

University of Tartu  
Department of Semiotics

Jorge González Fernández

**Semiotic Aspects of  
Historical-Comparative Linguistics  
and Linguistic Paleontology**

Master Thesis

Supervisors:

Prof. Peeter Torop (University of Tartu) and  
Prof. Vartan Kazaryan (Moscow State University)

Tartu  
2015

I have written the Master Thesis myself, independently. All of the other authors' texts, main viewpoints and all data from other resources have been referred to.

Author: Jorge González Fernández

*Signature:* .....

*Date:* .....



## Table of contents

<b>INTRODUCTION.....</b>	<b>3</b>
Overview and objectives of work.....	3
Historiography and theoretical framework.....	5
<b>PART I: Semiotic aspects of the investigations in historical and comparative linguistics .....</b>	<b>8</b>
1. The discipline of historical and comparative linguistics.....	9
1.1. The historical-comparative method.....	11
1.2. Problematic questions concerning the investigations by historical and comparative linguistics.....	13
2. The <i>syntactic dimension</i> of linguistic semiotic systems.....	20
2.1. The structuralist approach in linguistics.....	21
2.2. The semiotic approach to natural language.....	24
Conclusive remarks: the object of study and the ultimate scope of historical and comparative linguistics.....	31
<b>PART II: Semiotic aspects of the investigations in linguistic paleontology .....</b>	<b>34</b>
3. The discipline of linguistic paleontology.....	35
3.1. Linguistics as a discipline for investigating the past.....	36
3.2. Problematic questions concerning the investigations within linguistic paleontology.....	40
4. The informativeness of linguistic data about prehistoric past.....	45
4.1. The role of language inside culture from the point of view of the Tartu-Moscow school's semiotics of culture.....	46
4.2. Linguistic paleontology as a field for interpretative investigations.....	52
Conclusive remarks: the validity of linguistic arguments for the investigation of human past.....	54
<b>CONCLUSIONS .....</b>	<b>57</b>
<b>List of references .....</b>	<b>59</b>
<i>Ajaloolis-võrdleva lingvistika ja lingvistilise paleontoloogia semiootika aspektid</i> Resüme .....	64
<i>Семиотические аспекты сравнительно-исторического языкознания и лингвистической палеонтологии</i> Резюме .....	66

## INTRODUCTION

### Overview and objectives of work

The present work will discuss from a semiotic theoretical approach the investigations within the disciplines of *historical and comparative linguistics* and *linguistic paleontology* with the aim of explaining certain problematic questions that are a source for debate inside these disciplines.

The discipline of *historical and comparative linguistics* investigates the diachronic dimension of languages and proceeds to the comparison of languages with two purposes: establishing kinship relations between them and reconstructing protolanguages. The *historical-comparative method* developed within this discipline seeks for systematicity in the correspondences observable between kinship related languages, explains them in terms of exceptionless laws, and attempts to deductively reconstruct ancestor protolanguages. Because of its aim for systematicity, it is claimed to be a scientific and objective methodology that leads to verifiable conclusions, however, despite this claimed objectivity, many aspects of the explicative laws and the hypothetical reconstructions are a matter for debate between comparatists and they are under constant reformulations. On the other hand, the application of the *historical-comparative method* is not equally efficient in reaching exhaustiveness and systematicity at all the different levels of languages, being the most effective at the phonetic level, less effective at the morphemic level and much more problematic at the lexical level and at the syntax level.

In its part, the discipline of *linguistic paleontology* assumes the informative capacity of languages about the extra-linguistic reality and attempts to investigate non-linguistic aspects of the past of human cultures by analyzing the linguistic data provided by historical and comparative linguistics. The arguments resulting from this kind of investigations are a matter of constant debate between scholars and, in addition, this discipline has been a profuse source of pseudo-scientific theories about the past of societies. This calls into question the ultimate validity of the arguments resulting from the investigations of linguistic paleontology.

In view of the presented epistemological difficulties that concern the ultimate informativeness about human past of these two disciplines, our work will discuss some important aspects of their methodological and theoretical approaches from the point of view of semiotics, with the conviction that the discussion of the semiotic aspects involving these investigations is elucidative of the problematic questions presented above. Our discussion concretes in three main points:

- 1) We argue that it is the structuralist approach to the study of languages what enables objectivity and verifiability in the historical-comparative method. Based on this postulation, we will discuss what features of linguistic semiotic systems are actually object for systematic comparison in the *historical-comparative method*, what is the ultimate nature of the reconstructions of proto-languages resulting from this comparative work and what is the explanation for the unequal effectiveness of such methodology at the different levels inside linguistic systems.
- 2) We turn to the Tartu-Moscow School's semiotics of culture for discussing the informativeness of linguistic data about the extra-linguistic reality as ground for the approach of linguistic paleontology in its investigation of human past.
- 3) And in the last place we will expose the implications of our semiotic theoretical approach for delimiting the ultimate entity of the knowledge about ancient and prehistoric past provided by the disciplines of *historical and comparative linguistics* and *linguistic paleontology*.

Our theoretical approach is a semiotic one and departs from the postulation that, when conceiving natural languages as semiotic system, the systematicity and structuredness that is reached by the structuralist linguistic approach to languages scope exclusively what we call the *syntactic dimension* of linguistic semiotic systems. This means that only the relations of linguistic signs with other linguistic signs inside a linguistic semiotic system is describable in terms of systematicity and structuredness, while the *semantic* and the *pragmatic dimensions* of the system are not equally systematically and structurally describable. Based on this initial postulation, we will argue that if the *historical-comparative method* developed within *historical and comparative linguistics* can be claimed to be a scientific and verifiable one is because its ultimate object for comparison are the systematic and structured descriptions of natural languages provided by the structuralist linguistic approach, and we will conclude that the investigations within this linguistic discipline are suitable for verification and objectivity insofar as its object for study consists on the *syntactic dimension* of linguistic semiotic

systems, being this explicative of the fact that the *historical-comparative method* is unequally efficient in its approach to the different levels of linguistic systems.

With regard to *linguistic paleontology*, we cannot accept that its investigations are suitable for verification and objectivity in itself because they consist on an interpretation of linguistic data, and for that reason they are easily suitable of misinterpretation and overinterpretation of the linguistic evidence. We find nonetheless a theoretical legitimation of its methodological approach when it assumes a capacity for informativeness in linguistic data about the past in the proposal of the Tartu-Moscow School's semiotics of culture and its understanding of language as a semiotic system that operates inside culture, in interrelation to all other cultural semiotic systems, and occupies a central role inside the whole system of culture. In spite of this ultimate legitimation at the theoretical level, and due to the inevitably interpretative character of its investigations, the arguments provided by linguistic paleontology are not a source of objective and irrefutable knowledge about the past, but a source of plausibility in our hypothesizing about it; it needs confrontation with the data provided by other disciplines that also deal with the investigation of human past in order to provide reliable knowledge.

### **Historiography and theoretical framework**

Several authors will be covered along our work inasmuch as they stand as main referents in the linguistic disciplines that are the matter our discussion. The author that we must mention in the first place is Vyacheslav Ivanov, who is one of the most outstanding scholars within the disciplines of historical and comparative and linguistics and linguistic paleontology nowadays, as well as one of the founding members of the Tartu-Moscow school of semiotics. His work in theoretical semiotics is influenced by his investigations within these disciplines, and his contribution to the field of the Indo-European historical and comparative linguistics is a fundamental one, being the developer – together with Tamaz Gamkrelidze – of the *glottalic theory*, which is a proposal that has changed the whole state of the arts of the reconstruction of the Proto-Indo-European consonant phonetic system (see Gamkrelidze and Ivanov 1995 [1984]). Another reason for its importance as referent for our work is his protagonist role in the initial development of the Tartu-Moscow school *semiotics of culture*, which postulations are the main theoretical framework of our discussion.

A direct historiographical precedent to our work is the introductory and compelling volume by Raimo Anttila *Historical and comparative linguistics* (1989 [1972]), which also carries out an implementation of semiotic theories and terminology for explaining the methodological approach of that discipline. His approach shares therefore a similar purpose to the one of our present work, but he focuses on the Peircean triad of kinds of referentiality – *iconicity, indexicality* and *symbolicity* –, while our approach focuses on Charles W. Morris' distinction (on the basis of an also Peircean triad) of three dimensions in semiotic analyses – *syntactical, semantical* and *pragmatical*. We cannot compare our work to the one offered by Raimo Anttila because his offers much more detailed and illustrated exemplifications and a much more extension and scope of topics. We find in his work a supporting complementary approach to the one implemented in our discussion that focuses on different semiotic aspects of these disciplines but leads to conclusions that are compatible with the ones offered by us. His analysis will be for this reason a main referent in our work, especially as a source of more competent support of our theoretical postulations.

As have already been mentioned, the discipline of *semiotics of culture* developed by the Tartu-Moscow school will be primarily important as the theoretical background to be implemented in our discussion. The understanding of *culture* and *language* as semiotic systems, the conception of the role of language inside culture and the theorization on semiotic systems developed by the scholars of this school guide our theoretical discussion of the disciplines of historical and comparative linguistics and linguistic paleontology. In concrete, we will pay special attention to the semiotic theory of Yurij Mihailovich Lotman as the most important author within this school, taking into account as well the contributions relevant to our topic of other important historical members of the school: the already mentioned V. Ivanov, B. Uspenskij, I. Revzin and V. Toporov, who was as well one of the main investigators of the Indo-European culture. The *Theses on the semiotic study of cultures (as applied to Slavic texts)* (Ivanov; Lotman et al. 1973) will be taken as the main concretion of the contributions of the Tartu-Moscow School's *semiotics of culture*.

The last main referential framework in our theoretical approach will be the semiotic understanding of natural language and the theories on linguistic semiotic systems carried out by Soviet Semiotic scholars: the semiotic theory of language of Sebastian Shaumyan, who would also be a main reference for the Tartu-Moscow scholars, and the contributions of several Soviet scholars concerning the characterization of natural language as a semiotic system, as crystalized in the compendiums: *Симпозиум по структурному изучению*



*знаковых систем. Тезисы докладов (Rylina (red.) 1962), Лингвистические исследования по общей и славянской типологии (Nikolaeva (red.) 1966), Структурная типология языков (Ivanov (ed.) 1966), Материалы к конференции «Язык как знаковая система особого рода» (Serebrennikov (ed.) 1967), and Принципы типологического анализа языков различного строя (Uspenskij (red.) and Revzina (comp.) 1972).*

As a last preliminary remark, we must add that the actual examples of theories illustrative of the work within the disciplines under discussion are taken from the concrete case of the investigation of the Indo-European linguistic family because of our lack of detailed knowledge about such investigations for other linguistic groups. However, the general remarks and the considerations about the methodology of both disciplines are anyway applicable to the whole of their investigations regardless of the concrete linguistic family object of study.

## **PART I: Semiotic aspects of the investigations in historical and comparative linguistics**

In the present chapter we will discuss some relevant aspects concerning the object of study, methodology and final results of the investigations in the discipline of *historical and comparative linguistics* from a semiotic theoretical approach. Our initial postulation is that the historical and comparative study of languages is heir of the structuralist approach in linguistics, and that the scientific nature of the historical-comparative method is possible due to the systematicity and structuredness that this structuralist approach achieves in its apprehension of language as an object for study. We argue that these systematicity and structuredness are possible in the description of only the *syntactic dimension* of languages when understood as semiotic systems, and therefore the effectiveness of the historical-comparative method occurs specifically within the limits of that dimension. As a demonstration of our claim, we turn to the fact that the highest success within the historical and comparative linguistic approach has occurred for the levels inside language that don't scope outside of the *syntactic dimension*, namely the phonetic level and the morphemic level of grammatical meanings, while it has not been equally effective for the level of lexical meanings, which already scope the *semantic* and *pragmatic dimensions* of linguistic systems.

## 1. The discipline of historical and comparative linguistics

Before proceeding with the discussion of its problematic aspects, we will briefly describe what does the investigation within *historical and comparative linguistics* consist on. This preliminary information will be necessary for our subsequent theoretical discussion, and it is especially relevant for the readers that are not familiar with this field of studies.

*Historical and comparative linguistics* is the discipline within linguistics that investigates languages diachronically; it studies their evolution through time<sup>1</sup>. This field of studies originated within the investigations of the Indo-European linguistic family that started at the end of the XIXth century with the aim of explaining scientifically the similarities perceived in languages very distant both in time and space of Europe and Asia. At the very first stage of these comparative linguistic studies, the languages under investigation were Sanskrit, Greek and Latin, but very soon similarities between these ancient languages and others, including modern ones, would be noticed and the latest would be thus added to the analysis, conforming the Indo-European linguistic family. The investigations within this linguistic field are *historical* because take as object for study the changes occurring in linguistic systems through time with the aim of explaining the diachronic evolution followed by a certain language to its present stage, and are *comparative* because comparison is the way of establishing genetic and kinship relations between languages.

The comparative approach confronts elements of languages that are attested in historical sources, and permits to grasp which of them are genetically related, conforming linguistic families. But comparison has as final goal the reconstruction of features of hypothetical protolanguages that are unknown to us but which existence has to be postulated as ancestors

---

<sup>1</sup> Although the discipline is almost exclusively referred as either *historical and comparative linguistics* or *comparative-historical linguistics*, also the term *genetic linguistics* has been proposed by Raimo Anttila: “Genetic linguistics is a cover term for both historical and comparative linguistics because both deal with languages showing genetic affinity: historical linguistics treats linear relationships, and comparative linguistics treats collateral relationships” (Anttila 1989: 19).

We will stick to the most widespread between scholars denomination of *historical and comparative linguistics*.

<sup>2</sup> «Основой для установления родства двух и более языков [...] является выявление систематических отношений между множествами лексических и грамматических морфем и их сочетаний в словах. В Genetic linguistics is a cover term for both historical and comparative linguistics because both deal with languages showing genetic affinity: historical linguistics treats linear relationships, and comparative linguistics treats collateral relationships” (Anttila 1989: 19).

We will stick to the most widespread between scholars denomination of *historical and comparative linguistics*.

of the attested and genetically related ones. As stated by Ferdinand de Saussure – who was one of the main contributors to the investigations of the Indo-European linguistic family – “the sole means of reconstructing is by comparing, and the only aim of comparison is a reconstruction. Our procedure is sterile unless we view the relations of several forms from the perspective of time and succeed in reestablishing a single form” (Saussure 1959 [1916]: 218). This comparative work with kinship related languages permits the reconstruction of (parts of) the lexicon, of the phonetic system and of grammatical features of their lost common linguistic stage or protolanguage.

The historical-comparative linguistic work must detect systematic correspondences between the features of the languages under comparison both for demonstrating their kinship relation and for attempting reconstruction: “it is fundamental for establishing the kinship of two or more languages the detection of systematic relations between sets of lexical and grammatical morphemes and of their combinations in words attested in these languages” (Ivanov 2009: 1)<sup>2</sup>. Only if there is systematicity in the correspondences observable between attested languages their genetic kinship can be scientifically demonstrated<sup>3</sup>. What’s more, the higher or lesser degree of systematicity in their correspondences delimits the plausibility reachable at the reconstruction of protolanguages. Those reconstructions may in some cases become a matter of empirical verification (under the appearance of new linguistic data, for example) and this way the previously established hypotheses can be confirmed or rejected, but for most of the cases they remain hypothetical, being the degree of systematicity the only

---

<sup>2</sup> «Основой для установления родства двух и более языков [...] является выявление систематических отношений между множествами лексических и грамматических морфем и их сочетаний в словах, в этих языках засвидетельствованных». (Ivanov 2009: 1) [Translated by the author]

<sup>3</sup> The assumption that the structural and systematic correspondences between languages are the proof of their genetic kinship is rarely called into doubt and stands as the main conviction inside historical and comparative linguistics. However, it has also been called into question. The most critical and pessimist point of view in respect to the reconstructibility of a protolanguage is the one of Nikolaj Trubetskoj, who denies the existence of any proto-Indo-European language and explains the similarities between Indo-European languages as result not of a process of “divergence” from a common ancestor proto-language, but of “convergence”, due to repeated and long contacts between them during their evolutions:

“There is actually no reason that obliges to assume a unique Proto-Indo-European language from which supposedly all Indo-European languages developed. With the same justification we can assume the opposite view about this development, and that means, that the ancestors of the Indo-European branches were initially languages, which were non-similar to each other and only in the course of time, thanks to continuous contacts, got significantly similar to each other by mutual influences and loans, without fully coincide with each other. The history of languages knows both divergent and convergent developments.” (Trubetskoj 1987 [1958]: 45); «Таким образом, нет, собственно, никакого основания, заставляющего предполагать единый индоевропейский праязык, из которого якобы развились все индоевропейские языки. С таким же основанием можно предполагать и обратную картину развития, то есть предполагать, что предки индоевропейских ветвей первоначально были непохожи друг на друга и только с течением времени благодаря постоянному контакту, взаимным влияниям и заимствованиям значительно сблизились друг с другом, однако без того, чтобы вполне совпасть друг с другом. История языков знает и дивергентное и конвергентное развитие.» (Trubetskoj 1987 [1958]: 45) [Translated by the author]

criterion for their plausibility. In fact, in absence of systematicity in the similarities observable between related languages no kind of plausibility or verisimilitude can be assigned to the resulting conclusions concerning their past evolution and their protolanguage.

## 1.1. The historical-comparative method

At the beginning of the development of the discipline of historical and comparative linguistics for the concrete case of the Indo-European linguistic family, a methodology was developed that fixed several criteria for reaching systematicity and accuracy in the comparisons and assure the verisimilitude and plausibility of the resulting conclusions: phonetic laws, laws of change and reconstructions of protolanguages. “After the more than two centuries of existence of comparative linguistics, it has been accumulated not only a huge material of facts related to the plausible establishment of kinships between languages, but also a varied set of theoretical assumptions and presumptions that permit the foundation of this scientific field” (Ivanov 2009: 1)<sup>4</sup>. The resulting methodology is known as the *historical-comparative method*<sup>5</sup>.

The historical-comparative method delimits the procedure for the scientific demonstration of the kinship and genetic relation of a group of languages and for the reconstruction of their hypothetical protolanguage. It requires the detection of systematic correspondences between the related languages and explains them exhaustively in terms of exceptionless *phonetic-laws*. Its aim is to describe the evolution followed by these languages back to the hypothetical and unknown-to-us protolanguage in terms of *laws of change*. Fulfilled this requirements, it is ultimately able to postulate hypothetical reconstructions of protolanguages in a process of logical deduction from the previously established laws.

We must remark that the methodology labeled here as the *historical-comparative method* consists on several concrete procedures. The most important are the ones of *internal* and

---

<sup>4</sup> «За более чем два века существования сравнительного языкознания накоплен не только огромный фактический материал, относящийся к надежно установленным родственным связям между языками, но и достаточно разнообразный набор теоретических предположений и допущений, позволяющих ориентироваться в основаниях этой области науки». (Ivanov 2009: 1) [Translated by the author]

<sup>5</sup> In the history of the discipline the method has been called both *the comparative method* and *the historical-comparative method*, so in this work we will use both names interchangeably. The use of the term “method” indicates the scientific validity recognized for this methodological approach.

*external reconstruction: internal reconstruction* consists on investigating the most ancient forms attestable for every linguistic system under the assumption that the oldest forms are the most informative of the previous non-attested stages of evolution of that language, while *external reconstruction* proceeds to actually confronting these attested forms with their correspondent cognates in other languages. Inside these concrete procedures, certain secondary methodologies are implemented by way of tools for investigation: statistical methods and analogical methods, among others. Having made this clarification, we will further on refer to the whole of such methodological approach as the *historical-comparative method*.

The first methodological approach that fulfilled the scientific requirements of systematicity and exhaustiveness for the comparison of languages was developed by Franz Bopp (1791-1867), who is considered for this reason the founder of the discipline of historical and comparative linguistics, although the main objective of his study of languages was not related to the establishing of such methodology but the demonstration of a concrete thesis concerning the origin of languages (see Bopp 1816 and 1833)<sup>6</sup>. Another contribution to the development of the historical-comparative scientific approach was made by the work of the Neogrammarians at the end of the XIXth century, which put the emphasis on the accumulation of data for every attested language and in the systematic confrontation of these enormous quantities of data, fixing the methodological approach of the discipline.

After its initial development within the Indo-European linguistic family for the reconstruction of the hypothetical Proto-Indo-European language, the historical-comparative method was subsequently implemented and adapted to non Indo-European languages for the establishing of other linguistic families, the description of their evolution and the investigation of features of their protolanguage: “Classical Indo-European comparative grammar created and developed in the 19th century was successful in reconstructing a common ancestor of a whole family of languages. In the late 19th and 20th centuries these methods were applied to most of the languages of the world that gave a picture of their history in the last millenia” (Ivanov 2008 [2007]: 197). The way of proceed is the same for every linguistic family, and the method operates in the same way, although it may adapt to the peculiar typological and structural characteristics of the concrete language family, focusing its comparative approach on some or others linguistic features.

---

<sup>6</sup> A detailed explanation of the contribution of Franz Bopp for establishing the historical and comparative methodology, as well as about the also important contribution of Friedrich Schlegel (1772-1829) can be found in Voloshina 2014.

## 1.2. Problematic questions concerning the investigations by historical and comparative linguistics

The achievement by means of the *historical-comparative method* of, on the one side, systematicity in the comparison and, on the other side, exhaustiveness in the explanation of correspondences between linguistic systems has made of it the prototypical example of a scientific methodology applied to the study of cultural phenomena, and this is proudly claimed by linguists, especially the ones involved in the discipline of historical and comparative linguistics, as we can appreciate in the following statement by Vyacheslav Ivanov – repeated, by the way, in several of his works: “Comparative linguistics still remains the main example of an exact humanitarian semiotic science as its postdictions (conjectures about the history) can be falsified (in Popper’s sense)” (2008: 223 and 2004b: 131-132<sup>7</sup>).

The main case that is always argued as the example of the verifiability in the investigations carried out by historical and comparative linguistics is the discovering in 1927 by the linguist Jerzy Kuryłowicz of a element in Hittite language systematically equivalent of the *sonant coefficients* that had been postulated by Ferdinand de Saussure 50 years earlier – in 1887. On the basis of a systematic comparison employing algebraic functions of the vocalic systems in attested Indo-European languages, Saussure postulated the existence of a phonetic element in Proto-Indo-European language whose concrete phonetic characteristics are not precised but which occupied certain determined positions in cognates and explained the evolution of the vocalic system in the different Indo-European languages. This phonetic element was called by Saussure *the sonant coefficient* in his famous work *Mémoire sur le système primitif des voyelles dans les langues indoeuropéennes* (1879), which was published 50 years before the deciphering of Hittite language and the discovering of its kinship relation with the Indo-European linguistic family. Kuryłowicz discovers in his study of Hittite language the graphic element *h* (transliterated this way from the original cuneiform Hittite writing system), whose concrete phonetic character has not been univocally established but would correspond most likely to an aspirated sound. The importance of his discovering was that Hittite *h* occupies systematically the positions in Indo-European cognates for which Saussure had postulated his sonant coefficient and it was found in this way the empirical evidence that supports the

---

<sup>7</sup> «Реконструкция или “постсказание” [...] может быть проверена или фальсифицирована (в попперовском смысле) при открытии таких новых данных, которые не были известны в то время, когда осуществлялась реконструкция.» (Ivanov 2004b: 131-132) [Translated by the author]

hypothesis that had been reached by Saussure by means of solely the historical-comparative method.

Despite of the claimed scientific nature of the historical-comparative method, the linguistic hypotheses resulting in the form of *sound-laws*, *change-laws* or reconstruction of proto-forms are in most of their aspects a matter of disagreement between comparatists. In fact, the resulting *sound-laws* proposed as exhaustive explanations of the correspondences between languages have been a matter of constant debate and continuous reformulation. We can illustrate this fact with the example of one concrete sound-law, the so-called *Grimm's law*, explanatory of one of the changing processes hypothetically occurred in the evolution of the Indo-European linguistic family. The *Grimm's law* explained a series of changes in stops in Proto-Germanic during the process of dialectization from the common Proto-Indo-European stage: Proto-Indo-European /b/, /d/, /g/ > Proto-Germanic /p/, /t/, /k/; however, several exceptions existed for such rule and, with the aim of explaining them, the *Verner's law* was proposed. This reformulation achieved to explain such exceptions by taking into account accentuation: in non-stressed positions, *Grimm's law* took place, while in stressed positions there were another correspondences, resulting the process of change PIE /b/, /d/, /g/ > PG /p/ or /β/, /t/ or /ð/, /k/ or /γ/<sup>8</sup>. Both laws were broadly accepted and guided the investigations of the Indo-European linguistic family. However, in the formulation of *the glottalic theory* by Tamaz Gamkrelidze and Vyacheslav Ivanov in the 1970s, the previously established *Grimm's law* ceases to be taken as an actually occurred phonetic shift, and under this new paradigm the whole explanation of the dialectalization of the Proto-Indo-European phonetic system changed.

We may argue that such reformulations in the explanatory hypotheses are the result of just a more careful work of comparison with available data, and even of the appearance of new linguistic data; actually, for the concrete example given above, it is very important to acknowledge that the reformulation of the question made by the glottalic theory was developed when Hittite language had been already deciphered and recognized as an Indo-European language, being Hittite an important testimony about Proto-Indo-European consonantic system that the first formulations of *Grimm's* and *Verner's law* could not take into consideration. But they are not uncommon the cases of different theories coexisting for

---

<sup>8</sup> A more detailed explanation about *Grimm's* and *Verner's laws* can be found in *Gamkrelidze and Ivanov 1995 [1984]: 31-36*.



the explanation by different authors of the same linguistic fact. By way of example we can consider *the laryngeal theory*:

The *laryngeal theory* aims to explain the Proto-Indo-European vocalic system, more concretely the appearance of the differentiation of vowels by length – the existence of short and long vowels in Indo-European languages – and the phonetic phenomenon of *ablaut*, which is the regular alternation of vowels employed as a mechanism for word derivation and word formation attested in Indo-European languages. In order to explain systematically these phenomena concerning Indo-European vocalism it was postulated the existence in Proto-Indo-European of laryngeal sounds that occupied the positions of the *sonant coefficients* established by Ferdinand de Saussure<sup>9</sup>. The first problematic question concerning such hypothetical explanation is the number of laryngeals posited by different authors in Proto-Indo-European, which varies from three to four or even six<sup>10</sup>; but the theory itself is a matter of disagreement, with scholars who deny the existence of such laryngeals in Proto-Indo-European. They defend that there is no evidence for postulating such kind of phonemes in the Indo-European family and that the explanation of Indo-European vocalism should rather pay attention to accentuation and to prosody, specifically to the structure of *moras*<sup>11</sup>.

---

<sup>9</sup> More detailed explanations of the laryngeal theory and its historiography can be found in *Mallory and Adams 2006: 48-50*, *Gamkrelidze and Ivanov 1995 [1984]: 131-185* and *Clackson 2007: 53-61*.

<sup>10</sup> Gamkrelidze and Ivanov accept the existence of divergences in the hypothetical number of laryngeals in Proto-Indo-European and take all of them as valid as reflecting different chronological stages in the evolution of Proto-Indo-European, through which the number of the hypothetical laryngeals may have varied: “Many structural properties of Proto-Indo-European which are reconstructed in classical Indo-European studies as static schemas can be broken down into chronological stages. [...] This explains the frequent debates in Indo-European comparative grammar over linguistic structures that appear mutually exclusive; an example is the discussion of the number of laryngeals in Indo-European, where each of the several incompatible solutions has good evidence in its favor. In such instances the various solutions can be associated with different developmental stages of PIE [Proto-Indo-European], which permits us to regard many of the proposed structures as chronologically complementary and datable to different stages.” (Gamkrelidze and Ivanov 1995 [1984]: cii)

<sup>11</sup> The consideration of prosody for explaining the evolution of Indo-European was first proposed by Leonar G. Gertsenberg (1978: 5 [as referred in Belov 2011: 1]) and the theory of *moras* as explainable of it has been developed and is defended nowadays by the linguistic Aleksej M. Belov (see 2008 and 2011). This explanation enters in confrontation with the widely accepted theory of laryngeals, although they could be not exclusive of each other:

“The moras-based character of Indo-European protolanguage is confirmed by the results of my investigation. All this brings to think that not attending to the question of a possible relation between moras and laryngeals would be incorrect [...]. In conclusion, I remark once again that the explanation presented here doesn’t pretend to fullness nor to a status of fully completed (anti-laryngeal) theory, as well as doesn’t demand the negation of “laryngeals” in other cases. Apart from that we can appreciate that the analysis of the question of the oppositions of length in the protolanguages requires obligatory to take into account not only the laryngeal theory but also the theory of moras, whose harmonious combination could in the future explain certain obscure questions of Indo-European prehistory.” (Belov 2011: 1-6);

«Моросчитающий характер и.-е. праязыка подтверждается и результатами моих исследований. Всё это заставляет думать, что не рассмотреть вопрос о возможной связи моры и ларингала было бы несправедливо [...]. В заключении сделаю ещё раз оговорку о том, что приведённое объяснение не

In view of such debates, we face a problematic epistemological situation concerning the historical and comparative study of languages: despite of the fact that it possesses a systematized and structured object for study and has developed a scientific methodological approach, many of the resulting linguistic conclusions – in the form of sound-laws, change-laws and reconstructions of proto-languages – are polemic and are a matter of disagreement between scholars. It is not only that the appearance of new linguistic data modifies the theories explaining the diachronic evolution of the languages under investigation, but also that sometimes the same observable phenomenon in the diachronic evolution of attested languages doesn't lead to an univocal explanation agreed by all linguists but, on the contrary, receives different hypothetical explanations. This calls into question what is the ultimate nature of the object for study in the discipline of historical and comparative linguistics and, in relation to it, what is the source for objectivity in its approach. These problematic questions are to be discussed further in our work.

Besides, we face one more epistemological problem when noticing the unequal effectiveness of the historical-comparative methodological approach for each of the different hierarchical levels distinguishable inside natural languages: the phonological level, the morphological (inside which we need to distinguish between grammatical and lexical meaning), the lexicon level (understanding by that the level words and separable units of speech) and the level of syntax.

Already at the first stages of the development of historical and comparative linguistics, it was recognized that the comparison should not operate with whole words, because it was fruitless for reaching systematicity, but with morphemes. It is in morphemic cognates that phonetic systematic equivalences between languages may be found because it is the morpheme and not the word the minimal unit inside languages that carries meaning, whether this meaning is grammatical or lexical. Morphemes are therefore the “lowest” unit for which is possible to establish correspondent cognates between different languages and only afterwards can both “lower” and “higher” levels be compared. For the reconstruction of a lower level inside any linguistic system it is necessary, as has been explained by Toporov and Ivanov (cf. 1997: 64), the attendance to the immediately upper level: phonemes are reconstructed by comparing morphemes, morphemes in turn by comparing words, etc.

---

претендует на полноту и статус полноценной (“антиларингальной”) теории, равным образом как и не требует отмены “ларингальных” в других случаях. При этом мы видим, однако, что рассмотрение вопроса о количественных противопоставлениях праязыка требует обязательного учёта не только ларингальной теории, но и теории мор, гармоничное сочетание которых могло бы со временем прояснить некоторые тёмные вопросы индоевропейской предыстории.» (Belov 2011: 1-6)

However, it is evident that the effectiveness of the comparative method is not the same for each of these levels. In concrete, the exhaustiveness of the reconstruction decreases as we move “upwards” in the different levels of language: it is the most successful at the phonological level and at the morphological level with grammatical<sup>12</sup> meanings, but the comparison is much more problematic when concerning lexical meanings – at the levels of both lexical morphemes and whole words – and also when dealing with the level of syntax.

We can confirm this unequal effectiveness by taking a look at the current state of the investigations of the Proto-Indo-European language. The phonological system – the “lowest” level in linguistic semiotic systems – of the hypothetical Proto-Indo-European stage of language stands as the best example of exhaustive and successful reconstruction by means of the historical and comparative method, although some of its features are still matter for debate: “It is true that some uncertainties in the IE correspondence sets remain, particularly for sounds in languages which are not well-attested or for which written records do not go back very far, but for the most part the work of finding which sounds are cognate in different languages has been done” (Clackson 2007: 40).

Moving “upwards” in the hierarchy of levels inside languages, at the morphological level the reconstruction of the Proto-Indo-European linguistic system results more problematic than the one at the phonetic level. In concrete, the historical-comparative approach at the morphological level turns out effective for the plane of form of the identified cognates, but problematic for dealing with its plane of content or semantic value, in other words “semantic reconstruction is complicated by a lack of formal criteria for positing source semantemes, i.e. for positing a content plane for the reconstructed protoforms” (Gamkrelidze and Ivanov 1995 (1984): ci). What is more, the effectiveness of the reconstruction at the morphological level is not the same when dealing with grammatical meanings than when dealing when lexical ones. The only exception is the case of terms with a specially restricted and delimitable lexical value, such as numerals or terms of kinship (“mother”, “father”, etc.), which are actually

---

<sup>12</sup> “As demonstrated the experience of the detailed study of the Indo-European linguistic families and their individual branches, at least for languages of such type (synthetic with a well developed morphology, a significant role of fusions in the structure of the word, that is normally multi-morphemic and multi-compound) the more clear results in diachronic investigations are obtained in the comparison of grammatical morphemes.” (Ivanov, V. V. 2004a: 2);

«Как показал опыт детального изучения индоевропейской семьи языков и ее отдельных ветвей, по крайней мере для языков такого типа (синтетических с хорошо развитой морфологией, значительной ролью фусии в структуре слова, обычно многоморфемного и многосложного, и наличием системы морфонологических правил, в том числе акцентуационных) наиболее отчетливые результаты в диахронических исследованиях получены при сравнении грамматических морфем.» (Ivanov, V. V. 2004a: 2) [Translated by the author]

effectively compared and reconstructed by the historical-comparative method. But in spite of such concrete exceptional cases, in general the form of morphological cognates (regardless of whether they are roots, affixes or any other kind) are the most successfully confronted and reconstructed when they possess either grammatical meaning, while for the cases of morphemes with lexical meaning and of whole lexemes, the concretion of a semantic value for the reconstructed proto-form is highly problematic:

The existence of so many potential, and so few established, reconstructed lexemes causes difficulties for research into the semantic structure of the PIE [Proto-Indo-European] lexicon. It is largely impossible to discern where the boundaries between the meaning of one reconstructed root or lexeme and another lie, since we are so unsure about how much can be securely reconstructed (Clackson 2007: 158).

The problematic aspect for dealing with lexical semantics in the historical-comparative approach derives from the difficulty of concretion of lexical values for linguistic forms as well as on the sometimes contradictory testimonies of the lexical value of a certain form in its different cognates in related languages:

A second major catch to our recovery of the Proto-Indo-European lexicon concerns the reconstructed meaning of a word. Sometimes there is uniformity across all or almost all the groups offering cognates. [...] In some instances the level of ambiguity appears truly perverse, especially when the cognates suggest what might seem to be diametrically opposed meanings. [...] A third type of problem is when the range of meanings is obviously related but so disparate that we can only hazard a vague proto-meaning which might underlie the original word. (Mallory and Adams 2006: 110-111)

Lastly, and moving “upwards” in the hierarchy of level, the reconstruction of the syntax of the Proto-Indo-European linguistic system has not achieved further from a few aspects of the word order in simple sentences or certain considerations about interrogative ones:

[I]t is not clear that any substantial reconstruction of syntactic patterns, with the exception perhaps of elements of word order, can be achieved without recourse to morphology, so that the study of the syntax of IE [Indo-European] can appear to be essentially the study of the function of forms, and whether a theoretical linguist of today would accept this as an adequate approach to syntax must be open to doubt. (Penney 2000: 35 [as cited in Clackson 2007:157])

The unequal effectiveness of the historical-comparative method<sup>13</sup> for the different levels of linguistic systems that we have just exposed implies a second problematic epistemological

---

<sup>13</sup> James Clackson, in his overview of Indo-European historical and comparative linguistics accounts for this unequal effectiveness of the method for the different levels inside language with the following words: “Here we would like to stress that syntactic reconstruction is a different type of enterprise from phonological reconstruction, and it is not possible to compare the two directly. Phonological reconstruction is a ‘first-order’ reconstruction: it is not reliant on any other reconstructed data, and the establishment of correspondence sets in phonology is normally sufficient evidence to justify the supposition of a genetic relationship between languages. Syntactic reconstruction is a ‘second-order’ operation, as is morphological reconstruction and the reconstruction of lexical fields; all rely upon both the reconstructed phonology and the knowledge that the

question concerning what is the ultimate scope of its scientific approach, provided that it is able to operate with systematicity and exhaustiveness at the phonological level and at the morphemic-grammatical level, but not to the same extent at the levels with lexical meanings.

To sum up, we face two problematic questions concerning the field of historical and comparative linguistics: in the first place the contradiction between the scientificity of its approach and the debatable character of its conclusions, and in the second place the unequal effectiveness of its method for the different levels inside languages. For solving such questions, we must precise what is the ultimate nature of the actual object for study in this discipline. This will permit us to delimit what is the source for objectivity and what is the source for uncertainty in its methodological approach, and to explain what is the reason for the uneven effectiveness of its method.

---

comparanda come from genetically related languages. Syntactic reconstruction, like morphological reconstruction, has to go beyond simple comparison and pay special attention to the most archaic forms to establish what is the most likely scenario to explain the data in the daughter languages. Syntactic reconstruction consequently often involves weighing up two or more rival hypotheses [...]. Sometimes [...] there may be no means of deciding between two alternative reconstructions, or indeed there may be no means of identifying any satisfactory reconstruction.” (Clackson 2007:158)

## **2. The *syntactic dimension* of linguistic semiotic systems**

In this chapter we aim to consider certain semiotic aspects of the investigation within historical and comparative linguistics for discussing the problematic questions that arose concerning its epistemological approach and have been previously exposed: namely what is the ultimate nature of its object for study, what is the source for objectivity in its approach and what is therefore the ultimate epistemological scope of the historical-comparative method.

When considering the structuralist approach in linguistic from a semiotic theoretical approach, it becomes evident that just the relations of linguistic signs with other linguistic signs inside a linguistic semiotic system are suitable for description in terms of systematicity and structuredness, while the *semantic* and the *pragmatic dimensions* of the system are not to the same extent systematically and structurally describable. For this reason, our discussion will depart from the postulation that, when conceiving natural languages as semiotic systems, the systematicity and structuredness that are reached by the structuralist descriptive approach in linguistics scope exclusively what we call the *syntactic dimension* of linguistic semiotic systems. That notion refers to the relations of linguistic signs that are elements of the system strictly between themselves and will be the central one in our discussion.

We will expose that the ultimate object for comparison in the *historical-comparative method* developed within *historical and comparative linguistics* are the systematic and structured descriptions of natural languages provided by the structuralist linguistic approach, and our conclusion will be that the investigations within this discipline are suitable for verification and objectivity only as long as they concern the *syntactic dimension* of linguistic semiotic systems as object for study. Finally, we will present our conclusion as the explanation of the fact that the *historical-comparative method* is the most efficient for the phonological level, whose range of referentiality scopes strictly the *syntactic dimension* of linguistic semiotic systems.

## 2.1. The structuralist approach in linguistics

The historical and comparative approach to the study of the diachronic evolution of languages is closely related to *the structuralist approach* in linguistics. The object for comparison and reconstruction is given by the descriptive structured systems that result from that linguistic approach and, as we will see, this has important implications for delimiting what is the source for systematicity in the *historical-comparative method* and therefore for objectivity and scientificity in the diachronic study of languages carried out by the discipline of *historical and comparative linguistics*.

Inasmuch as the development of the historical-comparative method is the result of such approach, we must in the first place explain what is understood as *the structuralist approach* in linguistics. In his book *Modern Structural Linguistics. Problems and Methods*, Isaak Revzin (1977) – one of the founding members of the Tartu-Moscow school of semiotics – presents his understanding of what do the following approaches in linguistics consist on and what are the main conceptual and methodological differences between them: *traditional linguistics*, *structural linguistics*, and *generative linguistics*. The aim of his book is to propose a new linguistic approach labeled as *modern structural linguistics*, but the aspect of his work that is most interesting for our approach is the characterization that he offers of the *structural linguistic* approach. Revzin defined *structural linguistics* as the branch in linguistics that “consciously seeks for the accurate detection of the semiotic aspects of the sign (in abstraction of its material embodiment) and of its relations with other signs in the system and in speech, and that means, of the structure of languages.” Connected to this, the structuralist approach “seeks for the application of concepts and methods of modelling” (1977: 46)<sup>14</sup>. The structural linguistic approach is therefore a descriptive one, it studies the *structure* of languages at the level of abstraction and its final outcome is therefore abstract models that are descriptive of the *structure* of language (cf. Ibid: 45).

One of the consequences of such approach is a markedly formalistic methodological approach in the sense that it is bounded to focus on the *plane of form* of the sign system under consideration, because it is the *plane of form* what presents a predisposition to be apprehended as *structure*, at least to a much higher degree than the *plane of content* does. In

---

<sup>14</sup> «[И]од структурной лингвистикой понимается направление, сознательно стремящееся к четкому выявлению семиотических аспектов знака (в абстракции от его материального воплощения), его отношений к другим знакам в системе и в речи, т. е. структуры языков, и в связи с этим к применению идей и методов моделирования.» (Revzin 1977: 46) [Translated by the author]

Saussurean terms and with regard to the linguistic sign, the plane of form equals to the plane of *signifiers* and the plane of content to the one of *signifieds*: “Even though the concepts of signifier and signified were introduced by Saussure (in connection with a tradition existing from Saint Augustine) for words and morphemes of natural language in the first place, structural linguistics in its descriptive version [...] investigated only signifiers and didn’t therefore have to do with signs” (Ibid: 35)<sup>15</sup>. The aim for systematic description of a structure in language brought along the focus on the plane of form of the system, and that is, the mutual relations between signifiers, reaching only partially the whole of the linguistic sign and the whole of the linguistic system.

The structural and descriptive linguistic approach is the heir of both the Neogrammarians’ methodological approach to the study of languages and the ideas of Ferdinand de Saussure. The Neogrammarian linguistic approach took along the search for scientificity in the study of diachronic linguistics, the aim for massive compilations of data (together with an inductive approach to data for the elaboration of linguistic hypotheses), and the understanding of language sound change as a systematic and exceptionless process, suitable therefore of explanation in terms of laws and by certain principles. In his part, Saussure’s approach established the conception of language as a *system* that would lead the subsequent linguistic work to put the emphasis on the concept of *structure* in order to satisfy the search for scientificity and method. The origin of the structuralist approach in linguistics was explained by Émile Benveniste as follows:

“In the course of the last twenty years, the term "structure" has undergone considerable extension after acquiring a doctrinal and somewhat programmatic value. Moreover, it is not *structure* that seems to be the essential term now so much as the adjective *structural*, used to qualify a kind of linguistics. *Structural* rapidly led to *structuralism* and *structuralist*. [...] The principle of "structure" as a topic for study was asserted a little before 1930 by a small group of linguists who proposed to react thus against the exclusively historical concept of language, against a linguistics that broke language down into isolated elements and was engaged in following the changes that took place in them. It is agreed that this movement had its origin in the teachings of Ferdinand de Saussure at Geneva, as they were put into writing by his students and published under the title *Cours de linguistique generale*. Saussure is rightly called the precursor of modern structuralism. He certainly was, except for the term. It is important to note, for exactitude in describing this movement of ideas which must not be simplified, that Saussure never used the word "structure" in any sense whatever. In his eyes, the essential notion was *system*. In that was the novelty of his doctrine, in the idea – so full of implications that it took a long time to perceive and develop

---

<sup>15</sup> «Хотя понятия означающего и означаемого были введены Соссюром (в соответствии с традицией, идущей ещё от Блаженного Августина) в первую очередь для слов и морфем естественного языка, структурная лингвистика в её дескриптивном варианте, а, следовательно, отчасти и математическая лингвистика, исследовали лишь означающие, а, стало быть, ещё не имели дела со знаками.» (Revzin 1977: 35) [Translated by the author]



– that language forms a system. [...] Thus the notion of language as system was accepted long ago by those who were taught by Saussure, first in comparative grammar and then in general linguistics. If one adds to this two other principles which are equally Saussurian: that language is form, not substance, and that the units of language can only be defined by their relationships, one will have indicated the fundamentals of the doctrine which some years later was to show the *structure* of linguistic systems.” (Benveniste 1971 [1966]: 79-80).

The ultimate necessity for the rise of the structuralist approach in linguistics is to be found in the nature of language as an object for study. Being language a cultural phenomenon, the approach to it as object for scientific knowledge cannot be automatically the one of natural sciences and thus the need for verifiability and objectivity has to be fulfilled on different basis. The notion of *structure* is precisely what makes possible the application of a systematic method for the apprehension of natural language:

Linguistic facts, under their appearance of heterogeneity and independence, subordinate as a whole to a structure. The notion of “law” of natural sciences is substituted by the notion of “structure” in human sciences. [...] The rank of “scientific object” arises in cultural facts from their “structural” character [...]. Cultural facts are dynamical, historical structures. Language can become the object of a cultural science, which has as purpose not the formulation of laws without exceptions, but the discovery of the structural relations that support the system. [...] The functionalist method is ineffective in the sphere of culture if we try to reach to exact formulations or functions. However from the moment when we admit, as an assumption, that the phenomenological data in the world of culture keep certain relations between themselves, that are organized in a higher structured unit, we can deduce that the functional method is also useful for such realities. It won’t uncover exact relations, which don’t exist, but rather structural relations of the data ordered within a system. In this sense we believe that the structural interpretation has outlined the adequate assumptions for a scientific knowledge of language. [...] On the changing value that is human language, whose nature is the one of a structure, the functional method is applied and linguistic science is obtained. (Bobes Naves 1973: 29 - 59)<sup>16</sup>

To summarize, the *structure* of language is the object for study in the structuralist approach, but that object for study arises only at the level of description in a process of abstraction by which the regularities that are observable at the level of use of the linguistic

---

<sup>16</sup> [L]os hechos lingüísticos, bajo su apariencia heterogénea e independiente, se subordinan en conjunto a una estructura”. La concepción de “ley” de la ciencia natural, es sustituida por la concepción de “estructura” en las ciencias humanas.” [...] La categoría de “objeto científico” proviene en los hechos culturales de su carácter “estructural” [...] Los hechos culturales son estructuras dinámicas, históricas. El lenguaje puede constituirse en objeto de una ciencia cultural, cuya finalidad no será la formulación de leyes sin excepciones, sino el descubrimiento de las relaciones estructurales que mantienen el sistema. [...] El método funcional es ineficaz en el ámbito de la cultura, si pretendemos llegar a fórmulas o funciones exactas. Pero desde el momento que se admite, como un presupuesto, que los datos fenoménicos del mundo de la cultura mantienen unas determinadas relaciones entre sí, organizadas en una unidad superior estructurada, podemos deducir que el método funcional es válido para esas realidades. No descubrirá relaciones exactas, que no existen, sino relaciones estructurantes de los datos, en orden a un sistema. En este sentido creemos que la interpretación estructural ha señalado los presupuestos válidos para un conocimiento científico de la lengua. [...] Sobre este valor cambiante, que es el lenguaje humano, cuya naturaleza es ser una estructura, se aplica el método funcional y se consigue ciencia lingüística. (Bobes Naves 1973: 29 - 59) [Translated by the author]

sign are extracted and systematized in the descriptive metalanguage<sup>17</sup>. It is because of that systematization of the linguistic phenomenon to turn it into an object for study that the conception of language in terms of *structure* takes along the notion of *system*; “the systematicity of natural language reveals itself in the fact that every (significant) of its elements can be explained through others. In this sense language forms a system” (Nikolaeva and Uspenskij 1966: 71)<sup>18</sup>. The achievement for systematicity in the structuralist approach to language allows for a scientific methodological approach to language that fulfills the need for objectivity and verifiability but we cannot ignore that it is only at the abstract level of metalinguistic description that *systematicity* is reached.

## 2.2. The semiotic approach to natural language

*Semiotics* as an academic field deals with the theorization in abstraction of the notion of sign (categorizations and typologies of signs), of the notion of *semiosis* (a process of mediation by signs, as proposed by Charles S. Peirce, see 1931-1958 5.484), with the study of sign systems and with certain notions closely related to these ones such as *code*, *referentiality*, etc., as well as with the empirical analysis of concrete processes of *semiosis*, the description of sign systems, etc. Natural language falls inevitably within the field for study of *semiotics* and the semiotic approach to it brings along the recognition of the ultimately sign nature of linguistic phenomena, so language is not conceived anymore immanently<sup>19</sup> *in itself and by itself*<sup>20</sup>, but

---

<sup>17</sup> “the linguist deals constantly not with the facts of language as such, but with a given model of these facts, as presented by one or another description.” (Revzina, O. G. 1972: 4);

«лингвист постоянно имеет дело не с фактами самого языка, а с определенной моделью этих фактов, представленных в том или ином описании.» (Revzina, O. G. 1972: 4) [Translated by the author]

<sup>18</sup> «системность естественного языка проявляется в том, что каждый (значащий) элемент его может быть объяснен через другие. В этом смысле язык образует систему» (Nikolaeva and Uspenskij 1966: 71) [Translated by the author]

<sup>19</sup> On the immanent approach to language: “One of the possible approaches to language implies the approach to it as an immanent system, “in itself and for itself”. Provided that this approach is fundamental for linguistics, we will refer to it as properly linguistic. For its development, as is well-known, did a lot F. de Saussure, L. Hjelmslev and others.” (Vardul’, I. F. 1967: 5);

«Один из возможных подходов к языку предполагает рассмотрение его как имманентной системы, «в себе и для себя». Так как этот подход является основным для лингвистики, будем называть его собственно лингвистическим. Для его разработки, как известно, много сделали Ф. Де Соссюр, Л. Ельмслев и др.» (Vardul’, I. F. 1967: 5) [Translated by the author]

<sup>20</sup> The foundation of linguistics as a field of study includes the assumption that the unique object of study for linguistics is language studied “in and for itself”, which is explicitly elaborated for the first time in the thought

in relation to the notion of *sign*, assuming that “the essential facts of language can be discovered and explained by inferring them from the principles of the linguistic sign” (Shaumyan 2006: 263). The recognition of language as a semiotic phenomenon together with the very first explicit proposal of a new unified science concerning the study of such phenomena are to be found in the well-known words by Saussure:

“Language is a system of signs that express ideas, and is therefore comparable to a system of writing, the alphabet of deaf-mutes, military signals, etc. But it is the most important of all these systems. *A science that studies the life of signs within society* is conceivable; it would be part of social psychology and consequently of general psychology; I shall call it *semiology* (from Greek *sēmeíon* ‘sign’) *Semiology* would show what constitutes signs, what laws govern them.” (Saussure 1959 [1916]: 16)

Actually, it may even seem paradoxical that despite of his recognition of the semiotic nature of language, it was precisely Saussure’s structural conception of language would lead to the structuralist immanent approach to language that would leave out its ultimate semiotic nature to pursue an immanent a systematic study of it. Leaving paradox apart, it must be mentioned that the semiotic theoretical understanding of language initially proposed by Saussure was continued and further elaborated by Louis Hjelmslev (see 1961 [1943]) and Émile Benveniste (see 1981 [1969]). As a continuation of their approaches, a semiotic theory of natural language was developed by Sebastian K. Shaumyan (see 1987), who will be one of the referents in our approach.

A very first consequence of the consideration of language as a semiotic phenomenon is that it is recognized as of ultimately a same nature as any other cultural phenomenon; all cultural phenomena are semiotic phenomena because they consist on processes of mediation of signs. In the words of Sebastian K. Shaumyan, the common semiotic nature of these phenomena lies on the fact that they are “genetically the product of human consciousness, but ontologically are independent of human consciousness” (Shaumyan 1998: 2). He also points out the relevance of a semiotic theoretical approach to language for understanding its essence:

“[L]anguage can be studied from different points of view, but a semiotic approach is crucial for comprehending the essence of language. Since natural languages are only a subclass of possible sign systems, a study of natural languages in a context of a general theory of possible sign systems is indicated. This theory is called semiotics. The principles and laws of semiotics apply to the study of natural languages, as well as to other sign systems” (Shaumyan 1987: 17-18).

---

of F. de Saussure, eminently as compiled by his students in the *Course in General Linguistics* (Saussure, F. de 1959 [1916]). Apparently, this wide-spread notion of Saussure’s thought that closes the *Course* was an addition of Bally and Sechehaye in the elaboration of it (cf. Saussure, F. de 1991: 306).

Because of this switch in recognizing languages as part of the semiotic continuum that is human culture, the semiotic theoretical approach to linguistic semiotic systems has a broader scope than the strictly linguistic one, and it even takes along a reevaluation of the linguistic task, leading to important implications for all branches of linguistics, including the discipline of historical and comparative linguistics, which is the one under discussion in our work. Our following theoretical approach to linguistic systems will be in fact oriented to the discussion of the problematic epistemological questions concerning that discipline.

As the starting point of our semiotic approach to language, we will turn to the distinction between *syntactics*, *semantics* and *pragmatics*. Charles W. Morris established – on the basis of the well-known triadic notion of the sign established by Charles S. Peirce – the distinction between a *syntactical*, a *semantical* and a *pragmatical dimension* in the analysis of any semiotic phenomenon:

In terms of the three correlates (sign vehicle, designatum, interpreter) of the triadic relation of semiosis, a number of other dyadic relations may be abstracted for study. One may study the relations of signs to the objects to which the signs are applicable. This relation will be called the *semantical dimension of semiosis* [...]; the study of this dimension will be called *semantics*. Or the object of study may be the relation of signs to interpreters. This relation will be called the *pragmatical dimension of semiosis* [...] and the study of this dimension will be named *pragmatics*. [...] Since most signs are clearly related to other signs, since many apparent cases of isolated signs prove on analysis not to be such, and since all signs are potentially if not actually related to other signs, it is well to make a third dimension of semiosis co-ordinate with the other two which have been mentioned. This third dimension will be called the *syntactical dimension of semiosis* [...], and the study of this dimension will be named *syntactics*. (Morris 1938: 6-7)

For the case of *semiotic systems*, and that includes languages – which from now on will be referred to as *linguistic semiotic systems* –, that distinction implements in the following way:

- a) The *semantic*<sup>21</sup> dimension: the relations of that particular sign system with the extra-systemic reality.
- b) The *pragmatic* dimension: the relations of that particular sign system with the participants of the act of semiosis and their praxis.
- c) The *syntactic* dimension: the relations of the signs that are elements of that sign system with other signs that are also elements of it. This means, the relations of linguistic signs between themselves.

---

<sup>21</sup> Although Morris employs the terms *syntactical*, *semantical* and *pragmatical*, in our work we will employ the variants *syntactic*, *semantic* and *pragmatic*. We don't imply any different understanding of these three semiotic dimensions and consider both variants as fully equivalent and interchangeable. The preference for the second option is just a matter of shortening the terms.

We must precise that this distinction is valid only at the theoretical level because “it becomes increasingly questionable whether or not the isolation of the syntactic, pragmatic, and semantic branches of semiotics is indeed possible and viable” (Ivanov 2008 (2007): 226), and actually in any analysis of either an act of semiosis or a whole sign system, these three dimensions reveal themselves as interdependent to such a degree that is it not possible their full isolation. Nevertheless, in spite of them not being isolatable at the empirical level, the differentiation of these three dimensions stands as a useful assumption for the very first approach at the theoretical level<sup>22</sup>.

We have previously exposed that the structuralist linguistic approach carries out an abstract metalinguistic description of linguistic semiotic systems, models them as a structure and establishes systematic relations between linguistic elements, what confines it to deal with the *level of form* of the system. Keeping that in mind, and on the basis of the triadic semiotic distinction, we claim that the aspect of linguistic semiotic systems that is described in the structuralist approach to linguistic semiotic systems is specifically their *syntactic dimension*, and that means, the mutual relations between the linguistic signs that are elements of that system, including both paradigmatic and syntagmatic relations. The aim for systematicity and structuredness that characterizes the structuralist linguistic approach is reachable only when modelling such relations between linguistic signs in abstraction, but is not equally reachable for giving account of their reference to the extra-linguistic reality – *the semantic dimension* – and their performative aspects – *the pragmatic dimension*.

This fact was actually already noticed by Charles W. Morris when firstly proposing his triad of dimensions: “Syntactics, as the study of the syntactical relations of signs to one another in abstraction from the relations of signs to objects or to interpreter, is the best developed of all the branches of semiotics. A great deal of the work in linguistics proper has been done from this point of view, though often unconsciously and with many confusions” (1938: 13). And also the linguist and semiotician Vyacheslav Ivanov reports on it:

In the twenties and thirties it appeared to most scientists that a purely syntactical analysis without the consideration of meaning might be sufficient for a description of a sign system. The notion of meaning as well as pragmatic context of the signs use had escaped the attention of scholars as well as of the avant-garde artists and art historians to whom the

---

<sup>22</sup> The same remark is made by the Spanish semiotician María del Carmen Bobes Naves:

“At the level of facts, the three primary elements of subject-sign-object are inseparable, and only at the theoretical level, as a method for study, can we admit as valid a binary relation” (Bobes Naves 1973: 122);

“A nivel de hechos son inseparables aquellos tres elementos primarios sujeto-signo-objeto, y únicamente a nivel teórico, y como método de estudio, se puede admitir como válida la relación binaria.” (Bobes Naves 1973: 122) [Translated by the author]

internal structure of an object seemed the only relevant object of study. In all the fields of semiotic activities, [...] the internal (purely syntactical) relations among the elements seemed much more important than their semantic interpretation or pragmatic use. (Ivanov 2008 (2007): 225)

Indeed the description of grammar, syntax, phonetics and morphology of natural languages as carried out in traditional linguistics deals with relations of the different signs that operate inside linguistic systems between themselves. It may seem paradoxical, but also such tasks as the elaboration of dictionaries and thesauruses restrict to the description of mainly the *syntactical dimension* of linguistic semiotic systems. Such tasks seem to deal with the description of the relation between the linguistic system and the extra-linguistic reality and therefore with the description of the *semantic dimension* of the linguistic system, but the actual range of it is very short. It actually results quite the contrary, the semantic work within traditional linguistics in dealing with the lexicon in dictionaries and thesauruses establishes relations between elements of a linguistic system and other elements of that same system.

At this point, a terminological clarification will be appropriate in order to clearly differentiate the terms *syntactic*, *semantic* and *pragmatic* as understood in the linguistic sense and as understood in the semiotic sense. As the reader has already noticed, we reached a point in our discussion when we do not mean the same by these terms anymore depending on whether we are speaking from the linguistic point of view or from the semiotic one. In the linguistic approach, by *syntax* we mean either the relations between words in the context of a sentence or the description of the mechanisms for constructing sentences that a given natural language offers. However, in our semiotic theoretical approach by the *syntactical dimension* of a linguistic semiotic system we take a broader scope and mean all kind of relations between the signs that are part of that linguistic system. This notion includes not only the relation between words in forming a sentence, but also the abstract paradigmatic relations between linguistic signs in a given language, the relations of opposition between phonemes and all other possible relations between elements of a linguistic system<sup>23</sup>. In order to avoid any confusion, we will stick to the term *syntax* for referring to the traditional sense of the term

---

<sup>23</sup> We must remark that our conception of “the syntactical” has nothing to do with its notion in the widespread linguistic theory of Noam Chomsky presented in *Syntactic structures* (1957). Although Chomsky’s theory also employs the term “syntactic”, there is no implied connection between our and his notion; in the first place because he is referring to basically the traditional linguistic understanding of syntax, while we are referring to *syntactics* in the semiotic sense, and in the second place because his theory of language focuses on describing the syntagmatic axis, while ours refers to both syntagmatic and paradigmatic relations. His definition of *the syntactic* is the following one, and is obviously not coincident with ours: “Syntax is the study of the principles and processes by which sentences are constructed in particular languages. Syntactic investigation of a given language has as its goal the construction of a grammar that can be viewed as a device of some sort for producing the sentences of the language under analysis.” (Chomsky 2002 (1957): 11)

employed within linguistics and will employ the terms *syntactics* as originally established by Charles Morris (cf. 1938: 7). The term *syntactic dimension* will be employed for referring to our semiotic conception of it.

Having made this terminological clarification, we will try to support our claim that the aim for systematicity and structuredness of the structuralist linguistic approach are reached for specifically the *syntactic dimension* of linguistic semiotic systems but are not reachable for the *semantic* and *pragmatic* ones by briefly commenting the most well-known attempts for the structural and systematic study of the latest, which are demonstrative of the limitations of the structuralist approach. For the case of semantics a metalanguage for description of the same systematicity and fixity is hardly reachable because the ultimately necessary formalistic approach doesn't work for the apprehension of lexical meaning:

In linguistics is not clarified enough what should be the object and the method of semantics, that is to say, the scientific investigation about the level of meaning. [...] Provided that meaning doesn't seem to be reducible to form, as phonetic, morphologic and syntactic are, positivist and formalist schools cannot admit the possibility of any objective and empirical structuration of semantic oppositions. (Bobes Naves 1973: 157)<sup>24</sup>

As noticed also by Benveniste, semantic notions are “more difficult to objectify and especially to formalize, since they are entangled in the extralinguistic “substance”” (Benveniste 1971 [1966]: 264). Nevertheless, an attempt for a structural study of the semantic dimension of linguistic semiotic systems was developed by Julius Greimas, where he tries to overcome this initial impossibility and implement the same structural approach to the analysis of meaning in search of a *structural semantics* (see Greimas 1966). However, this attempt still lacks of a successful implementation in the sense that hasn't been able of the degree of systematization of its object for description. We quote here the critical words of the Spanish Indo-Europeanist Francisco Adrados concerning this question:

Structural Semantics is still more a program than a doctrine, although many of its fundamental basis are clearly established. It lacks, mainly, of concrete work on data, and presents to many ideas that are vague and too general. We believe that Semantics can do a lot to restore the link between Linguistics and the rest of human sciences, as well as to seek again for a global conception of language, both in its formal aspects as in its aspects of content. (Adrados 1971: 34)<sup>25</sup>

---

<sup>24</sup> “En la lingüística no está aclarado suficientemente el objeto y el método que debe tener la semántica, es decir, la investigación científica sobre el nivel del significado. [...] Puesto que la significación parecía no poder ser reducida a forma, como los valores fonéticos, morfológicos o sintácticos, las escuelas positivistas y formalistas no pueden admitir la posibilidad de una estructuración objetiva y empírica de las oposiciones semánticas.” (Bobes Naves 1973: 157) [Translated by the author]

<sup>25</sup> “Presentadas así las cosas, la Semántica Estructural es todavía más un programa que una doctrina; aunque algunas de sus bases fundamentales están claramente establecidas. Falta, principalmente, el trabajo concreto sobre los datos, sobran ideas vagas y generalizantes. Creemos que la Semántica puede hacer mucho para

For the case of pragmatics, the systematic structuralist approach faces the same limitation. The great attempt to fulfill the analysis of the pragmatic dimension of natural language is the *speech acts* theory developed by Austin and Searle (cf. Austin 1962 and Searle 1969). Also the linguistic approach by Wittgenstein and the Oxford Circle puts the emphasis on pragmatics. In the linguistic study of pragmatics occurs a turn to the context and the performativity of verbal utterances, which used to be ignored; “the speech acts theory helped to find adequate ways of description of such pragmatically important units of natural language as performative sentences” (Ivanov 2008 (2007): 225-226), but what is important for us is that such attempts had to stray from the structural and formalistic approach and this demonstrates the already mentioned limitations of this approach for the non-syntactical dimensions of natural languages.

As a last remark, we point to the fact that the systematicity, exhaustiveness and objectivity reached by structural linguistics in its description of the syntactic dimension of linguistic semiotic systems has turned linguistic methodology into a model for the systematic description of other cultural sign systems that operate as well with certain regularity and stability in actual use. The spread of structuralism to other humanitarian disciplines in, the works by, for example, Uspenskij for the study of myth, by Lévi-Strauss in anthropology, by Propp for the analysis of the narratives of the folk tale, and by the Russian formalist school in its study of poetics and literature, stand as good examples of it.

---

restaurar la conexión de la Lingüística con las demás ciencias humanas y, también, para que se vuelva a intentar sentar una concepción global de la lengua, tanto en sus aspectos formales como en sus aspectos de contenido.” (Adrados 1971: 34) [Translated by the author]



## **Conclusive remarks: the object of study and the ultimate scope of historical and comparative linguistics**

Natural language is a phenomenon that due to its complexity is suitable for different and complementary approaches and reluctant to any holistic one. For the case of *historical and comparative linguistics* in its diachronic study of languages, the object for study is specifically the structural metalinguistic descriptions that result from the structuralist approach in linguistics. Its object for study does therefore not exist a priori, but is constructed in a process of abstract description and structural modelling. It is because of the systematic nature of that structurally constructed object for study that the *historical-comparative method* can deal with objectivity in a way comparable to the method of natural sciences. We cannot ignore that that systematicity arises only at the abstract level and as property of the metalanguage for description, and not of the object itself.

The paradoxical coexistence of both objectivity in the method and, as we previously exposed, controversy and debate about the resulting theories and hypotheses lies on this constructed nature of its object for study, which behaves structurally and systematically only, as we have said, at the metalinguistic abstract level. The abstract and descriptive nature of the structural objectification of language cannot fully correspond to the actual occurrence of language at the level of speech, in its concrete realizations in specific contexts, and this lack of full correlation is manifested already in the synchronic description of languages and therefore also in the diachronic one. This gap is, from our point of view, the source of constant reformulation and debate of the results in both synchronic and diachronic studies of language. For the latter case, that is the one under discussion here, the constant reformulations and debates are due to either the appearance of new empirical evidences questioning the accuracy and completeness of the previously established structural model, or the different subjective valuations of the validity of the models by different scholars.

As systematicity and structuredness are reached at the level of description and only for the case of the relations between the linguistic signs that conform the system, we must admit that the *historical-comparative method* cannot operate effectively outside of the syntactic dimension, and its ultimate scope is thus compelled and limited by it. When comparing

linguistic systems in the historical-comparative approach, what we are actually confronting are abstract metalinguistic structures that model specifically *the syntactic dimension* of each of them.

This fact concerning the ultimate nature of the object for study in *historical and comparative linguistics* is explicative of the unequal effectiveness of the *historical-comparative method* for each of the levels inside linguistic semiotic systems. Its success when operating at the phonological and morphological levels is due to the fact that the phonemes and the morphemes with grammatical meaning make reference to strictly other linguistic signs, to other elements of the system, and therefore they have no scope of referentiality out of the system: they are confined to *the syntactic dimension* of the linguistic system. However, for the levels that possess lexical meaning, and that means morphemes with lexical meaning and lexemes, the historical-comparative method cannot operate with the same degree of systematicity and we argue that this is so because the content of those linguistic signs concretes only in actual use and they scope thus the *semantic* and *pragmatic dimensions* of the linguistic system. In its part, the poor reconstructability of the level of syntax is also due to that same fact; although it may seem contradictory, the syntax level of a linguistic semiotic system doesn't scope only the syntactical dimension but, given that involves both grammatical and lexical meanings and is built only upon them<sup>26</sup>, it involves necessarily also the *semantic* and *pragmatic dimensions*<sup>27</sup>.

When attempting to reconstruct proto-languages, what we are doing is working on the *syntactic* logic of abstract structured models of linguistic systems, and, as they belong to the

---

<sup>26</sup> To support our argument about the dependence of the syntax level on the lexical content, we provide the following conclusions by Raimo Anttila and Henrik Birnbaum:

“In the diagrams of syntax (the syntactic patterns, sentence schemes, or whatever we call them) lexical items act as diacritics.” (Anttila 1989: xii).

“[S]yntax constitutes an independent component of linguistic structure with operates on an autonomous level (or, to be precise, on two interrelated, specifically syntactic levels). Syntax, though separate, interacts and correlates with other components of the linguistic system, or (if viewed in the purely theoretical terms of a construct) of the linguistic model. The syntactic component is regarded, then, to be directly related to and informed by lexical semantics, on the one hand, and, on the other, by morphology, i.e., the theory of the formation and variation of minimal meaning carriers (morphemes).” (Birnbaum 1984: 26)

<sup>27</sup> There are another two reasons for the poor reconstructability of the level of syntax and for the lack of effectiveness of the historical-comparative approach at that level. The first one is that syntax is a feature of language highly suitable for change – “highly susceptible to foreign influence” (Birnbaum 1984: 41) – and the second one is that it doesn't change in such a systematizable way as the phonetic level, what makes it a subject of not successful explanation by means of the comparative method: “[W]hile in phonology (except in stylistically conditioned free variations), once a gradual shift is completed, one entity perceived as distinctive (a feature, a phoneme) fully replaces an earlier one so that on the perceptual (acoustic) level and either/or situation obtains, in syntax a both/and situation (syntactic synonymy) may prevail for a considerable period of time. In fact, to some degree this is the rule, not the exception: language at all times avails itself of a variety of formal (structural) means to convey one and the same grammatical meaning (separate function or sentential meaning)” (Birnbaum, H. 1984: 40-41).

level of abstraction, the actual power of resemblance of the resulting hypothetical reconstructions about actual stages of language having existed at some point in the past is very restricted. By analogy, to the same extent as we recognize that the structural descriptions of languages in synchrony model actually observable features of the level of speech, we recognize that the hypothetical reconstructions of proto-languages also reflect, at least to some extent, general features of the linguistic stage having actually existed in the past, although we cannot take them as actual realizations of that language.

## **PART II: Semiotic aspects of the investigations in linguistic paleontology**

In the present chapter we will discuss some relevant aspects concerning the field of studies of *linguistic paleontology* that consists on the investigation of human past from the analysis of strictly linguistic data. The main assumption lying behind the procedure of these investigations is that linguistic data are informative of non-linguistic aspects of past societies. We will discuss the ground for such assumption from the semiotic approach to culture provided by the Tartu-Moscow school's *semiotics of culture*.

The constant debates about the ultimate validity of the linguistic arguments about human past and the profusion of pseudo-scientific theories arisen from this discipline put into question the validity of the conclusions accomplished by this field of studies. After discussing the problematic epistemological status of the discipline and recognizing the interpretative character of its investigation, we will conclude that *linguistic paleontology* stands as a source for plausible arguments but not of objective knowledge about human past, and that its conclusions are dependent of the evidences accomplished by other disciplines in order to provide reliable knowledge about that past.

Besides, we will specify the different degree of plausibility reachable by *linguistic paleontology* depending on the kind of linguistic data under consideration. In concrete, linguistic arguments are to be recognized more validity and plausibility when extracted from ancient attested linguistic stages than from hypothetically reconstructed protolanguages and, concerning the latest case, the linguistic arguments about the prehistoric societies carrier of protolanguages are more solid when extracted from the reconstruction of *the syntactic dimension* of linguistic semiotic systems, and much more a source for uncertainty when extracted from the consideration of the lexicon.

### 3. The discipline of linguistic paleontology

The investigations within *historical and comparative linguistics* and their resulting conclusions – sound-laws, change-laws and reconstructions of protolanguages – deal with strictly linguistic data, but they are relevant also outside linguistics. As these investigations deal with past linguistic stages that are to be inevitably link to human societies, the initially linguistic data and conclusions become a source of information about the history and the prehistory of those societies:

The picture of areal affinities drawn from the study of grammatical, phonological, and lexical isoglosses among dialects of a proto-language provides a way of determining how dialectal differentiation of the proto-language proceeded and hence of establishing what contacts took place among speakers of the dialects. In other words, purely linguistic facts make it possible to establish extra-linguistic factors, such as historical interactions among speakers of the linguistic systems and their achievements in the domain of material and spiritual culture. (Gamkrelidze 2008: 143)

The most striking case is the informativeness about prehistoric societies that are unknown-to-us but which existence is revealed by the linguistic evidence of a stage of protolanguage, which implies logically the existence of a society carrier of it.

The field of study so-called *linguistic paleontology*<sup>28</sup> undertakes the investigation of non-linguistic features of the prehistoric societies carriers of protolanguages on the basis of the linguistic conclusions previously accomplished by *historical and comparative linguistics*. It was the Swiss linguistic Adolphe Pictet who coined the term *linguistic paleontology* to refer to this kind of investigations in his work *Les origines indo-européennes ou les Aryas primitifs: Essai de paléontologie linguistique* published in three volumes from 1859 to 1863 (see Pictet 1859-1863).

---

<sup>28</sup> In certain bibliographies about the issue this discipline is referred to as *paleolinguistics*. Both terms refer to the same kind of investigations and are therefore equivalent. We will preferably employ the term *linguistic paleontology*. In order to underline the ultimate non-linguistic object for investigation of these investigations, the discipline has been also called *linguistic cultural paleontology* and *linguistic paleontology of culture* by Ivanov and Gamkrelidze:

“This trend in linguistics might better be called ‘linguistic palaeontology of culture’ since its object of investigation is not the proto-language but the proto-culture of the speakers; what is reconstructed is not so much the language itself as the extra-linguistic world reflected in the linguistic data” (Gamkrelidze and Ivanov 1995 [1984]: 380; and also Gamkrelidze 2008: 144).

### 3.1. Linguistics as a discipline for investigating the past

The work within *linguistic paleontology* turns linguistics into a source of information about the past of human societies. The data and conclusions accomplished by the discipline of *historical and comparative linguistics* become a tool for investigating ancient and prehistoric cultures:

Linguistics, the scientific study of language, can reach more deeply into the human past than the most ancient written records. It compares related languages to reconstruct their immediate progenitors and eventually their ultimate ancestor, or protolanguage. The protolanguage in turn illuminates the lives of its speakers and locates them in time and place. (Gamkrelidze and Ivanov 1990: 110)

In concrete, *linguistic paleontology* aims for the description of the cultural system of prehistoric societies (beyond their linguistic system), the delimitation of their geographical location and the establishing of the chronology of their existence. With those purposes, the analysis focuses on three different kinds of lexical data:

- 1) The *reconstructions of proto-languages* accomplished by historical-comparative linguistics, with special attention to certain semantic fields of the reconstructed proto-lexicon:
  - a) The lexicon referring to environment, geography, flora, fauna and landforms is studied as informative of the initially unknown geographical location of the “proto-societies” under consideration, which is postulated by tracing the characteristics of landscape reflected in such lexicon.
  - b) The lexicon referring to livestock, agricultural and economical activities and work tools is analyzed as informative about the material conditions of life of those societies.
  - c) The lexicon referring to “immaterial culture” is analyzed as informative of features of the cultural system beyond language of that prehistoric society, such as their religious beliefs, their myths or their modes for social organization<sup>29</sup>.

---

<sup>29</sup> Not only linguistic data are analyzed for investigating unattested cultures, for that same purpose also the historically attested cultural systems of the societies carriers of languages that are descendant of that protolanguage are compared. In a similar procedure as when confronting linguistic systems in historical-comparative linguistics, the mythological and religious systems of ancient societies undergoes comparison in order to reconstruct previous unattested cultural systems. One of the main investigators in this field, specially devoted to the Proto-Indo-European ancestor mythological system, has been Vladimir Toporov, who was also one of the members of the Tartu-Moscow school of semiotics. His work has been compiled in two collections: *Toporov 2004-2010* and *Toporov 2010*.

- 2) *Loanwords* between linguistic groups are analyzed in search for clues about contacts between different populations. They sometimes allow establishing a relative chronology for those contacts between populations.
- 3) *Toponyms* – place names of geographical locations of any kind: rivers, mountains, cities or any other. It exists the assumption inside historical and comparative linguistics that toponyms present an especially accurate tendency to conservation, even in cases of changings of the linguistic identity of the locators of a given geographical location, and that it is therefore possible to uncover the linguistic identity of ancient inhabitants of a given geographical location by tracking the etymology of the toponyms spread over that region and assigning it to a concrete linguistic group.

Besides, also approximate chronologies for the period of existence of a given proto-language are attempted by means of methodologies that try to fix rates of language change for concrete intervals of time, such as the methods of *glottochronology* and *lexicostatistics*, which were established by Morris Swadesh (1909-1967) and have been nevertheless matter of strong criticism and several reelaborations (see Swadesh 1955).

As can be seen, the lexicon is the linguistic feature on which *linguistic paleontology* focuses the attention, but it must be mentioned that inferences about non-linguistic aspects of the past are not only made on the basis of lexicon. The phonological, morphological and typological features of languages are not explicitly included in the work by *linguistic paleontology*, but actually inferences about human past are implicitly made from this kind of data. Already the postulation of a proto-society is made upon the observable typological proximity between languages forming a family, and further on, dialectal groups inside that family are inferred in relation to the more or less typological proximity between the languages of the family and these dialectal groups are taken as informative of population sub-groups, the identified processes of divergent dialectalization are, at least to a certain extend, informative about migratory processes, while the attested processes of linguistic convergence are in turn informative of contacts between different populations that cohabitated at some point.

In this way *linguistic paleontology* turns linguistics into one more of the disciplines for investigating human past: “The historical-comparative investigations of linguistic families and macro-families eventually permitted to confront linguistic conclusions with biological, anthropological or archaeological ones, and this would lead to a revolution in the

understanding of the past” (Ivanov 2004b: 11)<sup>30</sup>. In fact, the procedure of *linguistic paleontology* has to confront the data provided by other disciplines that also deal with the investigation of the ancient and prehistoric past of human populations. Especially important is the confrontation with the data provided by archaeology<sup>31</sup> and by the studies on the genetics of human populations<sup>32</sup>. Each of the different disciplines involved in the investigation of human past generates data that are complementary to the ones provided by the others and needs confrontation in search for complementary support and for “reciprocal verification”:

‘The time machine’ built by historical and comparative linguistics can serve not only as a methodological (still very often unattainable) model for other sciences that deal with similar tasks in their material, but can also allow for the unification of different reconstructions of the past – linguistic, archaeological, anthropological. There exist many fields in the dictionary of ancient proto-languages, where is necessary a confrontation with data of other sciences (for example, in the study and reconstruction of the names for metals, domestic animal and plantations). In this process is possible not only the reciprocal verification of conclusions that have been obtained independently, but also the determination of the ancient settling of the carriers of the proto-language and of the directions of their different migrations that lead to the dialectal division. (Ivanov 2004b: 134)<sup>33</sup>

In outline, archaeological investigations identify a geographical location, can offer accurate chronologies and provide evidences about the material culture of a given prehistoric society. Linguistic paleontology in its part establishes a linguistic identity of a certain prehistoric population and that permits its genetic connection to societies of different periods that share that same linguistic identity. It provides as well informative clues about the material and non-material culture of that population and offers relative and approximate chronologies that are nevertheless much more imprecise than the ones offered by archaeology. For its part, the studies on the genetics of the population indicate migratory movements of prehistoric populations and inform about kinship relations between different populations. Despite of the

---

<sup>30</sup> «Сравнительно-исторического исследования семей и макросемей языков в конечном счете позволило бы сопоставить языковедческие выводы с биологическими, антропологическими, археологическими и привело бы к революции в понимании прошлого.» (Ivanov 2004b: 11) [Translated by the author]

<sup>31</sup> Some of the main works on archaeology devoted to the clarification of the Indo-European question can be found in *Renfrew 1987* and *Gimbutas 1956* and *1963*.

<sup>32</sup> A complete presentation of the conclusions of the investigations of the genetics of the population is provided in *Cavalli-Sforza et al. 1994*. An important study of genetics concerning the entrance of Indo-European languages in Europe can be found in *Haak et al. 2015*.

<sup>33</sup> «‘Машина времени’, построенная сравнительно-историческим языкознанием, не только может служить методологическим (часто пока недостижимым) образцом для других наук, решающих на своем материале сходные задачи, но и в принципе может способствовать объединению разных реконструкций прошлого – лингвистических, археологических, антропологических. Существуют многие области словаря древних праязыков, где сопоставление с данными других наук необходимо (например, при изучении и реконструкции названий металлов, домашних животных и растений). На этом пути возможна не только взаимная проверка выводов, полученных независимо друг от друга, но и определение древнего расселения носителей праязыков и направлений их ранних миграций, приведших к диалектному членению.» (Ivanov 2004b: 134) [Translated by the author]



complementarity of the data of these three fields for investigation, the actual bringing together of their conclusions turns out problematic because in most of cases it is not possible a univocal correlation of a given archaeological culture with a fully correspondent linguistic identity and with a certain genetic filiation. Whether coincident or not in their hypotheses, the result is that we face three different kinds of arguments for investigating human past – linguistic arguments, archaeological arguments and genetic arguments – and they have to be brought into concordance in order to provide reliable and solid knowledge about that past.

We will briefly give account of the present state of the investigations about the past of the Indo-European populations for illustrating the problematic complementarity of the several disciplines dealing with the investigations of ancient and prehistoric past, although we must remind that these investigations are not restricted to Indo-European studies. The same way as the data of Indo-European linguistics are employed by means of linguistics paleontology for investigating the prehistory of the Indo-European population groups, the linguistic data concerning other families such as the Altaic, the Uralic or the Semitic one, are employed for investigating the ancient and prehistoric evolution of the populations carriers of their languages.

Nowadays inside Indo-European studies three best-founded theories are the most broadly accepted for the explanation of the geographical origin and of the process of dispersion of the Indo-European languages: *the Kurgan theory* postulated by the archeologist Marija Gimbutas, *the Neolithic theory* posited by Colin Renfrew and *the Near-Eastern theory* proposed by Tamaz Gamkrelidze and Vyacheslav Ivanov. *The Kurgan hypothesis* by Marija Gimbutas (see Gimbutas 1956 and 1963) is well attested by archaeological data and also by genetic studies (see Haak et al. 2015), although doesn't coincide that well with the characteristics of the Proto-Indo-European landscape demanded by the linguistic reconstruction of Proto-Indo-European lexicon. *The Neolithic expansion theory* by Colin Renfrew (see Renfrew 1987) matches very well the map of migrations provided by genetic data but goes very far back in time for locating the period of existence of the proto-Indo-European common state of language, what implies a lengthy chronology that gets difficult to accept in terms of language change (along such a long period of time would be expectable a higher degree of differentiation between the different Indo-European ancient historical languages than the one actually attested in them). The theory proposed by Gamkrelidze and Ivanov (see 1995 [1984])

for its part is the most satisfactory for the demands of the linguistic arguments and doesn't contradict the genetic evidence but is not well attested by archaeological finds.<sup>34</sup>

### **3.2. Problematic questions concerning the investigations within linguistic paleontology**

The underlying theoretical assumption in the general procedure of *linguistic paleontology*, by which linguistics becomes a source of information about the past, is that linguistic data are informative of the extra-linguistic features of the society in which they are employed and, more concretely, that the reconstructions of hypothetical protolanguages carried out by *historical and comparative linguistics* are informative of non-linguistic aspects of human societies. Such theoretical assumption is taken for granted as an obvious fact from within the discipline, but it may be called into question and we will discuss the ground for such assumption. A semiotic theoretical approach is necessary in which the main question will be the role of language inside culture, which is a fundamental question for establishing the theoretical ground of the informativeness of language about non-linguistic aspects of cultures. With that purpose, we will turn to the semiotic understanding of *culture* developed by the Tartu-Moscow school's semiotics of culture.

Regardless of whether the principle of informativeness of linguistic data about the past of human societies is actually legitimate or not, the fact is that the arguments, hypotheses and theories actually developed in the field of linguistic paleontology have been a source for constant debate not only between linguists but also between archaeologists and historians. A common criterion for delimiting the ultimate validity of the linguistic arguments about the human past has not been reached. A very good illustration of that lack of agreement can be found in the academic discussion carried out in the review of the work *Archaeology and Language: The Puzzle of Indo-European Origins* by the archaeologist Colin Renfrew (1987), published in the journal *Current Anthropology* (Renfrew et al. 1988).

We have already referred to Colin Renfrew as the developer of one of the main hypotheses about the geographical origin and the mode of expansion of the Indo-European languages: *the*

---

<sup>34</sup> More detailed information about the mentioned theories and the current state of the investigations of Indo-European prehistory can be found in the works cited along the paragraph and also in *Mallory 2013*.

*Neolithic theory.* Although it is not relevant for our main discussion here, we may add that his theory locates the Proto-Indo-European *urheimat* in Anatolia and associates the expansion of the Indo-European languages along Europe with the expansion of the agriculture and the neolithisation of the continent. In the work that we just mentioned he not only develops his theory but argues the correlation of archaeological and linguistic data for investigating the prehistory of Indo-European populations<sup>35</sup> and in the review of that mentioned work, several authors argue the theory and the arguments elaborated by Renfrew and along the discussions there is the constant questioning of the ultimate validity of the conclusions of linguistic paleontology to serve as arguments for investigating the prehistoric past, which varies depending on the criterion of each author. The anthropologists David W. Anthony and Bernard Wailes criticize the lack of attention paid by C. Renfrew to linguistic paleontology's arguments:

If even a small part of the Proto-Indo-European lexicon can be reconstructed with some degree of confidence, it can provide the equivalent of a fragmented text – a “window” through which we may see at least some aspects of Proto-Indo-European beliefs, economy, technology, and sociopolitical systems. It seems odd that anthropologically oriented archaeologists should either ignore this potential source of interpretation or, as with Renfrew, treat it in such cavalier fashion. He appears to us too casual in his discounting of the ability of comparative linguistics to discern the semantic content of Proto-Indo-European. [...] To dismiss a whole range of reconstructed terms (floral and faunal, hence environmental; technological, e.g., “wheel”; social, e. g., \**rēg*) as having little or no interpretive significance comes close to denying any historical value to historical linguistics. (Renfrew et al. 1988: 441-442)

In his intervention, the linguist P. Bali presents Renfrew's position on linguistic paleontology in the following way:

The traditional notion to which most Indo-European linguists are exposed is that the early Indo-European-speakers were horse-riding independent nomads who moved out in pulseline thrusts from their original home in the Eurasian steppe. This may be romantic fantasy based on misuse of linguistic paleontology (as Renfrew argues) and unrealistic demographic models, but it does not conflict with the evidence and goals of a formal linguistic analysis. [...] Renfrew criticizes linguists for their willingness to accept migrationist theories and their overreliance on the findings of linguistic paleontology. He criticizes archaeologists for their traditional overemphasis on pottery types and their general ignorance of and disregard of linguistic facts. (Ibid: 445-446)

---

<sup>35</sup> Colin Renfrew comments in this way his objective in *Archaeology and Language: The Puzzle of Indo-European Origins* (1987): “The time has surely come for some fundamental rethinking not only of the Indo-European problem but of the whole relationship between historical linguistics and prehistoric archaeology. Such a reassessment is necessary, I believe, if we are to escape from the myths which currently imprison us within a framework of 19<sup>th</sup>-century origin, constructed at a time when both archaeology and linguistics were in their toothless infancy.” (Renfrew et al. 1988: 441)

Also attending to Renfrew's critique of the validity of the linguistic arguments, the linguist Robert Coleman explains in this way the problematic question concerning linguistic paleontology that leads Renfrew to such critique:

From the range of cognate semantic equivalents in different languages it is possible to build up a picture of the proto-lexicon and from it to infer something of the physical surroundings and culture of the Proto-Indo-European-speakers. The arbitrary and unrigorous methods that have characterized much of this linguistic paleontology certainly deserve Renfrew's scepticism. Nevertheless, some progress has been made, and more is possible. (Ibid: 450)

He further adds:

Most of the lexemes that can be confidently assigned on the basis of widespread attestation (e.g., \**g<sup>w</sup>her-* 'hot', \**ped-* 'foot', \**penk<sup>w</sup>e* 'five', \**en* 'inside', \**bher-* 'carry', \**swep-* 'sleep') do not tell us much. Where the same meaning is carried by two or more lexical roots, it is often difficult to decide whether (a) one is original and the other innovatory, (b) more than one belonged to the proto-lexicon but there were differences in meaning, or (c) all are innovatory, either because the original word has been lost or because the thing itself did not belong to the Proto-Indo-European "world" before the dispersal. [...] Lexical arguments for the reconstruction of Proto-Indo-European cultural institutions and concepts are even more hazardous. The triadic speculations of Dumézil and his disciples deserve even more sceptical treatment than Renfrew gives them, and even Benveniste's work in this field, mentioned more approvingly, is often more stimulating than persuasive. (Ibid: 450)

His reference to Dumézil and Benveniste's works is due to the importance of their contribution to Indo-European studies within the field of linguistic paleontology, even though their conclusions are often, as observed here, put into doubt. Coleman concludes that "Indo-European archaeology and comparative Indo-European linguistics can never work for long in isolation from each other; they are in many ways interdependent" (Ibid: 453).

The discussion carried out along the review of Renfrew's theory that we are quoting here is especially interesting because also Marija Gimbutas participates. She is, as we have already mentioned, the author of another of the most solid theories concerning the origins and dispersion of Indo-European populations: *the Kurgan theory*, which is in many aspects difficultly compatible with Renfrew's one. Her theory locates the Proto-Indo-European *urheimat* in the Eurasian steppe (nowadays Russia), links the Proto-Indo-European community with the nomad and warrior *Kurgan* culture and explains the spread of Indo-European languages as the process of conquest and expansion of this culture (see Gimbutas 1956 and 1963). Her intervention in the discussion referred here exemplifies the variable strength assigned by the investigators to the linguistic arguments, depending of whether they are more or less supportive of their main theories: "If the earliest agriculturalists of Europe were Indo-European-speakers, why are basic agricultural terms non-Indo-European? And why is the terminology associated with religious worship, especially the names of goddesses,

notably non-Indo-European? Renfrew does not ask these key questions” (Renfrew et al. 1988: 453). In concrete, she is denying the equation of the spread of agriculture (neolithisation) in Europe with the expansion of Indo-European languages with the argument that it contradicts the lack of common Indo-European agricultural lexicon, and that means, because it contradicts the testimony of linguistic paleontology<sup>36</sup>:

That agriculturalization was not Indo-Europeanization is shown first of all by the non-Indo-European belief system found among the earliest agriculturalists of Europe. [...] Furthermore, the linguistic substratum of the Europe attested to us is non-Indo-European. [...] For example, most of the names for cereals and legumes cannot be accounted for by Proto-Indo-European as we know it but were taken over by the Indo-Europeans from the local European substrata. Thus the common names for rye, barley, oats, beans, lentils, peas, vetch, poppies, flax, and hemp are known only to the European branch of Indo-Europeans. Some of them, such as names for beans, peas, and vetch, are known only in the south, in Greek, Latin, and Albanian. The word for “apple” is also non-Indo-European substratum in the south. Significant for substratum studies is the fact that in the Indo-European languages of Europe domesticated animals and birds often have two sets of names, non Indo-European and Indo-European. [...] Contrary to the existing linguistic data, Renfrew declares, “If the Proto-Indo-Europeans were familiar with domesticated sheep, goats or cattle, then they must certainly also have been acquainted with wheat, barley and peas and also a range of other animal species”. Except for *\*mel-i-*, ‘a “ground” cereal, millet’, *\*yewo-*, ‘a cereal used for fermentation’, and *\*pūro-*, ‘a grass or spelt wheat’, there are no well-attested Proto-Indo-European words for cereals. There is no linguistic evidence for proto-Indo-European arboriculture and viticulture. [...] This evidence cannot be applied to early European and Anatolian farmers [*as Renfrew proposes*], who were full-fledged agriculturalist. (Ibid: 454)

However, C. Renfrew in his response to Gimbutas’ reliance on the linguistic evidence, plays a completely opposed criterion upon the same question giving more reliance to the archaeological evidence and questioning the validity of the linguistic argument:

In response to my remark, which she quotes, that “if the proto-Indo-Europeans were familiar with domesticated sheep, goats or cattle, then they must certainly also have been acquainted with wheat, barley and peas,” she says that “there are no other well-attested proto-Indo-European words for cereals” (beyond the three which she quotes). That this should be so simply emphasises the weakness of a rigid “protolexicon” approach. The archaeological evidence makes it unlikely that herders of sheep and goats would be unfamiliar with wheat and barley. (Ibid: 465)

The lack of agreement between scholars reflected here on the actual relevance of the capacity of linguistic paleontology for providing knowledge about prehistoric past and the

---

<sup>36</sup> Marija Gimbutas also recognizes the informative capacity of Indo-European mythological comparative studies about a common Proto-Indo-European mythological system of beliefs, and is convinced that such evidence has to be taken into account into the elaboration of theories: “Dumézil’s life work laid the foundation for comparative Indo-European mythology. His studies have shown that mythology reflects an ideological structure that is also manifest in social organization. It cannot be doubted that the proto-Indo-Europeans shared a mythology, social forms, and institutions, and this shared construct is not of recent formation as Renfrew tends to see it. It has been shown that the specific correspondences between many Indo-European legends and myths are far too detailed to be explained by a phylogenetic theory or by one of intracultural loans.” (Ibid: 454-455)

fact that the informative capacity recognized to linguistic arguments ultimately depends, as we have seen, on the interpretative will of the investigator, call into question the ultimate validity of the linguistic arguments resulting from the work within linguistic paleontology about prehistoric past.

Such a doubtful validity gets especially compromising because of the continuous occurrence of abundant implausible hypotheses that has taken place within the investigations on linguistic paleontology during the two centuries of historical-comparative studies, where many of the theories proposed about the societies carrier of protolanguages have been proved false. That implausibility arises from an overestimation of the informativeness of these linguistic data about prehistoric past, which leads to wrong assumptions about the prehistoric past that lack of objectivity and fundament, and therefore provides nothing but pseudo-science. Indeed, as a result of the unstable fundamentals of the linguistic paleontological approach, the investigation of the origin of the Indo-European languages and of the characteristics of the Proto-Indo-European society has witnessed many examples of pseudo-scientific hypotheses<sup>37</sup>.

---

<sup>37</sup> Even tendentious theories have arisen from these field of studies, such as the appropriation of the term “Arian” by Nazi propaganda in Germany, which was primarily a term reconstructed as the self-denomination of the hypothetical Proto-Indo-European society merely because cognates of that term are attested with that same purpose of social self-denomination in different branches of the Indo-European linguistic family. The pseudo-scientific conviction of an ever existing Proto-Indo-European Aryan society that expanding from Central/Northern Europe ruled by elite-dominance vast lands from Western Europe to India, that was based on arguments of linguistic palaeontology not duly attested, was employed as the legitimation of ethnic superiority. In fact, after this appropriation of the term “Aryan” by the Nazi propaganda, it is avoided within Indo-European studies, despite of being the most appropriate one to refer to the Proto-Indo-European society for that very reason of its testimony in many branches of the family. On this issue: “The problem of Indo-European origins is one which has underlain much of the development of prehistory in Europe from the common beginnings of prehistoric archaeology and comparative linguistics in the Romantic movement of the early 19<sup>th</sup> century. Many of the earlier comprehensive accounts of European prehistory were written with this problem in mind. In the last 50 years, however, the explicit racism which in Germany was the ultimate outcome of the Romantic search for ethnic origins led many prehistorians to avoid this issue, in reaction to the political abuses of archaeology under the Nazis.” (Andrew Sherratt in Renfrew, C. et al. 1988: 458)

#### **4. The informativeness of linguistic data about prehistoric past**

We confront three problematic epistemological questions concerning the discipline of linguistic paleontology. The first one is the theoretical assumption of the informativeness of linguistic data about prehistoric past, which is the ground for these investigations. We will find the theoretical legitimation for such assumption in the semiotic understanding of culture and in the understanding of the role of linguistic systems inside cultural systems carried out by *semiotics of culture* and developed within the Tartu-Moscow school of semiotics.

The second question is the one about the ultimate validity of the linguistic arguments about the past of human populations, which is of capital importance considering the controversial character of the resulting conclusions provided by linguistic paleontology, and the third question is the profusion of pseudo-scientific arguments and theories concerning human past that have arisen from within linguistic paleontology. In view of such problematic questions, we will discuss the epistemological status of linguistic paleontology in order to concrete the ultimate value of the testimony about human past that it provides and this will lead us to recognize the interpretative character of this field of investigation, that opens the door for implausible conclusions in spite of the ultimate and previously mentioned theoretical ground for its approach. The rising of pseudo-scientific theories within linguistic paleontology is to be explained in terms of overinterpretation and misinterpretation of the linguistic data available. This will lead us to conclude that the knowledge about past reachable by means of this discipline cannot seek for objectivity, but for plausibility and that it is irremediably dependent on the data provided by other disciplines that deal as well with prehistoric past.

#### 4.1. The role of language inside culture from the point of view of the Tartu-Moscow school's semiotics of culture

The aim of our work at this point will be to describe the functioning of linguistic semiotic systems inside human culture in order to discuss the capacity for informativeness of linguistic systems about non-linguistic aspects of culture, which is the ground on which *linguistic paleontology* bases its investigational approach. The ultimate nature of the question demands a semiotic theoretical approach<sup>38</sup>, and we will implement concretely the semiotic conception of *culture* as a semiotic system that is built upon a hierarchy of different semiotic systems are interrelated and in functional correlation<sup>39</sup>, which is offered by the Tartu-Moscow school's *semiotics of culture* (or *cultural semiotics*)<sup>40</sup>.

The main general assumptions of the Tartu-Moscow scholars about *culture* and its dynamics crystallized in the publication of the *Theses on the semiotic study of cultures (as applied to Slavic texts)* in 1973 (Ivanov; Lotman et al.), which “marked the real birth of the research field called the *semiotics of culture*” (Salupere et al. 2013: 5). There we find the following definition of *culture*: “From the semiotic point of view culture may be regarded as a

---

<sup>38</sup> On the ultimate semiotic nature of cultural phenomena – including natural language – and on the pertinence of semiotics as theoretical approach for its study, we compile the following words by Peeter Torop: “[T]he essence of culture is semiotic by its very nature, since its foundation is information and communication. On one hand, the study of culture would be possible via the semiotisation of culture-studying disciplines, which would bring them closer to the essence of culture. [...] cultural semiotics offers a systematic approach to culture and creates a complementary methodology, which ensures the mutual understanding of different culture-studying disciplines.” (Torop 2006: 305-306)

<sup>39</sup> The origin of such conception lies at the very beginning of the semiotic work of Tartu-Moscow scholars: “The different modelling semiotic systems form complex hierarchical series of levels, where the system of the lowest level (as for example, natural language) serves for the codification of signs that belong to systems of the higher level (for example, the sign systems of art and science): in its turn, each of the sign systems that belong to these hierarchical series can form an ordered (or partially ordered) sequence of levels. The most developed ones are the methods for investigation of the relations between different levels of natural languages (within structural and mathematical linguistics).” (Ivanov 1962: 6); «Различные моделирующие семиотические системы образуют сложные иерархические ряды уровней, где система низшего уровня (например, естественный язык) служит для кодирования знаков, входящих в системы высшего уровня (например, знаковых систем искусства и науки): в свою очередь, каждая из знаковых систем, входящих в эти иерархические ряды, может образовывать упорядоченную (или частично упорядоченную) последовательность уровней. Наиболее разработанными являются методы исследования отношений между разными уровнями в естественных языках (в структурной и математической лингвистике).» (Ivanov 1962: 6) [Translated by the author]

<sup>40</sup> The immediate predecessor of Tartu-Moscow *semiotics of culture* is to be found in the Prague Circle's work. This influence has been well explained by Thomas Winner: “[T]he Prague Linguistic Circle made the first steps toward the rich development of the semiotics of culture in its application of modified Saussurian concepts, in its development of the concept of function and of polyfunctionality of human behavior, in its position of the interrelatedness of systems, and finally, in the person of Bogatyrev, in its view of the opposition of the inner vs. outer sphere of culture in which the former is seen as organization whereas the latter is regarded as entropy and chaos” (Winner, T. 1979: 80-81). Further explanations of the most important external influences on the Tartu-Moscow school can be found in Salupere and Torop 2013: 16-20 and Hymes 1978: 399-404.



hierarchy of particular semiotic systems, as the sum of the texts and the set of functions correlated with them, or as a certain mechanism which generates these texts” (Ivanov; Lotman et al. 1973: 68). From the semiotic point of view *culture* is therefore to be understood as a complex organization of sign systems. In the works of Yuriy Lotman we find the following conception of culture:

“Culture presents many different aspects, and its apprehension is, to a great extent, defined by the meta-position of the researcher. From the semiotic point of view, culture can be conceived as a complexly organized sign mechanism that permits the existence of one or another group of people as a collective identity with a certain common supraindividual intellect, a common memory, a unity of behaviour, a unity of modelling for itself the world surrounding and unified relations towards that world.” (Lotman 1977: 138)<sup>41</sup>

It is “a historically existing bunch of semiotic systems (languages) that can arise as one hierarchy (superlanguage) or appear as a symbiosis of independent systems. It includes not only a certain combination of semiotic systems, but also the totality of messages that have historically been produced in those languages (texts)” (Lotman 1970: 6-8, 28.0)<sup>42</sup>.

A very first implication of the switch in the theoretical conception of *culture* from the point of view of *cultural semiotics* is that it permits the recognition of verbal language as a phenomenon of the same entity as any other semiotic system operating inside culture. We may refer to all cultural semiotic systems from a similar stance and with a common terminology (and a common modelling), and we may as well identify similar dynamics of staticity and changeability in all of them. But the most important implication of the approach by *semiotics of culture* for the issue under discussion here is that it postulates that linguistic semiotic systems occupy a “central position” inside culture. In fact, culture is claimed to be “a system of systems based in the final analysis on a natural language” (Ivanov; Lotman et al. 1973: 71). Yuriy Lotman and Boris Uspenskij, in their article *On the semiotic mechanism of culture* (1978) expose this idea and conclude that linguistic semiotic systems play a central role inside the whole system of culture because of the high degree of structuredness that they

---

<sup>41</sup> «Культура многоаспектна, и определение ей, в значительной мере, определено метапозицией исследователя. С семиотической точки зрения культуру можно представить себе как сложно организованный знаковый механизм, обеспечивающий существование той или иной группы людей как коллективной личности, обладающей некоторым общим сверхличным интеллектом, общей памятью, единством поведения, единством моделирования для себя окружающего мира и единством отношения к этому миру.» (Lotman 1977: 138) [Translated by the author]

<sup>42</sup> «Итак, культура - исторически сложившийся пучок семиотических систем (языков), который может складываться в единую иерархию (сверхязык), но может представлять собой и симбиоз самостоятельных систем. Но культура включает в себя не только определенное сочетание семиотических систем, но и всю совокупность исторически имевших место сообщений на этих языках (текстов).» (Lotman 1970: 6-8, 28.0 [as referenced in: Levchenko and Salupere (eds.) 1999: 121]) [Translated by the author]

possess, which make them especially significant in the structuring function of the whole culture:

As a methodological abstraction, one may imagine language as an isolated phenomenon. However, in its actual functioning, language is molded into a more general system of culture and, together with it, constitutes a complex whole. The fundamental "task" of culture, as we will try to show, is in structurally organizing the world around man. Culture is the generator of structuredness, and in this way it creates a social sphere around man which, like the biosphere, makes life possible; that is, not organic life, but social life.

But in order for it to fulfill that role, culture must have within itself a structural "diecasting mechanism." It is this function that is performed by natural language. It is natural language that gives the members of a social group their intuitive sense of structuredness that with its transformation of the "open" world of realia into a "closed" world of names, forces people to treat as structures those phenomena whose structuredness, at best, is not apparent. Indeed, in many cases it turns out not to matter whether some meaning-forming principle is a structure, in a strict sense, or not. It is sufficient that the participants in an act of communication should regard it as a structure and use it as such for it to begin to display structure like qualities. One can well understand how important it is that a system of culture has, at its center, so powerful a source of structuredness as language. (Lotman and Uspenskij 1978: 213).

From our point of view, the "high degree of structuredness" of linguistic semiotic systems pointed out by Lotman and Uspenskij stems from the high degree of structuredness and systematicity observable at specifically *the syntactic dimension* of linguistic semiotic systems. The inner highly systematic structure of linguistic semiotic systems turns them into the primary structural modelling tools inside culture.

These two aspects concerning the functioning of natural language inside a cultural system may seem contradictory at first sight but are actually compatible: on the one hand, verbal language inside culture is a semiotic system of the same entity as other codes and conventions that take part in it: the codes or conventions for clothing, feeding, acting in public, political organization, etc., because all of them are ultimately semiotic systems and their dynamics inside culture and society are not essentially different from the ones of verbal languages. This permits an equation in conceiving language and the rest of cultural facts inside a society. But at the same time linguistic semiotic systems occupy a central role in the system of culture and that confers them an especially significant position in its interrelation with the rest of systems inside a given culture.

The semiotic understanding of culture has important implications for explaining the investigation of past cultural stages. Yurij Lotman in his article *Problems in the typology of culture* (1988) defines culture as "the totality of nonhereditary information acquired, preserved, and transmitted by the various groups of human society" and distinguishes three aspects of it: culture as a *structure*, as *information* and as a *hierarchy of codes* (cf. Lotman,

Yu. 1988: 213-214). In relation to the two first aspects, when facing a cultural *text*<sup>43</sup> (and that means not only a verbal message in written form but any kind of material or immaterial fact distinguishable inside a cultural system), two dimensions are to be recognized in it: its *content* and its *structure*. The *content* possesses the value of being informative about the cultural system it belongs to; it is the information it presents by itself. But provided that culture is always a system of codes that intertwine with a certain degree of systematicity and structuredness, from a given cultural text may be grasped not only this *content* but also other aspects of the cultural system that it belongs to. A cultural text belongs to the *structure* of culture and it is necessarily related to a language or code inside that culture, which is in turn interrelated with the rest of codes or systems that operate inside that same culture. That is the reason why that cultural text is informative of more aspects of the culture than strictly the primary *content* that it transmits.

In constructing a typological and structural history of culture, we must necessarily base our analysis on a separation of the content of cultural texts from the structure of their “language.” In considering the sum total of facts available to the historian of culture, we must also distinguish between the system that can theoretically be reconstructed (a culture’s “language”) and the way in which the culture is realized from the mass of material external to the system (a culture’s “speech”). In this way, we can examine all the facts in the history of culture from two points of view: as significant information, and as the system of social codes that permits the expression of this information with signs in order to make it the patrimony of a human collectivity. (Lotman, Yu. 1988: 214)

Lotman illustrates the situation with the case of material cultural objects; they don’t only serve to a practical purpose inside the cultural system in which they are produced but, precisely because of being products of the cultural system, they possess informative value about that culture. From the inside of a cultural system, a given object serves to a concrete

---

<sup>43</sup>The concept of *text* is especially important in the semiotics of the Tartu-Moscow school and needs to be clearly distinguished from the traditional notion of text in linguistic and literary studies. Conceived in a semiotic sense, the concept of text “is applied not only to messages in a natural language but also to any carrier of integral (“textual”) meaning” (Ivanov et al. 1973: 58) and more concretely in *cultural semiotics* is conceived as text any “message which is constructed within the given culture according to definite generative rules” (Ivanov; Lotman et al. 1973: 63).

On the switch of attention to *text* in the Tartu-Moscow semiotic school: “What is the relation between sign and text (which is the main object for investigation of the members of the Tartu-Moscow School)? Can we consider the text (and not only word signs, which is what it is made of) a sign of a peculiar kind? Modern semiotics, unlike the first attempts of the continuators of Saussure and Peirce [...] study not only signs, but also their combinations, regardless of their length or complexity. The text stands as a succession of signs that plays itself the role of sign in certain cultural traditions” (Ivanov 2010: 33-34);

«Каково соотношение знака и текста (являющегося главным объектом исследования представителей московско-тартуской школы)? Можно ли считать текст (а не только слова знаки, из которых он состоит) особым рода знаком? Современная семиотика, в отличие от первых опытов последователей Соссюра и Пирса [...] изучает не только знаки, но и их сколь угодно длинные и сложные сочетания. Текст представляет собой последовательность знаков, которая в определенных культурных традициях сама выступает в роли знака.» (Ivanov 2010: 33-34) [Translated by the author]

purpose, but from the outside of that cultural system, that same object is informative further from its primary purpose inside it, about the cultural system in general:

In fact, even when faced by the so-called monuments of material culture (such as the means of production) we must bear in mind that these objects perform a double role in the society that creates and uses them; although they serve a practical purpose, they also concentrate in themselves the experience acquired during past working activity and ultimately become instruments for preserving and transmitting information. For contemporaries, who can receive this information from numerous more readily available sources, the first of these functions is the fundamental one; however for their successors, who may be archaeologists or historians, the second function will be the only one that matters. (Lotman, Yu. 1988: 213)

The third and last aspect of culture pointed out by Lotman – culture as a *hierarchy of codes* – refers to the already mentioned conception of *culture* as a hierarchical interrelation of semiotic systems (which here are referred as *codes*<sup>44</sup>). The existence of such interrelations is what permits the “double informativeness” of *cultural texts*: about its proper *content* and about the *structure* of its cultural system. We want to bring up the connection between Lotman’s explanation and the understanding by Ivanov and Toporov of the investigation of ancient history and prehistory as a process of deciphering (or decoding), in which the message to be restored – and that is the ancient or prehistoric culture under investigation – reaches us via different kinds of signals:

“The possibility of mutual translations between different semiotic modelling systems and the existence of many of these systems in any human collectivity turns unlimited and reliable the transmission of the content-plane of the message, which is related to the culture of that collectivity. It is for this reason that from the point of view of semiotics we can understand the task of such complex studies as, for example, the investigation of Slavic antiquity, as the restoration at different levels of a message on the basis of signals, which arrived to it by means of different channels of communication. Belong to those signals the existence of archaeological monuments and evidences of material culture, which can be interpreted as messages in a definite code, as well as data obtained by cultural anthropology. [...] All the complex of signals that are described in different historical and ethnographical disciplines and interpreted mostly by means of typological confrontation, can serve both for the reconstruction of the message itself, which is related to that given culture (for instance, proto-Slavic culture)<sup>45</sup>, as well as for the reconstruction of the social background of that collectivity, the knowledge of which permits to establish the semantic

---

<sup>44</sup> For the topic under discussion here, the distinction between the terms *semiotics system*, *code* and *modelling system* is not relevant and therefore, they are just treated as equivalents, but they should not be taken as equivalent within the works of the Tartu-Moscow school. The term *modelling system* will be especially relevant within the school, namely for the distinction between *secondary* and *primary modelling systems*: “language as a primary modelling system is the human being’s main means of thinking and communicating. As a secondary modelling system, language is the preserver of the culture’s collective experience and the reflector of its creativity” (Salupere and Torop 2013: 26).

<sup>45</sup> The investigations by V. V. Ivanov and V. N. Toporov about Proto-Slavic linguistic and cultural semiotic systems can be found in *Ivanov and Toporov 1963* and *1965*.

and pragmatic aspects of the reconstructed message.” (Ivanov and Toporov 1997 (1966): 73)<sup>46</sup>

How does this theoretical understanding of cultural facts translate to the investigation of prehistory and, more specifically, to the investigation of prehistoric societies past on the basis of linguistic data? In the first place because the claiming by Lotman refers by way of illustration to the material objects that are studied in archaeological and historical research, but in fact the same double informativeness can be conceived about also the linguistic data concerning past periods provided by *historical and comparative linguistics*. The “double informativeness” of cultural texts described by Lotman and the “mutual translations” between semiotic systems inside culture pointed out by Ivanov and Toporov concern also linguistic data, because they are also cultural texts and because linguistic systems are also part of that mutual translation between systems. What’s more, they occupy, as we have seen, a central role in between the interrelation of semiotic systems.

Besides, we have to keep in mind that any semiotic system, and that includes linguistic semiotic systems, is both the inheritance of its previously stages and the effect of partial but constant modifications, presenting a tendency to stability and preservation together with a tendency to constant change and innovation<sup>47</sup>. As posited by Lotman and Uspenskij, “culture is memory or, in other words, a record in the memory of what the community has experienced” and therefore it is, “of necessity, connected to past historical experience.”

---

<sup>46</sup> «Возможность взаимных переводов между разными моделирующими семиотическими системами и наличие многих таких систем в любом человеческом коллективе делает помехоустойчивой и надежной передачу плана содержания сообщений, относящихся к культуре этого коллектива Поэтому с точки зрения семиотики задачу таких комплексных наук, как, например, исследование славянских древностей, можно понять как восстановление на разных уровнях таких сообщений на основании сигналов, дошедших до него по разным каналам связи. К числу таких сигналов относятся и различные археологические памятники и свидетельства материальной культуры, которые можно интерпретировать как сообщения в определенном коде, так же как и данные, получаемые культурной антропологией [...]. Весь комплекс сигналов, описываемых в различных исторических и этнографических дисциплинах и интерпретируемых преимущественно путем типологических сопоставлений, может служить как для реконструкции самих сообщений, относящихся к данной культуре (например, праславянской). Так и для реконструкции того социального фона и того коллектива, без знания особенностей которого нельзя восстановить семантическую и прагматическую сторону реконструируемых сообщений.» (Ivanov and Toporov 1997 (1966): 73) [Translated by the author]

<sup>47</sup> “The transmission of culture in time can be to a great extent described as the conservation of systems of signs, which serve to the control of behaviour. The systems of programmed sign-control of human behaviour, developed thanks to a switch of external signs to the inside, usually work automatically, and that means unconsciously (related to this is the conservation of surviving forms that permit the reconstruction of archaic types of behaviour).” (Ivanov 1962: 4-5);  
«Передача культуры во времени в большой степени может быть описана как сохранение систем знаков, служащих для контроля над поведением. Системы знакового программного управления человеческим поведением, выработанные благодаря переходу внешних знаков внутрь, обычно работают автоматически, т. е. бессознательно (с этим связано и сохранение многих пережиточных форм, позволяющих реконструировать архаичные типы поведения).» (Ivanov 1962: 4-5) [Translated by the author]

(1978: 214). Taking this fact into consideration and having reported on the hierarchical interrelation of semiotic systems inside culture, we can accept that the preservation of inherited features in any of the semiotic systems inside culture are informative not only of the past stages of evolution of that semiotic system but also of past stages of the whole cultural system in which it operates.

We can therefore conclude the linguistic data provided by means of *historical and comparative linguistics* are to be recognized a capacity for informativeness not only about past linguistic stages but also of the extra-linguistic characteristics of the societies in which they were in use. This stands as the theoretical recognition of the informativeness of linguistic systems about the cultural system in which they operate, which is the underlying assumption that leads to the investigating work within linguistic paleontology. In addition, the primary role of linguistic semiotic systems in the function for modelling and structuring characteristics of culture permits to talk about an especially accurate capacity for informativeness.

#### **4.2. Linguistic paleontology as a field for interpretative investigations**

The linguistic arguments concerning the ancient and prehistoric past of human societies provided in the work within *linguistic paleontology* are recognized different validity, as we have already witnessed, at the moment of establishing actual theories about that past. In addition, and as we have pointed out along our work, the investigation of human past by linguistic means has led to a profusion of pseudo-scientific theories and that had put into question the validity of such approach. Even the linguistic arguments that have been most solidly accomplished are present problems for delimiting what is the ultimate validity of the testimony they offer, and are ultimately accepted or rejected only on the basis of their adequacy or inadequacy to the theory to be defended. Facing such questions, the academic status of *linguistic paleontology* is rather weak.

The questionable value of linguistic arguments concerning human past cannot be blame on the total incapacity of informativeness of linguistic facts about that past because, as we have previously discussed, there is a theoretical legitimacy in assuming that informativeness. We argue that the uncertain validity of the arguments provided by linguistic paleontology derives

rather from the interpretative character of its investigations. The strictly linguistic work carried out by *historical and comparative linguistics* departs, as we have exposed, from an objectification of its object of study that permits the implementation of the scientific method and the implementation of mechanisms of deduction and induction. However, the inference of non-linguistic aspects of cultural systems carried out by *linguistic paleontology* is a task based on the mere interpretation of the linguistic evidence. As this field of studies deals mainly with the lexical meanings of the linguistic data available, which is not suitable of such systematic study and verifiability as the (previously described in this work) *syntactic dimension*, it cannot possess an interpretative character, with no ground for unequivocal deduction or induction:

When reconstructing the phonological aspects of the lexical entities, we have strict criteria that help us distinguish between genetically conditioned and typologically conditioned features. In the case of semantic reconstruction, these criteria are far more obscure. In general, linguists are guided by vague ideas of semantic similarity; at best, they rely on typologically similar cases of semantic change that are historically attested for different languages. Thus, the reconstruction of the so-called “world picture” for any proto-ethnos often places the researcher on shaky ground. (Dybo 2013: 70)

The ultimate dependence on the interpretative capacities or will of the investigator is what makes the resulting linguistic arguments a matter of debate between scholars and leads to the impossibility of delimiting clear criteria about the ultimate informative validity of their testimony about past facts and cultures. Another consequence of this interpretative character of linguistic paleontology is that there is always the possibility for misinterpretation and overinterpretation: “Comparison of vocabulary is less rewarding than might be hoped. [...] Where there are apparent matches in significant cultural vocabulary, we must be careful that our eagerness to find out about Proto-Indo-European society does not lead us to over-interpret the linguistic evidence” (Clackson 2007: 210). The misinterpretation or overinterpretation of the available linguistic data leads to implausible conclusions about the period of the past and the society under investigation and produce nothing but pseudo-knowledge about the past of human populations.

## **Conclusive remarks: the validity of linguistic arguments for the investigation of human past**

As assumed by *linguistic paleontology*, linguistic data possess indeed informativeness about human past. However, and due to the ultimately interpretative character of its investigations this discipline cannot seek for objectivity in its hypothesizing of ancient and prehistoric past, but can only seek for the plausibility of the linguistic arguments that it offers as testimony of that past. *Linguistic paleontology* is therefore not a suitable source of reliable and objective knowledge about human past.

This shouldn't be understood as the denial of any possible reliability of the knowledge about past acquired by means of linguistic analysis, but rather as a delimitation of its ultimate epistemological value<sup>48</sup>: “The whole doctrine of making cultural inferences from linguistic evidence, known as linguistic paleontology, has rarely enjoyed particularly high repute. But in any case it remains true that linguistic paleontology very neatly provides assumptions to be tested” (Anttila, R. 1989: 373). The linguistic arguments provided about human past must, as we argue, aim for plausibility in view of the data available but cannot be a matter for verification<sup>49</sup> in itself.

---

<sup>48</sup> In order to provide the reader with a concrete exemplification of the informative capacity of linguistic data in the investigation of prehistoric past, we will briefly refer to the work of the Russian linguist Anna Dybo who, in his article *Language and archaeology: some methodological problems. 1. Indo-European and Altaic landscapes* (2013) compares the reconstructed landscape lexicon for the Indo-European and the Altaic linguistic families, showing how different landscapes they reflect. The linguistic material of ancient etymology in both these families draws a conceptualization of landscape that is proper only of only that linguistic group, and therefore informative of the prehistoric geographical location of its speakers before its dialectalization.

<sup>49</sup> In the following words, Vyacheslav Ivanov argues that the conclusions of linguistic paleontology are suitable for verifiability, but he is at the same time recognizing that such verification is found only outside the linguistic arguments, in evidences of non-linguistic kind:

“The considerable success of modern linguistic studies in the field of studying the semantic spheres of languages makes possible and desirable a detailed study of the diachrony of more semantic fields. This task offers results that are verifiable and usually allow falsification (in the sense of K. Popper’s philosophy of science) by the confrontation of the denotata (referents) of words with data from adjacent sciences. But at the same time this task [the confrontation of data] is very time-consuming and may therefore slow down the run of the respective comparison of lexicons in whole. That is why this research should be encouraged on the later steps, after the solution in advance of the general establishment of the kinship.” (Ivanov 2009: 2);

«Значительные успехи современной лингвистической науки в области изучения смысловых сфер языка и референции делают возможным и желательным детальное диахроническое изучение больших семантических полей. Это занятие дает проверяемые результаты и часто допускает фальсификацию (в смысле философии науки К. Поппера) при сопоставлении денотатов (референтов) слов с данными



The lack of verifiability and objectivity makes of linguistic paleontology an insufficient source of information about the past of human populations, but rather a field dependent on the data provided by other disciplines that deal also with investigation of human past, namely archaeology and populations' genetics, whose data are complementary to the ones offered by linguistic paleontology and stand, because of the nature of its object of study, a source for objectivity and univocality. This being so, we conclude that linguistic paleontology is a source for plausible arguments about the past of human cultures that depends on the confrontation with evidences accomplished by other disciplines in order to provide reliable knowledge about that past. The importance and relevance of the linguistic contribution to the study of human past lies on the fact that the kind of evidences that is able to provide are complementary to the ones provided by other disciplines and would be unreachable by other means.

In the first chapter we concluded that the investigations of *historical and comparative linguistics* can provide verifiable and scientific knowledge for strictly – as we called it – the *syntactic dimension* of languages and not for the actual manifestation of language at the level of concrete use. This has important implications for discussing the ultimate entity of the informativeness of linguistic data about human past and about the ultimate validity of the knowledge provided by linguistic paleontology. To the same extend as we recognized at the synchronic level that the abstract metalinguistic description of a given natural language is representative of the regularities observable at the level of speech, although not identical to it, we can recognize a certain degree of representativeness in the sources attesting ancient languages and in reconstructed protolanguages about stages of language as actually manifested at the level of speech. However, if for the cases of synchronic description and ancient languages attested in written sources, the metalinguistic structural description is actually built upon the actual level of speech, for the case of reconstructed protolanguages the situation is the opposite one: we infer features of the abstract system without any further evidence of the actual level of speech. For this reason, the representativeness of protolanguages about the actual level of speech of the linguistic stage they reflect is lessened. What is more, the reconstructed protolanguages are linguistic hypotheses and are not suitable for verification – except for certain cases such as the discovering of a correlation in Hittite of

---

смежных наук. Но вместе с тем оно является трудоемким и может поэтому замедлить ход соответствующих сравнений словарей в целом. Поэтому такое исследование скорее должно поощряться на более поздних этапах, после предварительного решения общей задачи установления родства.» (Ivanov 2009: 2) [Translated by the author]

Saussure's sonant coefficient, described above, in which occurs an actual empirical verification of the linguistic hypothesis – and that turns the arguments of *linguistic paleontology* based on these reconstructions into hypotheses built upon previous hypotheses, what again weakens the validity of such arguments.

On the other hand, the fact that the knowledge provided by *historical and comparative linguistics* is the most scientific for strictly *the syntactic dimension* of linguistic semiotic systems means that the conclusions of *linguistic paleontology* that are based on that syntactic dimension can be recognized more verisimilitude than the ones that rely on the lexicon. We already pointed out the fact that upon phonological, morphological and typological features of related languages also conclusions about the non-linguistic aspects of the past of human populations are extracted. Such arguments are most of the times employed by scholars without recognizing them as linguistic paleontology, but just as implicit logical conclusions of the linguistic evidence, but they are also making use of the informativeness of linguistic data about the extra-linguistic space. In order to argue this statement, we will turn to the following words by the linguist P. Baldi for criticizing Colin Renfrew's implicit delimitation of linguistic paleontology to the lexical level:

The most serious error in the linguistic argumentation [of C. Renfrew] is, however, the lack of concern with any features of language beyond the lexicon. Renfrew repeatedly criticizes the notion of the protollexicon and the deficiencies of linguistic paleontology. Why not, then, make use of the deep, *structural* features of phonology and morphology? What of the shared *nonlexical* characteristics of different groups, such as the morphological systems of Greek, Latin, and Sanskrit, the morphological parallels between Germanic and Hittite, the conservative phonologies of Armenian and Germanic, the phonological parallels between Hittite and Tocharian, the rich inflectional systems of some groups (Greek, Italic, Indo-Iranian, Baltic, Slavic) versus the less dense ones of others (Germanic, Hittite)? What about evidence for linguistic subgrouping (e.g., Italo-Celtic, Balto-Slavic)? While Renfrew justly criticizes the lexical approach, he nowhere makes use of the phonological, morphological, or even syntactic correspondences that are of so much value in establishing the protosystems of Proto-Indo-European. (Renfrew, C. et al. 1988: 447)

The plausibility of the information extracted by means of *linguistic paleontology* about human ancient and prehistoric past and the validity of its arguments are, in conclusion, higher when refer to ancient attested linguistic stages than to hypothetically reconstructed ones and, concerning the latest case, the linguistic arguments about the prehistoric societies carrier of protolanguages are more solid when extracted from the reconstruction of *the syntactic dimension* of linguistic semiotic systems, and that means, their phonological and morphological levels and their typological features, while they are much a source for uncertainty and disagreement between scholars when extracted from the consideration of the lexicon.

## CONCLUSIONS

Two academic disciplines dealing with the investigation of human past have been the matter for discussion in the present work. On the one side, *historical and comparative linguistics* as the investigation of the diachronic dimension of linguistic systems and on the other side, *linguistic paleontology* as the investigation on the basis of linguistic data of non-linguistic aspects of the past of human populations. The consideration of certain semiotic aspects involving the work within such disciplines has led us to the following conclusions concerning the knowledge about past provided by them.

The object for study of *historical and comparative linguistics* is the structural metalinguistic descriptions of natural languages that result from the structuralist approach in linguistics. It is because of the systematic nature of that structurally constructed object that *historical-comparative method* can deal with objectivity in a way comparable to the method of natural sciences, inasmuch as only upon systematicity of the object for study an inductive and scientific methodological approach can be applied. Besides, in a semiotic approach to linguistic sign systems, we can distinguish in them three different dimensions: the *semantic dimension* – the relations between the system and the extra-linguistic reality –, the *pragmatic dimension* – the relations between the system and the user of it in actual situations, the dimension of praxis – and the *syntactic dimension* of the system – the relations of linguistic signs between themselves, and that means, internal relations between elements of the system. As systematicity and structuredness are only reachable at the level of abstract description and specifically for the case of the relations of linguistic signs that are elements of the system between themselves, we claim that the historical-comparative approach to languages scopes ultimately *the syntactic dimension* of linguistic semiotic systems and can provide verifiable and scientific knowledge for strictly that *syntactic dimension*.

In the approach by the *semiotics of culture* developed within the Tartu-Moscow semiotic school, human culture is described as a hierarchical interrelation of semiotic systems that models and structures the extra-cultural space. Linguistic semiotic systems play a central role in that hierarchy of systems because of its especially accurate capacity for structuration,

which – as we argue – arises from the high degree of structuredness observable at specifically their *syntactic dimension*. Because of that central role, linguistic semiotic systems possess high informativeness about non-linguistic aspects of cultural systems, and because of the fact that all cultural systems are heirs of their past stages of evolution, they possess also informativeness about non-linguistic aspects of the past of cultural systems. *Linguistic paleontology* investigates non-linguistic aspects of human past on the basis of linguistic data and upon the assumption of that informativeness of linguistic semiotic systems about the extra-linguistic space. However, the ultimate validity of its conclusions is a matter for debate and disagreement in order to provide reliable knowledge about the past of cultures.

The problematic status of *linguistic paleontology* is due to the fact it ultimately consists on an *interpretation* of the linguistic evidence, suitable therefore for *overinterpretation* and *misinterpretation* of data. We conclude that because of its interpretative approach, these investigations cannot seek for objectivity but only for plausibility in the linguistic arguments that they offer as testimony of that past. These linguistic arguments lack of verifiability by themselves and can only be discussed as more or less plausible in the light of other kinds of evidences. *Linguistic paleontology* in isolation is thus not a suitable source of reliable and objective knowledge about human past, but a source for plausible arguments, and it is dependent on the confrontation with evidences provided by other disciplines such as archaeology or genetic studies of populations in order to provide reliable knowledge about human past.

We may make some more concrete distinction inside the different possible investigations covered within this field: the information extracted by means of *linguistic paleontology* about human ancient and prehistoric past is more plausible and valid when it refers to ancient attested linguistic stages than when it deals with hypothetically reconstructed protolanguages. For the latest case, the linguistic arguments about the prehistoric societies carrier of protolanguages are more solid when extracted from the reconstruction of *the syntactic dimension* of linguistic semiotic systems, while they are much a source for uncertainty and disagreement when extracted from the consideration of the lexicon.

## List of references

- Adrados, Francisco Rodríguez 1971. La semántica estructural: estado actual y perspectivas. In: *Habis* N° 2, 9-34.
- Anttila, Raimo 1989 [1972]. *Historical and comparative linguistics*. Amsterdam and Philadelphia: John Benjamins Publishing Company.
- Austin, John Langshaw 1962 [2005]. *How to do things with words*. Cambridge: Harvard University Press.
- Belov, Aleksej Mihajlovich 2008. Вопрос о морях (оппозиция арифметической кратности в греческом и латинском языках). In: *Discipuli magistro (к 80-летию Н. А. Федорова)*. Moscow.
- 2011. Проблема моры в индоевропейских аблаутных чередованиях: корни с “тяжёлыми базами”. In: *Известия РАН. Серия литературы и языка*, 49-54.
- Benveniste, Émile 1971 [1966]. *Problems in General Linguistics*. Coral Gables, Florida: University of Miami Press.
- 1981 [1969]. The semiology of language. In: *Semiotica*, supp. (1981), 5-23.
- Birnbaum, Henrik 1984. Notes on syntactic change: co-occurrence vs. substitution, stability vs. permeability. In: Jacek Fisiak (ed.). *Historical Syntax*. Berlin: Walter de Gruyter, 25-46.
- Bobes Naves, María del Carmen 1973. *La semiótica como teoría lingüística*. Madrid: Gredos.
- Bopp, Franz 1816. *Über das Conjugationssystem der Sanskritsprache in Vergleichung mit jenem der griechischen, lateinischen, persischen und germanischen Sprachen*. Frankfurt am Main.
- 1833. *Vergleichende Grammatik des Sanskrit, Zend, Griechischen, Lateinischen, Littauischen, Gothischen und Deutschen*. Berlin.
- Cavalli-Sforza, Luca; Paolo Menozzi and Alberto Piazza 1994. *The history and geography of human genes*. New Jersey: Princeton University Press.
- Chomsky, Noam 2002 (1957). *Syntactic structures*, Berlin: Walter de Gruyter.
- Clackson, James 2007. *Indo-European Linguistics. An Introduction*. Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi: Cambridge University Press.
- Dybo, Anna 2013. Language and archaeology: some methodological problems. 1. Indo-European and Altaic landscapes. In: *Journal of Language Relationship* 9 (2013), 69– 92.

- Gamkrelidze, Tamaz Valeryanovich 2008. Linguistics and paleontology of culture. In: *Bulletin of the Georgian National Academy of Sciences*, vol. 2, no. 4: 143-147.
- Gamkrelidze, Tamaz Valeryanovich and Ivanov, Vyacheslav Vsevolodovich 1990. The Early History of Indo-European Languages. In: *Scientific American* (march 1990), 110-116.
- 1995 [1984]. *Indo-European and the Indo-Europeans. A Reconstruction and Historical Analysis of a Proto-Language and a Proto-Culture [Part I and Part II]*, Berlin and New York: Mouton de Gruyter.
- Gertsenberg, Leonard Georgievich 1978. Реконструкция индоевропейских слоговых акцентов. In: *Вопросы языкознания* No 6, 36–44.
- Gimbutas, Marija 1956. *The prehistory of Eastern Europe. Part 1 - Mesolithic, Neolithic and Copper Age cultures in Russia and the Baltic area*, Cambridge: Peabody Museum.
- 1963, The Indo-Europeans: Archaeological Problems. In *American Anthropologist* vol. 65 (4), 815-836.
- Greimas, Julien Algirdas 1966. *Structural Semantics: An Attempt at a Method*. Daniele McDowell, Ronald Schleifer, and Alan Velie (trans.). Lincoln, Nebraska: University of Nebraska Press.
- Haak, Wolfgang; Iosif Lazaridis; Nick Patterson; Nadin Rohland; Swapan Mallick and Bastien Llamas 2015. Massive migration from the steppe is a source for Indo-European languages in Europe. In: *Nature*.
- Hjelmslev, Louis 1961 [1943]. *Prolegomena to a Theory of Language*. Madison: The University of Wisconsin Press.
- Hymes, Dell 1978. Comments on Soviet Semiotics and Criticism. In: *New Literary History*, Vol. 9, No. 2, *Soviet Semiotics and Criticism*: 399-411.
- Ivanov, Vyacheslav Vsevolodovich 1962. «Предисловие». In: Rylyna, Yu. Pe (red.) *Симпозиум по структурному изучению знаковых систем. Тезисы докладов*. Moscow: Академия Наук СССР, 3-10.
- (ed.) 1966. *Структурная типология языков*. Moscow: Наука.
- 2004a. О взаимоотношении динамического исследования эволюции языка, текста и культуры. In: Vyacheslav Vsevolodovich Ivanov, *Избранные труды по семиотике и истории культуры* (Т. III), Moscow: Языки русской культуры, 122-141.
- 2004b. *Лингвистика третьего тысячелетия*. Moscow: Языки русской культуры.
- 2008 [2007]. Semiotics of the 20th century. In: *Sign Systems Studies* 36.1, 185-244.
- 2009. К исследованию отношений между языками. In: *Вопросы языкового родства* (1), 1-12.
- Ivanov, Vyacheslav Vsevolodovich; Lotman, Yuriy Mihajlovich; Pjatigorskij, Aleksandr Moiseevich; Toporov, Vladimir Nikolaevich; Uspenskij, Boris Andreevich 1973, Thesis on the semiotic study of cultures (as applied to Slavic texts). In: Silvi

- Salupere, Peeter Torop and Kalevi Kull (eds.), *Beginnings of the semiotics of culture*. Tartu: University of Tartu Press (2013): 51-77.
- Ivanov, Vyacheslav Vsevolodovich and Vladimir Nikolaevich Toporov 1963. К реконструкции праславянского текста. In: *Славянское языкознание. Доклады советской делегации. V Международный съезд славистов*. Moscow.
- 1965. *Славянские языковые моделирующие семиотические системы (Древний период)*. Moscow: Наука.
- 1997 [1966]. Постановка задачи реконструкции знаковой системы. In: Т. М. Nikolaevoj (ed.) *Из работ московского семиотического круга*. Moscow: Школа «Языки русской культуры»: 45-73.
- Levchenko, Jan and Silvi Salupere (eds.) 1999. *Материалы к словарю терминов Тартуско-Московской Семиотической Школы*. In: Tartu Semiotics Library vol. 2. Tartu: University of Tartu Press.
- Lotman, Yurij 1970. *Статьи по типологии культуры: Материалы к курсу теории литературы. Вып. I*. Tartu: Тартуский Государственный Университет.
- 1977. Место киноискусства в механизме культуры. In: *Труды по знаковым системам VIII*, 138-150.
- 1988. Problems in the typology of culture. In: Lucid, Daniel P. (ed.). *Soviet Semiotics: An Anthology*. Baltimore & London: Johns Hopkins University Press, 213-221.
- Lotman, Yurij Mihajlovich and Boris Andreevich Uspenskij 1978. On the semiotic mechanism of culture. In: *New Literary History* 9(2): 211-232.
- Mallory, James Patrick 2013. Twenty-first century clouds over Indo-European homelands. In: *Journal of Language Relationship* 9: 145-154.
- Mallory, James Patrick and Adams, Douglas Quentin 2006. *The Oxford introduction to Proto-Indo-European and the Proto-Indo-European world*. Oxford and New York: Oxford University Press.
- Morris, Charles William 1938. *Foundations of the Theory of Signs*. Illinois: The University of Chicago Press.
- Nikolaeva T. M. (red.) 1966. *Лингвистические исследования по общей и славянской типологии*. Moscow: Наука.
- Nikolaeva, T. M. and B. A. Uspenskij 1966: Языкознание и паралингвистика. In: Т. М. Nikolaeva (red.), *Лингвистические исследования по общей и славянской типологии*. Moscow: Наука, 63-74.
- Peirce, Charles Sanders (1931-1958): *The collected papers of Charles Sanders Peirce*. Charles Hartshorne, Paul Weiss and Arthur W Burks (eds.). Cambridge, MA: Harvard University Press.
- 1982. *The Writings of Charles S. Peirce: A Chronological Edition. [Volume 2]*. Bloomington: Indiana University Press.
- Penney, John 2000. Review of Crespo and García Ramón (eds.) 1997, *Kratylos* 45, 29–35. [as cited in Clackson, J. 2007:157]

- Pictet, Adolphe 1859-1863. *Les origines indo-européennes ou les Aryas primitifs: Essai de paléontologie linguistique*. Paris: Joël Cherbuliez.
- Renfrew, Colin 1987. *Archaeology and language: The puzzle of Indo-European origins*, London: Jonathan Cape.
- Renfrew, Colin; David W. Anthony; Bernard Wailes; P. Baldi; Graeme Barker; Robert Coleman; Marija Gimbutas; Evžen Neustupný and Andrew Sherratt 1988. Book review of 'Archaeology and language: The puzzle of Indo-European origins' by Colin Renfrew. In: *Current Anthropology* Volume 29, Number 3, June 1988, 437-468.
- Revzin, Izaak Iosifovich 1977. *Современная структурная лингвистика. Проблемы и методы*. Moscow: Наука.
- Revzina, O. G. 1972. «Предисловие». In: B. A. Uspenskij (red.) and O. G. Revzina (comp.), *Принципы типологического анализа языков различного строя*. Moscow: Наука, 3-14.
- Rylina, Yu. P. (red.) 1962. *Симпозиум по структурному изучению знаковых систем. Тезисы докладов*. Moscow: Академия Наук СССР.
- Salupere, Silvi and Peeter Torop 2013. On the beginnings of the semiotics of culture in the light of the Theses of the Tartu-Moscow School. In: and Silvi Salupere, Peeter Torop and Kalevi Kull (eds.). *Beginnings of the semiotics of culture*. Tartu: University of Tartu Press.
- Saussure, Ferdinand de 1879. *Mémoire sur le système primitif des voyelles dans les langues indo-européennes*. Leipsick: Imprimerie B. G. Teubner
- 1959 [1916]. Charles Bally and Albert Sechehaye (eds.), Wade Baskin (trans.), *Course in General Linguistics*, New York: The Philosophical Library.
- 1991. *Curso de lingüística general*. Madrid: Akal.
- Searle, John 1969. *Speech Acts*. Cambridge: Cambridge University Press.
- Serebrennikov B. A. (ed.) 1967, *Материалы к конференции «Язык как знаковая система особого рода»*. Moscow: Наука.
- Shaumyan, Sebastian Konstantinovich 1987. *A semiotic theory of language*. Bloomington and Indianapolis: Indiana University Press.
- 1998. Two paradigms in linguistics. In: *Web Journal of Formal, Computacional & Cognitive Linguistics* [September 19].
- 2006. *Signs, Mind and Reality: A theory of language as the folk model of the world*. Amsterdam and Philadelphia: John Benjamins Publishing Company.
- Swadesh, Morris 1955. Towards greater accuracy in lexicostatistic dating. In: *International Journal of American Linguistics* 21, 121-137.
- Торопов, Владимир Николаевич 2004-2010. *Исследования по этимологии и семантике* [4 volumes]. Moscow: Языки славянских культур-Рукописные памятники Древней Руси.
- 2010. *Мировое дерево: Универсальные знаковые комплексы* [2 volumes]. Moscow: Рукописные памятники Древней Руси



- Torop, Peeter 2006. Semiotics, anthropology and the analyzability of culture. *Sign Systems Studies* vol. 34.2, 285-315.
- Trubetskoj, Nikolaj Sergeevich 1987 [1958]. Мысли об индоевропейской проблеме. In: N. S. Trubetskoj. *Избранные труды по филологии*. Moscow: Прогресс (Языковеды мира), 44-59.
- Uspenskij, B. A. (red.) and O. G. Revzina (comp.) 1972, *Принципы типологического анализа языков различного строя*. Moscow: Наука, 3-14.
- Vardul', I. F. 1967. К вопросу о собственно лингвистическом подходе к языку. In: B. A. Serebrennikov (ed.), *Материалы к конференции «Язык как знаковая система особого рода»*. Moscow: Наука. 9-12
- Voloshina, Oksana Anatol'evna 2014. Фридрих Шлегель и Франц Бопп (флексия и агглютинация как две тенденции развития индоевропейского слова). In: *Вестник Московского университета. Серия 9. Филология N° 1*.
- Winner, Thomas 1979. Some fundamental concepts leading to a semiotics of culture: An historical overview. In: *Semiotica* 27 (1/3): 75-82.

## ***Ajaloolis-võrdleva lingvistika ja lingvistilise paleontoloogia semiootika aspektid***

### **Resümee**

Käesolevas töös käsitletakse semiootilis teoreetiliselt kahte eriala, mis käsitlevad inimmineviku uurimist: üheks erialaks on *ajalooline ja võrdlev lingvistika*, mis uurib keelte diakroonilist dimensiooni, ja teiseks - *lingvistiline paleontoloogia*, mis püüab uurida inimkultuuride mineviku mitte-lingvistilisi aspekte, analüüsides ajaloolise ja võrdleva lingvistika poolt tagatud lingvistilisi andmeid. Teatud problemaatilised küsimused on allikaks arutelule mõlema eriala sees, mille tulemuseks on epistemoloogilised raskused seoses nende poolt tagatud mineviku teadmiste ülima olemusega. Semiootiliste aspektide arutelu, kaasates neid uuringuid, selgitab selliseid probleemseid küsimusi.

Ühelt poolt, *ajaloolise ja võrdleva lingvistika* uuringu objekt on ehitatud strukturalistlikule lähenemise poolt lingvistikas. Selle struktuuriliselt ehitatud objekti süstemaatilise loomu pärast *ajaloolis-võrdlev meetod* saab tegeleda objektiivsusega. Teiselt poolt, semiootilisest lähenemisest kuni lingvistiliste märgisüsteemideni me võime eristada nendes kolme dimensiooni: *semantiline dimensioon* – suhted süsteemi ja extra-lingvistilise reaalsuse vahel -, *pragmaatiline dimensioon* - süsteemi ja selle tgelikus olukorras kasutaja vahel, praktilise dimensioon ja *süntaktiline süsteemi dimensioon* – lingvistiliste märkide omavaheline seos, ja see tähendab, süsteemi elementide vaheline sisemine seos. Kuna süsteemsus ja struktureeritus on kättesaadavad ainult abstraktse kirjelduse tasandil ja just eriti lingvistilistele märkidele, mis on süsteemi elemendid omavahel, me järeldame, et ajaloolis-võrdlev lähenemine keeltele haarab lõppkokkuvõttes lingvistilise semiootiliste süsteemide *süntaktilist dimensiooni* ja võib pakkuda kontrollitavaid ja teaduslikke teadmisi rangelt selle *süntaktilise dimensiooni* jaoks.

Kultuuriga vastavuses olevate semiootiliste süsteemide hierarhiliste omavaheliste suhete sees lingvistilised mängivad kesket rolli modelleerimise ja struktureerimise funktsioonis, kultuuri iseloomus. Selle keskse rolli pärast, lingvistilis-semiootilised süsteemid omavad kõrget informatiivsust kultuuri süsteemide mineviku mitte-lingvistiliste aspektide kohta. *Lingvistilise paleontoloogia* distsipliin jätkab selle informatiivsuse eeldusega, kuid selle järelduste lõplik kehtivus on pidevalt kahtluse alla seatud selle tõlgendava lähenemise tõttu,

mis aga sobib andmete ületõlgendamiseks ja väärtõlgendamiseks. *Lingvistiline paleontoloogia* eraldi ei ole sobivaks inimmineviku objektiivsete teadmiste usaldusväärseks allikaks, kuid usutavate argumentide allikaks. See on sõltuvuses konfrontatsioonist vastastõenditega teiste distsipliinide poolt, nagu arheoloogia või populatsioonide geneetilised uuringud inimmineviku kohta usaldusväärsete teadmiste pakkumiseks. Teiselt poolt, selle järeldused on seda rohkem usaldusväärsemad, mida rohkem nad on välja võetud lingvistiliste semiootiliste süsteemide *süntaktilise dimensiooni* analüüsist, ja nõrgemad, kui nad on tuletatud leksikaalsetest tähendustest.

**Võtmesõnad:** ajaloolis-võrdlev lingvistika, lingvistiline paleontoloogia, ajaloolis-võrdlev meetod, süntaktiline dimensioon, lingvistilised semiootilised süsteemid, kultuurisemiootika

***Семиотические аспекты сравнительно-исторического языкознания и  
лингвистической палеонтологии***

**Резюме**

Данная работа, основываясь на семиотическом теоритическом подходе, рассматривает две дисциплины, исследующие человеческое прошлое: *сравнительно-историческое языкознание*, работающее с диахронией языков, и *лингвистическую палеонтологию*, исследующую нелингвистические аспекты прошлого человеческих культур на основе анализа лингвистических данных. Определенные проблематические вопросы являются источниками споров внутри этих дисциплин, результатом возникновения эпистемологических трудностей, касающихся знания о прошлом, достигаемого ими. Раскрытие семиотических аспектов этих дисциплин выявит такие проблематические вопросы.

С другой стороны, объект изучения сравнительно-исторического языкознания основывается на структуралистском подходе в лингвистике. Сравнительно-исторический метод является научным и объективным, поскольку при его использовании объект изучения структурно выстроен и имеет систематический характер. С другой стороны, с позиции семиотики в лингвистических семиотических системах можно выделить три измерения: *семантическое измерение*, то есть отношения между системой и внелингвистической действительностью; *прагматическое измерение*, то есть отношения между системой и её пользователем в данном конкретном контексте (измерение действия); и *синтактическое измерение* – отношения между лингвистическими знаками, то есть отношения элементов системы. Так как систематичность и структурированность достигаются только на абстрактном уровне описания отношений между лингвистическими знаками, являющимися элементами системы, мы можем заключить, что сравнительно-исторический подход изучения языков оперирует *синтактическое измерение* лингвистических семиотических систем и, следовательно, предоставляет научное знание конкретно об этом *синтактическом измерении*.

Внутри иерархических связей семиотических систем в культуре, лингвистические системы играют центральную роль при выполнении функции моделирования и структуризации. Благодаря этой центральной роли, лингвистические семиотические системы несут большое количество информации о нелингвистических аспектах прошлого данной культурной системы. Лингвистическая палеонтология в свою очередь использует эту информацию для дальнейшего изучения, но впоследствии обоснованность ее выводов может подвергаться сомнению из-за применения толковательного подхода, допускающего *сверхинтерпретации* и неверное истолкование (*недоинтерпретации*) данных. Лингвистическая палеонтология сама по себе не является источником надёжного и объективного знания о человеческом прошлом, но является источником аргументов для определения правдоподобности. Она зависит от конфронтации показаний, получаемых другими дисциплинами, такими как археология и исследования по генетике населения, с целью предоставления надежного знания о человеческом прошлом. С другой стороны, заключения, предоставленные лингвистической палеонтологией, являются более правдоподобными, если они получены посредством анализа *синтактического измерения*, и слабыми при получении путем анализа лексических значений.

**Ключевые слова:** сравнительно-историческое языкознание, лингвистическая палеонтология, сравнительно-исторический метод, синтактическое измерение, лингвистические семиотические системы, семиотика культуры

## **Non-exclusive licence to reproduce thesis and make thesis public**

I, Jorge González Fernández,

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to:

1.1. reproduce, for the purpose of preservation and making available to the public, including for addition to the DSpace digital archives until expiry of the term of validity of the copyright, and

1.2. make available to the public via the web environment of the University of Tartu, including via the DSpace digital archives until expiry of the term of validity of the copyright,

*Semiotic aspects of historical-comparative linguistics and linguistic paleontology*, supervised by Prof. Peeter Torop (University of Tartu) and Prof. Vartan Kazaryan (Moscow State University),

2. I am aware of the fact that the author retains these rights.

3. I certify that granting the non-exclusive licence does not infringe the intellectual property rights or rights arising from the Personal Data Protection Act.

Tartu, 25.03.2015

