LINGUISTIC COMPARISON

By R. H. ROBINS

Some sort of comparison of languages follows as an early consequence of any study or even awareness of languages other than one's own ; and it may be made on any criteria, though the results will necessarily vary according to the criteria selected, and in scientific linguistics fruitfulness of the criteria of comparison is the principal metacriterion.

An early example of some sort of comparison between languages was their classification into four groups according to the type of word used for *God*, made by J. J. Scaliger, which gave approximately the groups known later as Latin and Romance, Greek, Germanic, and Slavonic, though further comparisons were not made at the time by Scaliger in any systematic way.[1]

A more refined and productive basis of comparison was that of Wilhelm von Humboldt early in the nineteenth century, with the three well-known language types : isolating, agglutinating, and inflecting,[2] which, despite its *naïveté* in some respects, is by no means without use, especially if the three types are regarded as forming a triangular figure within which the typological position of a language may be plotted, and as themselves being three cardinal points towards which languages approach with greater or less purity, rather than as actual classes of languages.

This sort of comparison was dismissed by Antoine Meillet as a mere *amusette*,[3] and after Humboldt and until fairly recently attention tended to turn from the purely comparative study of languages to genetically orientated comparative-historical studies, or, in the English tradition, Comparative Philology, which for a number of reasons dominated the field of nineteenth-century linguistics. The established Indo-European linguistic family, and the system of detailed historical relations between the languages thereof, is tribute enough to comparative-historical linguistics, and its methods have been successfully applied to other genetic groupings of languages.

Apart from the great success that IE comparative linguistics has achieved, the usefulness of this type of comparison is obviously not limited to linguistics alone ; cultural and political history, and in less known areas and periods the migrations and diffusions of peoples, are all subjects in which the findings of comparative-historical linguistics may and do throw light, just as, conversely, historical data may be brought in to assist the genetic grouping of languages.

---

[1] *Deus-, theos-, Gott-,* and *bog*-languages respectively : H. Pedersen, *Linguistic science in the nineteenth century*, 1931, p. 6 ; H. Arens, *Sprachwissenschaft : der Gang ihrer Entwicklung*, 1955, p. 59.

[2] *Über die Verschiedenheit des menschlichen Sprachbaues*, edition of 1949, pp. 114–26.

[3] A. Meillet, *Linguistique historique et linguistique générale*, 1948, pp. 76–7.

The very success of this type of comparative linguistics, developed within
languages, and most successfully exploited in this field, has led some to questi
its necessary utility elsewhere, and even its legitimacy as a purely linguistic activit
In particular one must recognize the specially privileged area covered by l
with copious texts of all kinds of the three classical languages, Greek, Lat
and Sanskrit, and records of many others, together with detailed histori
knowledge of the peoples speaking such languages over a long and continuc
period, privileges that are not available to the same extent elsewhere and a
wholly lacking in some language groups (e.g. Algonkian and Bantu).

Realization of these limitations and the development of general and descript.
linguistics in the present century have led to a return to the type of comparis
implicit in early efforts and in the work of Humboldt. The typological comparis
of languages may be made on the basis of phonological structuring, grammati
structuring, or the relationship between these two levels.

Typological classifications based on phonology alone are seen in the classifi
tion of phonological systems by Trubetzkoy and by Hockett [2] ; and the obvic
fact that phonetic and phonological features tend to spread over contiguc
areas irrespective of genetic relationships has been treated by Jakobson und
the title of ' affinités phonologiques '.[3] In this connection one may instance t
area of glottalized consonants in the Caucasus uniting IE Armenian with t
non-IE languages of that region, and the area of retroflex consonants in Inc
uniting the Sanskritic languages (IE) of the north with the Dravidian langua؛
(non-IE) of the centre and south ; and we may note that these areas are t
only ones wherein glottalized and retroflex consonants have come to play
extensive and prominent part in the phonological systems of IE languages.

Classifications can be based on grammatical characteristics alone, as 1
example on the proportions in which endocentric and exocentric constructic
and derivational and inflectional word formations are found, or the degree
which word order is syntactically relevant and with what specific syntac
relations specific word orders correlate.[4]

Humboldt's basis of classification is inter-level and rests principally on 1
different ways in which languages may represent grammatical relations in senten؛
by the phonological forms of the words therein. Such inter-level comparison m
be the most significant, in that it involves a greater range of formal linguis

[1] cf. F. Boas, ' The classification of American-Indian languages,' *Language*, 5, 1929, pp. 1-
W. S. Allen, ' Relationship in comparative linguistics,' *TPS*, 1953, pp. 52–108, though t
latter involves an extreme restriction on what is properly to be termed ' linguistic '.

[2] N. S. Trubetzkoy, *Principes de phonologie* (tr. Cantineau), 1949, pp. 93–246 ; C.
Hockett, *A manual of phonology*, 1955, pp. 42–143.

[3] R. Jakobson, ' Affinités phonologiques,' *Actes du 4me Cong. Int. de Linguistes*, 19
pp. 48–58, and in Trubetzkoy, *Principes*, pp. 351–65.

[4] C. E. Bazell, ' Syntactic relations and linguistic typology,' *CFS*, 8, 1949, pp. 5–2
L. Tesnière, *Eléments de syntaxe structurale*, 1959, pp. 32–3.

structuring. Among modern linguists Sapir took up and developed the Humboldtian classification, but without Humboldt's suggestions of value judgments to be placed on the different types as means of expression or as indications of the mentality of the speakers, and without the later evolutionary accretions of others.[1] Greenberg has suggested ways in which the Humboldtian comparative classification as adapted by Sapir may be further formalized and quantified, by the introduction of ten indices or proportions of formative elements or processes to words and constructions, resulting in five parameters by which, taken together, languages may be placed on the scales of analytic, synthetic (inflectional), and agglutinative.[2] A similarly based classification of languages has been outlined recently by Bazell, with reference to the morphological forms of words and their relative degrees of determinacy of grammatical segmentation and class indication.[3]

Quite obviously both types of linguistic comparison, which we may call genetic and synchronic respectively, are legitimate procedures if conducted with appropriate rigour and with an adequate body of material. What is not perhaps always sufficiently recognized is that inferences from one type of comparison to the other are not necessarily valid. Genetic relationship means, at the minimum, that two or more languages have during past centuries, through a gradual and unbroken tradition,[4] developed from what was once a relatively homogeneous language area ; and the principal criterion for such relationship is a number of words or root morphemes, too many to be dismissed as coincidences, of similar or related meanings, whose phonological forms exhibit systematic correspondences (not necessarily actual similarities)[5] of elements and are not explicable as onomatopoeic creations or as loans. This base statement, of course, is nothing new, and requires considerable elaboration ; but in a short paper such as this it is perhaps enough. Now it is clear from numerous examples that structural features of languages at all levels, such as are used as the material of synchronic comparison, may change in the course of time while the basic vocabulary continues to reveal indisputable genetic relationship. Within IE, English syntax and Latin syntax are very unlike, and Pedersen notes some agglutinative developments in Armenian reminiscent of Turkish structures[6] ; and if B. Karlgren is right, the present isolating structure of Chinese is the result of gradual change from an inflectional type of language.[7] Trubetzkoy's six 'structural marks' of IE status are a case in point here[8] ; he declared his intention of defining IE without positing

[1] E. Sapir, *Language*, 1921, pp. 127–56.

[2] J. H. Greenberg, 'A quantitative approach to the morphological typology of language,' *Method and perspective in anthropology*, ed. R. F. Spencer, 1954, pp. 192–220.

[3] C. E. Bazell, *Linguistic typology* (inaugural lecture), 1958.

[4] J. Vendryes, *Le langage*, 1921, p. 349 ; Meillet, *Linguistique historique et linguistique générale*, pp. 77–8.

[5] W. P. Lehmann, *IJAL*, 24, 1958, p. 143.

[6] *Linguistic science*, p. 100.

[7] 'Le Proto-chinois, langue flexionelle,' *JA*, 15, 1920, pp. 205–32.

[8] 'Gedanken über das Indogermanenproblem,' *Acta Ling.*, 1, 1939, pp. 81–9.

an *Ursprache* and of relegating the common vocabulary criterion to a ʋ
secondary position,[1] but his six ' marks ' show every sign of having been selec
to fit the pre-existing genetically established IE family as closely as possible,
another six equally fundamental features might have split it up and united p
of it with genetically unrelated languages elsewhere in the world. And it ha
fact been shown that at least one of Trubetzkoy's ' marks ', the absence oʆ
ergative-type construction, does not apply to all the accepted IE family.[2]
would, of course, have been entirely legitimate for Trubetzkoy to try to rede
IE on a non-genetic, synchronic, basis ; but if that had been his intention, tʰ
would be no need to choose the particular criteria that should keep the genetic
established IE family intact, in preference to a possibly more balanced and equ
basic set of criteria. The point is that in view of structural evolution, and struct
influences produced by prolonged language contacts, it is illicit to exploit
criteria applicable to synchronic comparison to produce or even to butt
historical genetic groupings.

In dealing with the languages of a geographical area (South East Asia in
present context) it must be emphasized first (though it is obvious enough real
that there is nothing necessarily linguistic in either of the senses discussed ab
about a group of languages defined by the location of their native speech c
munities. But *a priori* pointers to the fruitfulness of both synchronic and gen
comparison of these languages (apart from immediately obvious typolog
likenesses and phonetic-semantic lexical correspondences among some of th
are : (1) the tendency for phonological and grammatical (structural) feature
diffuse over geographically contiguous areas, and (2) the likelihood that som
the population of at least parts of the area have migrated from unitary centre

It would seem desirable to keep the following points in mind in a gen
(comparative-historical) classification :

1. The degree to which in certain areas and among certain languages
absence of written records of any great time-depth relatively weakens the effecꞇ
ness of genetic grouping, though this may be partly supplemented in due co
by the findings of glottochronology. If such groupings are to be supplementeᵈ
classifications based on other types of criteria the mixed nature of the resuꞁ
picture must be recognized. In this connection the procedure of M. Guꞇ
in his *Classification of the Bantu languages* is illuminating.[3] He sets up two
of what he calls principal criteria,[4] the first being a set of synchronic feaꞇ
mainly relating to the prefix concord systems characterizing the gramma
these languages, and the second being a genetic criterion, namely the posses
in a part of the vocabulary of the language of a stock of words referable to

---

[1] ibid., p. 85.
[2] W. S. Allen, ' A study in the analysis of Hindi sentence-structure,' *Acta Ling.*, 6, 19
p. 69.
[3] London, 1948.   [4] ibid., pp. 11–19.

of hypothetical roots or ' starred forms '. He requires both criteria to be satisfied (as well as some subsidiary, less easily applied criteria which may be ignored here) for him to recognize a language as Bantu ; but languages may satisfy only one of the principal criteria, and in this case the genetic criterion receives precedence, as languages having the vocabulary relationship but not exhibiting the prefix system are called ' sub-Bantu ' and included in his treatment of the Bantu languages, but those having the synchronic grammatical features of the Bantu languages but not a significant body of common Bantu vocabulary are called ' Bantoid ' and not further dealt with.

2. The recognition of different strata within one language with different genetic affiliations, e.g. Sanskritic words in a number of the languages concerned. Traditionally, minority elements of this sort are treated as loans, but where they form a considerable part of the general vocabulary of a language and all come from one or two sources, a separate genetic classification of the different layers or strata composing the language, at least for some speakers, may be needed.

3. More generally, the sources and directions of loanword acquisition, apart from those in whole strata of the type just mentioned.

The synchronic classification of the languages of the area would appear more fluid initially, as the choice of types of criteria is wider. The two major metacriteria of the usefulness of features at every level would seem to be :

1. Features that at the level concerned are more rather than less deeply embedded in the language system will make for a more compelling classification (e.g. features characterizing a whole phonological or grammatical system, as against features relating only to a single consonant or a single word class).

2. Features that together exhibit fairly distinct isoglosses around and within the area are likely to be more significant for areal groupings than features with independent and less distinct isoglosses.

The fact, already mentioned, that structural features do spread over contiguous languages, often across genetic boundaries, favours *a priori* the search for significant structural traits in a geographically defined region.

Types of articulation, syllable structure, vowel contrast systems, systems of consonants, and pitch exploitation in lexically distinctive tones as against grammatically and semantically significant intonation (or, as in several languages, a combination of each) are all possible lines of comparative investigation at the phonetic and phonological levels. Grammatical level and inter-level classification is possible in terms of word structure, types of affixation or other word-form paradigmatic alternations, and basic syntactic structures. And lexical comparison may be made by studying the way the different languages divide up and structure lexically certain universal and culturally important fields, e.g. kinship terms, colours, social status, natural phenomena, and the like.

It must be emphasized that isoglosses of these kinds will encompass linguistic

B

typological areas, and not as such geographically significant areas, though the
is likely to be a correlation between them. It may well be that the linguis
delimitation of ' South East Asia ' by major isogloss coincidences and likew:
the division of this area into sub-areas by similar means will define areas th
include parts of the world that are not usually classed as South East Asia ge
graphically and politically, and exclude some parts that are, as well as cutti
across recognized non-linguistic divisions within the area. In this one must
ready to follow the linguistically significant criteria where they lead. Any selecti
of criteria primarily on the basis of coinciding with boundaries and divisions r
themselves linguistically decided is likely to be, for the reasons suggested earli
less significant even if they provide some feasible isoglosses.

In an area wherein for many years a good deal of work has been carried on
various investigators and by various methods of grouping and classificatic
sometimes without it being made explicit just what methods are being employ
and principles adhered to, it may be of use at the outset of the present symposiu
to keep very clearly in mind what data and methods of classification we ha
available, and on exactly what basis we are working in any particular operation.

THE COMPARATIVE METHOD AS APPLIED TO NON-INDO-EUROPEAN LANGUAGES

By Nils M. Holmer

It has been said from time to time that comparative linguistics as a scientific method does not exist outside the Indo-European languages. It is further not seldom announced that comparative study in the field of non-Indo-European languages requires a different method from the one adopted for Indo-European. Both views are equally wrong. It is, however, easy to see how they have originated.

Regardless of the fact that the comparative study of language on a scientific basis began with the Indo-European languages—in connection with the acquaintance with Sanskrit—we find ourselves in a particularly favourable situation when it comes to the historical-comparative study of these languages. For comparative linguistics is basically historical linguistics. Even apart from the great advantages offered by the discovery of Sanskrit as an Indo-European language we should probably have got just as far, sooner or later, by means of the historical study of, for instance, Latin and Greek. Through it we should have been able to trace the evolution of language and observe the regularity with which it takes place. The idea of the *sound law*—one of the most fundamental principles of comparative linguistics—would no longer have been far-fetched. Incidentally, the existence of a regular correspondence between the consonant sounds in related languages was observed by a Danish scholar quite independently of the results gained by Sanskrit philology and within quite a different field, viz. that of the Germanic languages.

Comparing the evolution of comparative linguistics in, for instance, Semitic and Finno-Ugric, we find that within the former branch linguistics has tended to take the form of philology, owing to the fact that the integrant languages are too closely related and too little altered in the course of time to stimulate the study of a historical evolution (in the same way the closely related Romance languages have proved to be more interesting from the philological than from the linguistic point of view). On the other hand, the languages of the Finno-Ugric stock are a little too distant from one another to serve as a basis for the establishment of a comparative linguistic science in the same sense as for the Indo-European languages, and this becomes still more noticeable as we proceed further east with a view to including Samoyed, Yukaghir, and perhaps some other languages.

Before the rise of comparative Indo-European linguistics, comparative linguistics consisted in the confrontation of usually isolated words, having a similar sound and meaning, in the most different languages, presumably under the impression that all languages went back to one form of human speech, shattered and split up at the time of the great ' confusion of languages ' in Mesopotamia. It is quite surprising that still in our days there are linguists who