

GENERAL PROBLEMS OF CLASSIFICATION OF AFRICAN LANGUAGES

On principle, there are no language classifications which are right or wrong, but only classifications which are more or less useful or useless. This statement at the beginning of my paper is intended to indicate the teleological perspective in which I want to view the more general problems involved in the classification of African languages. I shall discuss these within a framework which I derive from the four main components of any language classification, namely:

1. the aims and objectives of language classification,
2. the basic units to be classified,
3. the criteria of classification,
4. the models of language classification.

Language classifications are scholarly constructs which are essentially determined by the objectives for which they are made. Many problems arise from the fact that there are often incompatibilities between the objectives on the one side and the

Classification of African Languages

methods applied in constructing the classifications on the other side. In my view, these incompatibilities can have three different sources: (1) the classifiers apply the wrong methods for the wrong objectives; (2) the users of a given language classification use it for different objectives than determined by the classifiers; (3) the classifiers or users try to make it serve several incompatible objectives. All these problems will be dealt with in the second section of this paper.

Every classification is based on a procedure which correlates certain elements or objects under the aspect of certain criteria. In a language classification, the elements to be classified are the individual language systems at the level of human groupings defined by social and/or geographical features. Until recently, the classifiers of languages in Africa evidently gave little thought to the question of how these elements should be defined. However, with the increasing knowledge of the actual linguistic facts in Africa, this innocent attitude towards the question has been lost. The definition of “languages” and “dialects,” the phenomenon of dialect continua or the selection of the units to be classified are more and more conceived as problems by the modern language classifiers. I shall deal with this complex in the third section of this paper.

The question has always been discussed as to which kinds of criteria should be used in classifying languages. Are only linguistic criteria allowed or non-linguistic ones too?—In my opinion, the nature of the criteria plays a minor role. Considering the past practice of language classification, it seems to me more important to ask whether the criteria selected for a classification do really exist in the units to be classified, and whether the actually existing criteria are also adequate for the objectives pursued with the classifications. Only when these basic questions are clarified, the nature of the criteria may play a role. The discussion of these questions will follow in the fourth section.

Most language classifications are interpreted according to specific models. Models in this sense are so to speak the prefabricated parts of interpretative thinking. They combine certain premises with specific structural criteria in a preconceived way. Just by the choice of a model, the classifier of languages automatically introduces all the silent and overt assumptions,

structural elements, etc. which are pertinent to such a model. This may be innocuous as long as the classifier is aware of this effect not only in the abstract, but knows in detail which premises and structural elements he is introducing in this way. In other words: there is a close relationship between the objectives of a language classification. Contrary to the belief shared by most experts, I think that it is the objective of the language classification which should determine the choice of the classificatory model. If one is ready to accept that the objectives should be given primacy over the choice of the model, one has to ask which models of classification fit best which objectives and which functions of classification. This controversial question will be dealt with in the fifth section.

I. OBJECTIVES AND AIMS OF LANGUAGE CLASSIFICATIONS IN AFRICA

1. *General view*

So far, language classifications in Africa have been constructed for four purposes:

1. to serve as referential systems,
2. to discover and reconstruct inherent cultural and historical relations,
3. to discover typological relations of linguistic and non-linguistic kinds,
4. to serve as a basis for language policy making, language planning and similar purposes in the applied field.

On principle, I assume that all objectives are equivalent and that this catalogue of classificatory aims is open for further aims which may come up in the future. In this context, I do not want to discuss primarily the contents of the objectives of language classifications, but rather the problems which are related with their contents. These can be subsumed under the following three headings: (1) constructing classifications under false labels, (2) using classifications for other objectives than they have been made for, (3) mixing incompatible objectives in one classification.

Classification of African Languages

2. The problem of using false labels

Among classification of African languages false labels are widespread. I think that, for instance, most of the genetic classifications are simply systems of reference based on synchronical data and not what they claim to be, namely classifications which can reveal real historical facts. A false declaration is harmless if all users know about this and deal with it accordingly. For instance, I would never use Greenberg's comprehensive classification (1955; 1963) for other purposes than for organizing a library or a language catalogue. However, the representatives of neighbouring disciplines, in particular ethnologists, social anthropologists and ethnohistorians, who cannot easily evaluate the ingredients with which Greenberg's classification was cooked, are sometimes inclined to draw important historical facts from this construct. Thus, the real problem of false labels seems to lie in the transdisciplinary use of given language classifications. In order to prevent others from using them in an inappropriate way, it would be a great help, if the authors of language classifications themselves would lay open the objectives, premises, material foundations and perhaps also the limits of their classifications instead of making excessive claims for them.

3. The problem of later changing the original objectives

The application of language classifications for purposes for which they were not made is not harmful of itself. For instance, any classification which classifies a large number of African languages may be used successfully as a referential system. Inversely, it is highly problematic, however, to use a referential classification for other thematic objectives. Guthrie (1967:84), for instance, when constructing his Proto-Bantu hypothesis, drew far-reaching historical conclusions from his earlier referential classification which, according to his own statement (Guthrie 1948:27), was also based on arbitrary premises. Most of the conclusions proved later to be inadequate (Möhlig 1976; 1979).

Beyond the general function to serve as a referential system, later

changes of the objectives of a given language classification will mostly not be possible, because the choice of its criteria and structural particulars is too much influenced by the original aims. For instance, the question of whether two genetically closely related languages can be served by one school language, cannot be answered on the basis of their genetic relationship, but only according to the criterion of linguistic proximity, by whichever historical process this proximity may have been created.

Apart from the incompatibility which on principle exists between different classifications with different thematic orientations, there is always the possibility of comparing their results and to draw new thematic conclusions from the results of this comparison. For instance, if certain typological criteria lead to a similar classification as certain genetic criteria, then it suggests that both results are somehow related. However, I think that the exploration of such relations needs its own method, the contours of which are at present as vague as those of an interdisciplinary or transdisciplinary method.

4. *The problem of combining several aims*

The question of combining several aims in one language classification is similar to that dealt with in the preceding section. It differs in so far as, in this case, the classifier himself plans such a combination before structuring his classification. Within which limits is such a combination possible?

The predominant interest in the classification of African languages, as in the comparative study of African languages in general, has always been directed towards historical aims and, among these, particularly towards genetic relationship. Some of the classifiers, such as Johnston (1922), Guthrie (1967-71) and Heine *et al.* (1977), went even beyond the aims of language history and tried to connect also ethno-historical objectives with their classifications. It is questionable whether these two thematic complexes are really compatible. According to my experience, the linguistic criteria which give evidence of ethno-historical or cultural-historical facts are not merely those which testify to genetic relations between languages.

Classification of African Languages

Apart from the cases in which linguistic and non-linguistic aims may be incompatible, there even exists often an incompatibility among different linguistic aims. For instance, it has been often neglected in the past that there is a difference between a genetic and a historical classification. The genetic relations are only one aspect of language history among others which are perhaps more important. What is to be understood by a historical classification has been defined by Meinhof and van Warmelo (1932:176) as follows: "... a linguistic classification should be historical, that is a classification that embodies as far as possible the actual history of the languages concerned. It would have to show how the proto-form of the family split up into its various branches and subdivisions, how languages have influenced each other, and where foreign influence has been at work." It is evident that this formulation combines several aims within the domain of historical linguistics, demanding not only a complex selection of classificatory criteria but also a model of classification which goes far beyond the necessities of a genetic classification. It is certainly true that the genetic aspect is contained in the overall view of language history in general and thus, as a classificatory aim, is compatible with the other aspects of language history. However, *vice versa*, the models which are usually used for genetic classifications, because of their inherent assumptions and structural characteristics, are inadequate for, not compatible with, the wider objectives of a real historical classification.

II. THE DEFINITION OF THE BASIC UNITS OF CLASSIFICATION

1. The principle of using selected units

The older language classifications in Africa are based on the principle of using selected units of classification. This implies that, out of a totality of individual language systems to be classified, only some are selected and submitted to classification, whereas the majority are neglected. Former classifiers of African languages had no other choice because of the defectiveness and incompleteness of the linguistic documentations at their disposal. With the increasing knowledge of the African languages during the last three

decades, the justification for the application of the principle of using selected units has however completely changed. Now its application is being justified by the masses of linguistic data available. Greenberg, in his "complete genetic classification of the languages of Africa" (1963:1 ff), does not explicitly deal with the problem of selecting the units to be classified, but uses in fact a selection of some 700 out of a totality of over 2000 languages. Guthrie (1962, 1967-71) who, in his genetic classification of the Bantu languages, also had to face a mass problem, explicitly did deal with the question of selecting the units to be classified. In order to overcome the mass problem, he applied the principle of "test languages" (1967:97). From among over 500 languages and an unknown number of dialects, he selected 28 test languages. On this basis, he constructed a classification which was intended to be valid for the whole of Bantu.

From the point of view of taxonomical methods, there are no objections to selecting the units of classification as long as the units selected are representative also for the units disregarded. Guthrie tried to solve the question of representativeness in choosing at least one language for each of the 15 zones of his earlier referential classification. Another 13 languages were selected "to include differing degrees both of geographical contiguity and of linguistic similarity." (1967:97). But in spite of this comparatively careful procedure, he could not avoid putting together a collection of units to be classified which, with respect to the aims of this classification, has to be evaluated as non-representative. The reasons for this judgement emanate from the unbalanced geographical distribution of the test languages and the divergent degrees of group status which these test languages have. As to the first argument: if one starts from the geographical rosters of Bantu which Guthrie himself set up either in his practical classification (1948) or in his topological analyses (1967:81 ff), the geographical distribution of the 28 test languages is easily seen to be unbalanced. Likewise, if one takes as the basis the inventory of 83 language groups of Bantu compiled by Bryan (1959), it turns out that, out of these, only 21 groups are represented in Guthrie's test languages, and among these 21 groups, there are four groups which are represented by more than one test language. As to the second argument: some of the test languages, such as Swahili, composed

Classification of African Languages

of many dialects and thus located at a comparatively high level within the hierarchy of language grouping, are counted as one test language along with languages like Rundi or Nyoro, which are units located at the low dialectal level.

The criterion of representativeness within the context of language classification remains an unsolved question, about which an adequate discussion is urgently needed.

1. The dialectological principle

The dialectological principle demands that the languages to be classified cover the whole area of classification, in other words, that no dialectal “no man’s land” is allowed between them. The application of this principle is appropriate particularly for language classifications with an applied perspective such as language planning or language policy making. It is also useful within the domain of truly historical classifications. If one follows the dialectological principle, one is quickly confronted with a lesser or greater mass problem depending mainly on the size of the area of classification. Apart from this, one is exposed to two further groups of problems. The first is connected with the phenomenon of “dialect continuum”, the other originates from the fact that the individual language systems of a coherent area to be classified, regularly show various degrees of linguistic nearness to such an extent that the divergencies usually cannot be expressed by the traditional parameters of “language” and “dialect”.

The mass problem is governed by the formula $n(n - 1) : 2$, where n stands for the number of languages which have to be compared in pairs. If there are, for instance, 10 languages to be classified, the number of pairs of comparison amount to 45. A number of 20 languages gives already 190 pairings, and 30 languages amount to 435 comparisons in pairs. To give a concrete idea of the masses: the last amount corresponds roughly to the size of Chaga dialects on the slopes of Mt. Kilimanjaro and Mt. Meru in the northern part of Tanzania. The numbers show that this sort of a mass problem can only be controlled by limiting the area to be classified. If the area is larger than the work capacity of the classifier and eventually of his computers, a technique of connecting several

micro-classifications to form one macro-classification within a larger area as required by the classificatory aims has to be resorted to (Möhlig 1980:39 ff).

Dialect continua are characterized by transitional areas between neighbouring language systems without showing well marked language boundaries, such as bundles of isoglosses. Except for particular questions of language planning, the eventual language boundaries or the areas of transition within a given region are of lesser importance. For most classificatory purposes, the problem of defining language boundaries can be avoided if we concentrate on the dialectal centers and take their linguistic criteria as representative also for the peripheries. It is evident that such a procedure involves already a certain amount of abstraction from reality. For referential, historical and typological objectives, this may be harmless. But for language classifications which are to serve language planning or similar purposes, it is important to take also into account the transitional areas, since these allow to draw conclusions on interdialectal communication and experience with neighbouring dialects. Such information can be essential for choosing, for instance, school languages or media for adult literacy programmes.

The problem of having only two categories to express hierarchical status, namely “language” and “dialect”, is almost as old as language classification in Africa. Already Koelle (1854) conscientiously tried to note also the local varieties of certain languages within his compilation, without forcing them into the traditional categorial system. A linguist who exposes himself to the dialectal reality of a given area, has soon to realize that the complex structure of dialectal relations can hardly be captured by “language” and “dialect”. Whereas the term “language” can be used autonomously, the term “dialect” denotes a relational concept which only makes sense with reference to other language systems at the same time. Usually this term can be used in two ways, either in a statement like “A is a dialect of B” or in a statement like “A and B are related as dialects.” In these, the term “dialect” can denote a very close relation with mutual understanding, but it can also mean that the units called “dialects” are only rather remote members of the same language family. Since classifications with the perspective of language planning as well as those with a

Classification of African Languages

historical perspective depend on the exact representation of the actual language situations in a given area, more parameters are needed which are capable of coping with the complexity at the dialectal level. The new methods of dialectometry (Séguy 1973; Möhlig 1974; 1980a, Goebel 1982) offer the possibility to measure the dialectal distances between the different language systems of a given area on linguistic grounds. As a result of such measurements, the hierarchical structure of the different levels of dialectal relationship usually becomes visible. This allows us to distinguish, in a specific case, between dialects of the first, second, third, etc. levels of relationship. Winter (1980) tried even to calibrate the different levels of relationship quantitatively as well as qualitatively by using such criteria as “still spontaneously interintelligible,” “requiring experience” and so forth. An international project sponsored by the C.N.R.S. in France in the years 1981-83 for the development of dialectometry in Africa produced some more techniques and devices to cope with the different degrees of dialectal relationship. These will soon be published (Guarisma and Möhlig forthcoming).

3. The comparability of the basic units of classification

Whether we proceed according to the principle of using selected units or to the dialectological principle, in both cases, we have to define the units to be classified within the context of other comparable units. To identify them only by their names has proved to be insufficient. Comparability means that the basic units of the classification be defined at the same or almost at the same level of the relational hierarchy. If individual language systems are classified along with dialect clusters or even hypothetical language families, as has been done in older classifications, internal inconsistencies will be the consequences independently of the model of classification chosen.

As to the question of which criteria have to be positively observed to guarantee that the units to be classified are comparable, we have to take into account that the relational hierarchies which provide the framework for locating the units of the classification depend entirely on the specific characteristics of

the complex network of linguistic relations among the languages to be classified. In other words, these hierarchies are abstractions from the actual networks of relation. Therefore, they cannot be imposed from outside as invariable parameters. From this follows that practically each relational or classificatory hierarchy differs from any other hierarchy of this sort as to the numbers of units to be located at the same level. In this perspective, it becomes evident that it is extremely difficult, if not impossible, to define the individual language systems at higher hierarchical levels. Since only the base of the hierarchies is accessible with empirical methods, I propose that the group languages at the level immediately above the idiolects be taken as the basic units of any language classification. The term "group language" includes sociolects and dialects. Whatever the case may be, these units are testable with empirical methods, whereas the units of higher relational levels are often a matter of belief and ideology.

III. THE CRITERIA OF CLASSIFICATION

1. *General features*

At a first glance, it may appear to be self-evident that language classifications are exclusively based on linguistic criteria. However, languages are not isolated phenomena. They are handled by human beings in many functions such as communication, self-identification, artistic expression and so forth. Linguistic nearness, the basic parameter of most language classifications, whether silent or overt, is therefore composed of many factors, among them non-linguistic ones, too. In particular, factors, such as age, sex, culture and space, have an immediate influence upon the communicational habits of the people and, through this, upon inter-dialectal communication. According to the teleological point of view adopted in this paper, the choice of the criteria of classification depends on the objectives for which the classifications are constructed. Since many language classifications also pursue non-linguistic aims, I feel that it is justified also to consider non-linguistic facts. Whatever the nature of the criteria may be, their validity has to be judged against two questions: first,

Classification of African Languages

whether the criteria are real and, secondly, whether they are adequate in view of the classificatory aims.

2. The feature of reality

The question of whether the classificatory criteria are real, first of all, is a matter of empirical testing. This has to be said, because in a comparatively young discipline like African Studies, it often occurs that certain criteria are reported to exist where they are in fact absent. With the increasing knowledge about the African languages, such mistakes may die out automatically.—But perhaps this view is too optimistic.

If we relate the question of real existence with the problem of how to select the units of classification, another dimension of empirical reality becomes visible. Instead of being positive or negative, the answer may also be one of partial reality. If we say, for instance, that in Swahili the concept of crocodile is represented by a form *gwena* (so Guthrie 1970:230), this statement is true only for the northern dialects of Swahili. In the central and southern dialects of Swahili the correct form is however *mamba*. Depending on the classificatory aims, such sophistication may be important.

The historically reconstructed proto-forms which are often used as basic criteria in language classifications are another field where the feature of reality is highly problematic. I do not want to deny that historical reconstructions can reflect some historical reality, but I see that, in practice, they are too often the products of science fiction and personal imagination.

3. The feature of adequacy

If we relate the classificatory criteria with the aims of language classification, the question of adequacy arises. Is it, for instance, correct to construct genetic classifications on the basis of taxonomical criteria? In this context, I recall the problematic results based on lexico-statistical methods. I am inclined to see

here a classical case of inadequacy between the criteria chosen and the classificatory aims pursued. In my opinion, the choice of adequate criteria is the most critical part of the whole classificatory procedure. It should therefore be given careful attention when planning new classifications. However, the scholarly discussion about this problem is still very much underdeveloped. For the time being, we can only draw attention to this problem and appeal to the scholarly ethics of the classifiers that they themselves disclose the limits of the functional capacities of their classifications in order to prevent the users, in particular those from other disciplines, from drawing false conclusions.

4. *Linguistic criteria*

Looking at the linguistic criteria, it seems to me that the main problem in this context is the mass problem. The more languages are included in a classification, the less criteria can be considered because of limited capacity. The computers have widened the capacity, but the problem as such remains. Greenberg (1963:1 ff) tried to make a virtue of it in creating the methodological principle of “mass comparison.” Fodor (1966; 1982) argued against it with reasons which have neither been substantially discussed nor refuted by others.

The question about kind and quantity of the linguistic features will have to be answered for each classification individually and in a pragmatic way with the classificatory aims in view. So far, lexemes and morphemes were the preferential criteria chosen by the classifiers of African languages. The application of a mixed set of criteria gave rise to the problem of “mixed languages,” the prominent example of it still is the old “Mbugu case” of Tanzania (Tucker and Bryan 1974). Phonological criteria also played a role in language classification, but only selectively via the principle of regular sound correspondences used for the reconstruction of proto-lexemes. If one takes into account all the phonological features of the languages to be classified in their systematic structural contexts (Möhlig 1980b:36 ff), the problem of mixed languages can be solved (as to Mbugu see Möhlig 1983:158 f).

Classification of African Languages

5. Non-linguistic criteria

Among the non-linguistic factors, the geographical factor is the most important one. Languages as such have no residence, only the people who speak them can be located. Thus far, geographical criteria are abstractions. However, the geographical matrix has its own heuristic values which have been recognized by linguistics for more than a century.

Another non-linguistic feature in classifications is the time factor. The boundaries between synchrony and diachrony are highly problematic, at least in practice. As has been said before, the main interest in the comparative study of African languages has always been historical, although for the majority of African languages, true historical sources do not exist. By necessity, the criteria, also for historical classifications, are therefore mostly synchronic. Against this, one can argue that the synchronical linguistic facts are the result of historical processes, it must therefore be legitimate to view these facts under the aspect of being pieces of historical evidence. The methodological discussions on how to exploit these pieces of evidence are just beginning, now when it is recognized more and more that the old mechanistic methods of linguistic historical comparison, if measured against historical reality or truth, are insufficient.

Other non-linguistic criteria of language classification cannot be excluded right from the beginning, but, in the past, they have proved to be of little efficiency and adequacy. In general, we may say that language classifications which depart from the purely linguistic facts, gradually get into the interdisciplinary field for which, in most cases, they are insufficiently equipped with appropriate methods.

IV. CHOICE OF THE MODEL OF CLASSIFICATION

1. Relationship between the models and the aims of classification

There exists a close connection between the models of classification and the classificatory aims. A model of classification is mainly an instrument of representation. It is meant to

emphasize, for the user, certain relations between the languages classified, which the classifier thinks to be important and to suppress eventually other relations, which the classifier deems to be less important. Already because of this feature, models of language classification never give a true picture of the reality, as unfortunately many people believe, but only a frame of interpretation for certain matters and facts.

In analogy to the bipartition of linguistics in general, one can distinguish between synchronic and diachronic models of language classification. One would think that, because of their nature, the diachronic models are used for all classifications which pursue somehow historical aims, whereas, for all other purposes, synchronic models are the appropriate means of representation. However in practice, diachronic models are often used to serve synchronic purposes and, *vice versa*, synchronic models are used to serve diachronic purposes. The widespread mixing of the temporal aspects is a serious problem which, in my view, can only be helped by making it overt.

2. *Synchronic models of classification*

The basic factors which regularly underlie synchronic language classifications are: linguistic proximity and areal distribution.

Depending on the specific classificatory aims, qualitative linguistic factors may be added.

On the basis of these factors, the multidimensional relations between the languages to be classified usually become visible. Although the traditional dendrogrammatic representations may suggest it, yet these multidimensional relations are often not structured according to the logic: if A equals B and C, B must also equal C. It frequently occurs that A is a member of two different subgroups. This clearly contradicts the dendrogrammatic concept. A synchronic model of language classification should be able to represent this feature of complexity.

The other feature to be represented by a synchronic classification is the hierarchical structure of linguistic relationship. In the tradition of classifying African languages, the *Handbook of African Languages* with its mainly referential aims, tried to express

Classification of African Languages

the hierarchical structure by using such concepts as “language,” “dialect,” “dialect cluster,” “language group” and “larger unit.” The introduction of this conceptual framework certainly helped to reduce the confusing multitude of basic units to be classified, but remained confusing enough so that languages which are not so well known can only be found with the help of the alphabetical index. Guthrie (1948) improved the system of hierarchical representation, at least for the Bantu field, in choosing the geographical position as the highest hierarchical level. If one has a rough idea of the areal structure of the “zones”, he or she can easily derive from the referential code where a language is spoken and in which dialectal context it is situated.

Dalby (1977), in his general classification of all African languages, tried to combine the feature of relational complexity with the feature of hierarchical structure by introducing three geographical units at the uppermost classificatory level: the “fragmentation belt” in the central area and two “areas of wider affinity”, one in the north, the other in the south. With the help of a mixed code of numerical and alphabetical symbols, each language is related with either of the two areas of wider affinity and, additionally, with the fragmentation belt, be it in the positive or in the negative. Dalby’s classification could not gain the broad recognition which Guthrie’s classification has for Bantu, in my view, for several reasons: First, the referential code system for the individual languages is little illustrative. Secondly, the contour-lines of the areas of wider affinity are rather vague since these referential units have been defined on the basis of the old historical grouping “Hamito-Semitic” and “Niger-Congo”. Perhaps the decisive deficiency is, thirdly, the reduction of the number of referential areas.

Fivaz and Scott (1977), with reference to Guthrie, introduced a more or less independent referential code system of letters and decimal numbers for each smaller African language family recognized as such since the days of Westermann (1927) and Westermann and Bryan (1952). For the external groupings of the language families, a separate number code is used. It is structured along the lines of Greenberg’s major genetic subdivisions. This dual referential system has contributed considerably in increasing the transparency of the classification. Yet it does not reach

Guthrie's degree of clarity, because the order by letters within the families follows a genetic system, which often contradicts the close relations between areal neighbours.

For smaller areas of the size of a modern African nation, I myself developed a synchronic classificatory system with an applied perspective (Möhlig 1980a). This takes into account, on the one hand, the multidimensional complexity of language relations, and on the other hand, the hierarchical structure of these relations. All hierarchical levels are defined according to the actual inherent parameters of linguistic proximity or distance of the languages to be classified. For representation I use multidimensional diagrams as in figure no. 1. In these, the complexity of the linguistic relations is put into prominence. If I want to focus on the areal aspect of relations, I use cartograms as in figure no. 2.

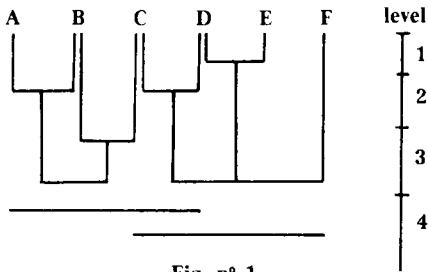


Fig. n° 1

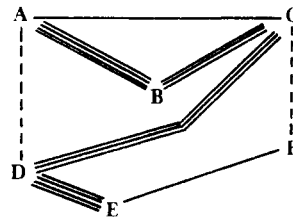


Fig. n° 2

3. Diachronic models of classification

The model of classification which is used most in the domain of African languages, is the “unilinear, monogenetic model,” which is normally rendered in a dendrogrammatic form. This model has the undeniable advantage of transparency. Its disadvantage is that its degree of reality decreases, the more languages are taken into account. If we correlate this model with a time factor which is derived from the attested language history in other parts of the world and project the historical implications of the model backwards into the past, we arrive at the conclusion that, in the whole of Sub-Saharan Africa, 5000 to 8000 years ago, there must

Classification of African Languages

have existed less than 10 languages. This simplistic picture is in sharp contrast with the complex and differentiated picture depicted by archaeology for the same era.

My dialectological studies give me a different picture. In the microcosm of a dialectal situation, many slightly divergent configurations of linguistic forms and patterns can be seen side by side, i. e. at the same time level. This is evidently due to the fact that, in different dialects, the clocks of historical development go at different speeds. Therefore, these dialectological configurations can be plausibly re-arranged in a sequential order and interpreted as the subsequent stages of the same chain of historical development. From such a chain, specific historical processes can be deduced. With reference to one language system, the totality of the historical process is often contradictory in itself in the sense that the direction of a certain chain of linguistic development may be reversed. For instance, there may be evidence that a certain class of distinctive sounds was deleted at an earlier period and, at a later period, it was re-introduced. In addition to such observations, the processes, when plotted on maps, show incongruent areal distributions. These and similar phenomena are the effects of stratificational processes, such as language shift, language mixing, superimposition, and others. The inclusion of such historical processes, besides the process of genetic evolution,

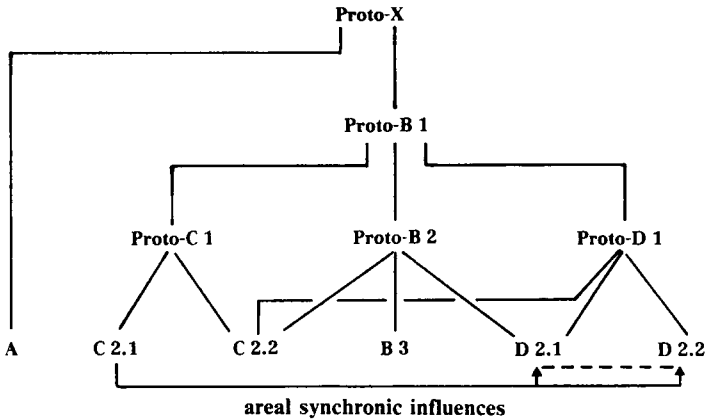


Fig. n° 3

certainly requires a more complex model of classification (Möhlig 1980b:48 ff). In figure no. 3, I give the outlines of such a stratificational model of language classification. Its effects, particularly on the classification of the Bantu languages, have been demonstrated in a separate article (Möhlig 1981).

V. Final remarks

For a long period in the history of African Studies, the discussion of the methods of how to classify African languages has been marked by stagnation due to the supremacy of the unilinear monogenetic model. Only when our knowledge about the African languages developed to the point where dialectological methods could be added to our arsenal, as in Europe already a century before, a new insight into the inherent structure of language relationships became possible. This gave an impulse not only to re-think the whole question of language classification, but also to recognize that there are difficult problems involved which we had not seen before.

In this paper, I have tried to discuss some of the major problems within a heuristic framework structured according to the main components of language classification, namely the objectives and aims, the units to be classified, the criteria of classification, and finally the classificatory models. The arrangement of these components expresses somehow the different degree of importance which, in my opinion, they have for the whole process of classification. This weightiness is however without any consequences for the arguments used in discussing the problems involved. More important is perhaps my teleological attitude toward all questions of language classification. In order not to mislead anybody, I made this clear right in the beginning. With a reminder of it, I end this paper.

Wilhelm Möhlig
(*University of Cologne*)

Classification of African Languages

REFERENCES

- BRYAN, M., *The Bantu Languages of Africa*, London - New York - Cape Town, Oxford University Press, 1959 (Handbook of African Languages).
- DALBY, D., *Language Map of Africa and the Adjacent Islands*, London, International African Institute, 1977.
- FIVAZ, D.; SCOTT, P. E., *African Languages. A Genetic and Decimalised Classification for Bibliographic and General Reference*, Boston, Mass., G. K. Hall, 1977.
- FODOR, I., *The Problems in the Classification of the African Languages*, Budapest, Center for Afro-Asian Research of the Hungarian Academy of Science, 1966.
- *A Fallacy of Contemporary Linguistics*, Hamburg, Helmut Buske, 1982.
- GOEBL, H., *Dialektometrie. Prinzipien und Methoden des Einsatzes der Numerischen Taxonomie im Bereich der Dialektgeographie*, Vienna, Verlag der Oesterreichischen Akademie der Wissenschaften, 1982.
- GREENBERG, J. H., *Studies in African Linguistic Classification*, Branford, Conn., Compass, 1955.
- *The Languages of Africa*, The Hague, Mouton, 1963.
- GUARISMA, G.; MÖHLIG, W. J. G., (eds.), *La méthode dialectométrique appliquée aux langues africaines*. (Forthcoming).
- GUTHRIE, M., *The Classification of the Bantu Languages*, London-New York-Toronto, Oxford University Press, 1948.
- "A two-staged method of comparative Bantu study." *African Language Studies* (London), Vol. III, 1962, p. 1-24.
- *Comparative Bantu. An Introduction to the Comparative Linguistics and Prehistory of the Bantu Languages*, Farnborough, Gregg Press. Part I, Vol. I, 1967. Vol. III, 1971. Part II, Vols. III-IV, 1970.
- HEINE, B.; HOFF, H.; VOSSEN, R., "Neuere Ergebnisse zur Territorialgeschichte des Bantu." In Möhlig, W. J. G.; Rottland, F.; Heine, B. (eds.) *Zur Sprachgeschichte und Ethnohistorie in Afrika. Neue Beiträge afrikanistischer Forschungen*, Berlin, Dietrich Reimer, 1977, p. 57-72.
- JOHNSTON, H. H., *A Comparative Study of the Bantu and Semi-Bantu Languages*, Vol. II. London, Oxford University Press, 1922.
- KOELLE, S. W., *Polyglotta Africana; or A Comparative Vocabulary of Nearly Three Hundred Words and Phrases, in More Than One Hundred Distinct African Languages*. London, Church Missionary House, 1854.
- MEINHOF, C.; VAN WARMELO, N. J., *Introduction to the Phonology of the Bantu Languages*, Berlin, Dietrich Reimer, 1932.
- MÖHLIG, W. J. G., *Die Stellung der Bergdialekte im Osten des Mt. Kenya. Ein Beitrag zur Sprachgliederung im Bantu*, Berlin, Dietrich Reimer, 1974.
- "Guthries Beitrag zur Bantuistik aus heutiger Sicht." *Anthropos. Zeitschrift für Völker- und Sprachenkunde* (St. Augustin), vol. 71, 1976, p. 673-715.
- "The Bantu nucleus: its conditional nature and its prehistorical significance." *Sugia. Sprache und Geschichte in Afrika* (Hamburg), vol. I, 1979, p. 109-141.
- "La dialectométrie: une méthode de classification synchronique en Afrique." In Guarisma, G.; Platiel, S. (eds), *Dialectologie et comparatisme en Afrique Noire*, Paris, Selafl, 1980, p. 27-44.
- Bantu Languages. In Heine, B.; Möhlig, W. J. G. (eds.) *Language and Dialect Atlas of Kenya*, Vol. I, Berlin, Dietrich Reimer, 1980, p. 11-53.

- “Stratification in the history of the Bantu languages.” *Sugia. Sprache und Geschichte in Afrika* (Hamburg), vol. III, 1981, p. 251-316.
- “Mbugu.” In Jungraithmayr, H; Möhlig, W. J. G. (eds.). *Lexikon der afrikanischen Sprachen und ihre Erforschung*, Berlin, Dietrich Reimer, 1983, p. 158 f.
- SEGUY, J., “La dialectométrie dans l’atlas linguistique de la Gascogne.” *Revue de linguistique romane*, vol. 37, 1973, p. 1-24.
- TUCKER, A. N.; BRYAN, M. A. “The ‘Mbugu’ anomaly.” *Bulletin of the School of Oriental and African Studies* (London), vol. 37, 1974, p. 188-207.
- WESTERMANN, D., *Die westlichen Sudansprachen und ihre Beziehungen zum Bantu*, Berlin, de Gruyter, 1927.
- WESTERMANN, D., BRYAN, M. A., *Languages of West Africa*, London, Oxford University Press, 1952 (Handbook of African Languages, Part II).
- WINTER, J. C. “Inter(natio)nal classifications of Kilimanjaro Bantu compared: towards an East African dialectometry.” In Guarisma, G.; Platiel S. (eds). *Dialectologie et comparatisme en Afrique Noire*, Paris, Sela, 1980, p. 101-132.