On the Question of Proto-Language Reconstruction

Following the discovery that the Indo-European languages are genetically related, it was seen relatively soon that Sanskrit does not represent the language state that is the source of the living Indo-European languages, neither could it be the common predecessor or at least very close to it; but rather the oldest known language. AUGUST SCHLEICHER, who — under the influence of Darwin's teachings — imagined the development of languages to be similar to the living world, set upon himself the task of deducing the Indo-European proto-language. This he represented in the form of a family tree. He thought that by comparing the data of related languages we could proceed back step by step in the history of the languages until we would be able to reconstruct, with the help of our knowledge of phonological rules, this Indo-European proto-form and proto-language from which the correspondences of the genetically related languages compared as well as the genetically related languages themselves are derived. SCHLEICHER saw the value of this procedure in the fact that such ancient language forms demonstrate in a concrete manner the results of research, facilitate an insight into the history of the daughter languages, and by means of them the erroneous nature of this view could easily be conceivable, the view according to which Sanskrit would also be the common predecessor of the non-Indian Indo-European languages.1 SCHLEICHER in his Compendium published not only the words and roots of the proto-language but paradigms as well: and moreover published a simple short tale constructed in the proto-language.2 SCHLEICHER was therefore not content with ascertaining genetic affinity, but he also wished, by disclosing the regularities, to reconstruct the common ancestor of the Indo-European languages. His procedure was not met with unanimous recognition: while the neo-grammarians were drawing upon Schleicher in their research on phonological rules and efforts in reconstruction, his pupil, JOHANNES SCHMIDT3 attacked the family tree theory by establishing his wave theory, and he classified the proto-language as a piece of scientific fiction, which naturally also meant denying the justification for reconstructing. In judging this question the

3 J. SCHMIDT, Die Verwandtschaftsverhältnisse der indogermanischen Sprachen. Weimar, 1872.
camp of comparativists was split into two parts: those who believed and were against reconstruction, and this debate has not died down even today. Karl Brugmann declared that the Indo-European proto-language reconstructed through the help of linguistics was not the kind of language that could have once been spoken, for the method of reconstruction can place beside one another forms derived from different ages and from different dialects of the proto-language. But Brugmann does not reject reconstruction itself, nor does he condemn it; on the contrary, he applies it. A. Meillet, for the most part, emphasized the unreality of reconstructions as well as their unverifiable nature and looked upon the disclosure of the system of correspondences as the task of comparative linguistics, although his point of view against it was not always unanimously unfavourable. H. Schuchardt and N. S. Trubetzkoj did not derive the Indo-European languages from one certain protolanguage, but explained them as the results of contacts and interrelationships between languages which were not in a genetic relationship with one another — cf. Trubetzkoj: Sprachbund — (in Finno-Ugric D. V. Bubrih asserted similar views which have however remained without reaction); after denying the existence of a proto-language, it would naturally follow that reconstruction would be refused as a method. In Italy among the representatives of the neolinguistic trend, G. Bonfante and V. Pisani conducted especially vehement attacks against reconstruction. Soliciting the help of linguistic geography teachings and referring to the not always clear-cut nature of the reconstructions, they deny the existence of a unified Indo-European, etc. proto-language, and see the task of comparative linguistics in the determination of isoglosses; according to them, only through a knowledge of these isoglosses can we arrive, in connection with outlining the temporal and spatial history of languages, at such stable and objective results which, as they say, can never be reached by adhering to the „Schleicherian myth” of the proto-language. According to them, we cannot understand the history of a language in the absence of a thorough and detailed knowledge concerning the history of the people speaking


that particular language. The American E. Pulgram also expressed his opinions concerning the complete unreliability of reconstructions, maintaining that two types of Indo-European proto-languages must be distinguished: one which actually existed but with which we can never be familiar; the other is the reconstructed Indo-European proto-language which is the sum of reconstructed forms not possessing actual values.

Others — e.g., F. de Saussure, R. A. Hall, H. Hoenigswald, A. Nehring, O. Szemerényi, R. Anttila — contrary to the „negativists” who were frequently under the influence of agnosticism, advocated reconstruction saying: the task of comparative linguistics is the possible reconstruction of proto-languages so that we can by this means obtain a more composite and concrete picture concerning the antecedants of the languages known today, their structure, the life of the society that spoke the proto-language, so that we can improve our knowledge concerning the historical linguistic events, etc. of the periods elapsing between the time of the proto-language and that of the daughter languages; and in accordance with this concept, the reality of the reconstructed unities does indeed exist, and these precisely reflect the prevailing results of comparative linguistics; reconstructions — just like the knowledge of natural laws — can be refined, and in broadening our knowledge they will also become more and more accurate.

In Finno-Ugric linguistics, beginning already with the first and at the same time serious consideration regarding the common word-stock of the genetically related languages, the (Magyar-ugor összehasonlító szótár) written by József Budenz, we can speak about the reconstruction of the Finno-Ugric proto-language forms. In the course of his university studies at Göttingen, Budenz became acquainted with the results and theories of German linguistics at that time, among other things, Schleicher's language theory. In the spirit of this, Budenz formulated the goal of Finno-Ugric comparative investigations in the preface to his dictionary: „... the Ugric languages demand from science to be compared above all with the aim that we can, so to say, revitalize their by-gone unity, that is the „Ugric proto-language”, determining with possible approximation its assumable phonetic character and with formal accuracy in conformity to past word-stock and parts of its formal grammatical structure”. Budenz attempted to reconstruct parts of

13 Bonfante: op. cit. 348.
18 A. Nehring, Zur „Realität” des Urindogermanischen. Lingua 10:357—68.
21 Cf. also Pisani: Lingua 4:350—1.
22 Cf. Szemerényi, Trends ... 9, Einführung ... 29.
24 MÚSz. III—IV.
the formal grammatical structure in his work entitled Az ugor nyelvek összehasonlító alaktana.\textsuperscript{25} With a short time delay another similar undertaking followed the Magyar-ugor összehasonlító szótár: O. DONNER's work entitled Vergleichendes Wörterbuch der finnisch-ugrischen Sprachen.\textsuperscript{26} DONNER likewise touched upon the means of reconstruction, but operated with word roots patterned after Indo-European\textsuperscript{27} and thus of course he undertook also the hypothesis of vocalic element in Finno-Ugric "word-roots" from nothing. Contrary to this, BUDENZ does not present his immense amount of etymological material in grouping according to roots, (this is alien to the Finno-Ugric languages), but according to the Hungarian member of the genetically related language correspondences — although he mentions the proto-word as the root-word equivalent\textsuperscript{28} —, or rather he notes that as a consequence of the complete unclearity of Finno-Ugric vocalism at that time, in hypothesizing the proto-forms he attempts a decision at best of the palatal or velar quality of the proto-language vowels.\textsuperscript{29} BUDENZ did not always attempt a reconstruction of the proto-word, saying that his work "in accordance with its detailed task is not an «etymological» dictionary... but simply a «comparative» one",\textsuperscript{30} and for the most part only then discloses the hypothesized forms if the proto-language word under consideration exists in the modern Finno-Ugric languages but only in its derivations. "He was the first who started to examine Finno-Ugric proto-forms clearly surpassing the comparativist attitude; but it is true that this was in one respect certainly early, for the history of the smaller language groups was not yet examined; but the initiative in any case remains the merit of Bundenz", wrote SETÅLÅ\textsuperscript{31} about him.

For more than half a century the BUDENZ type of reconstruction was not met with a follow-up in Finno-Ugric. We cannot look for the reason in the aversion to reconstruction but in the relatively poor results of comparative phonological research. From time to time in articles written by SETÅLÅ, PAASONEN, and later TOIVONEN dealing with the subject of etymology there appeared hypothesized proto-language forms, but more for the purpose of only summarizing the more complicated explanations and demonstrating the phonological changes described. PAASONEN\textsuperscript{32} and HANNES SKÖLD\textsuperscript{33} in their studies dealing with Finno-Ugric-Indo-European agreements, besides the Indo-European roots, and in the absence of reconstructed forms were compelled to list the Uralic or Finno-Ugric word-family members alleged to be reconcilable: the reason was again the lack of research in vocalism, as SKÖLD also pointed out.\textsuperscript{34} LEHTISALO's attempt into the historical phonology of Uralic, which was not successful,\textsuperscript{35} did not change the matter. STEINITZ's

\textsuperscript{25} J. BUNDEZ, Az ugor nyelvek összehasonlító alaktana. Budapest, 1884.
\textsuperscript{26} O. DONNER, Vergleichendes Wörterbuch der finnisch-ugrischen Sprachen. Leipzig, I. 1874., II. 1876., III. 1888.
\textsuperscript{27} Cp. eg. VglWb. V, VI.
\textsuperscript{28} MUSz. V.
\textsuperscript{29} MUSz. VIII.
\textsuperscript{30} MUSz. VII.
\textsuperscript{31} NyK 23:340
\textsuperscript{33} H. SKÖLD, Indo-uralisch. FUF 18:216—31.
\textsuperscript{34} FUF 18:218.
\textsuperscript{35} T. LEHTISALO, Zur Geschichte des Vokalismus der ersten Silbe im Uralischen vom qualitativen Standpunkt aus. FUF 21:5—55.
vowel theory\textsuperscript{36} and E. Itkonen's criticism connected with it, as well as his new vowel theory,\textsuperscript{37} brought about a sudden change. Following this, in the works of Björn Collinder proto-language reconstruction plays a particularly important role. In searching for a synthesis of the two theories, he worked out a theory with a rather new conception. Of course, we must not forget that Collinder already previously undertook in a few cases\textsuperscript{39} a hypothesis regarding the suffixes and words of the proto-language. In Hungary, however, large-scale etymological dictionary work produced a favourable possibility for reconstructing the phonetic shape of the Uralic, Finno-Ugric, etc. (proto-language) words.

Therefore, in Finno-Ugric research proto-language reconstruction was (again) established only through the efforts of Collinder as well as through the development of national etymological research,\textsuperscript{40} at least in practice. From the following it however appears that certain people feel the insistence of theoretical reservations almost obligatory: Collinder writes in the introduction\textsuperscript{41} to the section of his Uralic comparative grammar that treats reconstructed proto-forms that the readers can question them at their own discretion. In the chapter on historical phonetics he lists and examines as well-established correspondences a large number of the daughter-language data, which are derived from these proto-forms. But he does so with the firm conviction that the definite element of these proto-forms (vowel or consonant) is the sure continuation, present-day representative of a definite Uralic (Finno-Ugric) proto-language sound, reconstructed on the basis of reciprocal correspondences. It is somewhat strange that what we recognize as true and correct as an integral element, we must consider it in its totality to be already unreliable. Nor is it slightly contradictory, as György Lakó, the principal editor of A magyar szókészlet finnugor elemei, expresses his views about this same question: „As we continually go further back into the past, we can more or less finally reconstruct also the language form that was the common antecedent of all the Finno-Ugric languages, that is the Finno-Ugric proto-language,”\textsuperscript{42} or rather: ”The reconstructed forms require a clarification of the totality of our hypotheses drawn from the individual Finno-Ugric language (dialectal) forms and require them to be expressed then in the form of a concise symbol. While we do not attribute a real existence to them it is our conviction that with their applications as symbols we served the interest of those who use our dictionary.”\textsuperscript{43} In the „prospectus”\textsuperscript{44} of the other large etymological dictionary\textsuperscript{45} we do

\textsuperscript{36} W. Steinitz, Geschichte des finnisch-ugrischen Vokalismus. Stockholm, 1944.
\textsuperscript{38} B. Collinder, Comparative Grammar of the Uralic Languages. Stockholm, 1960.
\textsuperscript{39} B. Collinder, Indo uralisches Sprachgut. UUÁ 1934., Jukagirisch und Uralisch. UUÁ 1940.
\textsuperscript{40} A magyar szókészlet finnugor elemei. Budapest, I. 1967., II. 1970.
\textsuperscript{41} CompGr. 405.
\textsuperscript{43} MSzFE 17
\textsuperscript{44} A magyar nyelv történeti-etimológiai szótára. Budapest, I. 1967., II. 1970.
\textsuperscript{45} TESz. 30.
not find, in connection with the proto-language reconstructions in the word-
articles of words belonging to the so-called „ancient heritage”, any kind of
basic attitude, but the principal editor, LÓRÁND BENKŐ, in a university gram-
mar book expresses his opinion that the words reconstructed on the basis
of comparative linguistic results, can considered as only hypothetical ones, as
a „formula-type summary” about the earliest state of the Hungarian language.
In the previous case it is particularly striking that we can read about dictionary
reconstructions as such kind of symbols which were recognized out of necessity,
and still have not been perhaps properly established; while elsewhere this
same author even believes in the reconstruction of the Finno-Ugric proto-
language, if only „more or less”. Considering the not altogether consistent atti-
tude among Finno-Ugric historical linguists, I feel that RAIMO ANTTLA’S
remark is appropriate here: „...many of those linguists who are maintaining
a negativist position in theory are actually positivists in practice”.47

Those opposed to proto-language reconstruction now and then question
the existence of a proto-language (cf. above), saying that the present-day
languages are not descendents of one definite proto-language, but usually
came into existence through the mutual influence of more—whether genetically
related to one another or not — languages; or rather that we have scarcely
any or a very approximate knowledge concerning the temporal and spatial
localization of a given proto-language; we do not know the dialectal distribu-
tion of the proto-language, and the characteristics of the dialects, and how
they were related to one another, etc. The danger in a reconstruction is that
we might place beside one another forms derived from different ages or dif-
ferent dialects. Contrary to this, however, it would not show common sense to
deny the former existence of a common predecessor for the genetically related
languages known today, a (Uralic, Finno-Ugric; Indo-European, Germanic,
etc.) proto-language; naturally, however, the view according to which the
individual proto-languages would have been sort of undifferentiated forma-
tions and the languages derived from them, not being influenced by neighboring
genetically related and non-related languages (which would have devel-
oped according to their own internal laws) would be unacceptable. Therefore
we cannot in every case and generally regard the reconstructed proto-language
forms as being undoubtedly reliable. Taking the above also into consideration,
there could also be more reasons for this:

1. the more distant the age is in which the proto-language becomes
dissociated, the more difficult it is to determine its duration and to define it
in time, and the greater are the changes that took place in the daughter
languages, among them not a small number of „irregularities”, the reason or
conditions of which we do not know;

2. Just like the present-day living languages, the individual proto-
languages also evidently showed a more or less definite dialectal separation; but
for the time being, we can accept this circumstance in Finno-Ugric rather only
in theory than in practice; but if we wish to explain nowadays certain excep-

26—7.
47 R. ANTTLA : op. cit. 341.
48 P. RAVILA, Kantakieli kielihistorian peruskäsitteenä [In: Totuus ja metodi.
tional sound correspondences, "irregularities" by means of dialectal differences of the proto-language, than we would only give an opportunity for\textsuperscript{49} the arbitrary and unscientific etymologies; if we can obtain the completely reconstructed word-stock and grammatical framework of our language family, we will perhaps then have some hope of finding out something about the dialects of the proto-language; naturally we must imagine the dialectal distribution of the proto-language not or not only according to language groups known today, but former dialectal boundaries probably did not necessarily coincide with the boundaries or dividing lines between units of the population speaking the proto-language and moving in the path of separate language development;\textsuperscript{50}

3. If convergence development took place in the daughter languages in a certain area or areas, this cannot be clarified through comparative linguistic methods and it is lost for reconstruction;\textsuperscript{51}

4. Among languages that have continued independently, two or more living in the neighborhood of each other and coming into contact and creating areal contacts can form a system of secondary agreements;\textsuperscript{52}

5. Because of the possible lack of early linguistic monuments, a reconstruction that is built upon speculation must bridge over a larger temporal distance, as a result of which the chance of error is greater; this difficulty can be especially felt in the research dealing with the history of the Uralic languages;\textsuperscript{53}

6. The fact that our knowledge reflecting proto-language relations can change and grow supports already by itself its necessity that in each case we adjust the reconstructed forms to our more exact knowledge;\textsuperscript{54}

7. Finally, besides reconstructing the sound body of each proto-language word, the semantic hypothesis belonging to the word causes some problem if — as it sometimes happens — we come across divergent meanings or semantic ranges in the present-day languages which at first glance show a small number of common features; surmounting this calls for an investigation of meaning changes and an application of the hypotheses concluded from this investigation, because the reference to possible semantic parallels does not unconditionally lead to unanimously reliable and convincing results.\textsuperscript{55}

\textsuperscript{49} E. ITKONEN: Tietolipas 20:55, RAVILA, op. cit. 68—70.
\textsuperscript{50} H. PAUL, Prinzipien der Sprachgeschichte. 5. Auflage. Halle, 1920. 46—7.
\textsuperscript{51} MEIILET, Сравнительный метод... 20, HOENIGSWALD, Language Change... 135, 136.
\textsuperscript{54} Cf. G. BERÉCZKI: NyK 70:450—1, SZEMERÉNYI, Trends... 9, Einführung... 29.
According to GUY JUCQUOIS' assertions, reconstructed proto-languages contain a much more larger quantity of homophonous words than do natural languages; the reason for this can be seen in the methods of reconstruction. Thus, for example, in J. POKorny's work entitled Indogermanisches Etymologisches Wörterbuch, 95.5% of the reconstructed roots have at least one homophonous counterpart according to E. BENVENISTE's investigations. From the material of the Uralic Etymological Dictionary about 1200 word-articles (about two-thirds) have so far been prepared, and in it, among the reconstructed proto-forms, only about 250 (about 20%) have at least one homophonous counterpart; I regarded as homophones also those proto-forms in which the vocalic quality of the non-initial syllable could not be determined, e.g., *čęγke 'Dampf' FU — *čęγk3 'Keil' FU, *kuδ3 'Morgen' U — *kuδ3 'Schneefall' U, therefore, only potentially homophones; that is, if our knowledge concerning the proto-language would widen, we could count upon a 20% decrease of the above; but inasmuch as I, at the time the proportion mentioned above was established, regarded the verbal-nouns as one word, although it would have been more correct to accept them as homophonous proto-words, the 20% proportion would increase somewhat. But it could easily be possible that the decrease and increase already mentioned will be approximately of the same degree, it would therefore not bring about a considerable change; besides about 400 Uralic and Finno-Ugric word articles awaiting writing cannot to a significant degree alter this proportion; on the basis of this, we can therefore much more favourably judge the situation of Uralic etymological research — and together with this the situation of the reconstructions, as COLLINDER and A. JOKI referred to it. — By the way, the approximately 1600 reconstructable words of the Uralic-Finno-Ugric proto-language's word-stock (and then we considered the verbal-nouns to be only one and not two words!) contradicts AURÈLIEN SAUVAGEOT's thesis (which influenced perhaps a little strangely more than two decades) that the common word-stock of the Uralic languages known today is small; the predominant majority of these 1600 etymologies were already just as well known then, but this reasonably strange statement can be explained to a certain extent that the etymologies were published in several papers, sometimes in hidden places and so the revealed common word-stock could be surveyed and organized only with relative difficulty.

The proto-language reconstructions can be controlled not or only rarely, and in exceptional cases, but if the possibility arises, they turn out to be effectual and successful, and this naturally justifies those who believe in reconstruction

56 GUY JUCQUOIS, La théorie de la racine en indo-européen. La Linguistique 6/2 :69—102, 7/1 :73—91, 8/1 :73—103.
57 La Linguistique 6/2 :76.
and admit that the system of hypothesized forms is necessary for further research. The goal of comparative linguistic research is not only to ascertain the genetic affinity of the related languages and to determine the degree of relationship, but also that through the reconstructions we can obtain a possible maximum knowledge concerning the proto-language. If we know the two distant points in a developmental process (the possible system of the reconstructed proto-language and the present-day languages that developed from the proto-language), we can create also a picture about such circumstances of the development, upon which a mere comparison would not shed light. Of course, so that we can have a reliable picture of the split proto-language in a distant age, we must know the later developmental periods as well, the "intermediate proto-languages" also; thus, for example, a clarification of the Finno-Permic, Finno-Volgaic, and Ugric, etc. proto-language system would enlarge and make our knowledge more exact concerning the Uralic (Finno-Ugric) proto-language, although for the time being such a paradoxical situation took shape that we know more in phonological relation about the distant Finno-Ugric proto-language, than about the later Ugric proto-language or the early developmental periods of the separate lives of the single languages, as for example, old Hungarian. This is mainly so because Ugric vocalism research fell behind to a certain extent the Finno-Permic branch. The archaism of the vocalism of the Finnic languages (and congruent evidence with these of Lapp vocalism) facilitates, in a large degree, Finno-Permic and Finno-Volgaic, and further Finno-Ugric and Uralic reconstructions, thus the reconstructions of the Finnic languages play a key language role in this most problematic area of the reconstruction. Neither can we depend, in this area, on one of the Ugric languages as a key language. Another similarly not unimportant problem of Finno-Ugric is that for the time being we do not know how to separate from one another on the basis of phonetic points of view the Finno-Ugric and the Uralic proto-language, although perhaps a historical-comparative study of the up-to-now fairly neglected Samoyedic languages will bring the desired result.

In becoming frightened of the towering difficulties before us, we would give up on reconstruction, we would also give up revealing those historical processes which shaped the individual physiognomy of the present-day languages, however such investigations also mutually complement one another, and through this the reconstruction of the proto-language could be at the same time the goal and means as well. In addition to the phonological system of the proto-language, we can more or less reconstruct its morphological units: in relationship to the formation of the Finno-Ugric proto-language, see KÁROLY RÉDEI's paper; we can reason further the derivational and inflectional morphemes of the proto-language, but a complete reconstruction of the paradigm system of the proto-language would be an unfounded and illusory objective. In connection with the syntax of the proto-language, we can also arrive at certain cautious hypotheses, but we cannot scarcely know more

---

64 Cf. E. ITKONEN: FUF 37:40.
65 About this cp. HAJDÚ, BUNyt. 43.
66 Cf. SAUSSURE: op. cit. 271.
67 KÁROLY RÉDEI, Wortbildung in der PU/PFU Grundsprache.
about the syntax of the Uralic and Finno-Ugric proto-language than can be found in Péter Hajdú’s book.69 The practical and expedient nature of the reconstructions can otherwise be convincingly proven that the reconstructions of the Uralic etymological dictionary have already brought surprising and interesting results: through them and by the help of internal reconstruction we can obtain an insight into the secrets of the perhaps very early method of word formation in the Uralic proto-language, about which Marianne Bakró-Nagy’s paper70 gives account. With a knowledge concerning the complete reconstructable vocabulary we can become acquainted with the living conditions, social and economic relations, etc. of the people who spoke the proto-language. Nor could it be insensible that being in possession of the reconstructed vocabulary, we can know something sure and more about possible distant relations of our language family. The efforts directed towards the reconstruction of the proto-language are therefore useful and encouraging efforts accompanied by results, because the proto-language, like a mosaic hypothesized constitutes the object of further investigations.71

A re-establishment of the complete synchronic system of the proto-language cannot naturally be the goal of the comparison of the genetically related languages and the reconstruction of the proto-language,72 still less would be the useless construction and improvisation of proto-language texts. But our knowledge which is sometimes scanty does not authorize us to suppose that the proto-language could have been so „primitive” that the prestigious part of the grammatical categories could have been formed in the separate lives of the single languages or language groups.73

In the reconstruction of proto-language conditions and in clarifying language historical processes that took place, typological research, the discovery of universals, implicational connections, the application of general rules appearing in historical changes74 can offer a great help.

In Finno-Ugric studies at home — as I already mentioned — reconstruction because of etymological dictionary work, and more of all because of the Uralic Etymological Dictionary has become an especially actual question. In connection with this, I must remind the reader of the practical use of the

70 BUNyt. 80—4.
71 Cf. note 55.
73 This e. g. J. Berráé, Magyar történeti mondattan. Budapest, 1957. 43, 78—9, 175—6, but cf. Meillet, Einführung ... 20—1.
reconstructed proto-language forms, the proto-forms, but it will make possible an easily arranged publication of the dictionary material. As a result of an arrangement of word-articles according to the proto-forms in alphabetical order the Uralic Etymological Dictionary will be incomparably more easily treatable, then COLLINDER’s etymological dictionary of a much smaller extent which was assembled according to other principles of editing. Perhaps it is not an exaggeration to assert that for the comparative-etymological dictionaries that treat the vocabulary of more languages the publication method based on the reconstructions is the ideal one.

LÁSZLÓ HONTI

75 Eg. POKORNY, IdgEtwb., Cf. also E. BERNEKER, Slavisches etymologisches Wörterbuch. Heidelberg, 1908—14. 3.