Discussion Articles / Дискуссионные статьи

Jens Fleischhauer
Institute of Language and Information,
Heinrich-Heine University, Düsseldorf

A Phylogenetic Interpretation of the Comparative Method[1]

The Comparative Method is the central tool for classifying languages. Languages are classified according to their genealogical (genetic) relationships. This method is analogous to cladistics in biology. In the paper it is claimed that both methods are not just analogous, furthermore both are based on the same basic assumptions. These assumptions are (1) *monophyletic origin* and (2) *evolutionary development* of species/languages. It will be argued that the assumption of monophyletic origin is not valid in respect of languages, while an evolutionary account of language origin and development is not part of the Comparative Method and the associated tree model. So, a methodologically based criticism on Comparative Method and the tree model will be presented.

In the first part of the paper both accounts will be compared and the basic assumptions of them will be made explicit. In the second part of the paper some arguments against the monophyetic origin and development of languages will be given. It will also be discussed why an evolutionary account on language change is necessary for the interpretation and philosophical coherence of the Comparative Method and the tree model. But, as will be claimed at the end, such an evolutionary account is not compatible with the monophyly assumption, which results in a clash between both assumptions. A certain solution of this clash will be given, which results in skipping the monophyly assumption. This has consequences in respect of classification are now valid and which tools can be used for classifying languages.

Keywords: Cladistics; Comparative method; Language evolution; Tree model

1. Introduction

The Comparative Method is the central tool in historical linguistics for historical reconstruction and also classifying languages. A classification done with the Comparative Method is called a *genetic classification*. The result is that languages are arranged in language family trees. This means that languages are classified according to their genealogical relationships[2] and are interpreted as being in relation of child- or sisterhood to other languages. Such a way of classifying entities is called *phylogenetic classification* in biology; a classification by genealogical relationships. In linguistics it is sometimes also called a phylogenetic classification (cf. Whinnom 1971).

---

[1] I would like to thank Axel Bühler, Stefanie Schulze, Hakan Beseoglu, Adrian Czardybon and Daniel Schulzek for their helpful comments.

[2] I will use the terms *genealogical* and *genetic classification* interchangeably throughout this paper.

Journal of Language Relationship • Вопросы языкового родства • 2 (2009) • Pp. 115–138 • © Fleischhauer J., 2009

Several authors claim that the Comparative Method is analogue to biological cladistics (Platnick & Cameron 1977; Atkinson & Gray 2005), but it has never been explicated what this analogy really means. It is a further claim of Platnick & Cameron that both methods, cladistics and Comparative Method, are not only analogue, but furthermore that cladistics is the fundamental approach to all historical-genealogical sciences. However, there is not merely one cladistic method in biology; there are several rivaling interpretations of it. In Platnick & Cameron's view the Comparative Method is analogue to a certain interpretation of cladistics, namely the phylogenetic cladistics of Willi Hennig (1950, 1984).

The aim of this paper is to argue that Platnick & Cameron's claim is correct. Additionally I want to make explicit which basic assumptions are underlying such a cladistic approach. I will argue that there are two basic assumptions, which are inherent in biological cladistics as well as in the tree model associated with the Comparative Method. These assumptions are first, the *monophyletic origin* of species/languages and second, the *evolution* of species/languages. After arguing in favor of these assumptions, I will discuss their consequences in respect of languages. As a result of this discussion it shows up that the assumption of monophyletic origin is inadequate. This leads to the view that the tree model, which is intimately connected with the Comparative Method, cannot be used for an adequate representation of phylogenetic relationships of languages because languages do not necessarily evolve in a tree like pattern.

The argumentation in this paper is not directed against the Comparative Method, but against the tree model. As far as both are connected, the criticism also applies to the Comparative Method. I will use the terms *Comparative Method* and *tree model* synonymously and restrict the Comparative Method only to classifying aspects. Historical reconstruction, a further use of the Comparative Method, will not be dealt with. This paper is oriented towards the methodology of historical and classifying linguistics.

The paper is structured in the following sections. Section 2 is a short illustration of Hennig's cladistics. Section 3 shows the analogy between cladistics and the Comparative Method. In section 4 the basic assumptions of cladistics are discussed and it will be shown that they also hold for the Comparative Method. Problems regarding a phylogenetic interpretation of the Comparative Method, mainly based on the basic assumptions, will be discussed in section 5. At the end I will give a conclusion.

2. Cladistics

Platnick & Cameron (1977) point out that entities in biology and historical linguistics are classified in the same way. The phylogenetic relationships between these entities are used to classify. Both methods, cladistics and the Comparative Method, are not only analogue in these respect, they also use the same theoretical terms. This will be explicated in the following two sections.

In biology, species characters are used for classification. These characters can be taken from a wide range of data, e.g. morphological, genetic, ecological, behavioural, molecular or certain other data. But not every piece of data, which means not every character, can be used for inferring phylogenetic relationships. Only *homological characters* are relevant for such a classification. Homological characters are characters shared by different species due to common ancestry. The character state can be unchanged or changed in respect of the character state of the common ancestor (Hennig 1984: 37, Mayr 2003: 232). In one species the character can be unchanged, while in the other species it changed in a certain way. For example the shape of wings of certain birds can be changed or unchanged in respect of the shape of the ancestor species of them. What is only relevant, is that the changed or unchanged characters are inherited from a common ancestor. Central to this approach is a comparative definition of relationship, which is defined in the following way (cf. Hennig 1984: 19): *A species A bears a closer phylogenetic relationship to a species B then to any species X, if A and B are descendants of a stem species which is not also a stem species of X*. This means that two species are in a phylogenetic relationship if they share a common ancestor and there is no third species, which shares an ancestor with one of them but not the other. If two species share an immediate common ancestor both are called sisters of each other.

Homologous characters indicate that there is a certain phylogenetic relationship between two species. But it is not possible by using homologies to decide whether two species are sisters of each other or if it is a farer relationship. This problem will be solved by introducing certain types of homological characters.

According to the cladistics only those homologous characters which constitute evolutionary novelties can be used to reconstruct phylogenetic relationships. Homologous characters which are evolutionary novelties are called *apomorphies.* A character state that occurs in descendants of a species, but not in the species themselves, is called an *apomorphy*. It is a new characteristic feature of that character derived by independent evolution of the descendant species. The counterpart to this is called *plesiomorphy,* which is an unchanged character state in respect of the homologous character of the ancestor. Both terms describe the feature characteristic of a certain character in respect to the feature of the ancestor.

Apomorphies indicate the former split of a single species into daughter species. To a certain point *t* in time a species X splits up into two or more separate populations. By time there will be no further contact between the descendants of X, which means that they will evolve independently of each other. By losing contact, gene flow interrupts and the descendants of X will evolve independently of each other. Changes cannot be transmitted anymore between the now separated populations. The occurrence of an apomorphous character indicates that there is no more contact between two populations because otherwise the same novelty would show up in both of the populations and not only one of them. Therefore an apomorphy indicates that a splitting off of a species into two or more daughter species occurred, and the species are now separated from each other.

To know which characters are apomorph and which are plesiomorph is not enough to get the right classification. This merely gives the basic idea of cladistic classification, but much more is needed to get the right topology of a certain family tree. However I do not want to discuss this topic in detail; the aim is just to outline the basic ideas of cladistics.[3]

The terms *homology* and *apomorphy* are central to infer phylogenetic relationships between species. Apomorphous characters indicate a splitting off and such splits are represented in family trees. Each branch in such a tree represents a species, while every node stands for a splitting event. This means that the temporal beginning and the extinction of species are represented by the nodes at both ends of the branch. It is the process of branching which is represented in family trees.

There is a certain restriction on cladistic classification. According to Hennig every valid phylogenetic group needs to be a *monophyletic group*. If all and only all descendants of one single stem species are grouped together, this group is called a *monophyletic group*. Two species which stand in the relationship of sisterhood form a monophyletic group together with their stem species and their own descendants. For illustration a species tree is shown in fig-

---

[3] There is much debate about character weighting and algorithms to calculate the right trees, in biology and also in linguistics. But all these debates are focused in certain frameworks, like cladistics. This is the reason why I skip these discussions.

ure (1). Species A is the stem species of the group, whereas B-G are descendants of A. B and C are daughters of A and D-E and F-G are daughters of B resp. C. If A and all its descendants are classified together, such a classification is called *monophyletic*. Also B, D and E on the one side and C, F and G on the other side form monophyletic groups. If for example only A, B, D and E are grouped together, such a classification would be called *paraphyletic*, the stem species would not be grouped together with *all* of its descendants. A third type of classification is called *polyphyletic* and would be a grouping of e.g. E and F, two species without an immediately direct ancestor. In respect of cladistics, every valid classification needs to be monopyhletic, which is not the case for non-cladistic approaches, such as the one of Mayr (2001). So, this concept distinguishes cladistics from other approaches and it is a quite central theoretical term. Monophyly will be discussed in more detail later.

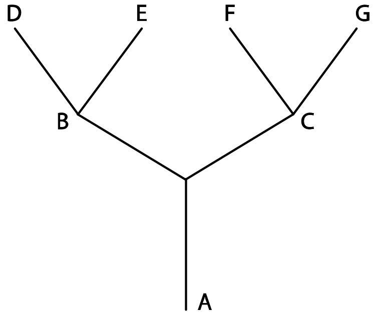

Figure 1: Species tree. Species are indicated by letters.
A is the stem species, while *B–G* are descendants of *A*.

While apomorphies are employed to indicate a splitting off of species, it cannot be used to infer monophyletic groups. A criterion to call a set of species descendants of a single ancestor is needed. In cladistics such a criterion are characters which are shared by the stem species of a monophyletic group and all its descendants. Such characters are called *synapomorphies*. They are shared by *all* members of the whole group of which X is the stem species and they are derived characters from this stem species. So a synapomorphous character is an evolutionary novelty of the stem species of the group. For the descendants of this stem species the character is plesiomorph.[4] On the one hand synapomorphies count as an apomorphy of the stem species and on the other hand they count as a criterion for the identity of the whole group. So synapomorphies indicate a line of descendant with common origin. Mammary gland and body hair are two examples of synapomorphies; they are important synapomorphies for the group of mammals. They showed up in the stem species of mammals and every mammal species shows some reflex of these characters.

This is a rather short description of phylogenetic cladistics, but it is sufficient for the aim of this paper. It is not important to get the method in full detail; the aim is only to illustrate the basics of the method so that the analogy to Comparative Method can be seen.

3. Comparative Method

The Comparative Method is the central tool in historical linguistics, which can be used for two different purposes. In the following I will mainly concentrate on the aspect of classifying languages and will say not anything relevant about the reconstruction of languages.

---

[4] This does not mean that the character feature cannot change in the single species, but the existence of the character is a common feature of the stem species and all its descendants.

The Comparative Method is based on two hypotheses (Jeffers & Lehiste 1979): the first one is called *relatedness hypothesis* and the second one *regularity hypothesis*. The relatedness hypothesis states that obvious similarities between forms belonging to different languages can only be explained by assuming that these languages descended from a common ancestor. While the regularity hypothesis expresses that related forms of two languages are not isolated, rather there are regular processes of sound change which affected a larger set of items. With these assumptions it is possible to reconstruct the stem form, if you can retrace the outcome of regular sound changing processes to the underlying common form. The regularity assumption is quite important, but it only captures sound changes and not every kind of language change. Both hypotheses together point out why it is reasonable to postulate genealogical relationships between languages: there exist regular similarities between different languages, which can only be explained by postulating a common ancestor of these languages.

A central concept in respect of genealogical relationship is *regular correspondence*. It is possible to define a regular correspondence as a not accidental phonetic/phonological and semantic similarity between forms in different languages. According to Roger Lass (1997) regular correspondences in languages with a common ancestor are called *cognates*. So, if there are regular correspondences between genealogically related languages, these are called cognates. The linguistic term *cognate* is equivalent to the biological term *homology*; both indicate certain similarities between forms in entities of common origin (cf. Altmann 1976, Lass 1997).

A justification for a claim of common descent of two languages is the possibility of reconstructing their common ancestor. Lass states: "[…] in linguistics the only solid definition of homology is 'regular correspondence', which is itself justified only by the possibility of reconstruction" (Lass 1997: 163). In linguistics regular correspondences are justified by the possibility of reconstructing the common ancestor of the languages, which show these correspondences. A reconstruction of proto-forms in linguistics can be used to justify a hypothesis of phylogenetic relationship between two or more languages (Lass 1997: 130, Dixon 1997). [5]

In Comparative Method, subgroups are established in the same way as in cladistics. Lass writes: "Taxa are defined […] by shared replicable apomorphies, not vague plesiomorphies. An apomorphous character can only be defined relative to an earlier, primitive one. And this means a reconstruction" (Lass 1997: 167). *Shared innovation* means nothing more than *synapomorphie*, an innovation of the stem species, shared by all the descendants of that species. So in linguistics and cladistics the same criterion is used. Greenberg (1957: 49) is consistent with Lass' statement: "The problem of subgrouping […] is the recognition of the existence of a set of common changes to a particular subgroup". Lass explicitly uses the term *synapomorphy*, while the term *shared innovation* is more widespread in historical linguistics (cf. Dyen 1953, 1973)[6].

This is quite enough to demonstrate the analogy between the Comparative Method and cladistics. The central term of both methods is *homology*, respectively *cognate*. One important characteristic is the use of synapomorphous characters and not the use of all cognate characters to classify languages. As in cladistics, cognates demonstrate a relationship between languages, but sole cognates do not indicate how close this relationship is. Lass (1997: 167) states explicitly that only shared apomorphous and not plesiomorphous characters can be used for classification. Aikhenvald (2001: 167) writes: "If a number of languages within a given family share retentions to the proto-language this does not require a period of shared development and does not constitute for subgrouping". This only means that plesiomorphous characters do

---

[5] There are also certain criteria for justifying homologies in biology (cf. Remane 1956)

[6] See also Chrétien (1963) for a short historical overview of the use of shared innovations for subgrouping in historical linguistics

not indicate a common origin of two languages, they can be accidental or the result of borrowing. But certain changes in different languages, for example a change from Proto-Indo-European (PIE) /p/ to /f/ in the Germanic languages, like in the Germanic consonant shift, cannot be attributed to accident and are only explained as a synapomophy of the stem languages of the Germanic branch. All Germanic languages inherited the changed forms, where a PIE /p/ was replaced by /f/, from their common ancestor Proto-Germanic.

If two languages do not share innovations, inferring a common phylogenetic history of these languages it is not allowed. This is typical of a cladistic point of view and distinguishes this method from other approaches of classification in biology, for example Mary's (2001) evolutionary systematics, in which both apomorphous as well as plesiomorphous characters for classification are used.

4. Basic assumptions

There are two basic assumptions underlying the cladistic approach of Hennig. In this chapter, I will discuss these assumptions and show that the same assumptions are underlying the Comparative Method as well. The assumptions are 1. *monophyletic origin* (or *monophyly assumption*) and 2. *evolutionary development*. Both assumptions will be discussed separately, in two different subsections.

4.1. Monophyly assumption

The *monophyly assumption* expresses the requirement that every valid classification only consists of monophyletic groups. In most accounts of the Comparative Method no requirement analogue to monophyly is explicitly stated. But I will demonstrate that this assumption also underlies the Comparative Method.

The term *monophyly* is in contrast with *polyphyly* and *paraphyly*, as discussed in section 2. In Hennig's view only monophyletic groups are valid for phylogenetic classification. But according to David Stamos there are two different interpretations of *monophyly*. On the one hand there is Hennig's view that a classification is monophyletic iff it includes a stem species and all and only all of its descendants. On the other hand there is a 'weak' interpretation which states that every species has a single origin (Stamos 2003: 271). The 'weak' interpretation focuses on the assumption that every biological species has only *one* ancestor. But this is just a convention (Stamos 2003: 309) and not valid for every biological species, for example there exist more than 700 naturally occurring plant hybrids (Judd et al. 2002).

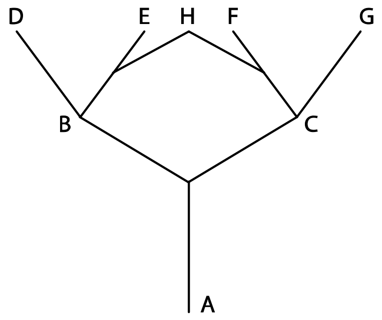

Figure 2: Species tree.
Species *H* is a hybrid species of their ancestors *E* and *F*.

Hybridization is a process of mixing different species so that they get a common offspring. In figure 2, the same species tree than in figure 1 is shown, beside that species H, a hybrid of E and F, is included. It is shown that hybrids are descendants of at least two different ancestors. McDade (1992: 1330) captures the problematic case of hybrids in the following terms: "Cladistic methods presuppose divergent evolution and cannot give the correct phylogeny for hybrids, which by definition have reticulating histories". By hybridization two phylogenetic lines converge together, so a hybrid does not come into existence by splitting off of the stem species into descendants. Furthermore hybrids are the result of converging different phylogenetic lines together. It is often claimed, for example by Hennig and Mayr, that hybridization does only occur in very limited cases between zoological species. But Arnold (1997) shows that hybridization is much more common than it has often been thought in biology.

As far as I know, there are only two uses of the term *monophyly* (Holden & Gray 2006: 23; Pagel & Meade 2006: 177) in historical linguistics. Some linguists write about the development of languages in terms, which seems to be equivalent to the weak interpretation of monophyly. Leonard Bloomfield (1933: 298) for example writes: "When we say […] that a resemblance between languages is due to relationship, we mean that these languages are later forms of a single earlier language." Beside such statements, the monophyly hypothesis is implicit in the arrangement of languages in language family trees. Every language in such a tree is shown to be the descendant of a single ancestor. In graph theoretic terms this means that every node in such a family tree has a single parent node (Warnow et al. 2006: 80).

Finally, the assumption of monophyletic origin shows up in the way how mixed languages are treated. Such languages have more than one ancestor and so it is reasonable to call them language hybrids. According to Thomason and Kaufman (1988) mixed languages cannot be genetically classified. An example of this would be the language Ma'a (or Mbugu), which seems to have a Bantu grammar and a Cushitic lexicon (Goodman 1971). Ma'a does not fit the definition of 'linguistic descent' which is given by Ringe et al., based on the work of Thomason & Kaufman: "A language (or dialect) Y at a given time is said to be descended from language (or dialect) X of an earlier time if and only if X developed into Y by an unbroken sequence of instances of native-language acquisition by children" (Ringe et al.: 2002: 63). This means that languages which do not originate or develop in an unbroken sequence of native-language acquisition do not have linguistic ancestors because they did not descend from a certain language.

The reason for this claim is the following. Thomason & Kaufman argue that there needs to be an uninterrupted transmission between two languages if one of them should be called the descendant (changed later form) of the other one. Languages with an interrupted transmission cannot be called changed later forms of a certain ancestor. Such an interrupted or imperfect transmission results for example when a language is acquired in a different way than by native-language acquisition (Thomason & Kaufman 1988: 9f.). The mixing of two languages, for instance caused by cultural pressure of a dominant group which results in bilingualism and shift in language use, would be an imperfect transmission. The new language, the mixed one, was not continuously acquired by children and changed via this transmission process. The mixing was a result of changes in the adult language. So, if two languages converge to a single language, this cannot happen via native-language acquisition, it happens by introduction of certain forms of a language X into an already acquired language Y of adult speakers.

A language needs to show reflexes of its genealogical origin in all it parts, according to Thomason & Kaufman's view. So, a mixed language such as Ma'a does not show such reflexes of one language in all its parts. Rather Ma'a shows in lexicon reflexes of Cushitic languages and in grammar of Bantu languages and this is the reason. This is the reason why one cannot say that there is a single ancestor of Ma'a, because different subsystems show reflexes to different languages. This demonstrates that there is an assumption of monophyletic origin underlying the view of genealogical relationships of languages.

As shown, the assumption of monophyletic origin is relevant both in biological and linguistic classification. In both disciplines mixed types, mixed languages or biological hybrids, are excluded from classification or are called 'bad' types. Arnold (1997: 13) states: "[…] all of the [species] concepts require "good" species not to hybridize and because this restriction reflects the underlying view that reticulate evolution is undesirable because it interferes with the process of divergence". Hybridization yields reticulate evolution and this is not compatible with the requirement of monophyletic origin. As hybridization interferes with the view of language origin by divergence, such languages are said to be genealogically unrelated. But, as Stamos said, the origin of a single ancestor is only an assumption and needs not to be valid for every species or every language.

4.2. Evolutionary development

After explaining the relevance and consequences of the monophyly assumption in respect of languages, I want discuss now the second assumption of cladistic classification, which is the theory of evolution. Mark Ridley explains why evolutionary theory is a basic assumption of cladistics: "If species did not originate in a branching hierarchy, there would be no phylogenetic hierarchy, and the philosophical coherence of cladism would be lost" (Ridley 1986: 86). The branching hierarchy is a result of biological evolution, so the evolution of species is necessary for such a branching order. Without evolution, there would be no branching, which means that no splitting off species would occur.

Other classification methods, for example phenetic ones, need not presuppose evolution. For such accounts it is irrelevant if species evolve or not. There is an interpretation of cladistics called *transformed cladistics* (e.g. Platnick 1979), which state that evolution is not a necessary requirement for classification. In this account it is not possible to decide whether a character is apomorphous or plesiomorphous. To state that one character is apomorphous is nothing more than to mention that the respective character shows up in an evolutionary novel way. So if one uses terms like *apomorphy* or *synapomorphy,* one needs an evolutionary theory, if not, it would be impossible to interpret this terms in an adequate way (Ridley 1986: 89). You cannot call a character an evolutionary novelty, if you do not presuppose evolution. Therefore, evolution is a basic assumption of cladistics in two ways: first, in cladistics the classification criterion is the splitting of species into daughter species, which is an evolutionary process; second, the theoretical terms of cladistics can only be interpreted for the background of evolutionary theory.

The Comparative Method also uses the splitting of languages as criterion for classification (Nakhleh et al. 2005a). Thus the evolutionary process of branching is also presupposed in historical linguistics. If languages did not evolve and if there is no branching of languages into daughter languages, the tree model would be philosophically incoherent.

It was shown in section 3 that the Comparative Method and the cladistic approach share the same theoretical terms. As far as I know, Roger Lass is the only who uses explicitly terms such as *apomorphy*, *plesiomorphy* and *synapomorphy* in historical linguistics. But as also shown in section 3, it is irrelevant if the theories differ in terminology. What is relevant is that the content of the theoretical terms is the same. As cladistics and Comparative Method classify entities in the same way, it is warrantable to call the Comparative Method a cladistic method, like for example Platnick & Cameron (1977), Pagel (2000), Atkinson & Gray (2005) do. This is the reason why I guess that the theoretical terms share the same content.

Furthermore, both methods share the same underlying basic assumptions. So the main point to capture is that the only reasonable interpretation of the Comparative Method is a phylogenetic interpretation. It follows, I guess, that it is also reasonable to claim that one needs an evolutionary theory of language change to interpret the Comparative Method.

5. Problems for the basic assumptions

As was shown in the last three sections, the Comparative Method is analogue to phylogenetic cladistics. Central is that both accounts share the same basic assumptions. Some problems which are based on the basic assumptions will be discussed in the following section.

At the end of the last section I claimed that an evolutionary theory of language change is necessary for the interpretation of the Comparative Method. There are different evolutionary theories of language change, for example Croft (2000, 2006) bases his account on David Hulls (1988) generalized evolutionary theory. While Ritt (2004) uses Richard Dawkins (1989) concept of *selfish genes* and constructs a similar account for language change. Both accounts on language change make different assumptions in respect of the evolution of languages, so they are not equivalent.

Ritt does not say much about classification of languages and compatibility between his evolutionary theory and the Comparative Method. Croft (2000) otherwise claims that the evolution of languages does not proceed in a tree-like pattern, but is reticulate. He explicitly states that his evolutionary account on language change and the tree model are not compatible because of the reticulate pattern of language evolution.

One process that is included in evolutionary accounts of language change and excluded by the Comparative Method is *borrowing*. There are different types of borrowing between languages and in its simplest case it describes the borrowing of a certain character from one language into another one. The biological equivalent to this is *horizontal gene transfer*. Jannig & Knust (2004) say about horizontal gene transfer that it is the transfer of genetic information between two individuals of the same generation. This type of transfer contrasts with the transfer of genetic information between parent and offspring. The transfer of genetic information between parent and offspring is *vertical gene transfer* because they do not belong to the same generation and it occurs be natural reproduction. In analogy borrowing can be called *horizontal transfer of linguistic material between languages*.

Wägele (2001) states that horizontal transfer in biology can be problematic for cladistics, because it might obscures phylogenetic relationships. But according to Wägele horizontal gene transfer is quite rare and so a minor problem which will raise no doubts concerning the adequacy of cladistics. According to his view, horizontal transfer leads only to minor mistakes in the reconstruction of phylogenetic relationships.

In cladistics only vertical relationships between species are regarded. Horizontal connections need to be excluded, because they violate the monophyly assumptions. So, hybridization or borrowing, as in the case of languages, cannot be captured in a tree-like model. The reason is that these trees only show one evolutionary path and this one is indicated by the vertical relationships between species or languages. But in the case of mixture there is more than one evolutionary path, which connects a species or language with several other species or languages (cf. Holden & Gray 2006).

There are differences between the borrowing of single forms and a mixing of languages as in the case of Ma’a. Chappell (2001) for example discusses hybrid relative clause construction in Sinitic languages, where one language uses the type of relative clause construction of a different language. Is the term *hybrid* adequate to describe such phenomena as Chappell does or would it be better to call it (grammatical) borrowing? This leads to the question which are the problematic cases in respect of the monophyly assumption.

I presume there is a fuzzy border between hybridization and borrowing, both are processes which lead to the same result: they mix different phylogenetic lines in a single outcome. The only difference lies in the amount of mixture. In the case of borrowing, there may be only a few characters (lexical forms or grammatical constructions) that are borrowed. The borrowing of a single feature can be seen as the extreme on the one side of the continuum. On the other side of the continuum, there are cases like Ma'a, in which languages of two different families are mixed together. The question is where to draw a line in this continuum to separate the problematic and the unproblematic cases from each other. A radical view would be to state that every single case in the continuum is a problematic one. On the other side, it would be also radical to consider only mixed languages and pidgins as problematic cases. Extreme borrowing would according to such a view be unproblematic. It is the case that extreme borrowing is a controversial case.

Bryant et al. (2005) discuss the case of the huge amount of borrowed lexemes in English and state that English needs to be separated from other Germanic languages in phylogenetic trees. English lost some of its similarities to the Germanic languages because of its extreme borrowing. These borrowings are conflicting signals, which indicate contradictory relationships between languages, in this case Romance and Germanic languages. It is not possible to decide which one of several languages is the unique ancestor of a certain language. In such cases no unambiguous family tree can be reconstructed and the evolutionary history needs to be described as a reticulate process. In Comparative Method this conflicting signals are ignored because borrowed characters are excluded from the analysis. But most analysts assert that English is a Germanic language (e.g. Thomason & Kaufman 1988). That the data are not as clear as the tree model indicates is shown by accounts like the one used by Bryant et al. However there are also conflicting signals in Australian languages (Dench 2001), Sinitic languages (Chappell 2001, Hamed & Wang 2006), African languages (Heine & Kutewa 2001; and the literature to Ma'a / Mbugu as a much debated mixed language, for example Greenberg 2005; Thomason 1983, Goodman 1971, Mous 2003). The reason for the existence of conflicting signals are extreme borrowing and dialect continua. So not only the extreme cases like mixed languages and pidgins are problematic for the assumption of monophyletic origin, also for example Bantu languages (Holden & Gray 2006; Marten 2006) are problematic and English could be a problematic case as well.

Beside the mentioned cases, Arnold says the following about the biological importance of hybridization: "[…] whether hybridization events are widespread or rare for a given species or species complex is not necessarily predictive of the degree of evolutionary importance" (Arnold 1997: 24). It is not important whether English is a problematic case or only mixed languages are, everything that is relevant is that hybridization and borrowing are naturally occurring phenomena of language evolution.

A differentiation between the origin and the development of languages need to be drawn. It is quite a difference if one speaks about reticulate origin or reticulate development of languages, such as Thompson (2002b) does. Borrowing and shift-induced inferences are processes of reticulate development. They influence an existing language and change its lexicon or grammar. These are neither exotic nor rare casis (Thomason 2002b: 20). But the origin of entirely new languages by such processes is a rare outcome, according to Thompson. I guess she is right. The frequency of reticulate development seems to be much higher than the frequency of reticulate origin. The drawn distinction is analogue to the distinction between hybridization and horizontal transfer. Hybridization in biology is the creation of a new species, whereas horizontal transfer may be caused by viral transmission. Nakhleh et al. (2005b: 800) say about hybrid speciation that in this case two lineages recombine to create a new species. In the case of horizontal transfer only (genetic) material from one lineage is transferred to another lineage, but this happens without merging two phylogenetic lines into one.

For biological analysis this means that species phylogenies and gene phylogenies are not identical, whether horizontal transfer occurs. There is a discrepancy between the history of the species and the history of the genes (Nichols 2001). The same happens in respect of languages as soon as reticulate processes are at work. Horizontal or reticulate processes and hybridization result in evolutionary histories, which cannot be adequately represented by trees. This means that not only reticulate origin, which is hybridization, is a problematic case for a cladistic method, but also reticulate development.

Critics could now replay that it does not really matter if there are single cases of horizontal transfer between two languages. They could ask: Do single occurrences of borrowing really show that the evolution of languages cannot be adequately represented in a tree-like model? The answer is yes and there are different reasons for this. First of all, language contact is not rare, so there is not only a single occurrence of borrowing. But should there be languages which borrowed only a very small fraction of features (lexicon or grammar), they would be the exception, not the norm. But more relevant is that the evolutionary history of a language is the sum of the evolutionary history of its parts. It is not the case that there is a certain entity called language which has an independent history of its parts. So, if a language developed in a reticulate way, but did not originate in a reticulate way, it would also be non-monophyletic and it could not be captured adequately in the tree-model. As a replay against the mentioned critics I would raise that reticulate origin and reticulate development are both problematic cases for the Comparative Method and the associated tree model.

Before finishing this section, I want to discuss a certain claim raised by Sarah Thomason. In a paper about creoles and the concept of 'genetic relationship' Thomason says that: "A claim of genetic relationship is not a generalized statement of historical connectedness, but rather a quite specific technical claim that a genetic hypothesis meets the rigorous criteria of the Comparative Method" (Thomason 2002a: 103). According to an evolutionary view, a claim of genetic relationship is more than a technical claim. The statement that two languages are genetically related is a statement with a certain historical content. It does not only mean that a hypothesis meets the criteria of the Comparative Method, it also means that languages evolution occurred in a certain way. To claim that two languages are genetically related is the claim that language evolution was branching and not reticulate; that no hybridization occurred; and also that in the process of evolution that languages shared a common ancestor. This is a claim with historical content, which can be right or wrong.

6. Conclusion

In this paper I followed the claim of Platnick & Cameron and others that the Comparative Method is a cladistic method. I explicated what this claim means in respect of the Comparative Method, mainly regarding the basic assumptions of cladistic method. The second part of the paper was a critical evaluation of the basic assumptions, which I will summary as follows:

1. The assumption of monophyly, which means the assumption of tree-likeness, is intimately connected with the Comparative Method. There are a lot of cases of single languages or whole families which are stated not to fit the requirement of the monophyly assumption.

2. The second basic assumption of cladistics is problematic for the Comparative Method. An evolutionary theory of language change needs to integrate horizontal relationships between languages, which is in conflict with the monophyly assumption.

Haspelmath's statement that "nobody ever claimed that family trees capture all historical relationships — they are intended to show relationships among languages that go back to a common ancestor" (Haspelmath 2004: 214) can be regarded as an argument against my view. According to him and many other representatives of Comparative Method, contact-induced changes should not be reflected in family trees, only language internal motivated changes. Surely, one needs to decide which aim a phylogenetic classification should serve. On the one hand it should be a natural classification of languages (Hempel 1965, 1969), and on the other hand it is used to investigate the evolutionary origin and development of languages. There is no need to connect these aims with an a priori assumption of tree-likeness. Hamed & Wang say: "There is indeed a tree-like signal, but it can be overlapped by conflicting signals which cause the hierarchical signal to fade and become ambiguous. Therefore, one should rather question *how* to justify that the tree model is adequate for a particular set of languages, and what alternative representations may be used" (Hamed & Wang 2006: 31). If Haspelmath is right in claiming that family trees should only represent genealogical relationships, it is a further question whether these relationships are monophyletic in every case. In the framework of the Comparative Method genealogical relationships are by definition monophyletic, but there are doubts if this is empirically adequate.

I agree with the claim of proponents of the Comparative Method that cognates need an explanation. Because of their regularity they are quite special, but borrowing on the other hand is not rare (Pagel 2000: 392). Borrowing is a natural process in the development of languages and to get a right picture of the evolution of languages, it is necessary to integrate this natural process into phylogenetic analysis. As Hamed & Wang say, cognates produce a tree-like signal, but borrowing results in conflicting signals. It would be wrong to concentrate on only one of them, what is needed is an approach which can combine both signals. There are a few accounts in biology to handle conflicting signals, some of them are also used in linguistics (e.g. Bryant et al. 2005; Hamed & Wang 2006).

The main point I want to raise is a critical comment on the Comparative Method. According to my analysis, there is a dilemma facing this method. On the one hand an evolutionary approach of language change is needed for interpreting the theoretical terms of Comparative Method. On the other hand this approach is not compatible with the monophyly assumption. The only reasonable solution is to abandon the monophyly assumption, but this would mean that the concept of linguistic descent explicated by Thomason & Kaufman and Ringe et al. is not adequate anymore. An explicitly evolutionary theory of language change is needed to underpin phylogenetic analysis and classification of languages and it is necessary to bring both, the evolutionary theory and the approach to classification, in harmony with each other. This would be the only possible solution to the dilemma, and it would be reasonable because the assumption of monophyly is only a convention, as Stamos said, and can be deleted if it is not adequate.

I guess that such a critic, as presented in this paper, is necessary if one wants to argue in favor of new phylogenetic methods which are sometimes used in linguistics like NeighborNet and so on. It is not just enough to present a new account on the phylogenetics of languages, it is necessary to show why such an account is reasonable. According to my view, new accounts a necessary because of the deep theoretical problems of the Comparative Method and the associated tree model.

Surely, my criticism raises questions about the status of classification of languages. The focus is on capturing phylogenetic relationships and not on using genetic relationships to get a unique assignment of languages into distinct classes. I would state that such an assignment would be quite artificial, and would not capture any relevant linguistic insights and so it is not a natural classification in the sense of Hempel.

Vladimir Belikov
Moscow State University

Reply to Jens Fleischhauer

The analogy between the classification of languages and biological objects has been repeatedly observed for a long time, but this is exactly what it remains: a mere analogy. The levels of abstraction in objects classified by biologists and linguists differ significantly, therefore the classification systems are also different.

It is evident that out of these two the biological classification is primary, whereas the linguistic one was constructed later in accordance with the former's pattern. Therefore, a short outline of the history of the biological nomenclature seems to be of importance.

The ‘*Systemae Naturae*’ by Carl Linnaeus (1758) is considered to be the seminal work for the biological nomenclature. The *Systemae* have introduced the *Genus* + *Species* binomial classification system. However, *Amphibia*, *Reptilia*, *Pisces* and other taxa are much older than that. Exactly one hundred years before the book by Linnaeus has appeared, the scientific knowledge of the 17th century has been summarized in the ‘*Orbis sensualium pictus*’ by Jan Amos Komensky. In this book, the category of *Amphibia* comprised not only the species ‘frog’, but also the crocodile, the turtle, the otter and the beaver who all live in two kinds of environments alternately, thus being *amphibious*.

The list of *Reptilia* represented in the *Orbis sensualium pictus* seems, at first glance, to better correspond with the modern views: out of eleven species mentioned there, only the scorpion was later reclassified as *Arachnida*, and the salamander was transferred to *Amphibia*; the rest of them remained being classified as reptiles. Today we find the *Draco* and *Basiliscus* genera among the scaled reptiles, but these are not the species which were referred to as *Reptilia* by the 17th century science. According to Komensky, ‘*Draco, serpens alatus, halitu, Basiliscus oculis ‹…› necantens*’ [Коменский 1941: 79], i.e. the “winged dragon” who kills with his breath, whereas the basilisk kills with his gaze. The illustrations supplied to the first edition correspond with this description fully.

What changes brought the 18th century? Taxonomy started to note the structure of the organisms, not just their appearances. Linnaeus considered the lack of molars (*dentes molares nulli*) to be one of the most characteristic features for referring a given species to the class of *Amphibia* (which according to him included both the reptiles and amphibians). In the subsequent 19th and 20th centuries, however, the external appearance attracted less and less attention: the lizards were considered to be more close to crocodiles than to snakes, yet it is with the latter that they are united within the same taxon. The most important point, however, was that the scientists have long since stopped accepting anything on the basis of mere trust: to include an organism into the system, the researcher had to see it with his own eyes and to analyze either the entire specimen or at least its significant fragment.

However, the taxonomic tree has begun to be interpreted as a genealogical one only after Darwin, and even then the scholars were far from consensus.

In Linnaeus’ taxonomy, living objects were divided into two kingdoms: Animals (6 classes) and Plants (24 classes). By the second half of the 20th century, the structure was work out in detail and became more complicated. New kingdoms of *Bacteria* and *Fungi* were singled out from Plants; then it appeared that some of the *Fungi* should be more correctly placed among *Protozoa*, which had been classified among Animals, while still another opinion excluded *Protozoa* out of Animals and joined them with **some** *Algae* (formerly a part of Plants) into the new kingdom of *Protista*, etc. The clarity and the strict order of the up-per taxa were ruined. 'Some authors make an attempt to find a way out of this situation by increasing the number of kingdoms (up to 13), but this complicates and hardens the classification', so an anonymous but, most probably, a prominent systematizer wrote around twenty years ago [Система организмов, 1989: 578; spacing is mine — *V. B.*]. Just like that: the right classification is the one that's more simple, enough said.

In the late 20th century, the readability of genomes allowed for the first time to construct the genealogical classification of living objects on a really objective basis. It transpired, among other things, that sparrows are more closely related to crocodiles than lizards. This fact is almost impossible to believe, yet it cannot be easily refuted. One can hardly think of more objective data on genetic relationship than molecules that codify inherited information.

Thus, for already several hundred years the biologists have been forming up the classification of living objects based on objective observation of tangible entities. Human individuals differ, but it is quite easy to step over these differences and to construct an invariable image of *Homo sapiens sapiens,* as well as — since recent times — to define the level of its relationship with the extinct *Homo sapiens neanderthalensis*, to prove that the chimpanzee is more closely related to man than to gorillas, etc. Within our species, the analysis of the mitochondrial DNA and Y-chromosomes allowed to start dealing with genetic relationships of various populations. All this is based upon an objective foundation.

Compared to this, what is the work object of linguists? Everyone speaks Russian (English, etc.) a little bit differently. Is it so hard to reveal a common invariant from these idiolects, to compare objective Slavic invariants (Russian, Bulgarian, Polish, etc.) with each other, with other Indo-European ones, etc.?

Comparative linguists work with the core lexicon — a very distinctive characteristic of the language that changes slowly through time. Phonetic laws have been discovered upon the birth of comparative linguistics. This seems to be a close analogy to inherited information in biology. But words are double-sided items: the laws of development of their external shape are well known (although particular cases of deviations are so numerous!), but we can only guess, to be honest, regarding their semantic development, so that certain etymologies differ in their level of probability. The 100-item wordlists are a good analogy for genomes, and the requirement that every word there represent an **exact** semantic correspondence to the standard is not accidental: in this way we can exclude problems of poorly formalized semantics. But it is no secret that the choice of an equivalent of a certain item in a well described language is sometimes arbitrary. Genetically modified organisms are an artificial phenomenon, whereas a loanword on the Swadesh list is a phenomenon that is quite habitual and often hard to identify.

The reality behind the divergence of living objects is quite simple — a trend towards splitting emerges within parts of a population that used to be joined; then, starting from a certain moment, two or more new populations appear instead, and their further contact will be exclusively within the framework of a food chain.

Contacts between languages may become stronger or weaker, may arise from scratch: this is common knowledge, yet essentially just a metaphor. People contact in different ways, including verbal communication (which can be bilingual), but languages themselves never contact with each other. Languages are abstract entities that cannot be seen or heard per se. Only texts, as a result of activities of language mechanisms, receive a "material embodiment".

By 'consuming' various incoming texts, a human being in his or her babyhood activates the inborn ability to analyze and generate other texts that will be appropriate in specific communicative situations. This ability transforms itself into an entity which we call communicative competence, which is more than mere knowledge of words and grammatical rules.

Texts of the same type (mutually understood ones) are provided by individual text generators of the same type. The invariant of such generators is a particular language: Russian, English, Chinese, etc. The non-professional views on what these invariants really represent are obscure, just since their direct observation is impossible.

Linguists themselves do not fully realize the structure of any particular language either, but at least they have learned that these are sign systems. In the process of researching upon them, they compile dictionaries and grammars which are considered 'incarnations' of these particular languages.

It is these, more 'tangible' objects, that serve as raw material for the specialists in comparative linguistics. Sometimes a scholar has to deal with inaccurate descriptions which are impossible to doublecheck, in which case one needs to reinterpret them on the basis of any accessible information and common sense. The 'basilisks killing with their gaze' and the 'salamanders living in the fire' are, fortunately, met rather rarely in the raw data that are used by comparative linguists; however, various mistakes in interpreting poor quality descriptions are quite possible. Dealing with 100-word lists where five or ten inaccuracies are present, one can expect rather sad consequences for the classification.

When analyzing complex cases of mixed languages, with more than one possible ancestor language, an interrupted language tradition is usually mentioned. In Jens Fleischhauer's article, the language of Ma'a (Mbugu) is described in certain detail as such having 'a Bantu grammar and a Cushitic lexicon'. It is argued that Ma'a cannot be considered as the descendant of a single language because its history does not fit the most important criteria: 'an unbroken sequence of instances of native-language acquisition by children'. The lack of historical sequence is posited on the basis of the analysis of synchronic comparison between lexicon and grammar. Comparative linguists always treat lexicon as a priority, but the Cushitic character of that language is not that simple: its non-Bantu part corresponds to different groups of the Cushitic branch, which made Viktor Porkhomovsky suppose that it is 'more grounded to refer Mbugu to Bantu' [Порхомовский 1982: 254]. In Ma'a grammar 'a notable proximity is found with Asu and Shambala' [Порхомовский 1982: 215]. To what extent the Bantu part of its lexicon is homogeneous is not yet clear, however, it is necessary to remember that no less than seven items in its 100-word list ('bite', 'come', 'dog', 'heart', 'know', 'rain', 'small') originate from Maasai, a Nilotic tongue [Thomason 1997: 475—476].

Is that enough to suggest that a homogeneous language tradition for Ma'a has been interrupted at some stage? The answer is negative. To make such a suggestion on reasonable grounds one should know the historical circumstances of the heterogenization of its lexical development. One can add that it is unlikely that they will ever become known. Bilingual communication competence is rather specific, and serious mutual influence of languages can be possible within an uninterrupted communicative tradition as well.

During a linguistic expedition to the village of Lyozhdug in the Komi Republic in 1968 I gained from my Izhma informant a written story of a fishing incident, when a teal duck with an attached ring got entangled in the nets. The last sentence was as follows: *Снимитім кольцосэ и узнайтім, что чирокыс зимуйтэма Францияын*. (The underlined morphemes are identical to Russian ones; the English equivalent will be something like *Takeoff-tim ring-se and findout-tim that teal-ys spend-winter-toma France-yn*). In this case of compound bilingualism the speaker retained the grammar of one language intact but used it to "wrap up" the lexicon of the other language, freely and unconsciously.

A scholar linguist would call such a phenomenon a code-mixing, however, the 'division of codes' is a researcher's abstraction. There was a certain communication mechanism in the mind of the speaker which generated texts of different types depending on the addressee. He has never communicated with monolingual Komi speakers, and with monolingual Russian speakers he would have simply said *Сняли кольцо и узнали, что чирок зимовал во Франции* ('We took the ring off and learned that the duck had spent the winter in France'). But if all the participants of a particular communicative act were bilinguals, he would freely vary synonymous Russian and Komi lexical items: *снять/босьтны, узнать/тӧдны, зимовать/тӧвйыны*.

There was no interruption in the transfer of the communicative tradition in that village fourty years ago: all local Izhma people talked like that. The quota of the etymologically Russian lexical items used within a particular act of communication depended on the age-group of speakers: the younger, the greater. It is true that communicative conflicts between the oldest and the youngest people could emerge occasionally, but such conflicts are regularly encountered under certain social circumstances even in monolingual environments (e.g., in modern Russian).

Not all bilingual cultures make an accent on the "purity of language". This happens not only in everyday communication, but in specially constructed texts as well. The examples are numerous, I'll give only one, an extract from a Russian Gypsy folk song[1] [Шаповал 2007: 289], its Russian translation below:

Ай ту, тэрнори, да не ломайся,
Мэ пхэнава дуй лава — собирайся.
Ада тэрнори на задыяпэ,
Мэ пхэндём лаворо — скэдыяпэ.

Ай ты, молоденькая, да не ломайся,
Я скажу два слова — собирайся.
Эта молодушка не задавалась,
Я сказал словцо — собралась.

Within one stanza we find two pairs of synonyms: Russian imperatives in the first two lines (*не ломайся*

---

[1] As an example from the other end of the literature genre axis, I would mention the poetic genre of *wakan renku* in medieval Japan, which alternated lines in Japanese and Chinese. In this case, of course, we speak about a specific type of elite bilingual communication, but there are examples when "naturally" mixed literary texts are generated in the environment of mass bilingualism as well.

'don't be shy, don't be opposed to my suggestion, don't hesitate', *собирайся* 'make up your mind, come on'), and their Gypsy preterit counterparts in the other two (*на задыяпэ* 'did not hesitate', *скэдыяпэ* 'made up her mind'). It is significant that the text, being bilingual, is nevertheless monocultural: the addressee of the first two lines is a Gypsy girl, which is seen from the stage of the poem: *шатрица рогожитко* 'a matted tent'.

It is also worth mentioning here that, in biological systematization, one usually resorts to similar or same principles when identifying higher taxa; such similarity is usually lacking in the genealogical classification of languages. In some survey works done on the world's languages, macrofamilies such as Nostratic and Sino-Caucasian are adjacent to such superphyla as Gé-Pano-Caribbean and Indo-Pacific. All these taxa are hypothetical; however, while the former are supported by well-grounded hypotheses, today are being refuted mostly by those who did not take the proper time to examine the argumentation behind them, the latter rather remind the classification of living objects in the 17th century.

It is quite natural to draw a boundary with regard to the place of habitat: some fish species can only live in fresh water, others live in the sea, so we divide them all into two groups. It would be logical to start the description of the latter with a phrase like '*Piscium marinorum maximus est balaena (cetus)*' [Коменский 1941: 87]. The translation might be 'the biggest sea fish is the whale'. In fact, Latin *balaena* (*cetus* also) is not just 'whale'. It is applied to any 'sea monster', which could be a shark or a whale, or for the late Roman period even the Leviathan as well.

From the point of view of the ethnotaxonomy, there is nothing wrong in the existence of a 'whale fish'. The Hawai'ians, for instance, apply the word *i'a* (quite approximately, 'fish') to both "large fish", such as whales, and "smaller" ones, such as oysters.

Natural sciences have already learned, a long time ago, to stay away from such "obvious" groupings. It is high time for the comparative linguistics equally to learn to stay away from the 'super-mega-unions' built on ethnotaxonomic principles, which are roughly defined as 'it is perfectly clear that …'or 'I know for certain that …'

Svetlana Burlak
Moscow, Institute of Oriental Studies

Reply to Jens Fleischhauer

Jens Fleischhauer's article deals with one of the most important aspects of historical-comparative linguistics, namely, genetic classification of languages. The author's goal seems to be a general comparison of the methods that are used to classify objects in biology and comparative linguistics. Such methods certainly have much in common, since in both cases the classification is based on historical relations of the classified entities — living species in biology and languages in linguistics.

Both methods distinguish between homologies and analogies. In biology, *homologies* are represented by different structures that go back to the same original structure of the ancestor species (e. g. human arms and birds' wings, or the "ritualistic" feather-cleaning among different species of ducks), whereas *analogical* structures are ones that are similar in aspect but have developed independently of each other (e. g., the обтекаемая body shape common for fish, ichthyosaurs and dolphins; the singing of birds and human language). In linguistics, one can define as homologies, for instance, etymologically equivalent words or parts of words in related languages. Typical homologies are Latin *q*, Greek *τ*, and Russian *ч* in the word '4' (Latin *quattuor*, Greek *τέσσαρες*, Russian *четыре*), or perfective forms with the auxiliary verb 'have' in English and German (it must be noted, though, that the author only adduces related lexical cognates as "homologies"). A basic criterion for recognizing "homologies" in words is the presence of regular phonetic correspondences. In contrast, "analogies" in languages represent coincidental resemblances, e. g. Quechua *ñuka* 'I' : Tocharian A *ñuk* 'I (fem.)'.

The general principle of building a genealogical tree is also common: the later the separation of the common ancestor, the closer the languages (or species) are related. Both in linguistics and biology the genealogical tree is primarily based on shared innovations (in biology called apomorphies), while common archaisms (plesiomorphies), inherited from a more remote ancestor, are not taken into consideration. An innovation that originated in the ancestor itself and is inherited (at least to some extent) by each single descendant is called a synapomorphy (of course, for all the descendants themselves this feature will be plesiomorphic). For Germanic languages, for instance, a typical synapomorphy is the First Consonant Shift.

Separation of a species or an ancestral language is usually understood to be brought about by a separation of populations. The joint group of an ancestral species or language with all of its descendants in biology is called 'monophyletic'. The group that joins an ancestral species with only a part of its descendants yields a paraphyletic classification; one that joins species or languages that do not share an immediate common ancestor is known as polyphyletic. Both in biology and linguistics the prevailing opinion is that the only correct classification is a monophyletic one.

Such a classification may adequately reflect real historical development of entities only in the case when said development represents only *divergence*, i. e. when the species (or languages) cannot acquire any features from neighbouring species (or languages). In biology, according to the author, this is reasonable, since cases of 'hybridization' between species are so scarce that they do not necessitate a shift of the current working theory. In linguistics, however, the situation is quite the opposite, since languages easily borrow both lexicon and grammatical models from one another, while in extreme cases of particularly intensive contacts we may witness the formation of a new language that has more than one ancestor — something that directly contradicts the monophyletic demand.

The author raises the complex question of what exactly are the relations between the tree model for linguistic relationship and actual language history. Normally, comparative-historical linguistics strives to exclude contact-induced phenomena from its scope — thus, in newer models of glottochronology borrowings are not taken into account already on the initial stage of calculations. However, as noted by J. Fleischhauer, borrowings on all levels of the language system are a natural phenomenon for almost any dialect. One can distinguish between different degrees of contact influence; thus, the author notes cases which he calls 'extreme borrowing' (such as English, saturated with Gallicisms), while even more intense influence can lead to the arisal of new 'mixed' languages from pigins (such as Ma'a with its Cushitic lexicon and Bantu-like grammar). The borrowing of grammatical elements is being treated by the author as 'reticulate development' which he distinguishes from network (i. e. mixed) origin — encountered much less frequently, as he writes. He also correctly notes that the tree-type model serves well only for those languages whose history represents an unbroken vertical transmission of the language from generation to generation by fully competent speakers.

Tree-type models are also problematic in the case of dialectal continua, when closely related languages 'mutate' into one another over a chain of intermediate idioms that are hard to assign to any one particular node on the genealogical tree.

It should, however, be noted, that the problems listed by J. Fleischhauer have already been considered in comparative linguistics, including Russian sources such as [Бурлак, Старостин 2005: 55–82, 129–170]. Thus, it is well known that dialectal continua not only exist, but also have a tendency to 'regroup', so that dialects that were once classified as belonging to one areal group can, after certain historical changes, be placed into a different one. Such is, e. g., the situation with certain formerly Northern Belorussian idioms that are now counted as Northern Russian [Беликов 1989]. The Old Novgorod dialect, formerly not belonging to either South, East, or West Slavic, eventually became a dialect of Russian [Николаев 1988, 1989]. A similar regrouping happened in Turkic, cf. "the presence of Karluk, Kypchak, and Oguz dialects of Uzbek that underwent certain common processes, despite the fact that genetically they belong to different subgroups of Turkic" [Дыбо А. 1996: 28]. Likewise, there have been discussions of different ways to deal with such situations within the framework of the tree-type model. A. B. Dolgopolsky, for instance, suggested abandoning any attempts to classify closely related languages into subgroups altogether. A. V. Dybo, on the contrary, considers it possible to construct a genealogical tree for such languages, admitting, however, that different classificatory standards (depending on the innovations taken into account) can yield different trees; the main task of the researcher is then to understand which of the resulting trees reflects the most archaic situation.

It should be noted that such 'reticulate evolution' is also encountered in biology [Иорданский 2001: 68]. Closely related species can (in natural conditions as well as in captivity) mate with each other and even give fertile offspring; this has been attested, e. g., for multiple species of macaques and some species of baboons [Бутовская, Файнберг 1993: 100–102 с лит.]. Situations where two obviously different languages ‘transmorph’ into each other through a chain of intermediate dialects also find their analogies in biology. Thus, populations of the great tit form an unbreaking continuum from Western Europe to the Far East, or, to be more precise, two unbreaking chains, one of which goes through temperate forests and the other one through Persia and tropical forests of India and Indochina. In each two adjacent populations, the species can mate with each other, but on the Amur river, where the two chains meet once again, tits from the northern chain cannot mate with those from the southern one (they are respectively called *Parus major major* and *Parus major minor*) [Иорданский 2001: 68].

Reticulate evolution (regrouping of dialects) confirms with the principle of language transmission from competent speakers to their children; nevertheless, it is still difficult to reconcile it with the tree-type model.

We also know the conditions under which dialectal continua can be subject to such evolution. This is possible only with closely related languages each of which lies within a zone of ‘potential intelligibility’ with its neighbour [Бурлак, Старостин 2005: 62, 156–162], since ‘bilingual situations make it hard for the speakers to be fully competent in both systems without mixing them up’ [Алпатов 1997: 111]. Likewise, in biology reticulate evolution is only possible for those neighbouring populations whose members can produce fertile hybrids. Moreover, it cannot be excluded that the initial stage of evolution both in biological species and in languages *always* has a reticulate nature, with the ‘tree’ observed only after natural selection or historical and cultural factors have led to the extinction of some part of the populations, with the others forming distinctly different entities.

The biological part of the article contains a certain number of flaws. Apart from not mentioning the possibility of reticulate evolution among subspecies and closely related species, Fleischhauer says nothing about the fact that *allopatric* formation of subspecies (divergence of species caused by geographical separation of populations) can, to some degree of probability, alternate with *sympatric* formation, when representatives of the same species, through different approaches to the exploiting of natural resources, can eventually give rise to several different species incapable of mating between themselves. Basically, the cladistic approach to classification, dealing with the splitting of one ancestor into several descendants rather than mere separation of one entity from another through branching, suits the needs of linguistics better than those of biology. Languages evolve constantly, regardless of whether they diverge into several descendants or not — differences between Old and Modern Greek are hardly any less significant than those between Latin and Italian or Spanish. If the ancestral language does split, all of its descendants evolve at generally comparable rates (especially if one chooses basic lexicon as the primary criterion, as is done in glottochronology). In biology, however, species can remain quite stable for extremely long periods of time — e. g., the tadpole shrimp (*Triops cancriformis*) that has been practically unchanged for the last 200 mln. years [Иорданский 2001]. This means that separation of certain populations with subsequent formation of a new species does not necessarily imply the loss of the ‘elder’ species or its transmutation into a qualitatively new entity, as the cladistic method would suggest. Well-argumented criticism of the cladistic principle in biology has been widespread in Russian science for quite a long time already [Принципы... 1989]; in the West, however, this approach still remains fashionable.

There are also certain misses in the comparative part of the article. Passages that deal with the construction of the genealogical tree do not mention the glottochronological methodology of S. A. Starostin (even though an officially published English translation of the principal work on this methodology [Starostin 2000] does exist). This approach includes, among other things, a way of filtering out phenomena that are due to contact between closely related languages [Бурлак, Старостин 2005: 164]. Fleischhauer says nothing about the methodology of separating borrowings from inherited strata in the language [Бурлак, Старостин 2005: 72–79], even though such separation allows for a much more precise application of the monophyletic principle — although, under the condition of particularly intensive contacts between languages, the dating of their divergence can be made somewhat more archaic [Бурлак 2000; Бурлак, Старостин 2005: 67]. One thing that is straightforwardly incorrect is inclusion of common archaisms (‘plesiomorphies’) among those units that are subject to borrowing.

The work states that a theory on the relations between inheritance and borrowing in language history is highly desirable; those wishing to get acknowledged with such a theory can be pointed towards works by V. I. Belikov [Беликов 1989, 2006], with a detailed and well-grounded classification of language contacts, as well as the handbook on comparative linguistics by S. A. Burlak and S. A. Starostin [Бурлак, Старостин 2005: 55–82] and the article [Бурлак 2007], where influence of contact factors on the historical development of languages is also discussed, and the tree model for language groups is deemed useful. It should be noted that creole languages — ones that do not, in fact, fit well within the tree-type model — are encountered unfrequently and have enough systematic and well recognizable differences from non-creole languages to be excluded from the tree-type model [Бурлак 2008]. It is also likely that proper conditions for the formation of creoles are forming only on 'advanced' stages of social development, characteristic exclusively of relatively modern times (A. Davletshin, p. c.).

Despite all this, I have read the article with interest, since any kind of publication dealing with issues of comparative-historical linguistics that have not yet been elaborated to everyone's satisfaction should be welcome, not to mention publications that also attempt to deal with the solution of similar problems in evolutionary biology.

Jens Fleischhauer
Institute of Language and Information,
Heinrich-Heine University, Düsseldorf

Reply to Vladimir Belikov and Svetlana Burlak

Vladimir Belikov and Svetlana Burlak express some quite interesting comments on my paper. Some of them I want to discuss — especially those that highlight the discrepancy between our views. I want to use this reply to put stress on some points of my paper that I consider as central.

In my paper I compare the comparative method and phylogenetic cladistics to explore the statement that there is an analogy between both methods. Belikov states that the analogy is merely an analogy, while my conclusion is that both are identical in methodological respect. So, I claim that it is more than just an analogy since both methods work in the same way. Based on this, I agree with Platnick and Cameron's claim that comparative method is a cladistic method. This also means that both face the same methodological problems.

Belikov claims that the classifications in biology and linguistics are different, because there are differences in comparability of objects in both disciplines. This is how I interpret his statement that "the level of abstraction in objects classified by biologists and linguists is quite different". I do not want to claim that biological and linguistic objects (e.g. species and languages) obey the same evolutionary regularities. Comparability of objects in the domains of biology and linguistics seems often to be assumed. If one uses biological methods to analyze linguistic data (e.g. Bryant et al. 2005; Hamed & Wang 2006), one presupposes that assumptions inherent in the biological methods apply well to linguistic objects. It is relevant to make a comparison between the domains of biology and linguistics if one wants to apply a method from one discipline in another discipline. But in my paper I neither compare the objects in the respective domains, nor do I transfer a method from biology into linguistics. All I do is to compare two methods and state that both share the same assumptions (monophyly and evolution). The argumentation merely relies on a comparison of the methods, it does not cling to a comparability between the respective objects classified in biology and linguistics. Just in facing the methodological problems, it is necessary to take the objects classified into account. To test if the basic assumptions of the methods hold, one has to take a look on the objects (species and languages) themselves. But this does not means that the objects have to be compared, they can separately violate the assumptions.

As should be clear (at least now), I do not want to defend cladism. My aim is to present methodological reasons against the view that an a priori assumed tree-model is adequate for representing language relationships. This should not mean that the monophyly assumption is wrong in any case, but that it has to be rejected as a solid basic assumption. That is also the reason why I do not discuss „the methodology of separating borrowing from inherited strata in the language", as Burlak says. It may be right that such an approach would allow a more precise application of the monophyletic principle. But again, I guess the discussion of the basic assumptions of cladistic methods shows that there is a problem with an a priori presumption of monophyly. For my argumentation it would not bring any new into the discussion in considering Starostin

(2000). Starostin takes language divergence and genealogical trees as axiomatic, so his account faces the same problems than any other account which assumes monophyly. In my view it does not help to increase the methodology for a more precise application of the monophyly principle, the only reasonable solution is the deletion of this as an basic assumption.

Burlak states that there are some flaws in the biological part of the paper. She mentions that I said nothing about *allopatric* and *sympatric* formation of subspecies. This is right, but I neither discuss speciation nor compare models of speciation (divergence) in biology and linguistics. In the light of my paper, the way speciation or branching occurs is only relevant in respect to those cases in which language origin is the outcome of an hybridization event. There are a lot of cases in which other models, analogous to an allopatric formation of subspecies, describe the process of language origin quite well. For my argumentation it is only relevant to show that there are cases that do not fit the branching model. I discuss such cases on two levels: first, on the level of origin, second, on the level of development. As Burlak states, a mixed origin is rather seldom, but despite this, it occurs as a natural process. There are such cases as mixed languages (other than Ma'a, there are still more cases regarded as mixed, cf. the discussion in [Matras 2003]), creole languages and dialect continua, so they have to account as conflicting with the monophyly assumption. The cases of the second type deal with (extreme) borrowing and areal features. Burlak and Belikov question the first cases, in saying that Ma'a can be reanalyzed in terms of unbroken transmission, while Burlak says that creoles are quite different from non-creoles. This difference justifies dealing with them in dissimilar ways. Given that Ma'a and creoles can be seen as controversial and, as such, not very well supporting my claim, there are also dialect continua and the second type of cases like 'extreme borrowing' or horizontal transmission of areal features as further problematic cases. Maybe Burlak and Belikov won't accept some of this cases, because they belong to the 'domain' of language development and not language origin. One could state that only language origin and not language development is relevant for classifying languages. As I argue in the paper, if one needs to rely on an evolutionary approach of language change, one needs to take account of the full evolutionary history of languages and cannot purely focus on language origin. An evolutionary account of language change, I argue, is one of the basic requirements of comparative (cladistic) method.

Maybe Burlak and I differ in respect of what should be grasped by the comparative method: 'merely' language origin or language development as well. I focus on the second alternative, which is why I do not discuss speciation or methods like glottochronology. If one accepts the basic assumption of monophyletic origin of languages, it would be quite important to refine your methods to improve your data base, so that it fits the monophyly assumption much better. But if one rejects this assumption, methods that improve your data so that it better fits the tree model do not seem to be so important any longer. Nevertheless, this makes the distinction between inherited and non-inherited language features not any longer superfluous.

The conclusion of my paper is that there is a dilemma in respect to the basic assumptions of the method. Both of the basic assumptions cannot hold together, but it is only one of them (monophyly) that can be abandoned. This means that procedures of analyzing relationships between languages should not take the tree-likeness as axiomatic. One cannot grasp the evolutionary process underlying the development and origin of languages if one puts languages into a certain model, namely a tree-model, if this model does not fit the historical development of the languages. The comments of Burlak and Belikov mostly lead into the direction of reanalyzing data and improving the methods, so that it could be said that the languages, if they can be genetically classified, eventually fit the monophyly assumption or no longer present any evidence against it. The main point of my argumentation was that one needs to take account of language evolution for the philosophical coherence of the method. But language evolution is reticulate, and the conflict that arises from this can only be solved by putting the monophyly requirement aside.

Literature

Aikhenvald 2001 — AIKHENVALD, A. (2001). Areal diffusion, genetic inheritance and problems of subgrouping: a North Arawak case study // AIKHENVALD, Alexandra, & R. M. W. DIXON (Eds.), *Areal Diffusion and Genetic Inheritance* (pp. 167–194). Oxford: Oxford University Press.

Altmann 1976 — ALTMANN, G. (1976). Homologiekriterien in der Linguistik und Biologie // W. SCHÄFER (Ed.), *Evolutionierende Systeme I und II*. Frankfurt am Main: Verlag Waldemar Kramer.

Arnold 1997 — ARNOLD, M. L. (1997). *Natural Hybridization and Evolution.* Oxford: Oxford University Press.

Atkinson, Gray 2005 — ATKINSON, Q. D., & GRAY, R. D. (2005). Curious Parallels and Curious Connections — Phylogenetic Thinking in Biology and Historical Linguistics // *Systematic Biology,* 54(4), pp. 513–526.

Bloomfield 1933 — BLOOMFIELD, L. (1933). *Language.* London: Georg Allen & Unwin.

Bryant et al. 2005 — BRYANT, D., FILIMON, F., & GRAY, R. D. (2005). Untangling our past: Languages, Trees, Splits and Networks // R. MACE (Ed.), *The evolution of cultural diversity* (pp. 67–84). London: UCL Press.

Chappell 2001 — CHAPPELL, H. (2001). Language Contact and Areal Diffusion in Sinitic Languages // AIKHENVALD, Alexandra, & R. M. W. DIXON (Eds.), *Areal Diffusion and Genetic Inheritance* (pp. 328–357). Oxford: Oxford University Press.

Chrétien 1963 — CHRÉTIEN, C. Douglas. (1963). Shared innovations and Subgroupings // *International Journal of American Linguistics, 29, pp.* 66–68.

Croft 2000 — CROFT, W. (2000). *Explaining language change: an evolutionary approach.* Harlow, Essex: Longman.

Croft 2006 — CROFT, W. (2006). The relevance of an evolutionary model to historical linguistics // O. NEDERGAARD THOMASON (Ed.), *Competing models of linguistic change. Evolution and Beyond* (pp. 91–132). Amsterdam: John Benjamins Publishing.

Dawkins 1989 — DAWKINS, R. (1989). *The selfish gene.* Oxford: Oxford University Press.

Dench 2001 — DENCH, A. (2001). Descent and Diffusion: The complexity of the Pilbara Situation // AIKHENVALD, Alexandra, & R. M. W. DIXON (Eds.), *Areal Diffusion and Genetic Inheritance* (pp. 167–194). Oxford: Oxford University Press.

Dixon 1997 — DIXON, R. M. W. (1997). *The rise and fall of languages.* Cambridge: Cambridge University Press.

Dyen 1953 — DYEN, I. (1953). Malgache et maanjan: Une comparaison linguistique // *Language, 29*(4), pp. 577–591.

Dyen 1971 — DYEN, I. (1971). The Impact of Lexicostatistics on Comparative Linguistics // I. DYEN (Ed.), *Lexicostatistics in Genetic Linguistics. Proceedings of the Yale Conference, Yale University, April 3–4, 1971*. The Hague: Mouton.

Goodman 1971 — GOODMAN, M. (1971). The strange case of Mbugu // D. HYMES (Ed.), *Pidginization and creolization of languages*. Cambridge: Cambridge University Press.

Greenberg 1957 — GREENBERG, J. (1957). *Essays in Linguistics.* Chicago: The University of Chicago Press.

Greenberg 2005 — GREENBERG, J. (2005). Are there mixed languages? // J. GREENBERG (Ed.), *Genetic Linguistics. Essays on Theory and Method*. Oxford: Oxford University Press.

Hamed & Wang 2006 — HAMED, M. Ben, & WANG, F. (2006). Stuck in the forest — Trees, Networks and Chinese dialects // *Diachronica, 23*(1), pp. 29–60.

Haspelmath 2004 — HASPELMATH, M. (2004). How hopeless is genealogical linguistics, and how advanced is areal linguistics? // *Studies in Language, 28*(1), pp. 209–223.

Heine & Kutewa 2001 — HEINE, B., & KUTEWA, T. (2001). Convergence and Divergence in the Development of African Languages // AIKHENVALD, Alexandra, & R. M. W. DIXON (Eds.), *Areal Diffusion and Genetic Inheritance* (pp. 393–411). Oxford: Oxford University Press.

Hempel 1965 — HEMPEL, C. Gustav. (1965). Fundamerntals of Taxonomy // C. Gustav HEMPEL (Ed.), *Aspects of Scientific Explanation*. New York: The Free Press.

Hempel 1969 — HEMPEL, C. Gustav. (1969). *Fundamentals of Concept Formation in Empirical Science* (10th ed.). Chicago: The University of Chicago Press.

Hennig 1950 — HENNIG, W. (1950). *Grundzüge einer Theorie der phylogenetischen Systematik.* Berlin: Deutscher Zentralverlag.

Hennig 1984 — HENNIG, W. (1984). *Aufgaben und Probleme stammesgeschichtlicher Forschung.* Berlin: Paul Parey.

Holden & Gray 2006 — HOLDEN, C., & GRAY, R. (2006). Rapid Radiation, Borrowing and Dialect Continua in the Bantu Languages // P. FORSTER, & C. RENFREW (Eds.), *Phylogenetic methods and the prehistory of languages*. Cambridge: McDonald Institute for Archaelogical Research.

Hull 1988 — HULL, D. (1988). *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science.* Chicago: The University of Chicago Press.

Janning & Knust 2004 — JANNING, W., & KNUST, E. (2004). *Genetik.* Stuttgart: Thieme.

Jeffers & Lehiste 1979 — JEFFERS, R. J., & LEHISTE, I. (1979). *Principles and Methods for Historical Linguistics.* Cambridge: MIT Press.

Judd et al. 2002 — JUDD, W. S., CAMPBELL, C. S., KELLOGG, E. A., STEVENS, P. F., & DONOGHUE, M. J. (2002). *Plant Systematics: A Phylogenetic Approach.* Sunderland: Sinauer Associates.

Lass 1997 — LASS, R. (1997). *Historical Linguistics and Language Change.* Cambridge: Cambridge University Press.

Marten 2006 — MARTEN, L. (2006). Bantu classification, Bantu trees and phylogenetic methods // P. FORSTER, & C. RENFREW (Eds.), *Phylogenetic methods and the prehistory of languages*. Cambridge: McDonald Institute for Archaelogical Research.

Matras 2003 — MATRAS, Yaron (Ed.) (2003). *The mixed language debate*. Berlin: Mouton de Gruyter.

Mayr 2001 — MAYR, E. (2001). The Biological Species Concept // Q. D. WHEELER, & R. MEIER (Eds.), *Species Concepts and Phylogenetic Theory. A Debate*. New York: Columbia University Press.

Mayr 2003 — MAYR, E. (2003). *The Grwoth of Biological Thought: Diversity, Evolution and Inheritance.* Cambridge: Harvard University Press.

McDade 1992 — MCDADE, L. (1992). Hybrids and Phylogenetic Systematics II: The Impact of Hybrids on Cladistics // *Evolution, 46*(5), pp. 1329–1346.

Mous 2003 — MOUS, M. (2003). *The Making of a Mixed Language: The case of Ma'a/Mbugu.* Amsterdam: John Benjamins Publishing.

Nakhleh et al. 2005a — NAKHLEH, L., RINGE, D., & WARNOW, T. (2005a). Perfect Phylogenetic Networks: A new methodology for reconstructing the evolutionary history of natural languages // *Language, 81*(2), pp. 382–420.

Nakhleh et al. 2005b — NAKHLEH, L., WARNOW, T., LINDER, C., & ST. JOHN, K. (2005b). Reconstructing Reticulate Evolution in Species — Theory and Practice // *Journal of Computational Biology, 12*(6), pp. 796–811.

Nichols 2001 — NICHOLS, R. (2001). Gene trees and species trees are not the same // *Trends in Ecology & Evolution, 18*(7), pp. 358–364.

Pagel 2000 — PAGEL, M. (2000). The History, Rate and Pattern of World Linguistic Evolution // C. KNIGHT, STUDDERT-KENNEDY, MICHAEL, & J. R. HURFORD (Eds.), *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*. Cambridge: Cambridge University Press.

Pagel & Meade 2006 — PAGEL, M., & MEADE, A. (2006). Estimating Rates of Lexical Replacement on Phylogenetic Trees of Languages // P. FORSTER, & C. RENFREW (Eds.), *Phylogenetic methods and the prehistory of languages*. Cambridge: McDonald Institute for Archaelogical Research.

Platnick 1979 — PLATNICK, N. (1979). Philosophy and the transformation of cladistics // *Systematic Zoology, 28,* pp. 537–546.

Platnick & Cameron 1977 — PLATNICK, N., & CAMERON, D. (1977). Cladistic Methods in Textual, Linguistic, and Phylogenetic Analysis // *Systematic Zoology, 26*(4), pp. 380–385.

Remane 1956 — REMANE, A. (1956). *Die Grundlagen des natürlichen Systems der vergleichenden Anatomie und der Phylogenetik* (2nd ed.). Leipzig: Akademische Verlagsgesellschaft.

Ridley 1986 — RIDLEY, M. (1986). *Evolution and Classification: The Reformation of Cladism.* London: Logman.

Ringe et al. 2002 — RINGE, D., WARNOW, T., & TAYLOR, A. (2002). Indo-European and computational cladistics. *Transactions of the Philological Society, 100*(1), pp. 59–129.

Ritt 2004 — RITT, N. (2004). *Selfish Sounds and Linguistic Evolution.* Cambridge: Cambridge University Press.

Stamos 2003 — STAMOS, D. N. (2003). *The Species Problem: Biological Species, Ontology, and the Metaphysics of Biology.* Lanham: Lexington Books.

Starostin 2000 — STAROSTIN, S. A. (2000). Comparative-historical linguistics and lexicostatistics // RENFREW, C.; MCMAHON, A. & TRASK, L. (Ed.) *Time Depth in Historical Linguistics* (Volume 1, pp. 223–265). Cambridge: The McDonald Institute for Archaeological Research.

Thomason 1983 — THOMASON, S. (1983). Genetic Relationship and the case of Ma'a (Mbugu) // *Studies in African Linguistics, 14*(2), pp. 195–231.

Thomason 1997 — THOMASON S. (1997). Ma'a (Mbugu) // THOMASON S. (ed.) *Contact languages: A wider perspective.* Amsterdam: John Benjamins.

Thomason 2001 — THOMASON, S. (2001). *Language Contact: An Introduction.* Washington D.C.: Georgetown University Press.

Thomason 2002a — THOMASON, S. (2002a). Creoles and Genetic Relationship. *Journal of Pidgin and Creole Languages, 17(1)*, pp. 101–109.

Thomason 2002b — THOMASON, S. (2002b). *Cladistics and Reticulate Processes in Language Change and Diversification.* University of Michigan ms.

Thomason & Kaufman 1988 — THOMASON, S., & KAUFMAN, T. (1988). *Language Contact, Creolization, and Genetic Linguistics.* Berkeley: University of California Press.

Wägele 2001 — Wägele, J.-W. (2001). *Grundlagen der Phylogenetischen Systematik* (2nd ed.). München: Verlag Dr. Friedrich Pfeil.

Warnow et al. 2006 — Warnow, T., Evans, S.; Ringe, D.; Nakhleh, L. (2006). A Stochastic Model of Language Evolution that Incorporates Homoplasy and Borrowing // P. Forster, & C. Renfrew (Eds.), *Phylogenetic methods and the prehistory of languages*. Cambridge: McDonald Institute for Archaelogical Research.

Whinnom 1971 — Whinnom, K. (1971). Linguistic hybridization and the 'special case' of pidgins and creols // D. Hymes (Ed.), *Pidginization and creolization of languages*. Cambridge: Cambridge University Press.

Алпатов 1997 — Алпатов В.М. *150 языков и политика: 1917–1997. Социолингвистические проблемы СССР и постсоветского пространства.* — М., 1997.

Беликов 1989 — Беликов В.И. Древнейшая история и реальность лингвогенетических дендрограмм // *Лингвистическая реконструкция и древнейшая история Востока*. Материалы к дискуссиям на международной конференции. Часть 1. — М., 1989.

Беликов 2006 — Беликов В.И. *Конвергентные процессы в лингвогенезе* (диссертация в виде научного доклада, составленная на основе опубликованных работ, представленная к защите на соискание ученой степени доктора филологических наук). — М., 2006.

Бурлак 2000 — Бурлак С.А. История тохарских языков в свете данных глоттохронологии // *Проблемы изучения дальнего родства языков на рубеже третьего тысячелетия*. Доклады и тезисы международной конференции. — М., 2000.

Бурлак 2007 — Бурлак С.А. Языковой сдвиг и теория компаративистики // *Языковые изменения в условиях языкового сдвига*. Сборник статей. Отв. ред. Н.Б. Вахтин. — СПб, 2007.

Бурлак 2008 — Бурлак С.А. Креольские языки и глоттохронология // *Otrientalia et classica: Труды Института восточных культур и античности*. Вып. XIX. Аспекты компаративистики. 3. (под ред. И.С. Смирнова). — М., 2008.

Бурлак, Старостин 2005 — Бурлак С.А., Старостин С.А. *Сравнительно-историческое языкознание*. — М., 2005.

Бутовская, Файнберг 1993 — Бутовская М.Л., Файнберг Л.А. *У истоков человеческого общества*. — М., 1993.

Дыбо 1996 — Дыбо А.В. *Семантическая реконструкция в алтайской этимологии. Соматические термины (плечевой пояс).* — М., 1996.

Иорданский 2001 — Иорданский Н.Н. *Эволюция жизни*. — М., 2001.

Коменский 1941 — Коменский Я. А. *Мир чувственных вещей в картинках или Изображение и наименование всех главнейших предметов в мире и действий в жизни*. М.: Учпедгиз, 1941. (Russian translation and the original Latin text of: *Joh. Amos Commenii Orbis sensualium pictus… Noribergae, 1658*).

Николаев 1988 — Николаев С.Л. Следы особенностей восточно-славянских племенных диалектов в современных великорусских говорах. 1. Кривичи // *Балто-славянские исследования 1986*. — М., 1988.

Николаев 1989 — Николаев С.Л. Следы особенностей восточно-славянских племенных диалектов в современных великорусских говорах. 1. Кривичи (окончание) // *Балто-славянские исследования 1987*. — М., 1989.

Порхомовский 1982 — Порхомовский В. Я. Проблемы генетической классификация языков Африки // *Теоретические основы классификации языков мира. Проблемы родства*. М.: Наука, 1982.

Принципы и методы зоологической систематики. Труды ЗИН АН СССР, т. 209. — Л., 1989.

Система организмов // *Биологический энциклопедический словарь*, М.: Советская энциклопедия, 1989.

Шаповал 2007 — Шаповал В. В. *Самоучитель цыганского языка (Русска рома: севернорусский диалект).* М.: АСТ; Астрель, 2007.

В рамках исторической лингвистики языки, как предписывает сравнительно-исторический метод, классифицируются на основании их генетического родства. Аналогичным инструментом в биологии выступает кладистика. В статье утверждается, что биологический и лингвистический методы не только аналогичны, но и основаны на одних и тех же исходных положениях: (1) монофилетическое происхождение; (2) эволюционное развитие видов и языков. Утверждается также, что положение о монофилетическом происхождении применительно к языкам неверно, а эволюционное объяснение происхождения и развития языка не является интегральной частью сравнительного метода

и связанной с ним «древесной» моделью. Исходя из этого, высказывается ряд методологически обоснованных критических замечаний относительно применения к языковому материалу такой «древесной» модели.

Основным выводом статьи можно считать то, что эволюционное объяснение развития языка несовместимо с положением о его монофилетическом происхождении; для разрешения возникающего противоречия от последнего положения следует отказаться. Это, в свою очередь, будет иметь важные последствия как для оценки достоверности имеющихся на сегодняшний день языковых классификаций, так и для разработки более точных методик классификации в будущем.