

CLASSIFICATION OF THE LANGUAGES OF CAMEROON AND EQUATORIAL GUINEA ON THE BASIS OF LEXICOSTATISTICS AND MUTUAL INTELLIGIBILITY

S. Beban Chumbow
University of Yaounde I
Gratien G. Atindogbe
University of Buea
Engelbert Domche
University of Dschang
Dieudonne Martin Luther Bot
University of Douala

ABSTRACT This work clusters genetically related speech forms in Cameroon and Equatorial Guinea, and determines to which speech forms within these clusters are sufficiently and mutually intelligible to be grouped together in order to ease harmonization and standardization. The analysis of the linguistic situation in the two countries revealed that while clustering the speech forms on the basis of genetic relations via lexicostatistics has been quite fruitful, clustering on the basis of mutual intelligibility of at least 85% does not seem to significantly reduce the number of speech forms. Intelligibility surveys and testing have not been carried out in many of the clusters. However, it is important to continue the exercise in order to ascertain the exact situation.

Key Words: Cameroon; Equatorial Guinea; Classification; Lexicostatistics; Genetic relation; mutual intelligibility.

INTRODUCTION

This work is carried out within the context of the project on “Harmonization and Standardization of African Languages” initiated by the Centre for Advanced Studies of African Society (CASAS), Cape Town, South Africa. The overall objective of the CASAS research project is “to cluster African speech forms into sets which display an 85% level of intercomprehension” as a first step to the development, based on the economics of scale, of large literate communities. More specifically, CASAS is of the view that given evidence of a high level of mutual intelligibility between several speech forms, it should be possible to harmonize the standardization of those languages, i.e. to “develop orthographies which can be accessed by various dialects in a cluster without attempting to tamper with the phonological divergences of the various dialects.”

Another important goal of the CASAS project is to correct the myth of a multiplicity of over one thousand five hundred languages for Africa, since figures advanced particularly by such authors as Greenberg (1970), Gregerson (1977), Heine (1993) and Brown (1995) among others, appear not only astro-

nomically high, but were also not always based on reliable research findings.

Our task was to examine the linguistic situation in Cameroon and Equatorial Guinea within the framework of the methodology adopted by the network of researchers for the CASAS project to arrive at clusters of genetically related speech forms and determine which speech forms within these clusters are sufficiently mutually intelligible to be grouped together for the ultimate goal of the CASAS Project: harmonization and standardization.

RESEARCH METHOD

The take-off point of the study is the listing of languages or speech forms of Cameroon and Equatorial Guinea in Grimes (1992, 1996), and Gordon (2005).

For Cameroon, the listing in Grimes (1992, 1996, 2002) and Gordon (2005) is checked against the information contained in Dieu & Renaud (1983). This led to a list of the languages in an alphabetical order. On the basis of information from Dieu & Renaud (1983) and our own field work, we established a list of the speech forms in terms of the genetic groups or phyla.

We grouped the languages in terms of geographical contiguity and genetic relatedness instead of alphabetical order. More importantly, information is provided on the location of each language in terms of provinces (which in Cameroon are essentially geographical descriptions such as West, South, North, North-West, etc.) as well as on the population of speakers and the year of relevant statistics (see the Appendices).

Languages identified in Dieu & Renaud (1983), Grimes (1992, 1996, 2002) and Gordon (2005) served as the basis of further field work where numerous researchers were assigned languages of specific genetic groups to come up with language clusters showing varying degrees of genetic relations as well as obtain information and data on intelligibility among the speech forms of the genetic group. The results are shown in the Appendices, where languages of Cameroon are organized in clusters and subclusters within each language family.

The field work methodology applied to this last stage consisted of the following.

I. Lexicostatistics

Data elicited on speech forms (or languages) in a particular group on the basis of a word list of 50 words typical of the Swadesh word list are recorded and transcribed. On the basis of observed phonetic similarity or phonetic distance the cognate values between various speech forms are calculated in some cases by WORDSURV (a computer program designed by Wimbish (1989), but mostly by standard statistical procedures to obtain a cognate matrix for each cluster.

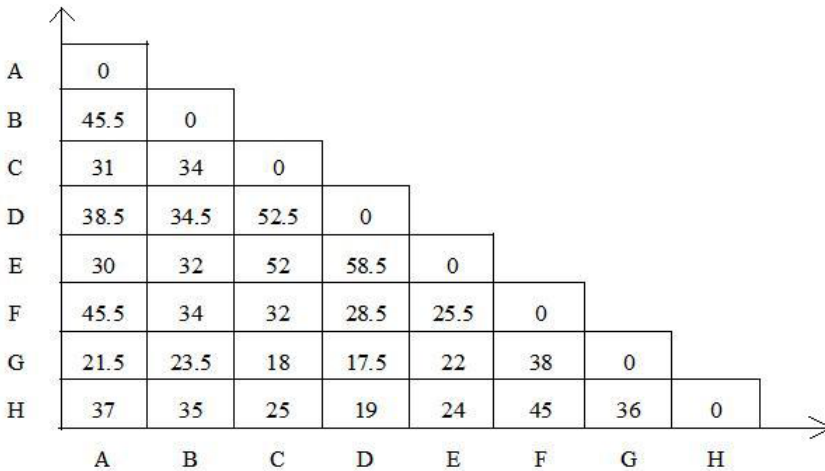


Fig. 1. Matrix of cognate values of the central Bamileke cluster.
 (The capital letters correspond to the languages identified in the text.)

Thus for instance, in the Wide Grassfield Bantu subgroup (Gordon, 2005) of the Niger-Kordofanian phylum, the Niger-Congo subphylum, the Atlantic-Congo family, the Volta-Congo subfamily, the Benue-Congo branch, the Bantoid sub-branch, there are the following Bamileke languages (among others):

- A. Ngombale B. Yemba C. Ghomala? D. Fe?Fe?
- E. Nda?nda? F. Gomba G. Megaka H. Ngyemboong

The cognate matrix established by our research following the lexicostatistic method is shown in Fig 1. These figures are a good indication of genetic relationship considering that only fifty words were used whereas the standard lexi-costatistic method uses more than double this number. It is also possible from this to determine the degree of relationship between each speech form by tracing the genetic tree indicating the depth of relationship between close neighbors as in Fig 2.

This means that from their common ancestor postulated as X, the eight speech forms branched out into two with Megaka constituting one group (Y), and Ghomala?, Nda?nda?, Fe?Fe?, Ngombale, Yemba, Gomba and Ngyemboong constituting the other (Z). Within the second group, Ngyemboong is very separated from the others which are more closely related.

On the basis of the same principles, the genetic relations between the languages of the groups can be captured. The tables in Appendices summarize these relationships in identifying clusters and subclusters within each linguistic subgroup.

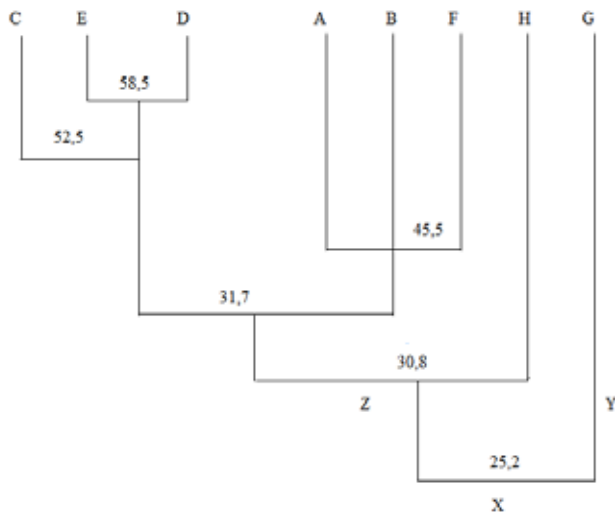


Fig. 2. Genetic relations among central Bamileke languages, see text for the meanings of the capital alphabet notation.

II. Intelligibility survey

The intelligibility survey questionnaires we designed were administered to at least three to five native speakers of each of the speech forms in the group or subgroup. Those who did not speak the language well were eliminated systematically from responses to specific questions of the questionnaire. The average score of bona fide respondents to specific questions of the intelligibility rating was retained as the mutual intelligibility score between the various speech forms or languages. These scores were then used to determine the Intelligible Survey Matrix (ISM) for the languages within the group. This exercise was expected to be followed by an actual intelligibility test. Following this method, the Intelligibility Survey Matrix for the speech forms of the Momo cluster, for instance, is presented as Table 1 below. (For the Momo cluster of languages, see 7 in Appendix 3).

On the basis of the ISM of the Momo cluster, we concluded that mutual intelligibility existed between Meta? and Moghamo because the intelligibility rating was significantly high and therefore Moghamo could be considered a dialect of Meta? and standardized as such. The other speech forms are, however, too distant from Meta? on the intelligibility scale to be considered dialects of the same language.

Table 1. ISM for the Momo cluster

comprehension of	by speakers of	average
Meta?	Moghamo	85 - 95%
Moghamo	Meta?	70 - 84%
Ngoshie	Meta?	50 - 69%
Ngie	Meta?	below 50%
Ngwo	Meta?	below 50%

III. Intelligibility Testing

Intelligibility testing is recommended for situations where information from intelligibility surveys is not conclusive or indicates a relatively good but not significant level of intelligibility (i.e. between 70-80%).

An abridged form of Casad's (1974) intelligibility test is used to verify intelligibility survey reports. The test, also known as the "Recorded Text Test (RTT)," has been perfected and extensively used by the linguistic survey teams of the Summer Institute of Linguistics (SIL) worldwide. The general principles of the RTT are as follows:

1. Two stories are recorded in two speech forms presumed to have a reasonable degree of intelligibility. One of the texts is about three minutes and the other about one minute.
2. The stories must be neutral in nature and character.
 - Stories are personal, i.e. actual experiences lived and not folklore or history.
 - Stories must not be such that the outcome is predictable and be guessed without linguistic inference or comprehension.
3. A translation of the recorded texts is made into the language to be used by the researcher/surveyor (English, French or some other language of wider communication).
4. Questions to elicit comprehension are formulated and interspersed throughout the story.
 - The questions are expected to be distributed evenly throughout the text.
 - They are expected to be varied in form so as to test different domains and levels of comprehension.
 - Yes/no questions are avoided.
 - Questions are limited to eliciting information explicitly stated in the story and not implied information.

5. Pre-testing of the RTT is conducted with about five speakers to eliminate potentially ambiguous and unproductive questions.

6. Questions are translated into each of the speech forms to be tested for mutual intelligibility.

7. In the testing exercise itself, efforts are made to ensure that only inherent intelligibility is tested and not acquired intelligibility, which is attributed to contact and indicative of bilingualism rather than mutual intelligibility. There are two ways of achieving this.

- (1) Testing begins with speakers of the dominant speech form(s), where this is discernable, being tested for their comprehension of the subordinate speech form under the assumption that generally, the dominant is less likely to acquire the subordinate dialect than vice versa.
- (2) Children are used to test comprehension of neighboring speech forms under the assumption that all things being equal, children are less likely to have had sufficient contacts with these other speech forms than adults and are therefore ipso facto free from the stigma of acquired intelligibility.

The standard deviation (SD) of all scores for the tested form may also be used to determine inherent intelligibility on empirical grounds and enables the researcher to determine whether the scores represent a learning process (acquired intelligibility) or inherent intelligibility.

8. After scoring the responses, the results represent a percentage of intelligibility between speech forms. The following needs to be noted:

- An intelligibility score of 85% and above between speech forms is considered adequate for the two speech forms to be standardized as one language for purposes of literacy.
- An intelligibility score of 70% to 84% is considered critical but not sufficient for standardization unless other sociolinguistic and extra linguistic factors are found to be crucial determinants.⁽¹⁾

It has been determined on empirical grounds that lexicostatistic scores (indicating lexical similarity) are related to mutual intelligibility. Thus, SIL in its surveys in Cameroon has established that if the percentage of cognates between two speech forms is 70% and above, they are likely to have a high degree of mutual intelligibility of the type that makes them dialects of the same language. If the cognacy score is less than 70%, it is most unlikely that the two speech forms can have an acceptable degree of mutual intelligibility to qualify as dialects of the same language (Starr & Stalder, 1990). However, it is not the case that a lexicostatistic score of 70% and above automatically qualifies the speech forms as dialects of the same language. The intelligibility test (IT) by means of the RTT must be administered to be sure.

What this means is that no two speech forms can be dialects of the same language unless they have a high degree of lexical similarity. However, two speech forms with a high degree of lexical similarity are not necessarily dialects of the same language because other relevant conditions (such as morpho-syntactic similarity, etc.) may not be met.

An Intelligibility Testing Matrix (ITM) for the languages of the Momo cluster (Table 1) was established by administering the RTT in the manner described above.⁽²⁾ The results are presented in Table 2.

Table 2. Intelligibility Testing Matrix (ITM) for speech forms of the Momo cluster

comprehension of	by speakers of	average
Meta?	Moghamo	91%
Moghamo	Meta?	76%
Ngoishie	Meta?	34%
Ngie	Meta?	19%
Ngwo	Meta?	0%

A comparison of the results of the intelligibility survey (Table 1) and the intelligibility test (Table 2) as evidenced by the ISM and ITM of these languages shows a striking positive correlation.⁽³⁾ The ITM proves conclusively that Moghamo is a dialect of Meta? and the two can be standardized as one language. The other speech forms are too distant from Meta? to be dialects of the language.

THE LANGUAGES OF CAMEROON ORGANIZED IN CLUSTERS

On the basis of the lexicostatistic analysis of related speech forms the languages of Cameroon are organized in clusters within each language family, sub-family, branch, etc. The detailed lists of language families, clusters and languages within each cluster are provided in the Appendices.

In this section, we summarize the language families, the name, the clusters and the number of languages within the cluster.

I. The Chadic Family (Afro-Asiatic)

There are 58 languages in addition to an extinct speech form, Zumaya. These languages are clustered into thirteen units (with many of them further subdivided into subclusters).

1. West 1
2. Margi 3
3. Gbwata 7
4. Daba 5
5. Wandala 8
6. Mafa 19

7.	Yedina	1
8.	Mandage	6
9.	Midaʔa	2
10.	Musgu	1
11.	Kada	1
12.	Masana	4
13.	Kera	1

Within each cluster shown in the Appendices, subclusters are indicated in columns.

II. The Adamawa Family (Niger-Congo)

There are 37 languages plus four extinct speech forms (Gey, Duli, Oblo, Nimbari) and two secret languages with no mother tongue speakers (Laʔbi, To). These are organized into five clusters as follows.

1.	Leko-Nimbari	5
2.	Koo-Dii	11
3.	Mbum	15
4.	Fali	2
5.	West Ubangian	4

III. Benue-Congo family (Niger-Congo)

There are 167 speech forms in this family, which can be divided into branches and groups as follows.

i) Non Bantu: 20 languages divided into four clusters

1.	Jukunoid	9
2.	Cross River	2
3.	Bendi	1
4.	Mambiloid	8

ii) Wide Bantu (Bantoid): 29 languages divided into five clusters

1.	Jarawan (including nearly extinct Ngong and Nagumi)	3
2.	Tivoid	12
3.	Ekoid	1
4.	Nyang	3
5.	Beboid	12

iii) Bantu Grassfield (Bantoid): 60 languages subdivided into seven clusters

1.	Momo	10
2.	Menchum	1
3.	Ring	15
4.	Ngemba	6
5.	Central Bamileke	10

- | | | |
|----|-------|---|
| 6. | Noun | 9 |
| 7. | North | 9 |

iv) Mbam Bantu (Narrow Bantu): 14 languages divided into six clusters

- | | | |
|----|--|---|
| 1. | Tikar | 2 |
| 2. | Tunen | 3 |
| 3. | Yambeta | 1 |
| 4. | Bati | 1 |
| 5. | Yambassa | 3 |
| 6. | Sanaga (including one dead language, Leti) | 5 |

v) Equatorial Bantu (Narrow Bantu): 44 languages and five clusters

- | | | |
|----|-----------------|----|
| 1. | Bafia | 4 |
| 2. | Coastal or Sawa | 15 |
| 3. | Basaa-Beti | 12 |
| 4. | Meka | 10 |
| 5. | Kako | 3 |

INTELLIGIBILITY TESTING AMONG CAMEROON SPEECH FORMS

In this section, results of attempts to determine the intelligibility between related speech forms in the various language clusters of Cameroon will be summarized. Generally, this exercise sought to answer the following questions: Are identified speech forms (through our own research and that of Dieu & Renaud (1983)) distinct languages or can some be shown to be dialects of the same language? In the latter case the endeavor will reduce the number of speech forms for which different and varying literature must be developed.⁽⁴⁾

I. Recent Intelligibility Survey Reports

Here we report results of our research based essentially on intelligibility surveys. Specifically, the report concerns only clusters we surveyed. Results of surveys on the same and other clusters by other researchers are presented in Section 4.2.

i) Chadic Family

Clusters ⁽⁵⁾		Intelligibility Survey Results
Margi	3 speech forms	Negative: separate languages
Daba	5 speech forms	Negative: separate languages
Wandala	8 speech forms	Negative: separate languages
Mafa	19 speech forms	The intelligibility is largely negative.

However, RTT is recommended for those pairs of speech forms where 2 out of 5 respondents claimed intelligibility. These are:

Mafa	-	Mefele
Mafa	-	Cuvok
Mofu South	-	Dugwor
Muyang	-	Melokwo
Mofu North	-	Merey

ii) Adamawa Subfamily

Clusters		Intelligibility Survey Results
Leko-Nimbari	4 speech forms	Negative: separate languages
Koo-Dii	11 speech forms	Negative: separate languages
Mbum	15 speech forms	Negative: separate languages
Fali	2 speech forms	Positive: one language
West-Ubangian	4 speech forms	Negative: separate languages

Note for Fali, that Dieu and Renaud (1983) had identified 6 speech forms subdivided into two groups: Fali North (3 speech forms) and Fali South (3 speech forms). Grimes (1992, 1996, 2002) and Gordon (2005) simply recognized these as two languages, Fali North and Fali South, with some dialects. However, there is evidence of widespread intelligibility between the two. SIL research actually proposes one standard Fali form for all varieties. This standard should be based on Fali Tinguelin dialect.

iii) Benue-Congo Family

(1) Grassfield Bantu (East Grassfield and Momo)

Clusters		Intelligibility Survey Results
North	9 speech forms	Negative: separate languages.
Central Bamileke	10 speech forms	RTT recommended for some speech forms where the intelligibility survey is inconclusive.
Ngemba	6 speech forms	<ol style="list-style-type: none"> 1. Intelligibility surveys have allowed the more than 15 speech forms to be reduced to six speech forms which can not be further grouped without serious consequences of loss of intelligibility. 2. RTT will be needed between: <ul style="list-style-type: none"> - Bafut and Nkwen, - Awing and Pinyin, to see if they could be considered as two languages rather than four. 3. Mundum, considered as a dialect of Mankon, has a relatively low level of intelligibility with Mankon.

Momo (West)	10 speech forms	Generally negative, except for Meta? and Moghamo which display a high degree of intelligibility and should be considered as one language.
 (2) Mbam Bantu		
Clusters		Intelligibility Survey Results
Yambasa	3 speech forms	Negative: separate languages
 (3) Equatorial Bantu		
Clusters		Intelligibility Survey Results
Bafia	4 speech forms	Negative: separate languages
Sawa	15 speech forms	Negative/positive: separate languages, except for Mokpwe and Wumboko for which RTT is recommended.
Basaa-Beti	12 speech forms	Negative/Positive: separate languages, except for Ewondo and Fang (Ntumu), Basaa and Bakoko between which RTT are needed.

II. Intelligibility Reports from the SIL Surveys

The Summer Institute of Linguistics (SIL) Branch in Cameroon established a Language Survey Department in 1987 and worked together with researchers from the Institute of Human Sciences and the University of Yaoundé in subsequent years to determine whether the various speech forms can ultimately use the same literature on the basis of evidence of intelligibility between them. The SIL linguistic surveys focus on three domains: dialectology, bilingualism and extra linguistic issues. The work used the RTT for testing intelligibility and the Second Language Oral Proficiency Evaluation (SLOPE) for determining levels of bilingualism. The SIL findings are presented below for each family, branch and cluster.

i) Intelligibility studies within the Chadic Family

Clusters		Results
Wandala	Hedi, Mabas, Gvoko:	Separate languages, confirmed by RTT (Stadler, 1993).
Daba	Daba, Buwal:	Separate languages (Seguin, 1991).
Mafa	Pelasla, Mbuko, Mefele, Muyang, Melokwo:	Separate languages (Harvey, 1983, 1991) and (Grant, 1991, 1994).

ii) Intelligibility studies within the Adamawa Family

Clusters		Results
Fali	Fali North, Fali South:	Two languages from six speech forms (Sweetman, 1981).
Mbum	Mbum, Nzakmbay, Kuo, Karang:	Separate languages, confirmed by Davis (1990).
Leko-Nimbari	Samba Leko, Nimbari, Wom:	Separate languages, confirmed by Griffin (1994).

iii) Intelligibility studies within the Benue-Congo family

(1) Non-Bantu

Cluster	Results
Mambiloid	Wawa, Kamkam: Separate languages, confirmed by Starr (1989).

(2) Wide Bantu

Clusters	Results
Tivoid	Iceve, Mesaka, Caka, Evand, Ipulo, Eman: Separate languages, confirmed by Regnier (1991).
Nyang	Denya, Kendem, Kenyang: Separate languages (Tyhurst, 1983).

(3) Grassfield Bantu

Clusters	Results
Ngemba	Mankon, Ngemba, Awing and Pinyin: Standardization proposed for as one language, on the basis of opinion surveys (Sadembouo, 1991). RTT not conducted yet. It is recognized that bridging materials would be needed.
North	Yamba, Mfumte, Dzodinka: Separate languages, confirmed by Grant (1994).
Central Bamileke	Yemba, Ngyemboon: Separate languages, confirmed by Starr (1990).

(4) Mbam Bantu

Clusters	Results
Yambasa	Bati, Nubaca, Dumbule: Separate languages, confirmed by Boone (1992a, b) and Boone et al (1992).
Sanaga	Elip, Mmaala, Yangben, Tuki: Separate languages, confirmed by Boone et al (1992).

(5) Equatorial Bantu

Clusters		Results
Bafia	Tibea, Bafia, Dimbong, Hijuk, Yambeta:	Separate languages, confirmed by Boone (1992c) and Bradley (1995a, b).
Basaa-Beti	Bebele, Bebil Ewondo, Bulu:	Separate languages, confirmed by Domche et al (1989) and Seguin (1990).
Meka	Makaa, Bekwel, Kol, Koozime, Mpyemo, So:	Confirmed as separate languages by Beavon (1989a, b).
Sawa	Batanga, Yasa:	Confirmed as separate languages by Davidson (1991).

LANGUAGES OF EQUATORIAL GUINEA

I. Basic Data

Table 3 presents the speech forms found in Equatorial Guinea along with the various dialects.

Table 3. Comprehensive list of languages found in Equatorial Guinea

	language	dialect
1.	Benga	
2.	Bube	Boombe, Bobe, Bubi, Ediya, Adija, Adeeyah, Booben Fernandian
3.	Fang	Pamue, Pahouin, Pangwe
4.	Ngumba	Mvumbo, Ngoumba, Mgoumba, Mabi, Mabea, Bisiwo
5.	Ngumbi	Combe, Kombe
6.	Yasa	Yassa, Lyaasa, Maasa, Bongwe
7.	Crioulo	A Spanish based Creole (pidgin?)
8.	Spanish	

II. Genetic Classification

Table 4 presents the speech forms of Equatorial Guinea organized in a genetic group of phylum along with some sociolinguistic information.

Table 4. The languages of Equatorial Guinea grouped genetically (Bantu Languages: A30 (Bube-Benga group), A70, A80⁽⁶⁾) by Guthrie 1967/1971)

	comprehensive list of languages	major town	number of speakers	year	extinct languages
1.	Benga	Corsico Island	2,000	--	--
2.	Bube	Malabo	21,780	1986	--
3.	Fang	--	75% pop.	1986	--
4.	Ngumba	--	7,833	1982	--
5.	Ngumbi	Batta	4,000	1972	--
6.	Yasa	Batta	1,500	--	--

III. The languages of Equatorial Guinea, organized in clusters

Table 5. Benue-Congo family: Equatorial Bantu Languages (Narrow Bantu)

	(A 30: Coastal Bantu)	(A 70)	(A 80)
1.	Benga	1. Fang	1. Ngumba
2.	Bube		
3.	Ngumbi		
4.	Yasa		

IV. Results of Intelligibility Surveys

The results of the intelligibility surveys reveal that the languages of Equatorial Guinea are separate languages, except for Yasa and Ngumbi for which RTT are needed. Three out of five respondents confirmed mutual intelligibility between Yasa and Ngumbi. They also recognized a close relationship between Yasa and Benga.

In terms of standardization we propose that Yasa and Ngumbi be standardized as one language to be called Yasa, if attitudinal surveys within the Ngumbi speaking community favor the use of Yasa. The choice of Yasa as standard or reference for standardization takes into account the fact that Yasa is spoken in Cameroon as well by a substantial number of speakers. As a transborder language the same norms of orthography and grammar could be used across the frontier from Cameroon to Equatorial Guinea, covering both the Yasa and Ngumbi communities (Chumbow & Tamanji, 1998).

The relation between Yasa and Benga is quite obvious in terms of genetic affinity but there is not enough evidence of mutual intelligibility to warrant the Yasa standard being extended to Benga. The RTT is needed to arrive at a definitive conclusion. All other speech forms do not evidence a significant level of mutual intelligibility and so must be considered separate languages for which standardization would be required in due course.

A summary of the languages spoken in Equatorial Guinea is provided as follows.

1. Yasa and Ngumbi
2. Benga
3. Bube
4. Fang
5. Ngumba
6. Crioulo

CONCLUSION

At the end of our exercise, a tremendous amount of energy has been exerted, efforts deployed and funds mobilized to identify, classify and determine the status of speech forms in Cameroon and Equatorial Guinea. While clustering the speech forms on the basis of genetic relations via lexicostatistics has been quite fruitful (Appendices), clustering on the basis of mutual intelligibility of at least 85% does not seem to have reduced the number of speech forms significantly. However, there are a few significant cases within the Adamawa family. The Fali cluster with six speech forms identified by Dieu & Renaud (1983) was reduced to two by Grimes (1996, 2002) and Gordon (2005) following empirical research on mutual intelligibility of these forms. We independently propose that there exists only one standardizable Fali language, based on the Fali-Tinguelin dialect (Sweetman, 1981). The situation can be summarized as follows:

Table 6. The Fali cluster

	variants	dialects	language
1.	Fali-Dourbeye		
2.	Fali-Bossoum	Fali-North	
3.	Bveri		Fali (Tinguelin)
4.	Kangu		
5.	Fali-Bele	Fali-South	
6.	Fali-Tinguelin	Intelligibility	SLOPE

In the Benue family on the other hand, a notable success of the mutual intelligibility testing exercise is recorded in the East Grassfield Bantu languages with the Ngemba cluster as presented in Table 7.

From the table, it is evident that if recent work with intelligibility testing has not reduced the number of languages significantly, it is precisely because previous work on intelligibility had already clustered the speech forms into languages. From the 286 speech forms or languages listed in Grimes (1992, 1996, 2002) and Gordon (2005) for Cameroon, the considerations of intelligibility discussed here reduce this number just slightly. This is because these authors have taken into consideration the results of the work on intelligibility carried out in the country, notably by SIL Survey Department. However, intelligibility surveys and testing have not been carried out in many other clusters. Despite the low yields, we believe that it is important to continue the exercise in order to ascertain the exact situation.

Table 7. The Ngemba cluster

	variants	dialects	languages
1.	Mankon		
2.	Mbatu		
3.	Chomba		
4.	Nsongwa	Ngemba or Mankon	
5.	Alatening		
6.	Akum		
7.	Mundum		Ngemba
8.	Ndjong		
9.	Nkwen	Nkwen	
10.	Mendankwe		
11.	Awing	Awing	
12.	Pinyin	Pinyin	
13.	Bafut	Bafut	Bafut
14.	Bambili	Bambili	Bambili
15.	Bambui	Intelligibility	SLOPE

ACKNOWLEDGEMENT We are grateful to the following students, who participated in the collection of data for this work: Madeleine Ngo Ndjeyiha (University of Yaoundé), Gerald Ngwa (University of Yaoundé), Michael Apuge (University of Buea), Ayu'nwi N. Neba (University of Buea).

NOTES

- (1) For details on the conception, design and administration of (mutual) intelligibility tests, see Casad (1974) and Starr and Stadler (1990).
- (2) See Chesley and Starr (1990).
- (3) The relatively high degree of correlation between the intelligibility survey results and results of (actual) intelligibility testing lends some measure of credibility to intelligibility survey (if it is properly and rigorously conducted). Given that actual intelligibility testing is very intensive, expensive and time consuming, an alternate intelligibility survey technique (relatively easier) can be used to estimate intelligibility before intelligibility testing is called in (if and when necessary).
- (4) The SIL survey teams in Cameroon have been equally concerned with this question and have done some fruitful work (on intelligibility testing) whose results are discussed in Section 4.2.
- (5) Only the number of languages or speech forms in each cluster is indicated. For the specific identity of the languages in the cluster, see the list in the Appendices.
- (6) The codes A30, A70, A80 etc. are labels assigned to groups of languages by Guthrie (1967/71) in his referential classification of Narrow Bantu languages. Indeed, on the basis of linguistics and geographical criteria the authors grouped those languages into 15 geographic zones labelled A, B, C, D, E, F, G, H, K, L, M, N, P, R and S. He numbered and proposed a language name for each group.

REFERENCE

- Beavon, K. & A.E. Johnson 1989a. *Sociolinguistic Survey among the Bekwel*. Summer Institute of Linguistics, Yaoundé.
- & ——— 1989b. *Sociolinguistic Survey among Mpyemo*. Summer Institute of Linguistics, Yaoundé.
- Boone, D.W. 1992a. *Sociolinguistic Survey in Mbola: Mbule. Survey Report*. Summer Institute of Linguistics, Yaoundé.
- 1992b. *Baca (Bongo) Survey Report*. Summer Institute of Linguistics, Yaoundé.
- 1992c. *Dimbong Survey Report*. Summer Institute of Linguistics, Yaoundé.
- Boone, D.W., D.P. Bradley & C. Grant 1992. *Central Yambassa Survey Report*. Summer Institute of Linguistics, Yaoundé.
- Brown, B.M. 1995. *Africa's Choice*. Thirty Years of the World Bank, Harmondsworth.
- Casad, E.H. 1974. *Dialect Intelligibility Testing*. Summer Institute of Linguistics of the University of Oklahoma, Norman.
- Chesley, W. & A. Starr 1990. *Linguistic and Sociolinguistic Survey in the Momo Division*. Summer Institute of Linguistics, Yaoundé.
- Chumbow, B.S. & P. Tamanji 1998. Linguistic Identity across the Border of the Cameroon Triangle. In (K.K. Prah, ed.) *Between Distinction and Extinction*, pp. 53-74. The Harmonization and Standardization of African Languages, Witwatersrand University Press, Johannesburg.
- Davidson, P. 1991. *Batanga Survey Report*. Summer Institute of Linguistics, Yaoundé.
- Davis, D.W. 1990. *A Linguistic and Sociolinguistic Survey of the Eastern Mbum Languages*. Summer Institute of Linguistics, Yaoundé.
- Dieu, M. & P. Renaud (eds.) 1983. *Atlas Linguistique de l'Afrique Centrale (ALAC), Atlas Linguistique du Cameroun (ALCAM)*. Direction Générale de la Recherche Scientifique et Technique (DGRST) / Centre Régional de Recherche et de Documentation sur les Traditions Orales et pour le Développement des Langues Africaines (CERDOTOLA), Yaoundé and Agence de Coopération Culturelle et Technique (ACCT), Paris.
- Domche, E.T. & L.M. Seguin 1989. *Sociolinguistic Survey of Yebekolo*. Summer Institute of Linguistics, Yaoundé.
- Gordon, R.G. 2005. *Ethnologue: Languages of the World. (15th edition)*. Summer Institute of Linguistics, Dallas.
- Grant, C.A. 1991. *A Sociolinguistic Survey of Pelasla (Rapid Appraisal)*. Summer Institute of Linguistics, Yaoundé.
- 1994. *A Rapid Appraisal Survey of Dzodinka and Mfumte*. Summer Institute of Linguistics, Yaoundé.
- Gregersen, E.A. 1977. *Language in Africa. An Introductory Survey*. Gordon and Breach Science Publishers, New York.
- Greenberg, J. 1970. *The Languages of Africa, (2nd edition, with additions and corrections)*. Indiana University, Bloomington.
- Griffin, M.A. 1994. *A Rapid Appraisal Survey of Mubako*. Summer Institute of Linguistics, Yaoundé.
- Grimes, B.F. 1992. *Ethnologue: Languages of the World. (12th edition)*. Summer Institute of Linguistics, Dallas.
- 1996. *Ethnologue: Languages of the World. (13th edition)*. Summer Institute of Linguistics, Dallas.
- 2002. *Ethnologue: Languages of the World. (14th edition)*. Summer Institute of Linguistics, Dallas.

- Guthrie, M. 1967/1971. *Comparative Bantu, Vols. 1-4*. Gregg International Publishers Ltd, Farnborough.
- Harvey, B.S. 1983. *A Sociolinguistic Survey of Hina Subdivision, Mayo-Tsanaga Division, Extreme North Division*. Summer Institute of Linguistics, Yaoundé.
- 1991. *A Sociolinguistic Survey of Mefe*. Summer Institute of Linguistics, Yaoundé.
- Heine, B. 1993. *Language as a Tool in Reconstructing the African Past*. Raymond Dart Lectures No. 28, Johannesburg.
- Regnier, C.D. 1991. *Tivoid Survey of the Akwaya Subdivision, Manyu Division, Southwest Province*. Summer Institute of Linguistics, Yaoundé.
- Sadembouo, E. 1991. *A Sociolinguistic Survey of the Ngemba Languages*. Summer Institute of Linguistics, Yaoundé.
- Seguin, L.M. 1990. *Survey among the Bebele*. Summer Institute of Linguistics, Yaoundé.
- 1991. *A Sociolinguistic Survey of the Buwal (Gadala)*. Summer Institute of Linguistics, Yaoundé.
- Stadler, J. 1993. *A Sociolinguistic Survey among the Xedi-Lamang*. Summer Institute of Linguistics, Yaoundé.
- Starr, A. 1989. Sociolinguistic Survey of Wawa, a Mambiloid Language of Cameroon. Summer Institute of Linguistics, Yaoundé.
- 1990. *Sociolinguistic Survey of Yemba and Ngyemboon*, Eastern Grassfields of Cameroon. Summer Institute of Linguistics, Yaoundé.
- Starr, A. & J. Stadler 1990. *Principles of Sociolinguistic Surveys: Actualities of Cameroon*. Summer Institute of Linguistics, Yaoundé.
- Sweetman, G. 1981. *Dialect Intelligibility Testing: The Fali Survey*. Summer Institute of Linguistics, Yaoundé.
- Tyhurst, J.J. 1983. *Linguistic Survey of the Nyang Languages*. Summer Institute of Linguistics, Yaoundé.
- Wimbish, J.S. 1989. *WORDSURV: A Program for Analyzing Survey Word List*. Summer Institute of Linguistics, Dallas.

——— Accepted April 25, 2007

Correspondence Author's Name and Adress: Gratien G. Atindogbe, P.O. Box 63, University of Buea, Buea, South West Province, CAMEROON.
E-mail: grat_atin@yahoo.com

Appendix 1

1. The Afro-Asiatic Phylum

Chadic Family (ALCAM zone 1 & 2): 59 languages (with 1 nearly dead language)

West (1)	Margi (3)	Gbwata (7)	Daba (5)	Wandala (8)	Mafa (19)	
1. Hausa	1. Psikye [3] 2. Hya 3. Bana [2]	<u>North</u> 1. Jimi [5] 2. Gude 3. Ziziliviken 4. Sarua 5. Tsuvan <u>Centre</u> 1. Nzanyi <u>South</u> 1. Bata [4]	<u>North</u> 1. Buwal 2. Gavar [1] <u>South</u> 1. Besleri [3] 2. Daba [4] 3. Mbedam	<u>East</u> 1. Wandala [9] 2. Glavda 3. Parkwa 4. Gaduwa <u>West</u> 1. Gvoko 2. Hedi [1] 3. Mabas [3] 4. Guduf [3]	<u>North-West</u> 1. Matal <u>North-East</u> 1. Pelasla [3] 2. Mbuko <u>South</u> 1. Wuzlam 2. Muyang 3. Mada 4. Melokwo 5. Zulgwa [2] 6. Gemzek 7. Mery [1] 8. Dugwor 9. Giziga, North 10. Giziga, South [3] 11. Mofu, North [2] 12. Mofu, South [3] 13. Baldamu 14. Cuvok 15. Mefe [4] 16. Mafa [4]	
Yedina (1)	Mangage (6)	Midaʔa (2)	Musgu (1)	Kada (1)	Masana (4)	Kera (1)
1. Buduma	<u>North</u> 1. Mpade [5] 2. Malgbe [4] 3. Maslam [2] 4. Afade <u>South</u> 1. Mser [4] 2. Lagwan [2]	1. Jina [4] 2. Majera [3]	1. Musgu [7]	1. Gidar	<u>North</u> 1. Masana [7] 2. *Zumaya 3. Mosi [1] <u>South</u> 17. Peve	1. Kera

Note: Figures in the round brackets indicate the number of languages, and those in the square brackets the number of dialects identified for the language. *denotes nearly defunct language.

Remark: In Table 1, the Mafa cluster is divided into 3 subclusters (North-West, North-East and South). The South cluster (16 languages) is subdivided into 3 manageable sub-clusters (South 1, South 2 and South 3), on geographical and linguistic criteria (contiguity and genetic relation) as seen below:

2. Mafa cluster

		Mafa		
North-West	North-East	South		
		South 1	South 2	South 3
1. Matal	1. Pelaska [5]	1. Wuzlam	1. Mery [1]	1. Cuvok
	2. Mbuko	2. Muyang	2. Dugwor	2. Mefele [4]
		3. Mada	3. Giziga, North	3. Mafa [4]
		4. Melokwo	4. Giziga, South [3]	
		5. Zulgwa [2]	5. Mofu, North [2]	
		6. Gemzek	6. Mofu, South [3]	
			7. Baldamu	

Also note that we shall use the acronym ALCAM and the title *Ethnologue for Dieu & Renaud* (1983) and Grimes (1992, 1996, 2002) and Gordon (2005) respectively in the tables.

Appendix 2**3. Adamawa Family (ALCAM zone 3): 37 languages**

Leko-Nimbari (5)	Koo-Dii (11)	Mbum (15)	Fali (2)	West Ubangian (4)
<u>Samba</u>	<u>North</u>	<u>North</u>	<u>Alcam</u>	1. Gbaya [21]
1. Samba Leko [7]	1. Mom Jango	1. Tupuri	<u>North</u>	2. Bangandu [2]
2. Wom	2. Koma [7]	2. Mundang [4]	1. Fali-Dourbeye	3. Baka
3. Kolbila	3. Gimnime [1]	3. Mambai	2. Fali-Bossoum	4. Gieli
4. Nyong	4. Gimme	4. Gidar	3. Bveri	
	5. Doyayo [2]			
<u>Nimbari</u>		<u>South</u>	<u>South</u>	
1. *Nimbari	<u>South</u>	1. Dama	1. Kangu	
	1. Peere [3]	2. Momo	2. Fali-Bele	
	2. Longto	3. Mbum [2]	3. Fali-Tinguelin	
	3. Duupa	4. Karang [4]		
	4. Pape	5. Kari [2]	<i>Ethnologue</i>	
	5. Dii [10]	6. Dek	1. Fali, North [3]	
	6. Saa	7. Kuo	2. Fali, South [4]	
		8. Pana [1]		
		9. Nzakmbay [1]		
		10. Pam		
		11. Ndai		

Note: Figures in the round brackets indicate the number of languages, and those in the square brackets the number of dialects identified for the language.

*denotes nearly defunct language.

4. Adamawa endangered and unclassified languages

Adamawa: extinct or nearly extinct	Adamawa unclassified
1. *Gey (extinct)	1. La?bi (language of initiation rites used by the Gbaya, Mbum, and some Sara-Laka)
2. *Duli (extinct)	2. To (ancient secret male initiation language of the Gbaya)
3. *Oblo (nearly extinct)	

NB: Gey is extinct, but the people are still ethnically distinct and speak Fulani.

*denotes nearly defunct languages.

Appendix 3

5. Benue-Congo Family

I) Non Bantu Languages (ALCAM zone 7): 20 languages

Jukunoid (9)	Cross River (2)	Bendi (1)	Mambiloid (8)
1. Mbembe, Tigon [5]	1. Korop	1. Bokyi	1. Suga
2. Jukun [2]	2. Efik		2. Mambila, Cameroon [3]
3. Kutep [5]			3. Kwanja [3]
4. Yukuben			4. Vute [8]
5. Akum			5. Kamkam
6. Beezen			6. Ndoola
7. Bikya			7. Twendi
8. Bishuo			8. Wawa [1]
9. Busuu			

Note: Figures in the round brackets indicate the number of languages, and those in the square brackets the number of dialects identified for the language.

6. Benue-Congo Family

II) Wide Bantu Languages (Bantoid) (Part of ALCAM zone 8): 29 languages with 2 dead languages

Jarawan (3)	Tivoid (12)	Ekoid (1)	Nyang (3)	Beboid (12)
1. *Ngong	1. Tiv [1]	1. Ejagham [3]	1. Denya [4]	<u>West</u>
2. *Nagumi	2. Iceve-Maci [2]		2. Kendem	1. Naki
3. Mbonga	3. Evand		3. Kenyang [3]	2. Bu
	4. Mesaka [2]			3. Missong
	5. Esimbi			4. Koskin
	6. Ipulo [2]			
	7. Eman [2]			<u>East</u>
	8. Caka [2]			1. Bebe
	9. Osatu			2. Kemezung
	10. Manta			3. Ncane
	11. Balo			4. Nsari
	12. Bitare			5. Noone
				6. Nsaa
				7. Cung
				8. Mungong

Note: Figures in the round brackets indicate the number of languages, and those in the square brackets the number of dialects identified for the language.

*denotes nearly defunct language.

7. Benue-Congo Family

III) Grassfield Bantu Languages (Bantoid) (ALCAM zone 9 & part of Zone 8): 60 languages

Momo (10)	Menchum (1)	Ring (15)	East-Grassfield (34)			
			Ngemba (6)	Central Bamileke (10)	Noun (9)	North (9)
<u>West</u>	1. Befang [6]	<u>West</u>	1. Bafut [2]	1. Ngombale [3]	1. Mungaka [3]	1. Limbum [3]
1. Ngwo [8]		1. Aghem [8]	2. Ngemba [11]	2. Megaka	2. Bamun	2. Dzodinka
2. Ngoshie [2]		2. Weh	3. Bambili [2]	3. Ngomba [3]	3. Medumba	3. Mfumte
3. Ngje		3. Isu	4. Mendankwe [2]	4. Ngyemboon [3]	4. Bamenyan [4]	4. Yamba [5]
4. Meta? [3]		4. Fungom	5. Awing [2]	5. Yemba [3]	5. Baba	5. Mbe?
		5. Oso	6. Pinyin [1]		6. Bafanji	6. Kofa
<u>East</u>		<u>Centre</u>		1. Ngwe	7. Bamali	7. Kwaja
1. Busam		1. Mmem		2. Ghomala? [4]	8. Bambalang	8. Ncha
2. Menka		2. Kom [1]		3. Fe?fe? [9]	9. Bangolan	9. Ndaktup
3. Atong		3. Bum		4. Nda?nda? [2]		
4. Ambele		4. Babanki		5. Kwa? [2]		
5. Mundan [8]		5. Kuo				
6. Njen		<u>East</u>				
		1. Lamnso?				
		<u>South</u>				
		1. Vengo				
		2. Kenswei				
		Nsei				
		3. Bamunka				
		4. Wushi				

Note: Figures in the round brackets indicate the number of languages, and those in the square brackets the number of dialects identified for the language.

8. Benue-Congo Family

IV) Mbam Bantu Languages (Narrow Bantu) (Part of ALCAM zone 5): 15 languages

Tikar (2)	Tunen (3)	Yambeta (1)	Bati (1)	Yambassa (3)	Sanaga (5)
1. Ndemli	1. Tunen [9]	1. Yambeta [2]	1. Bati	1. Nugunu [2]	1. Tuki [7]
2. Tikar [7]	2. Tuotomb			2. Nubaca	2. *Leti
	3. Noomande			3. Dumbule	3. Elip
					4. Mmaala
					5. Yangben

Note: Figures in the round brackets indicate the number of languages, and those in the square brackets the number of dialects identified for the language.

*denotes nearly defunct language.

9. Benue-Congo Family

V) Equatorial Bantu Languages (Narrow Bantu) (ALCAM zone 4, 6, and part of 5): 44 languages

Bafia (4)		Coastal or Sawa (15)		Basaa-Beti (12)	Meka (10)	Kako (3)
(A50)	(A10)	(A20)	(A30)	(A40) & (A70)	(A80)	(A90)
1. Tibeá	1. Balundu-	1. Bakole	1. Batanga [3]	1. Abo	1. Byep [2]	1. Pol [7]
2. Dimbong	Bima [4]	2. Wumboko	2. Yasa	2. Barombi	2. Makaa [3]	2. Kwakum [4]
3. Bafia [4]	2. Bakundu-	3. Mokpwe		3. Basaa [15]	3. So [2]	3. Kako [4]
4. Hijuk	Balue [4]	4. Isu		4. Bakoko [7]	4. Ngumba [4]	
	3. Bafaw-	5. Bubia		5. Beti	5. Koozime [4]	
	Balong [2]	6. Duala [4]		6. Bebele [2]	6. Mpongpong [5]	
	4. Bassossi	7. Malimba		7. Bebil	7. Mpyemo [2]	
	5. Akoose [6]			8. Bulu [5]	8. Bekwel [1]	
	6. Mbo [13]			9. Eton [4]	9. Bomwali	
				10. Ewondo [15]	10. Kol [4]	
				11. Fang		
				(or Ntumu) [3]		
				12. Mengisa		

Note: Figures in the round brackets indicate the number of languages, and those in the square brackets the number of dialects identified for the language.