New Linguistic Evidence and 'the Bantu Expansion'
Author(s): J. Vansina
Source: *The Journal of African History*, Vol. 36, No. 2 (1995), pp. 173-195
Published by: Cambridge University Press
Stable URL: http://www.jstor.org/stable/182309
Accessed: 22/09/2008 09:46

Journal of African History, **36** (1995) pp. 173–195

NEW LINGUISTIC EVIDENCE AND 'THE BANTU EXPANSION'

BY J. VANSINA
University of Wisconsin-Madison

MOST historians believe in the Bantu expansion: the Bantu languages spread over a subcontinent as the result of a single continuous migration or 'expansion'. This migration, was fueled by a population explosion which was produced by the introduction of farming, and later metallurgy. The migration rolled like a giant 'wave', divided in several 'streams', over the subcontinent and soon overwhelmed the autochthonous foragers.[1] Once linguists had unraveled the genetic subclassification of Bantu, it would be easy to identify the various streams and routes the Bantu migration took and then plot them on a map. The recent completion of a massive comparative study in lexicostatistics designed to establish this subclassification reached some unexpected results, which require a total revision of the whole hypothesis.

This article is based on the overall results of that research, which was carried out in collaboration with Y. Bastin and A. Coupez (treatment of the linguistic data) and with M. Mann (statistical treatment). With the approval of these scholars this article presents my personal view of the overall results. Before I report on these findings, however, it is necessary first to recall some basic definitions and models used by linguists, for historians who are unfamiliar with the field of comparative historical linguistics.[2]

SOME DEFINITIONS AND MODELS IN HISTORICAL LINGUISTICS

Speech forms a system of communication based on the utterance of arbitrary signs. It can be efficient merely if the linguistic code of a speaker is similar enough to be understood by others, for each speaker has his or her own peculiarities of speech. A single person has an idiolect. Members of a community also share features of speech among themselves which differ from those in use elsewhere: they speak a dialect. A dialect community is

[1] I am grateful to Y. Bastin, A. Coupez, M. Mann, J. Miller and T. Spear, who have commented on an earlier draft of this paper. Current views of the dynamics involved are still based on the scenario developed by R. Oliver, 'The problem of the Bantu expansion', *J. Afr. Hist.*, VII (1966), 361–76. See also P. D. Curtin *et al.*, *African History* (Boston, 1978), 25–30.

[2] The linguists concerned are preparing their own publications to present and discuss the new evidence. The computer inputs and outputs can be consulted at the Royal Museum for Central Africa in Tervuren. This article discusses merely its historical implications. W. Lehmann, *Historical Linguistics: An Introduction* (New York, 1962), is still excellent for beginners. Among the more advanced texts, H. H. Hock, *Principles of Historical Linguistics* (Berlin, 1986), and Raimo Anttila, *An Introduction to Historical and Comparative Linguistics* (New York, 1972), who relies less on Indo-European examples than other textbooks do, are recommended.

often territorial but can also be based on class, gender or even occupation. A language is the collection of all dialects which are mutually understandable. A language is therefore a dialect continuum. This definition is not always easy to apply, because speakers of dialects at opposite ends of the territory of a language may not be able to understand each other even though all adjacent dialects in the continuum are mutually understandable. It is well known that English dialects are not all immediately mutually understandable. Moreover, speakers of dialects in contact do influence each other, and some dialect communities influence others more than these influence them. This influence is not a consequence of any superior quality of the language but results from either the demographic advantage or the superior prestige of the group who speaks it, whatever the source of that prestige; this is a subject studied by sociolinguists.[3]

Change in speech occurs continually, as is evident to anyone who has left their homeland for a few years. Changes start in idiolects. Some of these are then adopted by a household or a few colleagues, and later some of them appear in the speech of a wider community: they seep from idiolect into dialect. Thus over time a dialect tends to become more and more different from others. And that differentiation more and more hinders efficient communication between communities. Yet people need to understand each other. Hence dialects also borrow new features from each other and thus counteract the differentiating tendency to restore efficient communication. Differentiation and borrowing therefore occur simultaneously. But they do not cancel each other out. Over time, the relative difference in the frequency of communication within a single community and between communities favors differentiation. Eventually dialects do become mutually unintelligible and differentiate into new languages. The dynamics of the spread of language change then are entirely social: relative isolation fosters rapid differentiation, while constant communication slows it down. If Icelandic is the most 'different' of the North Germanic languages, it is because its speakers have been the most isolated.

Moreover one language may gain at the expense of another. Usually the speakers of the losing eroding linguistic community become bilingual in a second language of higher prestige, and eventually their descendants may abandon their original language. This process is called language shift. Once again, this results not from any intrinsic linguistic qualities but from social factors such as demographic trends, conquest, trade, ritual use, immigration and so on. Thus, under the influence of conquest, the Celtic-speaking inhabitants of Antwerp first shifted to Gallo-Roman, and a few centuries later they changed again to a Low German dialect, this time as the result of Frankish immigration.

Linguistics has developed two different models to study the history of languages. A tree model, as in Figure 1, uses the analogy of a family: daughter languages split from mother languages. Thus most English dictionaries show the history of English as a family tree. The technique assumes, however, that every language can have only a single ancestor.

[3] On sociolinguists, see Hock, *Principles*, 627–61. Once a written form of a language is created from one or more of its dialects, this becomes the standard and is often called the language, while all the oral forms are then dubbed dialects.

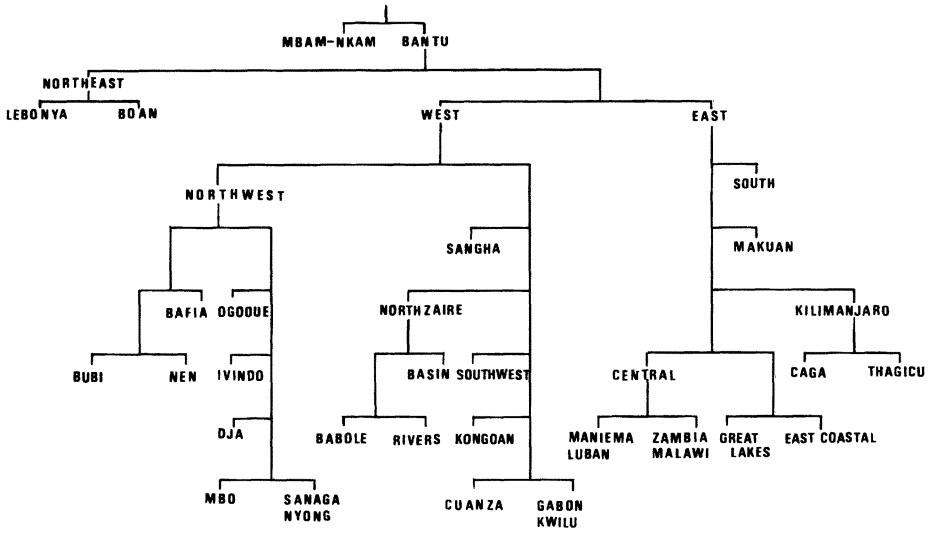

Fig. 1. Group average tree of the main subdivisions of Bantu.

Mixed languages with two ancestors are judged to be impossible. Because language is an arbitrary and fairly rigid system of communication, it is impossible to replace half of such a system at once by another one because the result would be unintelligible. Thus on a tree modern English stems from Old English, even though in fact Old English developed from a West Saxon dialect and modern English from a London dialect. Strictly speaking, Old English is not the direct ancestor of modern English. The family tree model conceals such multiplicity of origins.

The second model is the wave model. It starts from the observation that change spreads out from idiolects and dialects. It shows languages spoken side by side over a given area and portrays how changes introduced at one point spread outward like waves in a pond. Its focus on innovations allows it to show how changes in two or more different ancestral dialects can influence the production of a single new language: multiple parents become possible when the influencing languages are themselves quite closely related. In a wave model the history of Indo-European languages, for example, is shown as a series of overlapping circles representing greater and less similarity (Fig. 4). Such models illustrate how new languages differentiate within old dialect continuums.

The model one chooses influences the vocabulary one uses and the way one thinks about the social events which underlie language change. The tree model speaks of branches which split at nodes, while the wave model has clusters which differentiate. The tree model suggests that the daughter languages are isolated from each other and suggests group migrations away from the mother community as historical causes of change. At each node speakers of a branch leave those of the trunk to move away. One can locate such nodes on a map at the border between the present area of languages belonging to the branch and those which belong to the trunk. One then applies the principle called 'least moves' to plot the route taken by the

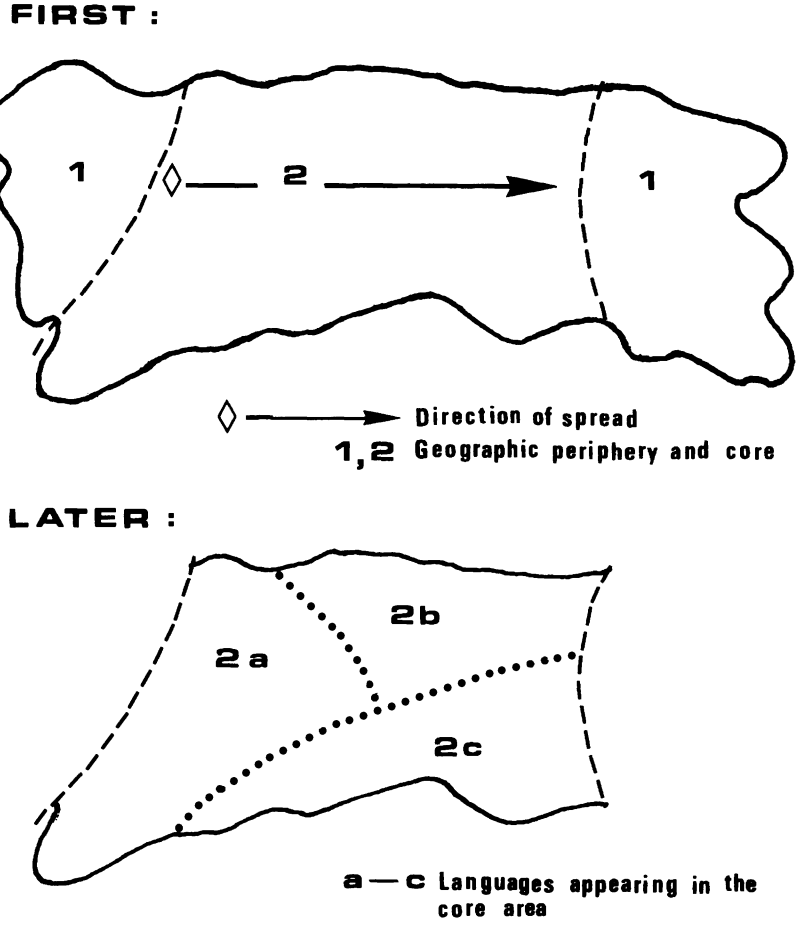

Fig. 2. Dialect differentiation (wave model).

speakers of the trunk, as they advance from node to node on the map. It is this procedure which allows us to locate the cradle of the original Bantu language and place it in westernmost Cameroun. It was a branch of a trunk whose numerous other branches are all located there. Bantu must have developed there, rather than, say, in Central Africa far to the south-east, because in the latter case one would have to explain how a dozen or so other branches had moved away from Central Africa to the north-west, all along implausible parallel routes.

Moreover, the tree model allows change within a stationary population only internally as idiolects change. The wave model does this by applying the principle of dialect differentiation. It explains how a stationary population will develop new languages, and it yields a very different representation of the movement of speakers. It holds that when speakers of a single language spread the increased distances between those who live at its peripheries and the bulk of speakers at the centre makes communication less frequent, compared to communication at the centre. Hence dialects at the periphery differentiate into languages first. Later, and without further population

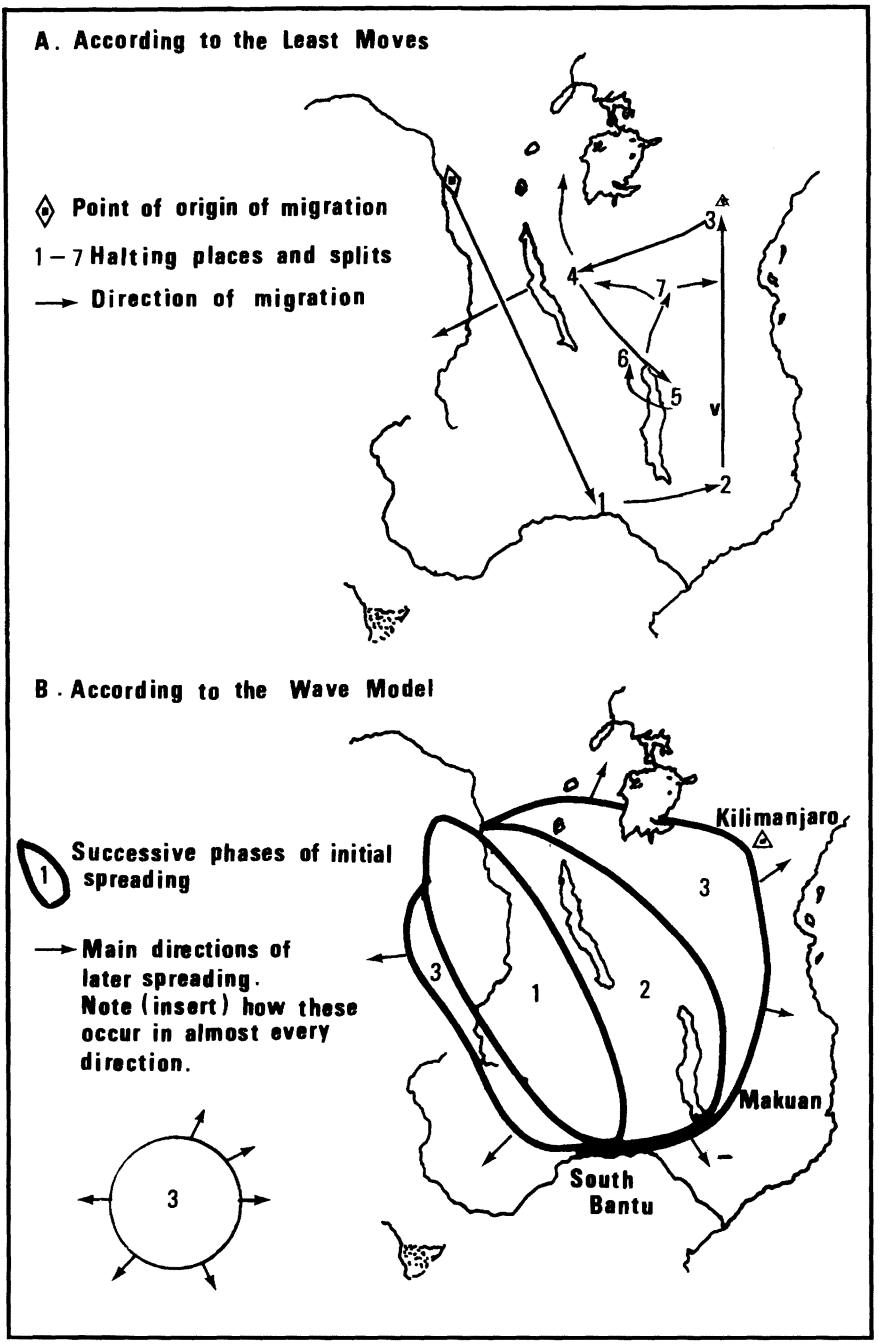

Fig. 3. The spreading of East Bantu: tree model vs wave model.

movement, the central or core area of dialects in the middle will eventually grow into new language areas as they eventually become mutually unintelligible. Figure 2 represents this model. At first the peripheral dialects (i) grow into languages. Later the core dialects (ii) follow.

While the principle of least moves is a valid guide to historical movements of people in some cases, it can be very misleading in others and should be used only as shorthand for the fuller wave model. The following comparison shows how different the results of the two models can be. Figure 3 (upper) plots the supposed migration of the bulk of East Bantu speakers from node to node according to the nodes of tree shown in Figure 1. The same data, even from this one tree, expressed in the wave model are then shown in Figure 3 (lower). They yield, first, various differentiations on the periphery and later a break up of the core area, as foreseen in Figure 2. The comparison is telling: the migration traced by the least moves model is patently absurd; the wave model is plausible. Incidentally, the results show that an original East Bantu spread first as a single language to cover much of the large area where its daughter languages are now spoken. That suggests that the spreading took several centuries at least during which the original language remained unified, which implies that even while expanding the communities speaking it maintained a high level of intercommunication with each other.

Language change can sometimes be documented directly from earlier written records or through irregularities within a single language, but in most cases changes are deduced from comparisons among modern languages.[4] Some among a number of genetically similar looking languages share more common features with each other than with others. Such features are innovations with respect to the whole number of them. When a set of similar looking languages is compared, a subset can be isolated on the basis of the similarity of its innovative features with regard to the rest. Some of these innovations presuppose earlier innovations in a wider subset, which is also identified through comparison with the whole set. Such successive innovations appear most clearly in phonetic change, that is, change in the qualities of the sounds used to produce words. This feature has consequently been used most frequently for setting up a succession of languages. Phonetic change is regular: it affects all (neogrammarian theory) or most words (other theories) in which the sound occurs. It usually affects a whole series of sounds at once, so that chance similarity can be ruled out. Finally, phonetic change is not totally arbitrary. Its direction is constrained by the physical features involved in making a sound. Thus it is very unlikely that a *b* made by rounding the lips will suddenly change to a *k* made at the back of the roof of the mouth, because in nearly all phonetic systems other phonemes are articulated at a number of other places between the lips and the back of the roof of the mouth. Finding regular correspondences across the phonemic inventory of the languages compared thus cannot be due to chance and proves a genetic link which is then further explored by looking at other phenomena. Documenting linguistic history by showing the sequences of phonetic, morphological, lexical and syntactic change which systematically explain the differences between the modern languages is the goal of all historical linguists.[5]

This procedure also leads to a reconstruction of most features of the now-lost ancestral language, including a portion of its vocabulary. Such

[4] Hence historical linguistics are often called comparative linguistics.

[5] D. Nurse and T. Hinnebusch, *Swahili and Sabaki: A Linguistic History* (Berkeley, 1993), is an excellent example of such an endeavor.

reconstructed languages are called 'proto' languages, and their words or morphemes are represented as starred forms because one always writes them out following a star (*).

A complete comparative study requires full descriptions for all the languages involved (and ideally for all their main dialects). With some 600 Bantu languages, for instance, this is a tall order. It will take generations before the raw data are assembled, if ever. To cope with this sort of situation Morris Swadesh developed a shortcut, through lexicostatistics applied to words. In studying words one must always remember that every word has both a meaning and an arbitrary utterance or 'form' which expresses the meaning. Thus the form *muntu* in most Bantu languages signifies the meaning 'person'. Swadesh knew that some words are more resistant to change than the general vocabulary. They designate basic concepts or actions and are used very frequently. In consequence their form is also more resistant to innovation. Only once in a while do people innovate new forms to express such old, core meanings. Out of the 600 Bantu languages, for instance, only a handful have dropped a variant of *muntu* for a new form such as, e.g., *oma* in Mpongwe (Gabon). Swadesh concentrated on such 'basic' words and by trial and error eventually developed a list of the one hundred words which were the most resistant to change. He also found that in the aggregate they changed at a rate that is constant over time, irrespective of language family.[6] It suffices to gather this list of words in the languages studied, compare each list with all others and express the comparison as a percentage of similarity. The figures can then be converted into a dendrogram, or family tree, using one of several possible statistical formulae.[7] One can then also define the difference between dialect and language by the number of similar forms, an index of retention. In Bantu lexicostatistical studies, for instance, that limit is set at 86 per cent similarity. Any higher percentage measures differences between dialects of the same language. Any lower number measures differences between languages.

However, because the comparison results in a tree, it shares the drawbacks of all tree models. Most linguists working in comparative Bantu studies have accepted the validity of lexicostatistics as a means to measure distance between languages and as a way to yield a preliminary genetic classification of the Bantu languages. But they are well aware of the drawbacks of the tree model. Moreover, because of the use of averages or medians to calculate relations between groups of languages the tree usually splits in only two branches, very rarely in three. It tends to render most situations in which three or more daughter languages appear as two or more successive splits of two branches each.

Swadesh, however, went further. Since the rate of change was constant in the aggregate, the index of retention could be converted into an absolute chronology. He developed a formula for this calculation and calibrated it against several language families for which approximate time depths were known: for instance, 30 per cent similarity between a present-day language and its ancestor comes out to 1860 B.C. This operation is called glotto-

[6] M. Swadesh, 'Towards greater accuracy in lexicostatistic dating', *International Journal of American Linguistics*, XXI (1955), 121–37, and earlier references there.

[7] A full discussion of lexicostatistical procedures is provided by S. M. Embleton, *Statistics in Historical Linguistics* (Bochum, 1986).

chronology. It has been applied to Bantu as well as to other languages. Most linguists however do not accept Swadesh's assumption that change in basic vocabulary is a constant, and hence they deny any validity to glottochronology, certainly over short runs of time. In addition, any single date attributed to language change remains problematic. Language differentiation is a process, not an event. Thus the separation of Dutch from German, for instance, lasted for a thousand years, and some dialects still 'belong' to both languages. Moreover the propensity of lexicostatistics to represent synchronic differentation as two close successive splits muddies the picture even more.

LEXICOSTATISTICS AND THE BANTU LANGUAGES

Successive attempts to reconstruct the genetic classification of the Bantu languages by means of lexicostatistics in the 1970s failed to produce reliable results, primarily because the resulting classification varied according to the size of the universe, i.e. the number of languages included in the comparison.[8] It was therefore necessary to obtain a very large sample of languages to obtain reliable results. By 1990 Y. Bastin and A. Coupez, affiliated with the Royal Museum for Central Africa (Tervuren, Belgium), who had started collecting data from informants in 1953, eventually had assembled 542 lists for 440 languages, out of some 600 total and established the etymology of the forms cited for each of the 92 items of the Swadesh list that can be applied to Bantu.[9] Because of the difficulties with the meaning of the other eight words, these cannot be used: for instance 'yellow' is expressed only by paraphrase in most Bantu languages. Given the inclusion of almost 86 per cent of all languages and the use of multiple trees discussed below, the resulting lexicostatistical classification can now be taken as nearly definitive. Future change stemming from the addition of new lists for languages which are not yet represented is expected not to affect the main divisions and subdivisions of Bantu.[10] Michael Mann, of the School of Oriental and

[8] Other factors involved are the exact list of rubrics used, the sources of the data and the counting procedures employed. The two published earlier attempts still in use are B. Heine, H. Hoff and R. Vossen, in 'Neuere Ergebnisse zur Territorialgeschichte der Bantu', *Zur Sprachgeschichte und Ethnohistorie in Afrika* (Berlin, 1977), 57–70 (on 147 languages and many lists drawn from dictionaries), and Y. Bastin, A. Coupez and B. de Halleux, 'Classification lexico-statistique des langues bantoues (214 relevés)', *Bulletin de l'Académie Royale des Sciences d'Outre-Mer*, XXVII (1981), 173–99. The Tervuren study produced an unpublished tree for western Bantu in 1983. The reason for the sensitivity of the results to the total number of languages included derives from the fact that the Bantu languages form a huge single pool of dialects. When only a fraction of the total number of languages is tested many places in the dialect network between the languages are unrepresented and the nature of the exact links between the languages tested may remain obscured.

[9] On the procedures to assess retention and the eight items on the Swadesh list that had to be discarded, Bastin, Coupez and de Halleux, 'Classification lexico-statistique des langues bantoues'.

[10] Most of the missing languages are spoken in southern Tanzania, south-eastern Angola and an adjacent area in Zambia between the Zambezi and Cuando Rivers. For southern Tanzania the comparative study of Derek Nurse allows one to foresee that their addition to this universe would probably introduce only minor and local adjustments. Cf. D. Nurse, 'The diachronic background to the language communities in southwestern Tanzania', *Sprache und Geschichte in Afrika*, IX (1988), 15–115. The situation in south-

African Studies in London, then programmed this huge volume of data for the computer. The resulting calculations produced two types of results. First the computer provided a table showing the distribution of each of the many forms which occur in the 542 lists to express the meanings of the 92 items included. These distributions allow one to reconstruct the proto-Bantu (starred) forms for each of these items and hence allow one to calculate how resistant to innovation (i.e. how 'basic') each individual item really is, compared to the others on the list.[11]

The computer also generated seven different family trees. They differ because of the different statistical procedures and assumptions used. Alternative assumptions include: (a) the use of either medians or averages to calculate percentages of cognation between groups of languages; (b) the recognition of smaller or larger groups as the units to be compared with each other in such calculations; and (c) assumptions involving the relative weight to be ascribed to the internal cohesion within an identified language group as opposed to the earliest appearance of a link between a single language within such a group and an outside language belonging to a different group.[12]

Nearly all previous publications had chosen only one and the same set of assumptions and procedures to establish a tree. They used averages rather than medians to arrive at single figures valid for clusters and assumed that languages were equally influenced by others inside the group and neighboring languages outside it. The 'group average' tree of this study is therefore the only one which is directly comparable to previous results and hence is represented here (Fig. 1). It divides Bantu into three main branches: north-east Zaire (called North Bantu), East and West Bantu. The figure also includes the major subdivisions in East and West Bantu but does not portray further subdivisions to the level of today's languages.

But is this the best among the seven trees? It is the most satisfactory when one compares all trees to the existing empirical classification elaborated by M. Guthrie and amended by E. Meeussen. That classification was based on a rough comparison of the major distinct phonological and morphological

---

western Zambia, known from unpublished data, is similar. More worrisome are the omissions of presumably Northern Bantu languages such as Homa (Sudan) and Kari, Boguru and Bali (north-east Zaire), for which only scanty and insufficient data are available.

[11] This in turn yields data about the relative stability of each of the 92 items compared to all the others.

[12] Assumptions, or the 'weights' given each possibility, involve priorities given either to the primacy of internal cohesion (connectivity) within identified language clusters (exclusivity) or to the primacy of connections between individual languages belonging to different clusters. Three figures are involved: (a) the lowest percentage of agreement within a cluster; (b) the percentage by which the cluster is linked to the most closely related cluster; and (c) the highest percentage at which any language within the cluster links up with a language outside of it. Figure (b) must always be lower than (a) and ideally higher than (c). In practice, borrowing from neighbors can render (c) higher than (a) or even (b), thus compromising the exclusivity from the cluster. In drawing a tree one must choose an intermediate between two extreme situations. These are the choice of not linking any cluster to any other before one has complete linkage, or total exclusivity (the statistical index of connectivity, VN = 100%) and the choice to link any language directly to its 'nearest neighbor' even if this neighbor does not belong to the same cluster (VN = 0%). The latter situation is called total connectivity.

features apparent in each of the groups they set up, i.e. on the intuitive experience of scholars about which differences were most likely to betray genetic links.[13] Yet a satisfactory correspondence does not justify the dismissal of all the other trees out of hand. A comparison between all trees soon shows that the three trees which are based on the most extreme assumptions, diverge markedly from the rest and also differ the most from the earlier empirical 'working' classification.

Eliminating them as implausible, however, still leaves four trees. A comparison between these shows that they all identify a number of nearly identical units, or blocks of languages, at various levels above individual languages. They also distinguish between an initial North Bantu branch and later West Bantu and East Bantu branches, and they attribute the same units to either West or East Bantu. But they diverge considerably in the attribution of units to North Bantu, in the sequence and numbers of splits, especially within West Bantu, in the arrangement of units, and in the attribution of some languages to one or another unit. Some trees do indeed elevate some individual languages to a position where they become a major branch by themselves, co-ordinate with West Bantu, East Bantu or even Bantu itself. Others with more extreme assumptions of connectivity or exclusivity eliminate a few units from Bantu altogether and link them with Bantoid languages into larger units of their own. This sensitivity of family trees to various statistical assumptions had hitherto escaped attention because all the earlier trees had been obtained by using the same 'group average' procedure.

The variability between all the trees is not random. A comparison between the four most plausible ones (i.e. the ones that used the less extreme assumptions) shows that differences in a sequence of splits are often due to the fact that the splits involved must have occurred almost simultaneously, because the percentage difference of retentions between each of the splits is very small. Take, for example, a split between A, B and C, shown in one tree as A + B and C first at 60 per cent and then C from B from 62 per cent, a second tree as A and B + C first at 60 per cent and B from A at 61 per cent and a third tree at A + C and B first at 60 per cent and then A from C at 62 per cent. The trees diverge because they cannot show simultaneity. A, B and C in fact split from each other roughly at the same time, in a three-way split. When different trees assign the same unit to a larger unit at the next highest level, the logic of this model forces a choice between only two larger units and always the same ones. One tree assigns A to a higher unit A + B and another to a higher unit A + C. No tree, however, assigns A to a higher unit A + X. Depending on the level of branching at which this arbitrary association occurs, the result can provoke substantial realignments in the overall shape of the tree.

These two types of differences derive from the lexicostatistical technique itself. The use of medians or averages for comparing groups with each other yields unrealistic numbers, unrealistic because they are calculated on a scale

[13] M. Guthrie, *Comparative Bantu* (4 vols.) (Farnborough, 1967–71). For the emendations, see Y. Bastin, *Bibliographie bantoue sélective* (Tervuren, 1975) (Musée Royal de l'Afrique Centrale: Archives d'Anthropologie 24), 1, and in the references there. For the linguistic criteria used, see M. Guthrie, *The Classification of the Bantu Languages* (London, 1948), and E. Meeussen, 'De talen van Maniema', *Kongo Overzee*, XIX (1953), 385–90.

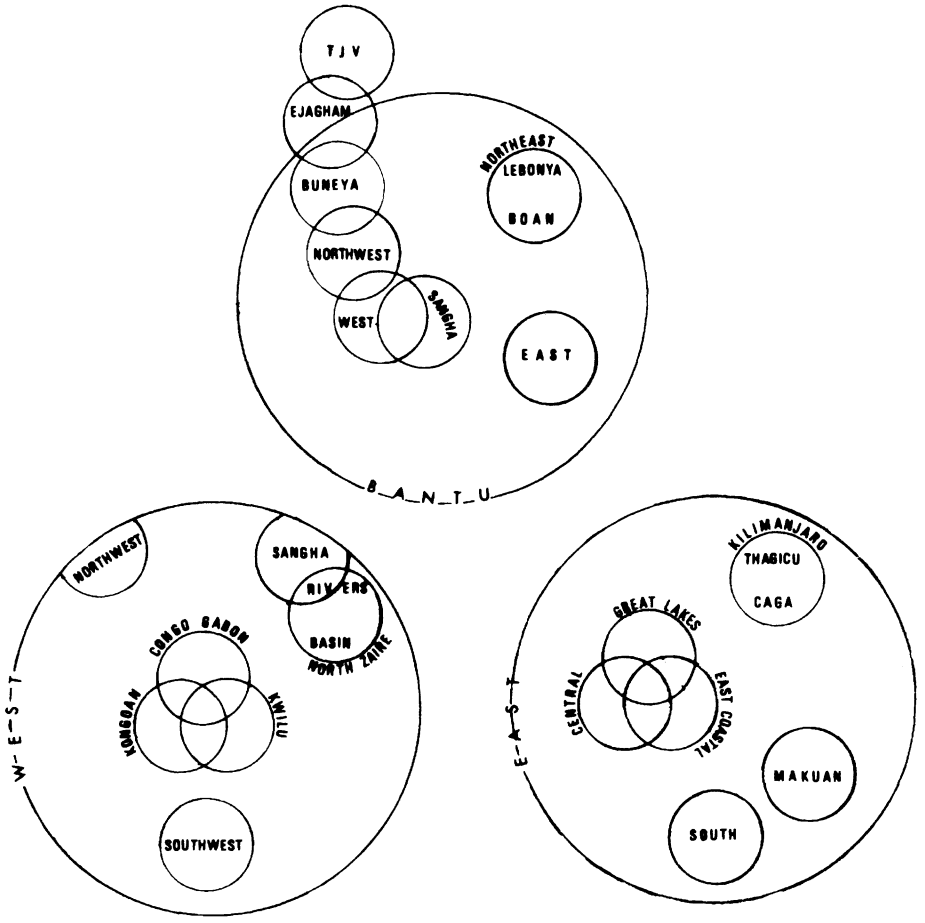

Fig. 4. Wave model of the main subdivisions of Bantu.

of two decimal places for five trees (0·01 per cent) and a quarter of a decimal place (0·4 per cent) for two trees – not on the real scale of the number of items available for each comparison, which was never higher than 92, usually in the high eighties and sometimes as low as 65.[14] The technique is therefore too sensitive. Such calculations will rarely yield percentages recognizable by the computer as identical, effectively linking three units together rather than two. By its very nature then the technique erroneously tends to show synchronic splits as sequences. In addition any tree model is inherently unable to attribute more than one genetic link to a unit. Each unit can have only one 'ancestor'.[15]

[14] M. Mann, personal information.

[15] In a situation where parents are dialects of the same language, mixed languages (*Mischsprachen*) are perfectly possible. To insist on a single ancestor becomes unrealistic when two neighboring related ancestral languages strongly influence the emerging new language. Such cases have often occurred in the Bantu linguistic continuum. One can come close to a mixed language situation even when one of the parent languages is very different from the other. The case of Mbugu in Tanzania is notorious. Cf. W. H. Whiteley, 'Linguistic hybris', *Afr. Studies*, XIX (1960), 95–7.

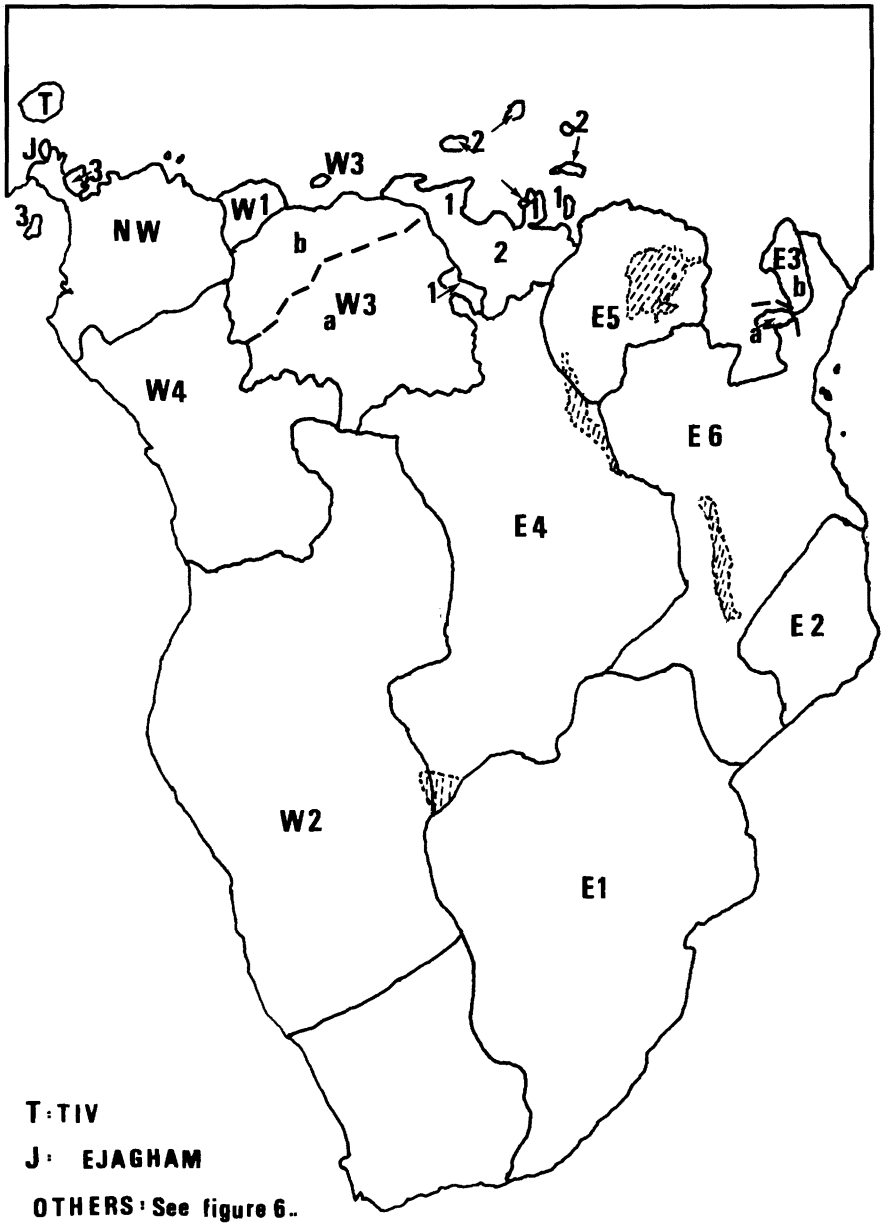

Fig. 5. The genetic classification of Bantu: map.

The data from all four trees can be used to set up a wave model which allows for multiple influences, as has been done in Figure 4. This figure is accompanied by a map (Fig. 5) showing the distribution of modern languages for each main subdivision and by a table of the subdivisions (Fig. 6). To establish this wave model, one first identified blocks which are identical in all trees from the level of individual languages to the level of East Bantu. Then one compared the different combinations in which they are presented,

I *The main subdivisions of Bantu*

1 Lebonya (Lengola, Bodo, Nyali)
2 Boan
3 Buneya (Bubi, Nen, Yambasan)
4 North-west [NW]
5 West [W]: subdivided into W1–W4
  W1 – Sangha
  W2 – South-west
  W3 – North Zaire: (a) Inner Basin; (b) Rivers
  W4 – West Coastal (Kongoan, Congo, Kwilu)
6 East [E]: subdivided into E1–E6
  E1 – South
  E2 – South-east
  E3 – Kilimanjaro: (a) Cagan; (b) Thagicu
  E4 – Central
  E5 – Great Lakes
  E6 – East Coastal

Note: This nomenclature is provisional. A commonly accepted standard remains to be developed.

II *Correspondences between 1-E6 and Bantu zones* (Guthrie, *Comparative Bantu*, and adjusted by the Tervuren group)

1 D12 (Lengola), D33 (Nyali, D35 (Bodo)
2 C42–45 (Boa group), D32 (Bira), D23 (Komo)
3 A31 (Bubi), A44 (Nen), A60 (Yambasan) and in some trees A50
4 A minus preceding, B10–B30; in some trees some B20 languages also in W4
5 W1 – No agreement on languages: some trees have C11 (Ngondi) – C12 (Pande), others C10 (Babole), C18 (Aka). Some of the latter place (C10, C18 in North Bantu
  W2 – K and R, L10, L22, L50, H41 but not K41, K42
  W3a – C33–C35, C50–C80, B82
  W3b – C31, C32, C36, C37, C41, C42; in some trees C11, C12
  W4 – B40–80, H, except for H41
6 E1 – S
  E2 – P30
  E3a – E60 (Chaga)
  E3b – E50 (Thagicu)
  E4 – D12–14, D20, L, M, K41, K42, M31 excepted; in some trees D25 (Lega) excepted; in some trees also N and D43
  E5 – Tervuren's J; Guthrie's D40–D60, E10–E40. In some trees also D25; in others D43 excepted
  E6 – E70, F, G, P10, P20, M31; in some trees N

Fig. 6. The genetic classification of Bantu.

including the percentages of retention proposed in each tree. These differences show the parameters of disagreement. Remembering that any tree is not intrinsically more likely to be more correct than any other, it follows that the disagreements have to do with the limitations of the model. By assigning simultaneity where required and multiple origins where indicated a single wave model results. The results can be compared with the tree shown in Figure 1, which is one of the four used in the wave model.

The comparison of the tree and wave models should make it evident that in this case the wave model is preferable to any tree, because the wave model does better justice to the data. It includes the information of all the plausible tree models, and it avoids the *a priori* limitations of the latter. The data, item by item, show that Bantu languages do form a huge dialect continuum, a fact which has profound implications for any interpretation of the dispersal of Bantu languages.

THE DISPERSAL OF BANTU LANGUAGES

The wave model shows a succession of three main phases in the dispersal of Bantu. As expected, within each phase dialects on the periphery grow into separate languages first while dialects in the core differentiate into languages later. In later phases the same process occurs over and over again within each of the territories of the new languages.

First, the original Bantu language became different from the Mbam-Nkam cluster of Bantoid languages in westernmost Cameroun. It then expanded *as a single language*, mostly eastwards towards the Great Lakes area with a secondary movement towards the lower Ogooué area, achieved in part by seagoing people.[16] Later, while the expansion was still going on, dialects at the opposite ends of the language continuum began to grow into languages of their own, no longer mutually understandable with the original language. This is very evident for the eastern part of the area. One suspects that this development began once the Bantu-speaking community had stretched eastwards of the great marsh and the inundated forests which divide Cameroon from northern Zaire. Henceforth constant communication between the easternmost lands occupied by Bantu speakers and the regions further west must have been difficult. Later, the core area occupied by the original language differentiated into several language clusters. Two among these were proto-West Bantu and proto-East Bantu. While East and West Bantu were differentiating from each other, a North-west cluster, still closer to West Bantu, was nevertheless already beginning to differentiate from West Bantu. At the same time, under the influence of nearby Bantoid clusters and of their relative isolation from the other dialects, a northern set of dialects within the North-west cluster itself had also started to differentiate from the other dialects within that group. No tree can possibly render such a situation of overlapping, near-simultaneous processes well. For instance, the tree of Figure 1, which sets up North-West Bantu as a first branch of West Bantu proper, is misleading. Only the wave model adequately represents this state of affairs. It follows from these data that proto-West Bantu was spoken in an area closer to the North-west cluster and the Atlantic Ocean than East Bantu was.

Next came a dispersal of both West and East Bantu. The proto-West Bantu language spread southwards at a time when the North-west cluster had not completed its differentiation from narrow West Bantu, i.e. West Bantu minus the North-west. This fact probably places the cradle of narrow

[16] This section is based on a detailed study of the massive new data available at Tervuren. The situation of Bubi, a language of the northern division of North-west Bantu on the island of Bioko, makes this clear. The position of Seki (Muni estuary) on the various trees can also be explained only by seaborne travel.

West Bantu somewhere in northern Congo. While this language was expanding, the dialects at the periphery of its territory once again grew into full-fledged languages. The earliest differentiation occurred on the northern edge of West Bantu territory in the inundated forests and swamps between the Sangha and Ubangi rivers. A single language ancestral to only one or a couple of modern languages was involved,[17] but the different trees disagree as to which ones they were. The next differentiation occurred at the southern periphery while narrow West Bantu was still expanding. There the ancestor of the languages spoken in the dry woodlands of south-western Zaire, Angola and Namibia appeared. Later the remaining common language at the centre also split into three. But this is not the whole story. It remains unclear when in this sequence the dialect ancestral to the languages spoken today in the rainforests in central and northern Zaire attained language status. According to the tree in Figure 1 it did *before* the ancestor of the south-western languages did. According to the other trees, the differentiation occurred *after* that event but before the core broke up into three languages. This ancestral language itself consisted of two blocks of dialects which, during their spreading, immediately turned into two languages themselves as their speakers became relatively isolated from each other. One dialect (Rivers) was the ancestor to the languages now spoken in the vicinity of the Zaire River itself; the other became the ancestor spoken in the whole interior equatorial basin of the Zaire River. The divergence between trees in placing this cluster is connected with its spatial position. It is peripheral with regard to the total north–south spread of West Bantu, but it is also the starting point for two separate diffusions eastwards beyond the lower Ubangi and the Congo/Zaire River: one north-eastwards along the main rivers and one due east into the interior Zaire basin proper. Moreover, the area bordered on the lands occupied by the earlier Sangha-Ubangi language, and the trees do detect influences from this source.

Meanwhile, proto-East Bantu also began to expand from the rainforests towards south-east Africa. The ancestor of the southern cluster separated probably somewhere near the middle Zambezi river. It was followed by proto-Makua in northern Mozambique. Then at the opposite end of the East Bantu area a Kilimanjaro cluster appeared, including Caga and Thagicu. Finally, the core area itself broke up into three units: East Central Africa (Shaba and Zambia), the Great Lakes area and eastern East Africa.

After the initial differentiations of Bantu the same process of fission was repeated within each main subdivision. The outermost dialects became languages before the core dialects did. This sequence has occurred again and again within each successive language until today. Some ancient languages have died out without posterity, however, so that in some cases the only survivor of an ancient cluster today appears to a single language. It then seems as if a single language differentiated early on and later was no longer subject to further dialect differentation. But that is only an illusion caused by lacunae.

The dynamics of differentiation also show new languages growing out of

[17] Some trees include only Babole and Aka and even have Babole and Aka split before West Bantu splits off from East Bantu. Others include only Ngondi-Pande and place Babole-Aka in a Northern Zaire cluster.

two different adjacent parent languages at once. The modern languages from Malawi are a good instance of such dual origins. They are the offspring of both the languages of Zambia and the Yao-Mpoto cluster in neighboring areas of Tanzania and Mozambique. Some trees portray their ancestor as a child of Zambian languages, while others link it to Yao-Mpoto. The language probably developed in the border area between the two clusters. Such complex situations were very common. On Figure 4 they are indicated by three overlapping circles. Notice how in the North Bantu diagram all the western clusters are thus linked in a single chain which actually extends into Bantoid. The Bubi-Nen-Yambasan unit, which belongs to the North-west cluster, is also linked to the Bantoid (not Bantu) Ejagham, which themselves are also linked to Bantoid Tiv. Such phenomena have been responsible for the difficulties in unraveling the genetic classification of Bantu and for the failure of single tree models.

In general the dynamics of language differentiation in Bantu do not support a scenario of 'migration' along specific routes. An application of the principle of 'least moves' yields absurd results in nearly all Bantu contexts. An example is illustrated by Figure 3. The supposed migration route of East Bantu traced by least moves yielding the following East Bantu speakers moved first from the Lualaba south of Kisangani to the middle Zambezi. There they shed southern Bantu languages. They then moved to northern Mozambique to leave Makuan before turning northwards to Kilimanjaro, where they left the Kilimanjaro languages. They then went westwards towards the vicinity of Kigoma, where they split up in three bodies which then traveled each in different directions: northwards (Great Lakes), west, south-westwards (East Central Africa) and south-eastwards (eastern East Africa)... and reoccupied territories already crossed by earlier migrants. The easternmost group then moved from Kigoma to the Upper Rovuma to shed the Yao-Mpoto group of languages, then turned around to the northern tip of Lake Malawi, where the migrants left the Nyakyusa group before reaching the Masai steppe and dividing there into one group that set off for the east coast and one into the plains of western Tanzania. In contrast, the wave model elegantly explains the sequence of differentiations, first at the periphery and later at the core, but it does not yield data about specific routes. To speculate about routes leads to unwarranted assumptions. Thus the usual view that East Bantu first spread into the Great Lakes area is not the most likely historical hypothesis. East Bantu may well have spread southwards west of lake Tanganyika, or on both sides of that lake, before it ever reached the Great Lakes area. One can say with confidence only that the East Bantu language spread from the rainforests of Zaire somewhere east of the Ubangi/Zaire confluence towards the middle Zambezi. In rare cases, however, the correlation of certain geographical features such as coasts, rivers or particular habitats with language distributions is suggestive, as we saw in the case of the diffusion of North Zairian languages into the rainforests along two different networks of rivers. The correlation between each network and a particular cluster of languages suggests that the languages spread by the major river routes.

When did the Bantu dispersals occur? The successive differentiation of languages yields a relative time sequence which has been used in the exposition of the diffusions. With regard to absolute chronology, one can use

indications from proto-Bantu vocabulary and later from the proto-West and East Bantu vocabularies and relate them to archaeological data. Because proto-Bantu attests to the making of pottery and the cultivation of rootcrops in fields, the first Bantu dispersal must have occurred after pottery and farming had been introduced. Archaeological evidence suggests that this situation first occurs near the Bantu cradle area from the fourth millennium onwards, and specifically around 3000 B.C., for the Bamenda highlands. The expansion of proto-Bantu therefore began after 3000 B.C. But proto-Bantu speakers neither smelted nor used metals, and neither did proto-West Bantu or proto-East Bantu speakers. The first sign of metallurgy in the Great Lakes region, now an East Bantu-speaking area, has been dated to *c.* 800 B.C.[18] It follows that by then the original North Bantu expansion was over and that the West and East Bantu dispersals had begun.

WAS THERE A BANTU EXPANSION?

Once the main outlines of the initial dispersal of the Bantu languages according to the new findings are known, one must ask what inferences can be drawn about the history of the speakers of these languages? Do the new findings support current views about the 'Bantu expansion'?

Scholars have been mesmerized by the huge extent of the present distribution of Bantu languages and could think only of a single process, an equally huge human migration, 'the Bantu expansion', to explain it. They interpreted the earlier and partial tree diagrams in this light and, using the principle of least moves, traced migration routes on the map. Tentative glottochronological dates lent verisimilitude to the whole. The unvarying success of these Bantu migrants came to be attributed to a vast technological differential: they were sedentary, they were potters, they were farmers, and later metallurgists, while the autochthons were just nomadic foragers. The migration was supposedly triggered by an overpopulation in the Benue–Cros area, itself the consequence of an immigration from the Sahara into West Africa caused by the desiccation of the Sahara *c.* 2500 B.C. This view was enthusiastically endorsed by most archaeologists, whose own interpretations became inextricably mixed up with this scenario.[19]

The scenario is fatally flawed however for two reasons. First it fell prey to the illusion that only human migration could fit the data. This is similar to the optical illusion of the child who throws a stone in the pond and watches the ripples spread. The water seems to move as the ripples spread out from the point of impact. Yet it does not. The water molecules are agitated up and down but stay in place: there is no flow of water. Similarly a language (ripple) can spread without involving the migration (flow) of any communities (water). The second fatal error was to collapse a history which encompassed the developments of one to several millennia into a single migration event.

[18] A. B. Stahl, 'Intensification in the west African Late Stone Age: a view from central Ghana', in T. Shaw *et al.* (eds.), *The Archaeology of Africa: Food, Metals and Towns* (New York, 1993), 261–73; A. Holl, 'Cameroun', in R. Lanfranchi and B. Clist (eds.), *Aux origines de l'Afrique centrale* (Libreville, 1991), 150–2.

[19] As in the syntheses of David W. Phillipson. Cf. his *The Later Prehistory of Eastern and Southern Africa* (New York, 1977), 210–30, and *African Archaeology* (Cambridge, 1983), 179–81.

The data show that many different discrete dispersals of single languages succeeded each other at different times, not continuously. Many stones have been thrown into different parts of our analogical pond.

Today this once-persuasive migration hypothesis is totally discredited, not only by the new linguistic data but also by the accumulation of archaeological evidence. It has become clear that there is no empirical support for a sudden huge emigration from the Sahara into the lands south of the Benue.[20] By itself the development of root-crop farming did not lead to a population explosion, if only because root-crop farming furnished only a small part of the food the speakers of proto-Bantu needed. The ways in which they obtained the bulk of their food supply were nearly indistinguishable from those of the autochthons.[21] The supposed technological superiority, especially of the earliest Bantu speakers, over the autochthons has all but vanished. Nor were all autochthons necessarily peripatetic nomads. The fishing folk among them were probably nearly as sedentary as Bantu villagers were. Moreover, autochthons may well have borrowed various elements of the technological package, one at a time as needed, which accounts for the presence of pottery on some otherwise typical Late Stone Age sites. Such a dynamic renders a hypothetical technological superiority of Bantu speakers even more moot.[22] After all, the spread of root-crop agriculture within West Africa or of grain farming within north-east and northern East Africa are not attributed to vast human migrations but to the diffusion of the technologies involved. In sum, there is no longer a cogent argument in favor of a Bantu migration.

How then should the new linguistic evidence be understood? There are successive spreads of a single language followed by differentiation into new languages, some of which later spread again. There are therefore many successive Bantu dispersals rather than a single continuous expansion. There is first an expansion of proto-Bantu, followed by one of a North-eastern language. Then follow simultaneously at first an initial North-western language spread and a narrow West Bantu one. Then comes a South-western language spread. Meanwhile, proto-East Bantu spread, then Southern

[20] Migrations from the Sahara into the Niger bend area and westwards are very well documented, but despite archaeological research no such evidence turns up in northern or central Nigeria.

[21] Jan Vansina, *Paths in the Rainforests* (Madison, 1990), 71–100. The proto vocabulary which documents this (pp. 267–301) is not based just on what is now known to be West Bantu but includes the north-eastern and the north-western clusters of North Bantu. It therefore refers to the period when North Bantu was spreading and before narrow West Bantu emerged. While the assumption that agriculture leads to spurts in population increase seems to be supported in some cases (e.g. in the lower Nile valley) it is not inevitably so and should be proven rather than assumed.

[22] B. Clist in Lanfranchi and Clist (eds.), *Aux origines*, 181–3, 225, 226, shows that new technologies, including metallurgy, diffused only very slowly. In Gabon, foragers, 'neolithic' food producers and iron users lived side by side for over four hundred years, which shows how relatively unimportant the innovations were. The presence of ceramics on Late Stone Age sites in Gabon, and especially on a site near Libreville dated to *c.* 3000 B.C. or even 4500 B.C. suggests an independent diffusion of ceramics among foragers. Clist also found that by 4200 B.C. foragers crossed the 25-km.-wide estuary of the Gabon river to fetch basalt stone (*ibid.* 166–7), which presupposes navigation equipment. Elsewhere in the Bantu-speaking area there are other cases of separate diffusions for pottery, herding, the farming of grain crops and metallurgy.

Bantu, Makuan and Kilimanjaro. That is a minimum of nine diffusions, at least, before communities speaking Bantu languages could be found here and there over most of the area now occupied by Bantu speakers, and those nine diffusions together may have taken up to two millennia to spread. Many other spreads follow later but involve internal colonization within the overall area.

These nine spreads were successive, but that does not mean that spreading was a single continuous process. Each spread involved different dynamics, and each had its own chronology. The assumption of a single large-scale migration by the original speakers of Bantu is extremely unlikely. In each case of spreading there are many possible alternatives between the extreme of large-scale immigration and the other extreme of language shift without any population movement at all, a scenario that is not more likely than migration. In the case of language shift the autochthons usually became first bilingual and then lost their original language only several generations later. This view implies that the communities using the Bantu language held either a decisive demographic advantage or enjoyed a huge prestige over the autochthon language or languages. Where does such prestige stem from? The technological differential, especially in early times, was small. Did Bantu-speakers have a monopoly on some form of trade? Were they conquerors, or somehow superior in religious matters? There is no evidence so far to support any of these suppositions, for the early dispersals at least. Much more complex hypotheses than the extreme assumptions of population replacement or language shift must be considered.

The only direct evidence that can help comes from comparison of the proto vocabularies of the languages that were in contact. While these can be reconstructed for various ancient Bantu languages, most if not all traces of the other, pre-Bantu languages have disappeared, at least outside of Eastern and Southern Africa. Thus one cannot often directly compare autochthonous communities with those of Bantu speakers. All one can do is to imagine the process which best fits the available proto-language on the Bantu side.[23]

As an exercise let us present a plausible scenario for the dispersal of proto Bantu itself.[24] Early Bantu speakers were organized in large households, several of which lived together in a village while various ties linked households and villages together in districts, the largest social unit, each district comprising perhaps 500 persons when its constituent villages had grown to full size. The area of original village settlement abutted on much less populated lands in the rainforests to the south and east, and the original Bantu speakers already were in contact with camps of nomadic foragers there. A few heads of households found out that there were patches of much

[23] In theory, careful archaeological excavation of both farming and contemporary foraging settlements should show the relevant differences. Even so, one may not retrieve many of the relevant but small differences between foragers and food producers.

[24] The scarcity of genetic data coupled to an *a priori* acceptance of the massive Bantu migration hypothesis renders the biological evidence inconclusive on this point. Cf. L. Excoffier *et al.*, 'Genetics and history of sub-Saharan Africa', *Yearbook of Physical Anthropology*, XXX (1987), 151–94, and especially 181–5; J. D. Irish, 'Biological affinities of Late Pleistocene through modern African aboriginal populations: the dental evidence' (Ph.D. thesis, Arizona State University, 1993). The scanty data would just as easily support the tentative scenario sketched here.

more fertile land, say up to 50 kilometers from their village, and decided to move their village. Or competition between such leaders may have led unsuccessful ones to move away. Or again, aspiring leaders may have gathered their followers and moved out. Leaders were soon followed, a few other villages were founded near the first settlers, and a district took shape. The district was in fact the unit of migration, although a migration was triggered by a single village. The villagers also made pottery and planted root crops. They were more sedentary than most of the earlier autochthons, and their settlements were larger. The maintenance of a single language over long periods of time implies that regular relations between the original districts and their colonies were kept up even beyond the district. In the new habitats of the immigrants the demographic size of their society, the district, is assumed to have been far larger than that of most autochthons, their societies being perhaps not much larger than a few camps of about thirty people each. The prestige of the Bantu language flowed more from their sedentarity and the size of their social organization than from their technology. To foragers a Bantu village was a large and exciting place. It easily became a stable point of reference for a large region all around it, a point often visited by foragers. Eventually many autochthons became bilingual as a result. Some among them may also have settled in the villages, perhaps as spouses, perhaps as hangers-on in the households of village leaders. The number of Bantu speakers thus rose much faster by co-optation than by natural demographic increase, for the input of farming was still too small to lead to great increases in natality. Still greater sedentarity was a factor which may well have favored a somewhat faster increase of population among villagers than among nomadic foragers. This effect became greater and greater as the generations passed, and in time the original languages died out.

Meanwhile tensions arose once again in the new Bantu-speaking settlements, either between different houses or between leaders. Again they were resolved by renewed emigration of Bantu speakers, this time accompanied by groups of the autochthons who had attached themselves to Bantu houses. Indeed, some groups may have consisted wholly of Bantuized autochthons. Be this as it may, eventually more and more migrants were physical descendants of former autochthons.[25] Hence the input of even acculturated autochthons into the life-style of early Bantu communities may have been far greater than has hitherto been acknowledged. In this scenario factors inherent to the social structure of early Bantu-speaking societies are responsible for the first dispersal of their language.

However, one must remember that there exists a strong probability that some autochthon settlements at least were perhaps as sedentary and as large as the Bantu villages were. One suspects this comparability especially in the case of fishing folk along the coast or along major rivers such as the Zaire itself. Such communities would retain their own speech, appropriate new technologies as they saw fit and hold up the diffusion of the Bantu language. But they were by-passed and eventually found themselves surrounded by Bantu speakers, whose numbers continued to increase faster over the generations by absorbing autochthon foragers in the hinterland. Finally, centuries later, the more sedentary autochthon settlements were also over-

[25] Vansina, *Paths*, 83–94, for the evidence which underlies the scenario sketched here.

whelmed and adopted Bantu speech. In this process the spread of the proto-Bantu language is then pictured as occurring in irregular patches which remained in good communication with each other. The diffusion was sometimes halted, but sooner or later it started again. It probably included reverses as well as advances. The final outcome, however, was the kind of spreading that linguistics documents. The whole conversion involved centuries rather than decades, and the largest numbers of agents involved were descendants of autochthons themselves. Thus major population replacements of autochthons by immigrant Bantu speakers may never have occurred.

This tentative scenario cannot explain all the numerous subsequent dispersals of Bantu languages later on. Relative local advantages in demography, in trade, in local resources, in plant lore, in technology, e.g. navigation and later metallurgy, even in military organization, or as a result of the perceived efficacy of Bantu-speaking healers or ritualists, any of these may have hastened or hindered the dispersion of a Bantu language. In some cases a Bantu language spread into territory already occupied by another Bantu language. It is known for instance that East Bantu speakers influenced all the West Bantu languages south of the rainforests (without replacing them) and introduced grain agriculture there. In some cases the spread of one Bantu language provoked language shifts from another Bantu language.[26] In some circumstances Bantu-speaking communities were handicapped relative to their predecessors. In Angola, for instance, the environments were so hostile to immigrants from the rainforests and so favorable to the autochthon foragers that the latter's role in the spread of South-western Bantu must have been significantly greater than has been pictured in the initial general model.[27] In the Great Lakes area a Bantu language spread into an environment where the autochthons were efficient farmers of grain crops and herders, both with extensive social organizations of their own, and a different sequence is again required to explain how Bantu languages eventually became dominant here.[28]

Nor has Bantu speech always been triumphant. There are well-known cases of language shift whereby Bantu speakers adopted languages of other families, and this may have been more frequent in the past than we now think. There are even Bantu languages which show such heavy non-Bantu

[26] Several well known cases of language shift occurred in the last millennium, e.g. the spreading of Fang or Ntumu, in Gabon. But there are indications that north-western languages lost ground to those of the narrow West Bantu core as early as the first centuries before our era in the coastal areas of southern Gabon and Congo, and it is evident that even earlier language shifts must have occurred from North to West Bantu in the region of the great bend of the Zaire river.

[27] In most of Angola neither domesticated yams nor oil palm trees will grow. Trapping is much less efficient there than hunting, and foragers relied heavily on the latter. Immigrant farmers from further north could settle in only a few localities very far apart from each other and had to be taught new hunting techniques by the autochthons.

[28] C. Ehret, in numerous publications since 1967. For a recent synthesis, see D. Schoenbrun, 'We are what we eat: ancient agriculture between the Great Lakes', *J. Afr. Hist.*, XXXIV (1993), 1–31, and references to C. Ehret, 9, n. 30.

influences that they possibly may result from language shifts first from Bantu to non-Bantu and later back from non-Bantu to Bantu.[29]

CONCLUSION

This tentative picture of the earliest dispersal of proto-Bantu is useful mainly because it contrasts so strongly with the 'natural' view that massive immigration was involved. It points to the need for careful archaeological excavations, not just of village sites, as has been usual to date, but also of foraging sites. Finally the scenario dissociates various technological innovations from each other and from a necessary connection with language spread, which implies that archaeologists should no longer extrapolate from the presence of one innovation to the presence of others.

Above all, the tentative scenario underlines the fact that the spreading of the Bantu languages was not a single, inevitable or continuous process and that, while straightforward migration may indeed have occurred in some cases, this must have been exceptional. To have believed in a vast migration now strikes one as quite far-fetched. Moreover, the tentative scenario allows one to question what hitherto were established verities. Thus it is still commonly held that the archaeological record for eastern and southern Africa tallies with the spreading of East Bantu languages and is proof of a vast migration. This idea is nonsense. The linguistic data show unambiguously that East Bantu spread as a single language from the rainforests to the middle Zambezi. Yet on the basis of the comparison of ceramic styles the archaeologists posit a Bantu migration southwards from the Great Lakes. In doing so they ignore the 'precursor' appearance of ceramics and/or cattle in Zimbabwe or near the upper middle Zambezi. They gloss over the fact that there are no fewer than three different spikes of diffusion southwards from southern Tanzania, respectively from the coast, from northernmost Malawi and from north-east Zambia, rather than a single 'East Bantu migration'. They ignore that while metallurgy was known in all three cases, this technology was unknown to East Bantu speakers.[30] The supposed continuity between all the ceramic styles involved in the three cases to make them part of one migration may well be illusory.[31] Even if they were, and despite claims to the contrary, one cannot automatically equate a ceramic style with a language.[32]

[29] Mangbele is an example of a Bantu language that died out as late as the nineteenth century through language shift in favor of a Mangbetu language. The Bergdamara and Coroka of Namibia and Angola may have abandoned a Bantu language for Khoi and San languages over a millennium ago.

[30] Proto East Bantu does not have a terminology for iron working. The distribution of the various forms for 'iron', 'iron ore', 'to forge', 'smith' and 'hammer' in the area support various interpretations but not a single origin. The question cries out for further study.

[31] D. P. Collett, 'The spread of early ironproducing communities in eastern and southern Africa' (Ph.D. thesis, University of Cambridge, 1988), concluded that the various ceramic styles cannot be grouped together into a single stylistic tradition.

[32] Despite the claims made, especially by T. Huffman, that language, group identity and ceramic style in small-scale societies correspond (cf. T. Huffman, *Iron Age Migrations: The Ceramic Sequence in Southern Zambia* [Johannesburg, 1989], 5–9), ceramic style and language are not always correlated. For instance, although all Bushong speak the

There never was a single Bantu migration, even if one calls it 'expansion'. Once this idea is banished it becomes possible to study almost from scratch a host of different historical phenomena which have occurred in different places of the subcontinent during the millennia, which had been collapsed into a single 'Bantu expansion'. But archaeological research and research in classical comparative linguistics, including vocabulary studies, will lead to a radical new and more realistic vision of the early history of the subcontinent.

SUMMARY

New linguistic evidence about the classification of the Bantu languages does not support the current view that these languages spread as the result of a massive migration or 'expansion' by its speakers. Rather the present geographic distribution of Bantu languages is the outcome of many complex historical dynamics involving successive dispersals of individual languages over a time span of millennia and involving reversals as well as successes. This is as true for eastern and southern Africa, where a close correlation between the archaeological evidence documenting the diffusion of basic food-related technologies, including metallurgy and the spreading of Bantu languages has become an axiom, as it is elsewhere. The linguistic evidence concerning the dispersal of Bantu languages in these regions of Africa is completely incongruent with the archaeological record. The existing Bantu expansion hypothesis must be totally abandoned. The scrapping of the hypothesis will make room for more realistic and quite different interpretations and research hypotheses. For example, it follows that the local or regional contribution of speakers of other languages, autochthons and others, to the development of later cultures and societies was probably considerably greater than has hitherto been acknowledged and that the continuities in historical dynamics of all sorts between the Bantu-speaking parts of Africa and areas further north and west are greater than has been hitherto realized.

---

same language the northern Bushong and their northern neighbors (who speak other languages) used household pottery made on the banks of the Sankuru river, whilst southern Bushong and their southern neighbors (who speak other languages) used household pottery of a different style made on the banks of the Lulua river.