussed

Altaic sound law, goes by the name of “the problem of rhotacism”. It makes the small Chuvash language (spoken on the middle reaches of the river Volga) a strange Turkic language in any respect. This is one of the key languages of the Altaic debate; one of the most clear-cut sound laws *within* the Turkic family of languages is the correspondence of Common Turkic (i.e., *any* Turkic language except Chuvash or its medieval ancestor, Bulghar) /z/ to Chuvash/Bulghar /r/, and of Common Turkic /š/ to Bulghar /l/. Note the following:

Common Turkic	Chuvash	English
<i>buzāgu</i>	<i>pǎru</i>	‘calf’
<i>ékkiz</i>	<i>jěķēr</i>	‘twin’
<i>öküz</i>	<i>vǎķār</i>	‘bovine’
<i>taš</i>	<i>čul</i>	‘stone’
<i>běš</i>	<i>pillěķ</i>	‘five’
<i>kümüş</i>	<i>kěměl</i>	‘silver’

This would not bother us much, if it were not for the fact that this sound law is not confined to Turkic alone, but pertains to other Altaic languages as well. A short list of cognate words in Mongolian (Written Mongolian and Khalkha) to the first four of the above mentioned is given below (there are dozens more):

WM <i>birayū(n)</i>	- kh. <i>bjaruu</i>
WM <i>ikir-e</i>	- kh. <i>ixer</i>
WM <i>üker</i>	- kh. <i>üxer</i>
WM <i>čilayun</i>	- kh. <i>chuluu</i>

Nobody today doubts that these etymological comparisons are valid as such, but the actual historical interpretation of the phonological processes underlying them differ as widely as one can imagine.

One approach is, of course, to view these languages as related, and these words as having a common heritage, which show a proto-phoneme \*r<sub>2</sub>, as it is usually labelled. This phoneme was retained in every daughter language (I do not give Tungusic data here, but some of these etymological sets do have a Tungusic witness, as well), and *only* changed in Common Turkic, the only innovative language in this respect:

Proto-Alt. \*r<sub>2</sub> => Proto-Turk. \*r<sub>2</sub> => Bolghar \*r, Common Turk. \*z  
=> Proto-Mong. \*r

Or, the alternative favored by many, though not all, critics of Altaic: Bolghar/Chuvash is the innovative language in this scenario:

Proto-Turk. \*z => Common Turk. \*z



=> Bolghar. \*r => loanwords in Mong. with /r/

If this is what has to be assumed, all the *r-* (and *l-*) words in non-Turkic languages suddenly look like loanwords from Turkic, and from a specific Turkic language at that, namely Bolghar. Needless to say, *all* proponents of an Altaic genetic unity (including Poppe, Ramstedt, R. A. Miller, and Sergei Starostin) insist on the first scenario.

Can we solve this problem? This remains difficult, and the amount of literature addressing this vexed question is gargantuan. However, it seems to be immediately clear that any possible solution may come from loanwords — this time from foreign loanwords which had already entered Proto-Turkic, and which did possess the required input-phoneme, which then would have undergone the sound change in the direction we want to clarify.

Perhaps the Chuvash words *pir* ‘cloth’ and *vyras-erni* ‘Friday’ (lit. ‘Russian Sunday’) given below may not be entirely without significance here:

čuv. *pir*: arab. *bazz*, Greek *bússos*

čuv. (*vyras*)-*erni*: Volga Bulgarian *erni* ‘week’, Persian *ādīneh*

The meaning of *pir* is ‘linen, cloth’ and it is clearly a non-Turkic word; it belongs to a widely attested etymon, a so-called Wanderwort, the ultimate origins of which may be sought in the ancient Mediterranean world. Its ultimate source seems to be Greek *bússos*, which found its way into Arabic *bazz*, from where it came to the Bolghars of the Volga, where it, and this is of course important, underwent the sound law  $z > r$ . Another possibly interesting word is the Chuvash term for ‘Friday’, *erni* (today only used in the compound *vyras-erni*, lit. ‘Russian Friday’, i.e. ‘Sunday’). This is also a loanword, this time from Persian *ādīneh*, with a spirantic *d*. Not *z*, to be sure, but we are lucky enough to possess a very precious source, the dictionary of the Karachanid prince Mahmud al-Kashgari (without doubt the first Turkologist). In his *Diwan Lughat at-Turk* of 1073 he, in words specifically labelled as “from the Bulghar dialect”, writes *z* where Old Turkic sources have *d*. If we understand Mahmud here correctly we have two clear loanwords which entered Bulghar at a time when it still had *z*. This, of course, would strongly favor the second of two alternative scenarios given above for \*r<sub>2</sub>.

It is of course possible to dispute every single one of these points, and I do not want to create the impression that I think the Altaic theory is over and done with just because of the weak points I just mentioned. But these points do, I hope, show at least two things, regardless of whether or not you are able to subscribe to my musings on “rhotacism”.

First, that there *are* after all, a large number of loan-words the three traditional Altaic languages have exchanged over the centuries, and that their number is greatly underestimated in the Ramstedt-Poppe version of comparative Altaic. And second, that this approach, notwithstanding all its defects, premature conclusions etc., *does* after all

deserve to be called scientific. It *is* based on the assumption that only the observation of rigid and regular sound correspondences may allow us to call a relationship proven, that surface similarities (let alone mere typological isomorphisms, etc.) of randomly chosen lexical items (and, to be sure, also of not-so-randomly chosen ones, e.g., the so-called basic core vocabulary) mean nothing for the serious comparativist,

With Gombocz, Ramstedt and Poppe Altaic linguistics begins its life as a science. From then on, it was no longer the playground of amateurs who — up to this day, of course — never cease to compare, say Turkish and Quechua, Basque and Armenian (bypassing the latter's Indo-European affiliation, of course) and over and over again Japanese with Sumerian. On the contrary, the mentioned scholars not only based their work on the firm ground of the neogrammarian comparative method, indeed, they did one thing which, to my way of thinking, can hardly be overestimated as a prerequisite of any scientific activity: they assembled a sizable number of facts, which manifestly are *in need of a scientific explanation*, and which seem to be amenable to such an explanation in a meaningful way. Whether the “yet-to-be-found”, or better, “yet-to-be-agreed-on” explanation will be more along the lines of common genetic descent or more of the areal linguistic type, is of secondary importance in this respect. I return to this important criterion of facts *in need of scientific explanation below*.

In the meantime, I have to keep track of the most important further development in Altaic linguistics: the revised Altaic theory as first espoused by the renowned Sinologist Sergei Starostin of Moscow in Starostin (1991). This approach differs from previous ones in many ways. One of its major differentiating traits is the systematic inclusion of Korean and Japanese in his concept of Altaic. Within the Ramstedt-Poppe model only Korean was systematically included, and Roy Andrew Miller is well known for his “Japanese and the other Altaic languages approach”. Starostin (1991) is also methodologically quite strikingly different from earlier traditions in the field, and to assess its methodological underpinnings, it is necessary to leap back in time a bit.

The history of attempts to classify the languages of mankind knows distinctive phases - fashions, if you want - which may be characterized as systolic and diastolic, respectively. I borrow this term from medicine, where “systole” refers to the contraction of the ventricles of the heart, and its opposite, diastole, denotes its relaxation and expansion. Actually it was Johann Wolfgang von Goethe, who borrowed this metaphor and used it for many purposes, so I may be allowed to steal it from him here.

To give a concrete example: the first historical phase of Indo-European comparative linguistics was characterized by a certain degree of *diastolic* (expanding) optimism. Scholars had yet to find out where the Indo-European family ends, and no less a figure than Franz Bopp experimented with adding South Caucasian (Kartvelian) and even Austronesian to his new discovery. This phase is, then, soon followed by a *systolic* (contracting) phase, in which scholars concentrated on secure findings, abandoned speculation and were busy to formulate these secure findings in ever increasing degrees of precision. The same may be said for “Ural-Altaic”. In Schott's work nearly all languages

of Northern and Central Eurasia form the object of its study, but at the beginning of the 20th century Uralic scholars had narrowed their focus on a rather small fraction of all these languages, but languages which turned out to be truly related and which repaid scholar's efforts by revealing more and more fruitful results of intensive linguistic work.

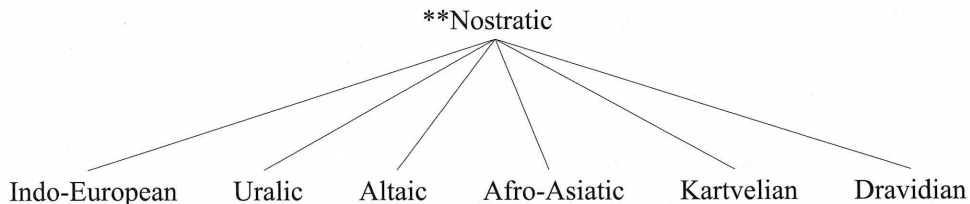
And now, at the present time, there are ample signs that the science of language-classification seems to be right in the middle of another diastolic phase. Of course, I am referring to the so-called "Nostratic" hypothesis, and before I finally turn to Sergei Starostin's Altaic I have to dwell on this for a few moments. The credit of being the founder of "Nostratic" usually goes to the eminent Indo-Europeanist Holger Pedersen, who writes as early as 1903:

Sehr viele Sprachstämme in Asien sind zweifellos mit dem Indogermanischen verwandt; (...) Ich möchte alle mit dem Indogermanischen verwandten Sprachstämme unter dem Namen 'nostratische Sprachen' zusammenfassen.

'Many linguistic stocks in Asia are doubtlessly related to Indo-European, (which) I would like to subsume under the name "Nostratic Languages"' (Pedersen1903, 561)

Pedersen himself did not elaborate this idea any further than that, but more than half a century later, a young and gifted linguist in Moscow, Vladislav Illich-Svitych, took up the challenge and compiled nothing less than a bulky comparative dictionary of the "Nostratic" languages - which from then on are known as a super-family consisting of six Old World linguistic stocks: Indo-European, Uralic, Altaic, Kartvelian, Afro-Asiatic, and Dravidian:

The Nostratic Macrofamily after Vladislav Illich-Svitych (1971):



Again, it has to be said, that Illich-Svitych *does* subscribe to the methodological postulates of the neogrammarians, that he *is* working to set up exact sound-correspondences, and that he, consequently, is not guilty of the largely methodless omnicomparativism of linguistic seven-day wonders like "Maya as Altaic" or "The Incas spoke Dutch" type attempts which every linguist constantly finds cluttering his mailbox.

It is, thus, at first glance, a *scientific* approach and it does merit a second look. I cannot go into the discussion of all the ramifications, problems, pitfalls, and possible merits of an endeavor as huge and encompassing as Nostratic here, but I may, rather summarily, mention that Nostratic *does* have its followers today — well-reputed linguists among them — but that it also meets with, sometimes fierce, criticism and overt rejection.

The more frequently heard objections against single Nostratic etymologies include:

- Nostratic linguists compare and reconstruct only words. Comparative morphology is greatly underrepresented in their work. Considering the major breakthroughs in language classification (the discoveries of Indo-European, Finno-Ugric, Uralic, Afro-Asiatic and many more are cases in point), the investigation of shared *morphology* is much more decisive for any proof of genetic relationship than any amount of lexical material could be.

- Frequent — too frequent — factual mistakes pertaining to the languages compared. To be sure, it is a bit unfair to criticize Illich-Svitych for that, since he almost single-handedly tried to detect an underlying system in the bewildering array of data from six distinct language families. *But*, unfair or not, absolute data accuracy is of course indispensable in order to convince the scientific community at large.

- Similar to the previous point, we often find single language items taken as representative for a whole family, without due observation of their history. This is true even when the history happens to be known: e.g., we cannot compare Nanai (Tungusic) *nasal* ‘eye’ directly with other words for ‘eye’ in other Altaic languages (like Mongolian *nidiin* and others), because the Tungusic sound laws simply do not allow us to take its initial *n-* as old. Had there been any initial nasal in this word in Tungusic, it would have been retained in other languages of the family as well, but all other Tungusic languages point to a proto-form *\*ya:sa*. The Nanai *n-*, then, can be very easily explained as due to an analogical influence from other eye-related words, like *ñamu-* ‘to weep’ and others.

- There are also wrong internal segmentation of morphologically complex words. As a simple example, I cite the following example, again from the Altaic part of Nostratic: we cannot meaningfully compare the second syllable of Turkic *yag-mur* ‘rain’ with Middle Korean *mul* ‘water’ to obtain a proto-Altaic ‘water’ - word. I cannot judge the reality of the Korean word, but the Turkic item, interpreted as ‘falling water’ by Starostin and others, is wrongly analyzed: first of all, *yag-* does not mean generally ‘to fall’, but only to fall when said of natural precipitations, i.e., ‘to rain’, ‘to snow’, ‘to hail’ and the like; you simply cannot, in Turkic, ‘yag- off your horse.’ Second, a compound like that is typologically alien to Turkic. And, last but not least, - *mur* is a fairly widespread Turkic suffix, denoting the result of a verbal action, *yag-mur* consequently simply ‘that, which rains, or drops from above’ Therefore, there is no falling *water* here, and consequently no Altaic water-word. Mistakes like this are aptly subsumed under the

label of ‘root-etymologies’, i.e., etymologies which only take account of one part of a given lexical item, hopefully the root (this Turkic example, however, shows that sometimes, when there is more than one element to choose from, the etymologist hits the wrong element and mistakes it for the root).

Some Nostraticists, who are fond of citing the Pedersen quote I gave above, would have been better advised if they had read a bit further, since Pedersen immediately goes on:

‘Bei dem Nachweis der Verwandtschaft der nostratischen Sprachen müssen nicht nur alle Wurzeletymologien und überhaupt alle etymologischen Spielereien fern bleiben.’

‘In the process of proving the relationship of the Nostratic languages all root-etymologies and all kinds of etymological games must be strictly avoided’ (Pedersen, *ibid.*).

While I hope that readers agree with me, and with Holger Pedersen, that mistakes like these should be avoided in linguistic comparisons at all cost, I am less sure whether the majority of readers will be able to share my opinion on a further methodological problem inherent in comparisons of this kind. This closely touches on the question of *why* we should try to do historical-comparative linguistics in the first place, why we should try to identify language families and *reconstruct* their proto-languages. I am not one of those politicians who keep asking scholars in the humanities what their work contributes to the gross income of the nation, or to national health, or whatever, but *within* the framework of the language sciences this is of course a legitimate question, and it is quite easily answered for traditional disciplines like Indo-European, Uralic, or, say, Austronesian linguistics. The recognition, and the comparative reconstruction of Proto-Indo-European, to take an example, is able to furnish nothing less than *explanations* of concrete, attested (not reconstructed) data which we find in the single languages. Of facts which are *in need of a scientific explanation*, since they often consist of irregularities, not explicable on the basis of the attested languages alone. It tells us nothing less than *why* on earth Greek, Sanskrit, or Tocharian look the way they do.

So far, I am unable to see how the Nostratic theory is able to *explain* any of the big problems of any of its constituent families, be it Indo-European ablaut, or its heteroclitic declension, or the problem of the original Indo-European verbal system, or the strange ablauting morphology of Semitic, Altaic and/or Uralic vowel harmony, and so on.

Another methodological problem, this time admittedly of a rather theoretical nature, concerns the very legitimacy of comparing proto-forms to reconstruct proto-proto-forms from them. I am not radically against this; if done carefully, this is well possible and sometimes inevitable to do. But it should be kept in mind, that proto-forms can never

have the same status as attested items: they are *constructs*, in the original sense of the word. They are products of scholars' minds, designed, set up, I even dare say *invented*, in order to *explain* concrete data. What we ask of proto-forms is that they are able to *explain* our concrete languages by being as simple as possible, and that they allow the derivation of the attested data by the assumption of an ideally small number of ideally simple rules. These rules, of course, should not contradict any known principles of language typology and how languages can change over time.

We can never know, whether such reconstructed forms are in any way close to the proto-historical *reality* of the assumed proto-language. In other words, we have no means whatsoever to know whether the 'horse' was ever referred to by a speaker of PIE as 'ek'wos'. The only thing we *can* say is that this reconstructed forms are a product of our minds, which, together with the sound laws, are able to *explain why* we find, in concrete IE languages, words like *equus*, *asvah*, *ech*, etc. Of course, it *can* be the "truth", but we have no means to assert that it is, since it is only there to explain the data we have (and, e.g., not the wealth of data we simply do not have, since the languages which might have furnished them died out unrecorded already centuries ago - a simple example is the well-known fact that we cannot reconstruct the morphology of Classical Latin from Romance languages).

Of course, it may be argued that any theory like the Nostratic hypothesis simply *has* to work with proto-languages, that it simply has no other choice. This is true enough. I do not want to raise any fundamental objection against such a procedure — actually I do this myself at times — but one note of caution is necessary: the very nature of reconstructions makes them "moving targets" so to speak. They are constructs which describe the current state of the research on a given language family. Further, it becomes obvious that the scholar who uses reconstructions for wider purposes has every right in the world to do with them whatever he or she likes, with *one* exception: there is no right to alter these forms for purposes of external comparison alone. And yet, this is what is often done. The Nostratic theory is quite full of such violations of this principle — it has been referred to as "reconstructing from above" — as my little example with Nanai *nasal* has shown.

After these lengthy excursus into more theoretical realms I think some more data are due to show you why I remain decidedly skeptical of the Moscow/Starostin version of Altaic. It would be easy, and obviously less than fair, to pick out at will a dozen or so etymologies from Starostin's 1991 book, point out some mistakes and declare the matter closed - in a way I did this above, but there I wanted to illustrate some typical mistakes. I will not choose the etymologies here; instead I let Starostin choose which etymologies should undergo my short test. I use the following criteria to single out my test candidates:

- The words will be from the well-known Swadesh 100 word list of basic core vocabulary. Personally, I do not tend to lay as much emphasis on basic core vocabulary as Starostin does; so-called basic vocabulary may get borrowed, just like everything else in language, but Starostin's theory emphasizes the

importance of this list greatly, so I will follow Starostin here for argument's sake.

- Since 100 words are too many to discuss here, I follow Starostin, who himself singles out **35** words from the Swadesh list, which he, following the Sinologist Jakhontov, views as the diachronically most stable lexical items on a world-wide scale. Again, I'm not necessarily convinced that this is correct, but, for argument's sake, I will follow Starostin here.

- For the purposes of this study 35 words are still too many, so I single out **eleven** items which fulfill the following additional criteria: they are not only supported - in Starostin's theory - by a subset of the five languages he views as Altaic (i.e., Turkic, Mongolian, Tungusic, Korean, and Japanese), but by *all five or* at least by four of them.

These, then, are the most widespread Altaic etyma in the whole of Starostin's book.

Since it may be justified to have a closer look at exactly these items to get a view of the whole theory, here is the list:

*Altaic	Turkic	Mongolian	Tungusic	Korean	Japanese	English
* <i>p'ieñV</i>	(j) <i>in-čik</i>	<i>ja-sun</i>	<i>pén-ɣen</i>	<i>s-pjé</i>	<i>pəniá</i>	bone
* <i>niā</i>	<i>jál</i>	<i>nidün</i>	<i>niā-sa</i>	<i>nún</i>	<i>màiN</i>	eye
* <i>sēri</i>	<i>ser'</i>	<i>seri-</i>	<i>sā</i>	<i>sari-</i>	<i>sír-</i>	know
* <i>birV</i>	<i>bī</i>	<i>būri</i>	-----	<i>pírí-</i>	<i>pitə</i>	one
* <i>tiōla</i>	<i>dāl</i>	<i>čilayun</i>	<i>žola</i>	<i>tōrh</i>	(d) <i>isi</i>	stone
* <i>mūri</i>	(-mur)	<i>müre-n</i>	<i>mū</i>	<i>mír</i>	<i>mí(-n-tú)</i>	water
* <i>gia-(wu)</i>	<i>nV</i>	<i>jaɣu(n)</i>	<i>gū</i>	<i>nú-</i>	<i>nV</i>	what
* <i>dila</i>	<i>jil</i>	<i>žil</i>	<i>dilačā</i>	<i>tolč</i>	<i>təsi</i>	year
* <i>siünV</i>	<i>sön-</i>	<i>sönü-</i>	<i>sī</i>	-----	<i>sín-</i>	die
* <i>p'örV</i>	<i>ört</i>	<i>för-de-</i>	-----	<i>pír</i>	<i>pə-i</i>	fire
* <i>siuwr</i>	<i>sirkä</i>	<i>sirke</i>	<i>sūra</i>	-----	<i>síram(u) i</i>	louse

I have copied this list strictly out of Starostin's books, with all diacritics as found there. It goes without saying that many single items (all of them are to be read as proto-forms) will be reconstructed differently - sometimes dramatically differently - by specialists on the single languages. Some cases in point I will mention myself, others, as far as Korean and Japanese are concerned, will doubtlessly raise your suspicion without my intervention. I will discuss these items as briefly as possible.

Let me begin with EYE, which I have to reject completely: I have already mentioned the mistake underlying the reconstruction of a proto-Tungusic word with an initial nasal here. Only Nanai has this nasal and all other Tungusic languages point to a proto-form \**ya:sa*, in Nanai a language-internal analogy was at work. As for the Mongolian form, it is interesting to note that, in connection with a different etymology, Starostin prefers to reconstruct \**nin-diün* (obviously in order to be able to compare it with Korean

\**nún*); here, we obviously have to depart from something like \**nil-dün* (there is no doubt that - *dün* is a suffix). However, this is wrong. Another Mongolian word, clearly etymologically related to *nidün*, *ni-gur* ‘face’, shows that the Mongolian root does not have a final consonant. Moreover, the combination -*ld-* would have remained stable in all Mongolian languages. For Turkic, I will not criticize Starostin’s reconstruction of -*l<sub>2</sub>*- here. I do not believe that it is correct, being an adherent of reconstructing -š- in these cases, but I will let this pass. Much more disturbing is the meaning of this word, it is ‘TEAR’ not EYE. Sure enough, tears are intimately connected with eyes, but there is more: the original meaning, widely accepted in Turkology, of this proto-Turkic item is: ‘moist, wet’. For example, in Mahmud al-Kashgari’s dictionary it is explicitly given as the opposite of ‘dry’ — hence also the widespread meaning ‘fresh’. On the Japanese word, I will say only so much that the correspondence of *n-* to Japanese *m-* is licensed in Starostin’s system only under the condition that a further nasal follows, but his own Japanese data do not support this: sure, there is Hateruma *min* ‘eye’, but, so I learn from the experts, this is an innovation, rather than an archaism. This leaves little room to accept any detail of this etymology.

WATER is equally garbled: I have already mentioned that Turkic \*\**-mur* is a complete lexical ghost; there is no such thing in Turkic. Mongolian *müren/mören* means ‘river’, admittedly rivers are watery things. With the Tungus etymon I have two phonological problems: first of all its vocalism - and the vocalism is the overall weakest point of the whole theory. Starostin does set up vocalic sound laws, but they are rarely ever observed, especially when he seems to like an etymology for different reasons: Northern Tungusic Ewen forces us to reconstruct a long *ö* here, rather than *u*:. Furthermore, there is, obviously, a final - *r* missing in the Tungusic form. For this, Starostin offers a sound law: -*r* is lost stem-finally, when originally an (equally lost) high vowel followed. Unfortunately, I cannot see such a high vowel in the data offered, not even in the semantically remote Mongolian word. Since this is clearly the case, the Tungusic word cannot have lost a final - *r*. The comparison with Middle Korean *mul* can also be dismissed. For Japanese, so I learn, PJ \**me-* has to be reconstructed, which turns the vocalism of the whole etymology into a complete mess. Too many problems for a valid etymology, I dare say.

The semantics of ONE, as far as Turkic and Mongolian are concerned, is hardly acceptable: Mongolian *büri* means ‘each, every, all’, which is too far-fetched to match ‘one’. The Korean-Japanese comparison, however, is acceptable.

For WHAT the problems are manifold. First of all, I cannot see any reason why Starostin reconstructs an unspecified vowel here for Turkic: the proto-form is firmly *ne*. Maybe the reason is that this vowel is not really the one needed for the comparison according to his own table of vowel correspondences (because we could not expect Mongolian *a*, then, nor Tungusic *ü*, much less Korean *u*). Furthermore, it must be noted that the Turkic pronoun is structurally very odd, it actually is the *only* proto-Turkic etymon with an initial nasal. Such nasals may only arise through regressive assimilation



by a following second nasal. This is lacking everywhere in Turkic, but the very fact should alert us to the strong possibility that this Turkic pronoun — like many pronouns in nearly all language families of the world — may have a much more complicated history, a history, however, which we might never be able to uncover. For the Japanese etymon, A. Vovin has recently supported a completely different etymology, an exciting proposal, as I want to say, and one which points into the direction of Austronesian, rather than Altaic.

DIE is no less problematic: the Turkic and Mongolian words probably do belong together, they both mean ‘to go out (of a flame)’. However, the Tungus word cannot belong here. First of all there is the issue of meaning, which is not DIE, but ‘to put (a fire) out’ (forms in single Tungusic languages which mean ‘to go out’ contain the Tungusic passive marker *-bu*). And, even more important, the vocalism is unreconcilable with the other words. Starostin is aware of that and knows how to get around this problem: the original vowel was neither *ɪ* nor *ö:* but - a diphthong: *-iu(:)*. Fair enough, but a long search in his etymological data revealed *this* example as the *only* one illustrating the badly needed sound law, which, then, is of course no sound law at all, but a completely ad hoc “solution”. I have to leave the question of an exclusively Japanese-Tungusic etymology open.

I can be brief on FIRE. The Mongolian form is a lexical ghost: there is *ör*, but it means ‘dawn, morning’, and it does not show, in the necessary Middle and Southern Mongolian texts and dialects, the *h-* phoneme, which is needed here to compare it to an initial *\*p-* in any other language.

For LOUSE, (actually, the Turkic word means ‘nit’, but this is fine with me), the Turkic and Mongolian words clearly belong together, though I prefer to view the Mongolian word as a loan from Turkic. However, I fail to see how the Tungusic word can belong here, again because of its vocalism. Here, as already in the case of DIE, Starostin reconstructs the diphthong *iu*; but, to be frank, I am a bit puzzled by this, since, according to his table of vowel correspondences, we should expect **Turkic** and **Mongolian** to show *o* and **Tungusic** to show *i* here, and not, as it happens, the other way round. I cannot comment on the Japanese word, I am afraid, but it does seem to contain at least one syllable too many, which the theory at least should try to account for. From A. Vovin I learn that Japanese *sirami* is actually to be segmented as *sira-* ‘white’ + *miy* ‘body’, which, if correct, helps to put this part of the etymology to rest.

BONE is, again, semantically quite garbled: the Tungusic word means rather ‘knee’, but the vocalism - again the vocalism - is incorrectly observed with respect to not only the Mongolian word but also to the Turkic one. For the latter, which means “lower leg” or sometimes ‘shin-bone’ or also ‘calf of the leg’, an entirely different etymology has been - in my view convincingly - proposed, which departs from the latter meaning and interprets the whole etymon as ‘a kind of swelling’, rather than a kind of bone.

I do not know how KNOW can be the original semantics of the Turkic and

Mongolian words; they clearly belong together - as loanwords, as I think, but their meaning is 'to feel' or 'to wake up'. The Korean word seems to be altogether a lexical ghost.

Finally, YEAR: again, the Turkic and Mongolian words clearly belong together, and Starostin devotes a lengthy passage to fight against the common opinion that the Mongolian word is a Turkic loan. It's original meaning, 'year of the 12-year animal cycle' makes it a rather unlikely candidate for being a Proto-Altaic etymon, which, according to Starostin, would have been used several millennia ago. The Tungus word means 'sun', not year. Further, it is attested *only* in Ewenki, and contains a suffix which usually only appears on verbal roots. I do not know such a verbal root, but there is, therefore, sufficient reason enough to mistrust this etymology. Again, I cannot competently comment on the Japanese and Korean words, but I have to mention that Korean *o* against Turkic, Mongolian, and Tungusic *i*, again, violates Starostin's own sound laws.

I spared one word: STONE. This is a famous Altaic etymology, one of the hallmarks of the theory, at least as far as the Western languages are concerned. I must admit that I, specifically for this word, stubbornly stick to the view that these words *are* related, but that they are quite clearly loanwords. The old alternative to Altaic, that many words, which Turkic and Mongolian have in common are simply Turkic — more specifically Bulghar-Turkic — loans in the latter, has often been attacked, sometimes simply denied. But here, in this word, we do not only see *one* typically Bulghar-Turkic feature in the Mongolian word, but *two*: the well-known *l* for common Turkic *š* *and* the initial affricate, also present in Chuvash *chul* (see above). It is not entirely clear, how the Chuvash word acquired this affricate, the position before old *\*a* is not sufficient. For this, it is often assumed that a — later lost — *i*-like vowel might have been responsible, and, for what it may be worth — this vowel *is* there in Mongolian. There simply is more than only rhotacism and lambdacism which can make a word look Bulghar — vocalism, *if* carefully observed for a change, surely is among this.

Many people instinctively reject the possibility that a word with a meaning like STONE can be borrowed at all. So does Starostin, and he summons Roy Andrew Miller to his support. Let me quote what Roy Andrew Miller has to say about that:

Why should any of these peoples ever have been under a compulsion to borrow from one another a word for something so common, ordinary, and well-known to their everyday lives as 'stone' (...) No matter how highly one may regard the Turks or how meanly one may think of the Mongolians and the Tungus peoples (...) the entire Altaic loan hypothesis quickly approaches the point of reducing itself to its own inherent internal absurdity when the semantic content of the lexical items involved is on this ordinary level, e.g., as in (this) example of 'stone' (Miller 1996: 03-04).

Well said, this is of course the end of "Anti-Altaic". The opponents of the theory

obviously do not know the first thing about freshman linguistics, Altaic exists, proven in a few words, words which sound apt to be written - I hope my penchant for cheap puns may be forgivable — in stone. Well, I may have some bad news for Miller, since I *do* know at least one historical people “primitive” enough to have felt the need to do such a thing: it was the Romans, who, for reasons only they know, borrowed Greek *petra* ‘stone’, which in the course of history, completely replaced their native term *lapis* and rules supreme in all Romance languages.

This closes my discussion of some of Starostin’s etymologies; I have not been able to destroy each and every comparison on this list, and I do not think this is necessary in order to drive my point home. Some readers may disagree with some of my criticisms, find them too harsh at times, maybe too weak at others.

What I wanted to show is that closer inspection of these etymologies *greatly reduces* the number of comparisons available for the demonstration of the relationship. And, what is more, with every single etymology which goes down the drain, the sound laws, which allegedly hold between the five Altaic language families lose one of their pillars. Some sound laws, to be sure, are illustrated by quite a number of etymological comparisons, others cannot stand losing more than one or two supporting proto-lexemes.

Again, I do not think that I have been able to destroy Starostin’s theory — let alone Altaic in general — with these critical remarks on some single etymologies. In his book, there are 386 of them, and his forthcoming Comparative Etymological Dictionary of Altaic is promised to contain hundreds more. We will see.

But in the meantime, I hope to have made some points clear. First, that the Moscow version of Altaic — and that is why I had to strain readers’ patience with my excursus into the Nostratic theory earlier — *does* repeat most, if not all, of the common misconceptions and mistakes — methodological and factual — of the Nostratic theory. And that this is the case *because* it is completely dependent on the acceptance of this — Nostratic — theory. And that, of all possible methodological mistakes — it repeats the *one* most salient, most fatal, of all vices of the Nostratic theory: that it does not search for *explanations* of problematic facts in the first place, of facts, as I have called them, which are manifestly *in need of explanation*. The existence of a Nanai word *nasal* ‘eye’ and of a Turkic word *yaš* ‘moist, tear’ simply is not a problem crying for an explanation. On the contrary: Nostratic, and the Nostratic-based theory of Altaic does something entirely different. In fact, it does exactly the opposite: while a healthy science sees itself confronted with such facts, this theory *constructs* them. Sure enough, the existence of a Proto-Tungusic *ñia-sa* EYE and a proto-Turkic *\*yal* EYE, paired with an observable correspondence of Turkic *y-* and Tungusic *ñ* in many examples and a Tungusic propensity to lose syllable-final *l* *would* constitute such a fact. But, as we have seen in this and other cases — and there are scores more like this — this *fact* is in fact a *non-fact*, a *construct* of the theory. And the biggest problem of them all is that the *explanation* of these constructed facts or problems is immediately offered by the very premise of the theory itself: that all these languages “sprung from some common source”. A theory, which has to

resort to constructing the very problems it should have been devised to solve in the first place is, *sit venia verbo*, a methodological freak.

My conclusion can only be that the Altaic question, whether Turkic, Mongolian, Tungusic, Korean, and Japanese form a valid genetic linguistic family, is as open as ever, even after the publication of Starostin's 1991 book, which has been widely, but as I hope to have shown, prematurely, acclaimed as a major breakthrough in the field.

But the work goes on. My personal view is that the alleged relationship between Turkic and Mongolian is the weakest part of the whole Altaic theory. After all I have seen, it is most likely, to my way of thinking, that most, if not all, lexical commonalities of these two families are due to contact and borrowing, rather than to common inheritance. But from numerous discussions with my colleague and friend A. Vovin I have learned that at the other end of the Altaic realm, exciting things are going on. The comparison of Japanese and Korean has been developed, and continues to be developed to ever higher levels of sophistication, and, as an outsider, I have to say that this endeavor is looking more promising than anything else in comparative Altaic linguistics. Once a clear picture of the Korean-Japanese proto-language emerges (if they form a taxonomic node), the next logical step will be to compare it to Tungusic. There is nothing wrong with that. If all the different kinds of mistakes and morphological misgivings I have been mentioning here are carefully avoided, there will be nothing to fear, but if nothing of substance emerges from this: why should we worry? The detection that a given language or family is *not* demonstrably related to a given other one — or to *any* other one — is not a nuisance, but a finding. But *if* positive results will emerge, I am sure that this work might eventually be able to fill Castrén's old term "Altaic" — which *is* after all dear to my heart — with new content and with new life. But unlike what Castrén and his contemporaries and followers thought, this possible revival of Altaic, whether it will ever cross the Eurasian continent to reach the Western languages that started this whole story or not, or whether it will reach its limits — the limits of language classification — in Manchuria, the putative homeland of Tungusic, it will originate in what we Europeans eurocentrically call the Far East, and it will bring, as it happened so often before, *ex oriente lux*.

## References

- Abel-Rémusat, Jean Pierre (1820) *Recherches sur les langues tartares, ou mémoire sur différens points de la grammaire et de la littérature des Mandchous, des Mongols, des Ouigours et des Tibétains, Tome Ier*. Paris: L'imprimerie royale.
- Castrén, Matthias Alexander (1850) *De affixis personalibus linguarum altaicarum*. Helsingfors (German translation: "Über die Personalaffixe der altaischen Sprachen" in: A. Schiefner (ed.) M.A.Castrén: *Kleine Schriften*, St. Petersburg 1862,151-222)
- Gombocz, Zoltán (1905) Az altaji nyelvek hangtörténetéhez. *Nyelvtudományi közlemények* 35: 241-282.
- Illich-Svitych, Vladislav Markovich (1971) *Opyt sravneniia nostraticheskikh iazykov (semitokhamitskii, kartvel'skii, indoevropeskii, ural'skii, dravidiiskii, altaiskii)*. *Vvedenie. Sravnitel'nyi slovar*.

Moskva: Nauka.

Miller, Roy Andrew (1996) *Languages and history: Japanese, Korean, and Altaic*. Bangkok: White Orchid Press.

Németh, Gyula (1912) Die türkisch-mongolische Hypothese. *Zeitschrift der Deutsche Morgenlandische Gesellschaft* 66: 549-576.

Pallas, Peter Simon (1776) *Sammlungen historischer Nachrichten über die mongolischen Völkerschaften*. 2 Bd. St. Peterburg: Kaiserl. Akad. d. Wiss. (Reprint: Graz: Akademische Druckund Verlagsanstalt 1980).

Pedersen, Holger (1903) Türkische Lautgesetze. *Zeitschrift der Deutsche Morgenlandische Gesellschaft* 57: 535-561.

Poppe, Nikolaus (1960) *Vergleichende Grammatik der altaischen Sprachen, Teil I: Vergleichende Lautlehre*. Wiesbaden: Harrassowitz.

Ramstedt, Gustaf John (1952-66) *Einführung in die altaische Sprachwissenschaft I-III*. Helsinki.

— (1914-16) Zur mongolisch-türkischen Lautgeschichte I-III. *Keleti Szemlé* 15: 34-150: 16: 66-84.

Schleicher, August (1850) *Die Sprachen Europas in systematischer Übersicht*. Bonn: König.

Schott, Wilhelm (1836) *Versuch über die tatarischen Sprachen*. Berlin.

Starostin, Sergei Anatol'evich (1991) *Altajskaja problema i proiskhozhenie japonskogo jazyka*. Moskva: Nauka.

# 日本語・アルタイ語族説・言語分類の限界

ステファン・ゲオルグ

ボン大学

キーワード：アルタイ語、言語分類、比較方法、ノストラティック言語学

この論文はアルタイ語族説をめぐる論争を簡潔に紹介している。アルタイ語族は、もともとチュルク語、モンゴル語、ツングース語からなるとされているが、近年では、それに日本語と朝鮮・韓国語がくわえられることが多い。アルタイ論争史を概観したあと、言語の系統分類の理論的基礎を論じている。そして、最後には、モスクワ学派のノストラティック言語学にみられるような、比較の方法について、詳細に検討をくわえ、説得力に欠ける方法論的、実証的問題点をあげている。

The present paper gives a concise presentation of the present state of the “Altaic” debate, which circles around the question, whether an language family consisting of Turkic, Mongolian and Tungusic exists and, in recent times, whether Korean and Japanese form part of this putative stock. After a condensed overview over the history of the debate, the theoretical foundations of genealogical language classification are discussed; finally a recently widely discussed approach to the comparison of these languages — namely that of the Moscow-based school of Nostratic linguistics — is examined in some detail and its methodological and factual shortcomings, which are found unsurmountable, are presented.