

THE GENETIC/HISTORICAL RELATIONSHIP OF FIVE BANTU
LANGUAGES OF EAST AFRICA¹

Casmir M. Rubagumya

1. Introduction

The aim of this research paper is to work out the genetic/historical relationship of the following Bantu languages: Kinyarwanda, Kihangaza, Kihaya, Kikerewe, and Kichagga (Kimachame).

So far, several linguists have tried to classify African languages in general. Others have paid particular attention to Bantu languages. In this paper, our interest is in the five languages mentioned above, which are spoken in East Africa. (For the purposes of this paper, Rwanda will be considered as part of East Africa)

Guthrie is one of the linguists who have done a lot of work on Bantu languages. In one of his works (Guthrie, 1948) the five languages that we are considering are classified thus:

<u>Zone D</u>	<u>Zone E</u>
D.61 Kinyarwanda	E. 22 Kihaya
D.65 Kihangaza	E. 24 Kikerewe
	E. 62 Kichagga

The implication of this classification is that Kihaya and Kikerewe are closer to Kichagga (i.e. more related to it) than they are to Kinyarwanda and Kihangaza. In my opinion, Guthrie's classification of these languages is misleading. I must hasten to add that I am aware that Guthrie's classification was geographically oriented. All the same, it gives the impression that Chagga/Haya-Kerewe are close genetically. It is this misleading impression that this paper sets out to correct.

By his own admission, Guthrie points out; "Zone D (in his classification) is of little linguistic significance..... Apart from group 60, our knowledge of languages of this zone is so fragmentary that even the grouping is in most cases very tentative". (Guthrie, 1948:40).

Here there is no dispute about the close relationship between Kinyarwanda and Kihangaza, as I intend to show later. What is disputed is when Guthrie says: "There are reasons for not placing any of these groups (of Zone D) in the neighbouring zones." (Guthrie 1948:40). Unfortunately, he does not tell us these reasons. What, one wonders, are the reasons for not placing group 60 of zone D (which includes Kinyarwanda and Kihangaza) in zone E (groups 10 and 20)?

My view is that Kinyarwanda and Kihangaza as a group are closer to Kihaya and Kikerewe (as another group) linguistically than Kichagga is to any of the two groups. In other words, I am arguing that Guthrie should have placed Kinyarwanda, Kihangaza, Kihaya and Kikerewe in the same zone, and Kichagga in a different zone.²

2. Historical Evidence

There is historical evidence (itself being partly based on linguistic evidence) to give weight to this view. Paradoxically, this evidence is based on Guthrie's theory! Although the question of the origin and the general direction of Bantu migrations is still a controversial one, several historians agree that it is very probable that the Bantu 'nucleus' was in the Katanga region. Basing his argument on Guthrie's theory of "common Bantu roots", Oliver (1966) reconstructs a map of likely stages of Bantu expansion (see Appendix).

In his discussion of the migrations of the Eastern Bantu speaking peoples, Kimambo (1974:196) argues:

In the case of Eastern Bantu speaking peoples, a south-north direction of expansion is demonstrated by the available linguistic, documentary and archaeological evidence... When all this evidence is put together, it logically indicates the possibility of Bantu speaking people expanding quickly not only along the coast but also along the highland areas immediately bordering the dry plain.

These two pieces of evidence taken together give us a fairly clear picture of the relationship of the languages under discussion.³ Several observations can be made:

- (a) It can be deduced from the evidence of the two historians that Kichagga is remotely related to the four other languages. They broke from each other during the second stage of migration, but Kichagga took a longer time to evolve as a language (4th stage).
- (b) Kirundi, according to Guthrie, has 44% of the common Bantu roots and evolved during the third stage of dispersion. From my own intuition as a native speaker of Kinyarwanda, I would suggest that Kinyarwanda and Kihangaza share a common development with Kirundi because they are very similar in terms of lexical items and syntax. In fact some people would go as far as suggesting that these three are dialects of the same language.
- (c) From Oliver's map, it seems that Kihaya/Kikerewe evolved from the same parent language as Kinyarwanda/Kihangaza/Kirundi, but at a later date. That is to say, while during the second stage of migration Kichagga had already broken away from

the other languages, the two other subgroups (Kinyarwanda/Kihangaza; Kihaya/Kikerewe) did not break from each other until the 3rd or 4th stage.

Commenting on Guthrie's classification, Nurse (1980) say: "Guthrie separates Western Highland (i.e. Kinyarwanda, Kihangaza, etc.) from Rutara (i.e. Kikerewe, Kihaya, etc.) - thus implicitly rejecting the notion of a lacustrine group including them - and puts them with other languages mainly spoken in Zaire. We cannot deny or support this idea, for we did not deal with data from these languages". (p.40) But, as pointed out earlier, Guthrie has not given us any reason why Western Highlands should be separated from Rutara. In fact, by his own theory of Bantu migrations, we have seen that these two sets of languages are most likely related to one another.

3. Methodology.

A word-list (100 words of 'core' vocabulary adapted from Swadesh: see Lehman 1973:108) was drawn up. Native speakers of the languages under consideration, were approached and asked to give (orally) the equivalent of the English gloss in their own language. To minimise error, at least two informants were consulted for each language, and their responses were cross-checked. Except for Kichagga where one dialect (Kimachame) was chosen, I did not take into consideration the question of dialects. For Kichagga, I had to choose one dialect because it is said that Kichagga dialects are very different from one another, sometimes to the point of mutual incomprehensibility. There was no criterion for choosing Kimachame instead of any other dialect; it was randomly picked because linguistically no dialect is superior to others or more representative of "the language".

In this research, the method of lexicostatistics has been used. Nurse (1976:1) defines this method thus:

The method of lexical comparison, known as lexicostatistics, is only interested in items which look alike because they are genetically connected and can therefore be derived from a single item in a hypothetical (or real) proto-language. When comparing two languages, the linguist takes a meaning and elicits the word of the meaning in the two languages. After eliciting enough information, he is then able to work out regular sound correspondences between the two languages.

Nurse has used a five-point scale to determine sound correspondence. Given the limited scope of this paper, I have judged Nurse's scale too complicated and time-consuming. Instead, a simpler scale has been used, where two lexical items are either in regular correspondence or similar (1 point), or they are not in regular correspondence (0 points). Below are some examples:

(i)

	<u>Kinyarwanda</u>	<u>Kichagga (Machame)</u>
ear	ugutwi	ya'wi
head	umutwe	m'we
heart	umutima	m'ima
five	-tanu	'anu

From the example above, /t/ in Kinyarwanda corresponds to /t/ in Kimachame. These pairs of items are therefore given 1 point each in the analysis of the data.

(ii)

	<u>Kinyarwanda</u>	<u>Kihangaza</u>
person	umunhu	umunhu
bird	inyoni	inyoni
child	umwana	umwana

The above pairs of items are similar (i.e. identical in sound and meaning). They are also given 1 point each.

(iii)

	<u>Kikerewe</u>	<u>Kimachame</u>
bird	enoni	ndeye
dog	embwa	kite
fly(n.)	ensohela	nghi

The above pairs of items are not in regular correspondence in any way, nor are they similar in sound. They are therefore given 0 points each in the analysis.

Since we have a 100 word-list, the points of each set of two languages are added to find out the percentage of relationship between those two languages. These percentages are then set in a table. Following Nurse's method, the two languages which share the highest percentage are collapsed into one column and this column is compared with each of the other languages. This "collapsing" continues each time taking the next highest percentage until we have the language which is the least similar to others being compared to them as a group. From these figures we can draw a geneological tree, which will in turn allow us to suggest a classification of the languages under consideration.

4. Findings and Discussion.

Findings from the data are presented below, in a table showing the percentages of relationship of languages studied.

(i) Nyarwanda

- 89 ---- Hangaza
- 54 ---- 56 ---- Haya
- 49 ---- 48 ---- 82 ---- Kerewe
- 35 33 ---- 35 ---- 30 ---- Chagga

The next step, as explained in the methodology, is to collapse two languages with the highest percentage and compare them with the other languages. In this case then, we will collapse Kinyarwanda and Kihangaza and tabulate their relationship to other languages thus:

(ii) Ny./Han.

55	-----	Haya			
48.5	-----	82	-----	Kerewe	
34	-----	35	-----	30	-----
				Chagga	

The next highest percentage is between Kihaya and Kikerewe. So, again we collapse them into one column and compare them with Kinyarwanda/Kihangaza on the one hand, and with Kichagga on the other. The table below emerges from the operation:

(iii) Ny./Han

51.75	-----	Haya/Ker
34	-----	32.5
		Chagga

This stage of our findings is very crucial to the assumption of this paper. We have argued right from the beginning that Guthrie's classification was misleading because it gives the impression that Kichagga is closer linguistically to Haya/Kerewe than Nyarwanda/Hangaza are. Now, from (iii) above, we can see that the degree of relationship between Chagga and Haya/Kerewe is 32.5%. On the other hand, the degree of relationship between Haya/Kerewe and Nyarwanda/Hangaza is 51.75%. This supports our hypothesis that Nyarwanda/Hangaza and Haya/Kerewe should be classified together as a distinct group from Chagga.

The next step is to collapse Nyarwanda/Hangaza and Haya/Kerewe into one column and compare them with Chagga.

(iv) Ny./Han./Hay./Ker.

33.25	-----	Chagga
-------	-------	--------

2. Nyarwanda/Hangaza

(i) Nyarwanda

(ii) Hangaza

5. Conclusion.

This has been a very modest attempt to look anew at Guthrie's classification of Bantu languages. Needless to say, the research has several limitations. The first limitation is the very nature and scope of this paper. It took Guthrie years of field research - I suppose - to arrive at his conclusions. It would be very pretentious of me to claim that I am in a better position to arrive at 'correct' conclusions, given time constraints within which research leading to this paper was done.

That apart, there are also limitations of methodology. Again, because of the restricted scope of the study, a 100 wordlist has been used. But as Batibo (1980:2) rightly argues, "a 100 word-list, however carefully compiled does not give out reliable results where historical interactions are so deeply and intricately embedded, and where factors like semantic change, ideophonic use, taboo words inhibition and chance elements are apt to distort the results".

The problem of determining which word are cognates must also be pointed out. Because I have used only a two-point scale, some kind of arbitrariness was (in a few cases) inevitable in deciding whether to give a point or no point to the lexical items.

Despite these limitations, I think that this study has all the same achieved something. By using linguistic and historical evidence, it has arrived at a classification of five Bantu languages, and has tried to show why this classification is considered

to have corrected the wrong impression given by Guthrie's classification.

There is so much disagreement in the field of African languages classification even among prominent linguists. As such factors like experience and time taken to do research do not necessarily rule out error. In this respect, I would suggest that I have the advantage of being a native speaker of one of the languages I have investigated. (Guthrie did not have that advantage. Note, for instance, a basic factual error he makes; he says Kinyakisaka is a dialect of Kihaya, when in fact it is a dialect of Kinyarwanda. (See Guthrie 1948:42).

One area which, in my opinion, needs immediate attention is the study of languages spoken in Zaire (especially eastern Zaire). Both Guthrie and Nurse indicate that they know very little about these languages; that is why they hesitate when it comes to classifying the 'Western Highlands' group (i.e. Kinyarwanda, Kihangaza, etc). Should this group be classified with languages spoken in eastern Zaire as Guthrie has suggested, or should it be included in the interlacustrine group of languages as Nurse suggests (and as I think it should)? This question can be settled if the languages in eastern Zaire are studied in more details.

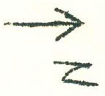
Another pertinent question concerns the notion of 'language' as opposed to 'dialect'. Are Kinyarwanda and Kihangaza different languages, or dialects of the same language? Is there such a thing as Kichagga language, or do we have different languages (i.e. Kivunjo, Kimachame, etc)?

These are some of the areas which I feel need to be looked into by those interested in Bantu linguistics.

Notes

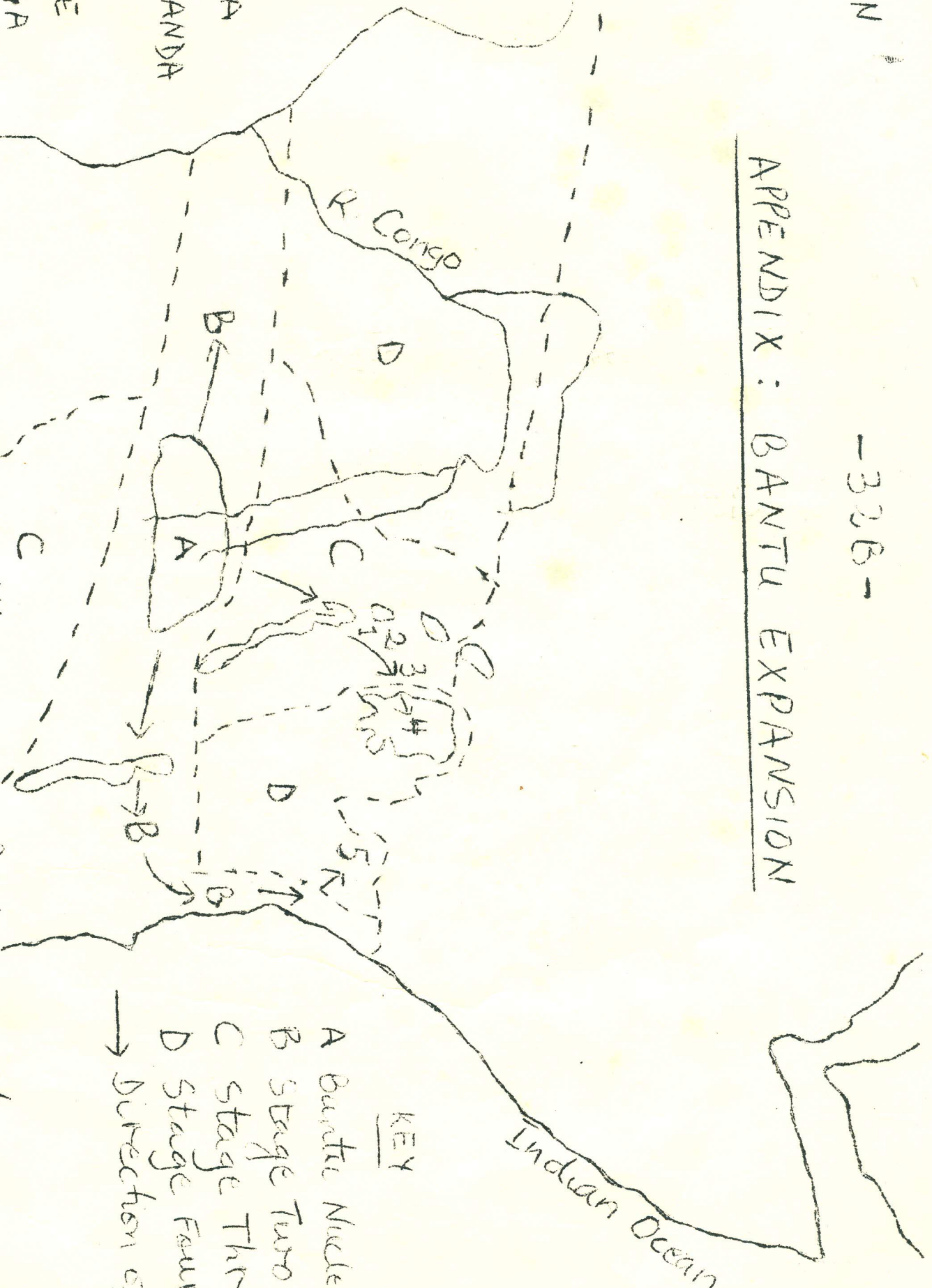
1. This paper was originally presented in 1982 as an undergraduate research paper (for LL 300). The present version ~~has~~ been revised and up-dated.
2. The Tervuren school (linguists like Meeussen, Coupez, etc.) classify Kinyarwanda, Kirundi, ~~Kichagga~~, Kikerewe and Kihaya under the same group, J, and Kichagga under a separate group, E. (See Bastin, 1975). This would seem to the author of this paper a more plausible classification than Guthrie's. Guthrie's classification is rather arbitrary, and it is only suitable as a referential classification rather than a genetic or typological one.
3. The discussion of Bantu migration as part of this study may seem far-fetched as the results of lexicostatistical studies show typological rather than genetic relationships. The linguistic closeness is not necessarily a reflection of genetic relationship, except where closeness is very high (Batibo, personal communication). However, the historical evidence presented would seem to reinforce the linguistic evidence in showing the relationships of the languages under discussion.

APPENDIX: BANTU EXPANSION



Atlantic Ocean

- 1 HANGAZA
- 2 NYARWANDA
- 3 HAYA
- 4 KEREME
- 5 CHAGGA



Source: Adapted from Oliver R. "The Problems of Bantu Expansion" in Journal of African History Vol VII

KEY

- A Bantu Nucleus
- B Stage Two
- C Stage Three
- D Stage Four
- Direction of Expansion

References

Bastin, Y. (1975) Bibliographie Bantoue Selective
Musée Royal de l'Afrique Centrale, Tervuren,
Belgique.

Batibo, H.M. (1980) Review of D. Nurse "Description of
Sample Bantu Languages of Tanzania" in Journal
of African Languages and Linguistics, Vol.2 no.2.

Batibo, H.M. (1985) "Some linguistic contribution to the
Cultural History of the people of Great Lakes" in
Journal of Language and Linguistics, Vol.2.

Guthrie, M. (1948) The Classification of Bantu Languages
Oxford University Press.

Kimambo, I.N. (1974) "The Eastern Bantu Peoples" in B.A.
Ogot (Ed.) Zamani, E.A.P.H.

Lehman, (1973) Historical Linguistics
Holt, Rinehart and Winston Inc.

Nurse, D. (1976) "Prehistory and Suggested map of the
Bantu Languages of East Africa" Mimeo.

Oliver, R. (1966) "The Problem of Bantu Expansion"
in Journal of African History Vol.VII.