



Language and Culture Archives

Bua bloc survey report

Douglas W. Boone and Kenneth S. Olson

©1995, SIL International

All rights reserved.

This document is part of the SIL International Language and Culture Archives.

More resources are available at: www.sil.org/resources/language-culture-archives.

BUA BLOC SURVEY REPORT

Douglas W. Boone
and
Kenneth S. Olson

May 1995

© 1995, Eastern Zaïre Group
Summer Institute of Linguistics

SIL (a.s.b.l.)
B.P. 872, Bunia
(av. Nizi, N° 3)
République du Zaïre

0.0 PREFACE

The Bua bloc linguistic survey was conducted by the Summer Institute of Linguistics with the authorization of the *Communauté Évangélique du Christ au Cœur de l'Afrique* (CECCA/16), the *Communauté Baptiste au Zaïre du Nord* (CBZN), and the *Communauté Évangélique en Ubangi et Mongala* (CEUM/51). We wish to thank the Presidents of these communities, as well as the pastors and preachers in the villages and cities that we visited.

We would also like to thank the *Commissaires de Zone*, the *Chefs de Collectivité*, and other officials with whom we met on the way, for their warm reception and their invaluable help.

1.0 INTRODUCTION

1.1 Purpose

For quite some time, people from the Bua language family of north-central Zaïre have shown interest in a mother-tongue Bible translation project. Members of the Summer Institute of Linguistics and Wycliffe Bible Translators have conducted at least five surveys in the past fifteen years in the Bua bloc attempting to ascertain the linguistic situation in the region. However, certain questions remain. The Bua bloc is spread out over two political regions, Equateur and Haut-Zaïre, and the research done in one region has not accounted for research done in the other region. Also, because of transportation difficulties, documentation of the bloc has been incomplete. With this in mind, the present survey had the following purposes:

- (a) Determine if the Bua bloc speech varieties spoken in Equateur (Pagibete) and Haut-Zaïre (Bua, Bati, Bengé, etc.) can use common literature,
- (b) Obtain more complete documentation of the Bua bloc dialects,
- (c) Build relationships with members of the local church to encourage their involvement in a Bible translation and literacy project, and
- (d) Assess more accurately bilingualism levels in Bangala and Lingala.

1.2. Motivation

The Congo-Western Zaïre Group of Wycliffe has been interested in placing a translation team in Pagibete; it was important to assess the relatedness of Pagibete to other Bua dialects before beginning this project.

DC3 and MAF/Caravan traffic between Haut-Zaïre and Equateur makes an allocation to Buta more feasible at present.

1.3. Review of Previous Research.

(Readers unfamiliar with the areas discussed are referred to the maps in Appendix A.)

1.3.1 Surveys in Haut-Zaïre

Depending on one's definitions, SIL had conducted between two and five surveys in the part of the Bua bloc which lies in Haut-Zaïre before the present survey. Contact was made with a pastor in Kisangani with a burden for Bua Scripture translation on three reported occasions: in 1977, 1979, and 1990. The 1979 visit included some translation consulting. Two survey trips were made to the Bua-speaking area, in 1986 and in 1991.

Bua is mentioned in the report of the first SIL exploratory survey in Zaïre (Keller 1977: Appendix 2, p. 28). Buta was identified as the centre, and the population was estimated by Gordon Molyneux of UFM as 25-35,000. The contact address was the church founded by UFM, then called CEHZ (now known as CNCA/21). Keller recommended that an SIL team resident at Kisangani could consult part-time with a MT translator, Pastor Masini (whose other name he doesn't mention). He reported that Past. Masini had already translated several Scripture tracts and was working on the gospels, but who hoped to use a committee approach in the future.

Elaine Thomas visited Kisangani in January 1979 and checked some of his translation of Mark's gospel (see Thomas 1979). By this time, Past. Masini Tozaza and his son, Past. Mongo Masini, had translated the four Gospels and Acts into Bua (she does not identify the

dialect). She specifically refers to their translation of the scripture booklet 'Listen' published by Scripture Gift Mission in 1978. In spite of their having had no linguistic or translation training, she considered the translation to be "good quality."

Thomas also referred to Buta as the centre of the Bua area, and notes that Buta is 340 km from Kisangani. Interestingly, she claimed that the majority of Bua churches (about 40 of them) were in the CEHZ area, while there were "some" in the CECCA/16 area.

While her report does not make an outright recommendation of SIL involvement in Pastor Masini's work, she points out his need for consultant help and financial support. In favour of such support she notes his clear interest, demonstrated aptitude (based on the quality of his work so far), and commitment (based on the fact that he has worked so long on translation in his spare time and without a salary). Possible barriers to project success included the distance between Past. Masini's residence and the Bua-speaking area, apparent lack of a local base of support, and the absence of UFM missionaries in Kisangani at the time.

In April 1986, Mike McCord and Rob McKee, SIL missionaries working with CECCA/16, conducted a short survey at Malingwia (east of Buta near Titule) to clarify the Bible translation need in Bua. They identified CECCA/16 church workers from a variety of Bua-speaking areas and interviewed six of them, chosen to represent the range of dialect and clan groups. They also interviewed the WEC missionary then residing at Malingwia. They also elicited a 150-item word list from each of the six Bua speakers, following the format of the list used for the researching of the *Atlas Linguistique de l'Afrique Centrale*. These data were elicited in the trade language, Bangala. Most nouns were elicited in singular and plural forms.

They concluded that Bua consisted of two mutually intelligible dialects: Yew (west of the Bima River, and including Bengé, Baati, Gbe, and Ngingita) and Bua (east of the Bima River). They also reported that Past. Masini's translation was in the Kiba dialect, spoken to the south of Yew and Bua. They concluded that there were so many differences between Kiba and the dialects they surveyed that comprehension of Masini's translation was inadequate. In their opinion, the linguistic centre of Bua was near Malingwia.

To supplement this information: the Bima river is the boundary between the Bambesa Zone to the east and the Buta Zone to the west. Kiba is spoken in the Banalia Zone, south of the Tele River. All the interviewees came from the Buta or Bambesa Zones.

McCord and McKee concluded that SIL should collaborate with CECCA/16 on a translation project in Bua. They pointed out that CECCA (and WEC, the founding mission) had a large work in the Bua area, that CECCA/16 showed much interest in translation, and that there were many well-educated young men who could help with a translation.

McCord and McKee judged that the largest protestant church work in the Bua-speaking area was that of CECCA/16, centred at Malingwia. (There were 150 local chapels in the Malingwia parish at the time; according to their own field notes, however, many of these were not in Bua-speaking areas. There were two administrative sections fully in the Bua area—Malingwia and Dingila-Tobola—two with Zande and Bua—Bambesa and Buta—and two non-Bua. CECCA's work has undoubtedly grown since then; also, according to new nomenclature introduced in 1993, Malingwia is a "District" and the former sections are now "Parishes".)

They mentioned two other protestant missions, the Norwegian Baptists and the Assemblies of God, with work concentrated toward Buta. Two large Roman Catholic missions are in the area surveyed, at Titule (on the west bank of the Bima river, a few kilometers from Malingwia) and at Bambili (near Dingila at the very northeast edge of the Bambesa Zone). Interestingly, though they challenged Miss Thomas' claim that the majority of baBua churches were in CEHZ, they did not mention UFM among the protestant missions with work in the area. UFM's (and CNCA's) work was of course primarily in the Banalia Zone, which they did not visit. There are CNCA chapels in Buta now (it is not known whether there were in 1986).

According to their field notes, McCord and McKee weighed Miss Thomas' suggestion that Past. Masini might relocate in Buta, until they concluded that the dialect distance was too great. They did recommend, however, that contact with him be renewed.

For reference: The Norwegian Baptists founded CBZN/14, which was called CBBU until recently. The Assemblies of God church in this part of Zaïre is CADZ/12.

The fourth documented "survey" in Bua was a visit to Past. Masini by Ed and Sheryl Mathis in Aug. 1990 (see Mathis 1990). He told them of six dialects, of which five, identified with Kole, Monganzulu, Bobati, Bobenge, and Titule/Bambesa, did not differ very much from one another. The sixth group, the baNgelema, was not easily understood by the others. Though they could not reconcile this accounting with that of McCord and McKee, it now would seem that Kole is Kiba, the next three are subsumed in "Yew", and Titule is the Bua dialect. Alternatively, Bobati and Bobenge could be Bati and Benge, spoken in the Aketi and Bongo Zones (thus he would have listed one dialect for each of the five politico-administrative zones in which Bua is spoken). By most accountings (including the Ethnologue), "Ngelima" (Ngelema) is a separate language.

The Mathises reported that at the time, Past. Masini was still committed to Bua translation, but that his son had gone blind and he himself was greatly hampered by failing eyesight. He was working on a revision of various published tracts, including the 'Listen' booklet. His translations of the gospels were still unpublished, apparently due to lack of consultant help.

In April of the following year, Tim Girard conducted the most thorough survey in Bua to date, examining four Bua dialects, which he identified as Mbili of Dingila, Yeu of Buta, Bongondja of Kole, and Benge of Likati (Girard 1990). The first two would be the Bua and Yew dialects surveyed by McCord and McKee; the Kole dialect is in fact Kiba; and Benge had not previously been surveyed. (Mbokpa, one of the CECCA evangelists interviewed by McCord and McKee, is an ethnic moBenge who grew up speaking "Bua of Buta". Upon inspection, the wordlist data he provided represent the "Buta" dialect. This same man, by now a pastor, also worked with Girard in 1991 and with the present survey team in 1994.)

In each of the four dialect locations, Girard applied two tests. To test "bilingualism" in Bangala, he used a method which he had developed, the Scripture Based Language of Wider Communication Comprehension test (SBLWCCT). To test dialect intelligibility, he used the Recorded Text Test method outlined in Casad (1974). He found that Buta (Yeu) was the "linguistically dominant" region, and though he found a high level of bilingualism in Bangala, he felt that the strong desire for Bible translation in the region was a more important factor.

According to the people he interviewed, the translation work by Pastor Masini was "just fine". In his analysis, the four dialects were "somewhat homogenous," with Benge being the "weak link." Girard encouraged the Bua to form a language committee, with representatives of CECCA/16, CNCA, and CBBU (CBZN), the three principal protestant churches in the Bua area.

From Girard's conclusions, we may infer that he considered the Buta dialect (or Yeu) to be the best choice of reference dialect, but that the existing material in Kiba was suitable. His recommendation that a committee be developed in Buta, with the participation of CNCA, while virtually guaranteeing that Kiba would not be the reference dialect, ensures that the interests of the baKiba are not neglected. He reconciled the conflicting conclusions of Thomas, on the one hand, and McCord and McKee, on the other, by demonstrating both that Buta is the linguistic centre (the dialect there being very much like what is spoken at Malingwia, though perhaps more cosmopolitan) and that Kiba is rightly considered a Bua dialect, and by recommending that three protestant churches collaborate in the oversight of the project.

Furthermore, Girard addressed other issues raised by McCord and McKee by extending the geographical scope of study (traveling both south of the Tele river and north of Aketi), by evaluating Bangala proficiency in a more objective way, and by demonstrating "inter-intelligibility" of the dialects. He did not survey Ngelema, nor did he determine the status of certain speech forms rumored also to be related to Bua (Bati and Benza in the Aketi Zone and Bango of the Basoko Zone).

Another valuable source, antedating SIL's research by over twenty years, is the report of the Linguistic Survey of the Northern Bantu Borderland (Van Bulck and Hackett 1956), which lists six groups in the "Bwa Bloc", namely, Apagibeti, Bengé-Baati, a transition group, Yewu, Bwa, and Ngelima. These are recognizable as Pagibete (not cited as "part of the picture" by anyone surveyed in Haut-Zaïre), Bengé and Bati taken together, Kiba (et al.), Yeu, Bua of Dingila, and Ngelema. The first and last of these may reasonably be called separate languages from Bua and the middle four correspond fairly closely to the four dialects surveyed by Girard.

For the record, Thomas had access to Keller's report; McCord and McKee cite both Thomas' report and Van Bulck and Hackett's classification; Mathis refers to the report by McCord and McKee and to Van Bulck and Hackett (but did not have Thomas' report); and Girard had access to McCord and McKee and the Mathis writeup.

The Ethnologue cited the 1977 population figure through the eleventh edition (1988). The twelfth edition (1992) cites the 1986 figure. Another change made from the eleventh to twelfth editions is a partial correction of identification of the area in which Bua is spoken (from "Buta, Bambesa, Poko, and Bondo Zones" to "Buta, Bambesa, Aketi, Basoko and Bondo Zones").

These changes were made on the advice of one of the present authors (Boone), who now realises that he guessed wrong on the status of the Basoko and Banalia Zones: a Bua dialect (Kiba) is in fact spoken in the Banalia Zone, as well as Ngelema, but Bua is not native to the Basoko Zone. (Apparently, Babango and Ngelema are spoken there, and Babango is not closely related to Bua. The inclusion of the Basoko Zone coincided with the identification of Bango as a dialect, a claim for which there was insufficient evidence.)

1.3.2 Surveys in Equateur

In 1985, Jim Fultz and David Morgan (1986) conducted a survey among the "Pagabete" (our term: Pagibete) people of Equateur. They identified three dialects in Equateur and recognised the relatedness of Bua, which they tentatively classified as a possible fourth dialect. The survey took place at about the same time as investigations in several other speech varieties, including the Bantu languages Ngombe, Budza, and Benza/Genza.

Their methods included a simple RTT and a Lingala Scripture comprehension test (described in Morgan 1989). Having observed inadequate comprehension of three versions of Lingala Scripture, particularly among women (and its distinctness from other languages in Equateur), they concluded that there was a need for Bible translation in Pagibete. Based on several considerations, they recommended that the dialect spoken at Ngakpo be chosen as reference dialect. It should be noted that in 1982, Mbangiye Mo Epolapola, a Pagibete-speaking CEUM pastor originally from Ngakpo, produced five booklets in Pagibete, with the advice of SIL members Margaret Hill and Elaine Thomas. These booklets included a Pagibete history, a book of folk stories, a "deliverance" book, and two books of scripture portions.

Fultz and Morgan also collected six 150-item wordlists, using an elicitation list in Lingala with only partial overlap with the one later used by McCord and McKee. Generally speaking, only the singular forms of nouns were elicited. There was one list each in Mongwapele (the western dialect, or Dialect I), the Ngakpo (central) dialect (or Dialect II), the same dialect as spoken in Butu in the Yakoma Zone, Egezo-Bauma (the eastern dialect, or dialect III, as spoken in the Bumba Zone), Egezo-Muma (unidentified location) and "Bowa" (apparently, as spoken in Titule). The Bua list was collected in Gemena.

According to their analysis of 1984 census figures, they concluded that there were about 25,000 Pagibete people in Equateur. This included about 4,200 residents of the Mongwapele *groupement* (I), 7,500 in the Ngakpo area (II), and 10,400 in the Bumba and Yakoma Zones (III), and supposing 2,900 in Businga town. Oddly, they state that the Dialect II population is "greater". As explained in Section 2.2.3 of the present report, it seems that the breakdown by dialect may actually be more along the lines of 4,200 (I), 6,000 (II), and 13,000 (III). Supposing 200-300 Pagibete speakers in Businga town, we arrive at a total of 23,500, of whom about half speak the eastern dialect. (With no up-to-date on-site census data, this conclusion is tentative.)

The surveyors recommended that further testing be done to ensure that speakers of the eastern dialect (and even of "Bowa") would be able and willing to use literature in the Ngakpo dialect. Since the 1985 survey, a Lingala Sentence Repetition Test (LSRT) has been developed (Radloff 1991, Phillips 1992). Such a test has the advantages of rapid administration (enabling surveyors to gather more data in a fixed period of time than other tests would) and calibration to a standard scale of second language proficiency (so that the results of different surveys can be compared, reducing uncertainty of interpretation).

In March 1994, therefore, JeDene Reeder and Sharon Stoothoff used the LSRT (as adapted to the Lingala of Equateur) to evaluate Lingala proficiency in Ngakpo (Reeder and Stoothoff 1994). They determined that levels of Lingala were insufficient for the Pagibete people to rely solely on Lingala Scripture. In addition, through sociolinguistic interviews, they found vigorous use of the mother tongue and a stated desire for mother-tongue Scripture translation. Since the present survey, they have gathered more information in Businga.

Our review of the literature concludes with a look at Guthrie's classification of the Bantu languages, according to which the Bua bloc is part of the Ngombe group (see Bastin 1978:141). This group comprises: Ngombe, Bwela, Bati/Benge, Bua (identified with Bali and Bango), and Beo (that is, Ngelema). Guthrie does not list Pagibete. The relationship of the Bua bloc (Pagibete, Bati/Benge, Bua, and possibly Ngelema) to Ngombe, Bali, and Bango was in question, however, at the time of the present survey.

A more complete review of the literature in the area of classification is in section 2.3.1.

1.4. Research Questions

In order to realize the purposes of the survey mentioned above, we formulated the following research questions.

Research questions for purpose (a):

- Can the Pagibete understand spoken Bua?
- How much contact is there between the groups?
- Are many changes needed to adapt text in Bua to Pagibete or vice versa? What kinds?
- What attitudes are there about different dialects?

Research questions for purpose (b):

- How many distinguishable dialects are there?
- How do the dialects differ?
- What attitudes are there about different dialects?

Research question for purpose (c):

- What churches work in the region?

Research question for purpose (d):

- Can the speakers adequately understand the trade language?

1.5. Venue and Scope

The survey was conducted from June 13 to July 15, 1994. We traveled from Isiro in Haut-Zaïre, to Ngakpo in Equateur, and then retraced our steps back to Isiro. Research points, in order of administration, were Buta, Aketi, Mondongbo, Ngakpo, Bunduki, Aketi, Buta, Malingwia, and Dingila. We wanted to travel to Likati, Kole, and Yandongi as well, but road conditions and time constraints prevented us from doing so. Contact was made with the protestant church communities on the way (CECCA/16, CNCA/21, CBZN/14, CROM/51, and CBFZ/13) in order to facilitate the trip.

1.6. Methodology

We used five principal methods for our research:

- (a) Sociolinguistic interview schedules,
- (b) the Rapid Appraisal Recorded Text Test,
- (c) Word lists and phrase lists,
- (d) Visiting church leaders and accepting their hospitality,
- (e) Rough adaptation of 'Wooni', the SGM 'Listen' booklet, from Kiba to other speech varieties.

Another survey instrument, the Gemena Version LSRT, was also considered with a view toward the fourth purpose. Because of time constraints, we did not evaluate Lingala ability, which had intentionally been placed fourth on the list of purposes. Since the test had already been done in a Pagibete area, and the Eastern Zaïre Group's next decision concerning a possible SIL project in Bua did not require that Lingala testing be done, it was decided that administering the LSRT was the "expendable" part of our work. Also, it was not clear whether the Gemena standard applies in Haut-Zaïre.

In each location, relevant sociolinguistic information was gathered through group interviews, conducted mainly in the local variant of Lingala or Bangala. The questions came from a prepared form (see Appendix B), but not all questions were used in every location. Data were also collected through conversations with interested and informed individuals, including pastors, government officials, church workers, students, and a missionary in Buta.

Dialect intelligibility in the Bua bloc was assessed using the Rapid Appraisal Recorded Text Test (RA RTT), at the suggestion of Jürg Stalder (the developer) and Ted Bergman (Africa Area Language Assessment Coordinator). The Rapid Appraisal approach to survey was introduced by Bergman (1991) and first implemented in Cameroon (Stalder 1993). This approach is characterised by its limited goals (to gain a general overview of the sociolinguistic situation in a particular area) and specific non-technical procedures. However, the need has been felt for a simplified version of the RTT (Casad 1974), suitable for use as part of a rapid appraisal survey. Stalder (to appear) has introduced the RA RTT to meet this need. A summary appears in Appendix C.

Numerous word lists, each consisting of between two hundred and three hundred items, were elicited using Lingala. Some lists were collected on successive days or from multiple sources, and some items were spot-checked on separate occasions. Each list is largely composed of core vocabulary, of which about 50% were nouns (including body parts, kinship terms, animals, plants, household objects, and natural phenomena), 20% were miscellaneous (numerals, adjectives, question words, pronouns, adverbs and statives (potentially adjectives or verbs)) and 30% were verbs.

A standard 200-item elicitation list was used not only in the principal dialects surveyed, but also in Ngelema, Kango, and Benza. The first 140 items were elicited in Ngombe. In 1995, Douglas Boone also elicited data from a Budza speaker studying in Bunia. Approximately eighty other items, many gleaned from other elicitation lists, including those of the 1985 and 1986 surveys, were elicited in the Bambesa and Buta dialects of Bua, in the three Pagibete dialects, and in Bati, Bengé, and later Kiba. Discussion of the development of word lists can be found in Blair (1990:27-33) and Boone (1989). (The two-hundred item list used was similar, but not identical to, the one suggested in the 1989 paper. The relative representation of semantic categories is very different from that in Blair's sample list.)

Similarity decisions in comparing the dialects to each other could usually be made by the inspection method, though many decisions made for comparisons with less-closely related languages were necessarily somewhat more arbitrary. The WordSurv computer program was used to tabulate the data.

When time allowed, "phrase lists" were also collected, using an elicitation list consisting of declarative and imperative statements, questions and answers, and additional grammatical paradigms.

Another method of shedding light on dialectological issues was an impromptu attempt in Ngakpo and Buta to adapt the 'Listen' booklet, of which the Congo-Western Zaïre Group had a copy on file (probably received from Past. Masini himself fifteen years before), into the local speech. The surveyors guided local speakers through between twenty and fifty verses of Scripture as translated in Bua-Kiba, instructing them to make the minimum changes necessary to produce natural-sounding text in their own dialect. This may be considered a Rapid Appraisal Text Adaptation Experiment—with untrained MT speakers and a linguist who does not speak any of the dialects in question.

Though the text adaptation experiment will not necessarily be a reliable indicator of the feasibility of the use of Computer Assisted Related-Language Adaptation (CARLA) for producing first draft text in related speech varieties, there are two benefits to its use in a

rapid appraisal survey. First, it gives a striking impression of the true degree of similarity or dissimilarity of two dialects, particularly for those analysts who relate better to written than spoken data. Second, if even the most well-traveled and Biblically-versed MT speakers cannot guess the sense of the source text with the help of a linguist, there can be no doubt that the translation in question will not be suitable for his people.

1.7. Definitions

The name "Bua" can be used at a number of levels of specificity. Most narrowly, it refers to the speech variety spoken in five collectivities of the Bambesa Zone. In this report, the term "Bua proper" is reserved for this dialect. "Bua proper" is considered to be a dialect of the Bua language; everyone agrees that one or more co-dialects are spoken in the Buta Zone, and some include the Kiba dialect, spoken in the Banalia Zone. In this report, we shall call this dialect cluster "Bua east".

"Bua east" stands in contrast to Bengge-Bati, a dialect cluster spoken in the Aketi and Bondo Zones. The term "Bua cluster" includes all the dialects mentioned so far, i.e. Bua east plus Bengge-Bati. This is the extent of Bua as listed in *Ethnologue* (Grimes 1992a:417). The present survey was concerned with the "Bua bloc", which we will define as the Bua cluster plus Pagibete, if only for the reason that a definitive judgement about the membership of other speech forms in the bloc cannot be made based on the data we collected.

It is probable that Kango and Ngelema could be included in the Bua bloc as well. For now, we will be content to use the term "Bua group" to refer to Pagibete plus Bua cluster plus Kango and Ngelema. This entity may be considered to have a status similar to Guthrie's "groups", of which C.40 and parts of C.30, D.20 and D.30 may tentatively be rearranged to form the Budza group, Ngombe group, Komo Group, and Budu group.

(This is supposing that Ngombe and Budza belong to separate "groups"; further study may support combining these. At present, Lika and Bali cannot confidently be attached either to the "Bua Group" or the "Komo Group", nor to a new group on their own. Lexically, Lika is equidistant from most of the Bua Group and from Bali (approximately 56% similarity), and Bali and the Bua Group have no more vocabulary in common than either of them (or Lika) has in common with Komo (approximately 45%). Lika and Bali have developed a nine-vowel system lacking in both the Bua and Komo Groups. There is apparently a functioning noun-class system in Lika and Bali, as in the Bua Group; in the Komo Group it has largely disappeared.)

The table summarises these relationships.

BUA PROPER	one dialect (leBoale)
+ other dialects	(e.g. Yew, Kiba)
= BUA EAST	Bua C.44 (not Bati C.43)
+ Bengge-Bati	(C.43)
= BUA CLUSTER	Bua (<i>Ethnologue</i> BWV)
+ Pagibete	(<i>Ethnologue</i> PAG)
= BUA BLOC	(what we surveyed)
+ Kango, Ngelema	(reportedly related)
= BUA GROUP	(not, e.g., the Ngombe, Budza, or Komo groups)

Some of these spellings have been chosen somewhat arbitrarily. In our analysis, there are seven vowels in these languages, in four heights. Orthographically, these may be written:

i e ɛ a ɔ o u.

This is the orthographic choice made for many other seven-vowel Bantu languages of Zaire, including Komo and Lingala (that is, when Lingala is written with seven vowels; otherwise, *e* and *ɔ* are written *e* and *o*). However, the *e* and *o* sound somewhat like *i* and *u*, the -ATR counterparts of *i* and *u*; furthermore, it is usual for the *i* and *u* of five-vowel languages (notably Swahili) to be realised as *e* and *o* in seven-vowel languages.

The first vowel in "Bua" is actually of the second degree of aperture and accordingly the name should be written "Boa" in the language itself. Thus, we write the dialect name "leBoa-le" (not "liBua-li"). However, when placed on a five-vowel grid, it is natural to

pronounce it "Bua" (if one gives equal weight to the two vowels), or "Bwa" (if one places more emphasis on the "a", which, in "leBoa-le" at least, bears a high tone). There is no reason to change the Ethnologue entry (now called BWA), but in this report we will write "Bua".

Similarly, in this report, we choose the spelling "Pagibete" (which is etymologically correct) over "Pagabete" (Ethnologue name), but do not recommend a change in the Ethnologue. Since in a sense, "Pagibete" is a nickname (see Section 2.1 "Glossonyms"), it may be preferable to change the Ethnologue entry to "Gezo", which is a true glossonym. This will depend on the preference of the speakers themselves.

Although the spelling "Ngelima" is usual in the literature, including the Ethnologue, we will spell it "Ngelema" in this report since none of our interviewees, which included a speaker of the language itself, called it "Ngelima".

2.0 FINDINGS

2.1. Glossonyms and Ethnonyms

Missionaries and government officials have traditionally used Swahili labels in referring to the Bua bloc people and languages. The prefix "ki-" refers to the languages, while "ba-" refers to the people. Thus, in Lingala and French, the baBoa (ethnonym) speak kiBoa (glossonym). Our interviews opened with inquiries into the local autoglossyms and autoethnonyms, the terms used to identify speech varieties and people groups when are people speaking their own language. (As we explained, we call our own language, «l'anglais», "English". Similarly, «le français» is an autoglossonym.) The results of this survey are listed in Table 1.

Table 1: Bua Bloc Glossonyms and Ethnonyms

reference name	Ethnologue code	Guthrie Code	autoglossonym (language name)	ethnonym (people name)
Bua	BWW	C. 44	leBoa, lèBòá-lè	ðàBòà
Yeu	"	(")	lèYéù	ðàYéù
(Nganzolo)	"			
(Kiba)	"			
Ngelema	AGH	C. 45	-Ngèlémà	
Bati	(BWW)	C. 43	èBátì	ðòBátì
Benge	(BWW)	C. 43	leBenge	baBenge
Pagibete	PAG	(?)	èGèzó	baPagibete

NB: e is between i and ε, o is between u and ɔ.
Transcription is tentative (e vs ε, b vs ð).

The Pagibete terms bear explanation. It appears that at some time the speech varieties have been distinguished by their different speech-initial formulas, essentially the different ways people say "I say (that)...". Several of these are cited by Van Bulck and Hackett (1956:78-80). Perhaps when someone in Equateur said "apa-gi beti" ("He says that..." according to Van Bulck's and Hackett's transcription, p. 78), speakers of neighbouring languages misinterpreted this as A-Pagibeti, the Pagibeti people, and by the Bantuisation process mentioned above, this became baPagibeti. Variant names such as "Pagabete" may have developed over time (through such processes as vowel assimilation?)

Ironically, the preferred lexical item (verb root) for 'to say' (Lingala 'koloba') is [-pag-] in Bua, Yew, and Kiba, [-pak-] in Bati and Benge (a regular phonetic difference), but [-kpet-] in all three "Pagibete" dialects. In other words, those who call themselves "bapagibete" are the ones who don't say 'say' "page"!

In Haut-Zaïre, several respondents recognised the expression "Pagibete" but took it as a general term for the whole bloc: those who say 'say' "page". Many of these were not even aware of the speech variety found in Equateur, or supposed it to be the same as a dialect spoken on the border between Haut-Zaïre and Equateur. Since the interviewees at Ngakpo also included the baBua among the baPagibete, but do not use the term baGezo, there

appears to be no ethnonym to distinguish the Equateur clans from their cousins in Haut-Zaïre.

Fultz and Morgan (1986:10) mention that dialect III is sometimes called egezo, but on the present survey it was rather speakers of dialects I and II who offered this glossonym. Interestingly, a speaker of the Bomokandi dialect of Kango (east of the Bua) from whom Douglas Boone collected a short word list in 1990 also cited "egezo" as her autoglossonym. Though one would expect the corresponding ethnonym "moGezo, baGezo", this is not used by any group surveyed.

Note also the suffix in le-Boa-le. The presence of suffixes in some noun classes is a distinctive of the Bua bloc languages. This fact is not easily recognised when only the singular form is elicited on wordlists. Lika has this feature, though in different classes, but Ngelema, Kango and Bali apparently do not. Neither, apparently, do Benza, Ngombe, and Budza.

2.2. Geography and Population.

2.2.1. Extent of the Pagibete and Bua areas

The home areas of speakers of the Bua bloc languages are found in eight zones of four subregions of the two northern regions of the Republic of Zaïre, as shown in Table 2.

Table 2: Where Bua Bloc Dialects are Spoken

Region	Subregion	Zone	Collectivity	Grpmts.	Dialect
Equateur	Nord-Ubangi	Businga	Businga	Mongwapele	(I)
				3 Omveda	(II)
		Yakoma	(?)	Pombi grpmt.	(II)
				2 others	(III)
	Mongala	Bumba	Yadongi	Ndundu-Sana	(III)
Haut-Zaïre	Bas-Uélé	Bondo	Mobenge-		
			-Mondila	unknown	(Benge)
		Aketi	Bodongola	all 4 grs.	(Bati)
			Bongi	5 grpmts.	(Bati)
			Kolóngwande	Baboa	(Bati)
			other grpmts are	Ngbandi	
			Mobati-		
			-Boyele	all 3 grs.	(Bati?
					or Benge?)
		Buta	Barisi-Mongengita		
			-Bakango		(Yew)
			Bayew-Bongongea		(Yew)
			Bayew-Mogbama		(Yew)
			Mobati		(may be ethnic Bati, but speak "Buta dialect")
			Monganzulu		(Monganzulu)
		Bambesa	Bakete	\	
			Bokapo	\	
			Bokiba	>	(Bua)
			Bolungwa	/	
			Mondongwale	/	
	Tshopo	Banalia	Babua de Kole	entire	(Kiba)

Based on our group interviews, the two next most closely related languages are Kango and Ngelema. This perception is supported by lexical similarity figures (see section 2.3.2.)

The baKango apparently don't even have a *groupement* of their own, but live along the (Uélé) river among the baBua and aZande. Evidently, another dialect of Kango is spoken further upstream, along the Bomokandi river in the Poko and possibly the Bambesa Zones. According to Ethnologue, Kango is "apparently a pidginized language". Ngelema is reportedly spoken in the Banalia Zone (Banalia-Bangba and Baboro collectivities) and the Basoko Zone (Wahanga and Bangelima de Mongandjo collectivities).

2.2.2. Neighbouring languages

Ngbandi is spoken to the north of the Pagibete area in the Businga, Yakoma, and Mobayi Zones. There are also some Ngbandi communities in the Aketi Zone, near the Bati. Mbanza is spoken in the western Businga and Mobayi Zones. Both are non-Bantu languages (Adamawa Eastern Branch of Niger-Congo family).

Zande is spoken in the northern part of the Bas-Uélé Region. It is the majority language of the Bondo Zone (to the north of the Mobenge-Mondila collectivity) and is the primary language of the two northeastern collectivities of the Aketi Zone (Avuru-Gatanga and Avuru-Duma). It is even spoken in one collectivity of each of the two major Bua-speaking zones: the Nguru collectivity of the Buta Zone and the Mange collectivity of the Bambesa Zone. It is the principal language of the Ango Zone, and is spoken in parts of the Poko Zone, including one collectivity which shares a border with the part of the Bambesa Zone in which Bua is spoken.

Thus there are Zande neighbours for each of the Bua bloc dialects spoken in the Bas-Uélé Subregion: on three sides of the Benge, between the western and eastern Bati areas, between the Bati and the Bua, to the north of the Bua in the Buta and Bambesa Zones, and to the southeast of the easternmost Bua. Barambo is spoken just north of the last-mentioned pocket of Zande, that is, in the Poko Zone to the east of the Bambesa Bua for over 100 kilometres along the road. Zande and Barambo, like Ngbandi and Mbanza, are Adamawa-Ubangi languages.

Another non-Bantu language in the area is Makere, a Central Sudanic speech variety (Nilo-Saharan language family) related to Mangbetu. Situated in the southern Bambesa Zone, it is south of the Bua dialect (Bambesa Zone) and east of the Yew dialect (Buta Zone).

According to our research, five Bantu languages are spoken in areas contiguous to the languages under study. Of these, three cannot reasonably be considered part of the wider Bua-Pagibete group: Ngombe (called Libale along the Dua river near the Pagibete-Omveda), Genza/Benza (using the former name in the Lisala and Bumba Zones, south of the Pagibete; using the latter name in the Aketi Zone, among and to the south of the Bati), and Budza (in the Bumba Zone, also neighboring Pagibete and Bati).

Two others seem to be the next-most similar languages to those in the Bua bloc: Kango (along the Uélé river among the Bua and the Zande) and Ngelema (to the south of the Bati and Monganzulu (Buta) dialects). Although word lists were collected in these speech forms, they were not in focus in the present survey.

According to the maps, languages neighbouring Ngelema include Bua and Benza to the north, Popoi (Central Sudanic) to the east, Bali (Bantu D.21) to the southeast, Mba or Manga (Adamawa-Ubangi) to the south, and Babango (related to, and possibly a dialect of, Budza) to the west. There may be dialect variation in Ngelema, reflected in the alternate names Hanga in the Basoko Zone (west) and Angba in the Banalia Zone (east).

2.2.3. Population

According to census figures from 31 Dec 1993 provided by the zonal offices in Bambesa, Buta, Aketi, and Bumba and an analysis of the data cited by Fultz and Morgan, we suppose there to be at least 20,000 (probably more like 23,500) Pagibete and at least 200,000 Bua in the broadest sense. The latter figure divides into approximately 70,000 of the Bua dialect (i.e. resident in the Bambesa Zone), 80,000 of the "Buta" dialect (Yew and Monganzulu), 25,000 Bati, and (with no data available from the Bondo and Banalia Zones) potentially 25,000 Benge and Kiba.

It is clear that the Bambesa and Buta dialects of Bua represent the overwhelming majority of the Bua population. Even if there are more than 25,000 Benge and Kiba speakers in the Bondo and Banalia Zones respectively, at an estimated 150,000, the leBoale and liYew-leNganzulu speakers far outnumber the speakers of liKiba and leBenge-leBati.

The survey team did not visit the Mongwapele *groupement* (where Fultz & Morgan's Dialect I is spoken), nor the villages of the Yakoma and Bumba Zone where Pagibete is spoken (home of Dialect III and location of the village of Butu). Furthermore, there is no recent population data for the Pagibete people except for the Bumba Zone (data collected

on the present survey) and Businga town (a total of only 200, according to residents interviewed on site by JeDene Reeder and Sharon Stoothoff (personal communication)). Nevertheless, I wish to suggest a reinterpretation of the data cited by Fultz and Morgan 1985, as follows.

It will be recalled (section 1.3.2) that according to Fultz and Morgan's analysis of 1984 census figures, there were about 25,000 Pagibete people in Equateur, of whom there were

- 4,200 residents of the Mongwapele *groupement*
(therefore speakers of their Dialect I),
- 7,500 in the Ngakpo area (centre of Dialect II),
- 10,400 in the Bumba and Yakoma Zones (Dialect III), and
- est. 2,900 in Businga town (no breakdown by dialect).

Because they break down the population figures by *groupement*, the data can be reinterpreted in light of other information. It appears that they have mistakenly overestimated the population in some cases, by counting a Ngombe *groupement* (Babale) as speaking Dialect II and by counting a Ngbandi *groupement* (Mongende-Sud) as speaking dialect III. They may have underestimated the population in another case, counting 4,481 Pagibete in one *groupement* of the Bumba Zone, whereas the Zonal office claimed there were not only 4,142 in that *groupement* but also 6,031 in four other Pagibete *groupements*, for a total of 10,173 Pagibete in the traditional areas of the Bumba Zone. Finally, it seems that they have misclassified the village of Butu (in the Pombi *groupement*), who confusingly, speak Dialect II though they live in the Yakoma Zone.

We propose the following revised population figures:

- 4,200 speakers of Dialect I, as stated in the 1985 report
- 6,000 speakers of Dialect II, the result of
 - 7,500 as stated in the 1985 report
 - 3,000 for the Babale, who speak a Ngombe dialect
 - +1,500 estimated Dialect II speakers in Yakoma Zone
- 13,000 speakers of Dialect III,
 - 2,800 in the Yakoma Zone (Bondalangi and Bodunga)
 - 10,200 in the Bumba Zone (1993 census data,
in five *groupements*, not 4,500 in one)
- 300 Pagibete speakers in Businga town
(to make total an exact number of five hundreds)

for a total of 23,500 Pagibete in the Equateur Region. If we retain Fultz and Morgan's data rather than mixing census data compiled nine years apart, we may retain the estimate of 4,500 in the Bumba Zone and suppose 500 in Businga (to make the total an exact number of thousands) for a total of 18,000 Pagibete.

The relevant figures for the Yandongi collectivity of the Bumba Zone follow:

Language	<i>Groupement</i>	1993 Population
Pagibete	Bauma	4,402 \
	Benzale-Akambu	633 \
	Bozogi	476 > 10,173
	Doko	520 /
	Ndundu-Sana	4,142 /
Genza	Bokoy	4,138 \
	M-B-A	8,952 \ 26,885
	Mombongo	4,027 > (cf.
	Yalisika	9,082 / 33,782
	Yasongo	686 / in 1984)
Budza	(the rest) total	20,537 in the collectivity (cf. 9,114 in 1984 census)

which makes a total of 57,595 in the traditional areas (*milieux coutumiers*); there are also 5,411 outside these areas and 3 foreigners, which means there were 63,009 inhabitants of the Yandongi collectivity in 1993.

Since the Genza and Budza data were not broken down by *groupement* in Fultz and Morgan's report, it is not clear whether the Bauma, Benzale-Akambu, Bozogi, and Doko *groupements* were omitted altogether or assigned to another language group, or whether they were then part of the Ndundu-Sana *groupement*. Note that Bauma is the hometown of one of their wordlist sources, Ndundu-Sana is the hometown of one of ours, and someone in Ngakpo said that there was a Bozogi *groupement* in the Bumba Zone, and that Pagibete was spoken there.

From the juggling of the 1984 figures only, Dialects I and II constitute at least 60% of the Pagibete population (that is, at least 12,000 of 20,000). However, they would constitute about half the Pagibete population if there are 13,000 speakers of dialect III in the Yakoma and Bumba Zones (as suggested by the 1993 data).

2.3. Classification and Dialects

2.3.1. Review of the literature.

2.3.1.1. Linguistic Survey of the Northern Bantu Borderland (LSNBB)

The term "Bua Bloc" comes from Van Bulck and Hackett (1956:74-81). In their view, the "Northern Group" of Bantu languages in the border area between Bantu and non-Bantu languages of what is now Zaïre includes: Kunda, the Ngombe Bloc, the Bua Bloc, and the Bakango (or fishers). The Ngombe Bloc is comprised of Ngombe, Budja (our term: Budza), and Binza. Lika and Bali, according to Van Bulck and Hackett, are "Extreme North Group transition languages".

From this expression we may gather that they saw Lika and Bali as sharing some traits with, for example, the Bua bloc languages, as well as with the languages they call the "Extreme North Group" (pp. 74f). These languages, many of which correspond to the "Kari group" in Bryan (1959:86f) and Voegelin and Voegelin (1977:60), would seem not to have been universally accepted as Bantu. Though still listed in the Ethnologue (Boguru, Gbati-Ri, Kari, Mayeka, Ngbee (Mangbele), Ngbinda and Nyanga-Li in Zaïre, Bodo in Central Africa, and Homa in Sudan), they were all very small groups in the 1950's and to all evidence, these speech varieties are now all extinct or practically so. One tantalising point is that some of these languages have/had "vestigial suffixes", as is evident in their names (liKari-li, Gbati-Ri, Nyanga-Li; compare "leBoa-le"). As has already been mentioned, some noun classes in the Bua bloc languages and Lika (but not Bali) still have apparent class suffixes.

The Bua Bloc, according to Van Bulck and Hackett, can be divided into six groups; Apagibeti, Bengé-Baati, a "transition group" (liNgengita-liNganzulu-Kiba), Yewu, liBwali, and Pseudo-Ngelima. (There are apparently no true Bangelima, only Pseudo-Bangelima!) Some of the other names which they cite in connection with these six groups are:

- for Apagibeti: eGulu-eBugbumba-Gezon in the northeast,
Bodjame-Bonzwaambi-Momongo-Mongwapere
- for Bengé-Baati [Napagisene]: Boganga, liGbaase, liGbe
- for the "transition group" [Napagibetini]: liNgingita,
leLisi, Gbaasa, liGanzulu, Kipa, other Bwa
- Yewu [= Napagibetini]
- for liBwali [= Napagitene]: Kete, Kapu, Kiba, Longwa,
Dongbale, Gongeya
- for Ngelima: leBoro, leAngba, leHanga, leBenja, leGenza

NOTE: Here I have changed the capitalisation but not the spellings used by Van Bulck and Hackett (pp. 79-80). Some of the names correspond to collectivity names (see below); elsewhere in the report, I follow the spellings of the collectivities.

Of the six groups, the first two and the last two are recognisable as what we are calling Pagibete, Bengé-Bati, Bua Proper, and Ngelema. It is hard to critique their grouping of the Buta and Banalia Zone baBua. Different interviewees distinguished the Buta Zone dialects differently, yielding no consistent division of what we finally simply called "Bua of Buta" into true dialects. It seems that a linguistic identity is imputed to clans, so that while not ruling out the existence of separate dialects (e.g. liYeu, liNganzulu, liGbe), we rearrange

Van Bulck and Hackett's third and fourth groups (both "called Napagibetini") according to Zone, i.e. we refer to "Bua of Buta" vs. leKiba.

Cope, for one, is impressed by the practice of designation (if not classification) by speech-initial formula as supposed evidence of the difficulty of sorting out the dialects of the area (1971:222). Fortunately, fewer difficulties were encountered on the present survey. Yet, as in the fifties (when the LSNBB was conducted), the only widely accepted cognomen for the westernmost group was a variant of the speech-initial formula.

There are five relevant collectivities in the Buta Zone. They are Barisi-Mongengita-Bakango, Bayew-Bogongia, Bayew-Bokwana, Mobati, and Monganzulu. Reportedly, Yew is spoken in the first three and Gbe in the Mobati collectivity. "Bogongeya" was one of the six peoples whom Van Bulck and Hackett said speak liBwali; the other five match up with the five Bua-speaking collectivities of the Bambesa Zone. By the way, the glossonym "Mbili", used by Girard (1991), does not appear elsewhere in the literature, and may be a back-formation from the town name "Bambili".

Pagibete is spoken in four *groupements* of the Businga Zone. They are: Bozame, Bonzwambe, Bomongo, and Mongwapele. Two *groupements* of the Mongwandi *groupement* of the Aketi Zone, bordering Equateur, are called Boguru and Bogboma. We did not get to visit this area, but were told that Bati was spoken there. Other Bati-speaking *groupements* include Bobwasa (in the Bodongola collectivity), and Boganga (in the Bongi collectivity).

It seems that three of the collectivities in which Ngelema is spoken are: Baboro and Bangba (presumably Ba + Angba) in the Banalia Zone, and Wahanga in the Basoko Zone. The area of the LSNBB map marked "Bindza" is a fourth collectivity, Bangelima de Mongandjo. The Bagenza mentioned by Van Bulck and Hackett supposedly live in the Mabinza collectivity of the Aketi Zone. These Bindza/ Benja/Genza should not be confused with their ethnic kin who speak a Ngombe-like language (op. cit.:77f; see also LIBINZA, LIGENZA, NGOMBE in Grimes 1992a:422f, 428).

2.3.1.2. Guthrie and Bryan

According to Guthrie's classification of the Bantu languages (as presented in Bastin 1978, pp. 140f), Bua, Bengé-Bati and Ngelema are part of the Ngombe group (C.40), together with Ngombe and Bwela. (Bwela, also known as Lingi, is a mysterious language which is not localised on any map, though it may be the same as Doko, found in the Lisala Zone according to the LSNBB map. Van Bulck and Hackett place it among the "Western Group transition languages", not with Ngombe). Budja is classed with Lingala in the "Bangi-Ntomba" group (C.30). "Bango" is listed as an alternate name for Bua and for Bali (D.21); this latter is also called "southeastern Bwa".

In Guthrie 1948, certain languages of Zones C and D are considered to be "Sub-Bantu". These are: Mangala, Bua, Angba; Bali, Amba (but not Komo!), Peri, Bira, and Huku. The reason for this label is "the deficiency of their grammatical agreement systems" (Cope 1971:221). It should be noted that in 1948, Guthrie had not yet attached the "Sub-Bantu" Bua and Angba to the Ngombe Group; on the other hand, Bati was classified as full Bantu language in the Ngombe Group. Years later (Guthrie 1971:42), Komo was duly characterised as Sub-Bantu, but Huku, now more appropriately called Nyali, is not so characterised (p. 43).

The case of Lingala is easy to understand: as a language of wider communication, especially among people whose first languages are non-Bantu, it is generally used in a reduced form. The others may be divided into those with several noun classes (but perhaps incomplete agreement systems) and those meeting Guthrie's stricter criterion (1948:19): that plurals are either identical to the singular, or formed by prefixation with "ba-". In the first category are Bua, Ngelema (Angba), Bali, and Nyali. Though not mentioned by Guthrie, Pagibete, (Benge-Bati,) Lika, Budu, Ndaka, Mbo, and Vanuma ("Southern Nyali"), would also fit in this category. In summary, in our terminology, these are the Bua Group, Lika, Bali, and the Budu Group. The second category apparently consists of the "Komo Group": Bira, Amba (or Humu), Bila, Kaiku, Bhele (Peri), and Komo.

Bryan (1959:88-91, 101f) groups these Sub-Bantu languages as we do: Bali Group (Lika and Bali), Bira Group (our Komo Group, although she says that it may also include Lengola and Mituku), and Nyali Group (our Budu Group). Unfortunately, the Eastern Zaïre Group Survey Department copy of Bryan 1959 is incomplete; her treatment of our Bua Group (as well as the other Bantu languages of Equateur) can only be inferred from Cope (1971:221). Apparently she follows Guthrie in counting Bua and Ngelema as part of the Ngombe Group. Concerning the conclusions which Cope draws from this, see the next section.

2.3.1.3 Doke (Cole) and Cope

Doke's classification, as reproduced in Cole 1959 (p. 201), places "Bua (Bwale)", with dialect names "Ngelima and Lebeo" (which others treat as one language, separate from Bua), and "Bira (Kumu)" in the Congo Zone. There are five other groups in the zone, which apparently include most of the languages of the Zaïre river basin. The two groups which neighbour the Bua and Komo groups are called Northeast (comprising Lokele and So, languages spoken downstream from Kisangani) and Middle Congo. Five clusters comprise the Middle Congo Group: Poto, Ngombe, Mongo, Ngala, and Bangi (Yanzi). "Buja" is a dialect of Ngala!

Amazingly, it seems that Bira/Komo are the only languages listed in Cole which Guthrie's classification places in Zone D. This is just one example of the gross incompleteness of the system. However, this classification properly shows an affinity between languages placed by Guthrie not only in different zones, but even in different areas. Ngombe and Budza are more closely associated in Doke's classification than either is with Bua.

Cope (1971:217-219) takes the liberty of revising Doke's zones and dubbing them "areas", to distinguish them from Guthrie's zones. His revision includes a rearrangement of the Northwestern and Congo Zones, resulting in the assignment of many languages, including Ngombe, Budza, Bua, and Komo, to the Northwestern Area and the creation of a Western Area (consisting of languages nearer the mouth of the river) in place of the former Congo Zone.

Cope appears to find great significance in the east-west split between the Bantu languages of northern Zaïre. He observes that Doke omits most of the languages classified "D" by Guthrie: Konjo (D.41) and several languages in the Ruandi-Rundi Group (D.60) are in the Northern Zone, and Bira (D.32), together with Komo (D.23, but now called D.37; Bastin 1978:142), are in the Congo Zone. Cope has the choice of placing the remaining "D" languages in the Congo Zone, with Bira (which is to the east of practically all of them), or the Northern Zone, with the interlacustrine Konjo, Rundi, and Rwanda (now classed as "J"; see Bastin). In spite of the fact that the Northern Zone in Doke's map (Cope 1971: 217) only just includes the Great Lakes area and leaves the interior of the country to the Congo Zone (or, further south, to the Central Zone), and his admission that we may be sure that Doke would have placed the languages of D.10 and D.20 and probably also D.50 in the Congo Zone (p. 219), he concludes that "these languages surely belong in the Northern Area" (p. 221).

The reason is apparent: he has assumed that Doke's incompletely catalogued zones should conform to Guthrie's areas. Guthrie's split between C and D is treated as authoritative, even though Doke's data contradict it. He finds great significance in the fact that Bryan 1959 does not revise this part of Guthrie's classification.

Bryan, by dealing with Zone C on pp. 33-55 and with Zone D on pp. 88-100, confirms by the gap from p. 55 to p. 88 the linguistic and geographic distinction not only between the two zones but also between the two areas. (Cope 1971:221)

Such a "confirmation" is essentially an argument from silence: she didn't challenge that part of Guthrie's system, therefore she must have had a good reason. The gap, by the way, is the reason the Eastern Zaïre Group does not have a copy of the part of Bryan's book which deals with the Bua bloc.

In fact, according to Bastin (1978:132), by the time Cope was writing, Meeussen and Doneux each had proposed a division of D— and E— similar to the ones suggested by Doke's map, by means of a new zone corresponding fairly closely to the western

(contiguous) part of Doke's Northern Zone. Thus, Cope missed a wonderful opportunity to suggest that the eastern third (or half) be stripped from Zone D and kept in the Northern Area, and that the remaining Zone D languages be counted (provisionally) as part of the Northwestern Area. Within a few years, says Bastin, the reality of Meeussen's and Doneux' proposed new zone was confirmed by statistical correlations by Coupez and Bastin. Formed from Guthrie's D.40,50,60 (with the exception of Nyanga, Bembe, and Buyu) and E.10,20,30, it is now called Zone J.

Cope (p. 222) admits that the northern Zone D languages and the north-eastern extremity of his newly-delimited "Northwestern Area" (C.43-45, i.e. our Bua Group) constitute "without doubt the most difficult in the field, because of the conglomeration of Sub-Bantu and transitional languages". However, he upholds the division between Guthrie's C and D, with the hesitant reassignment of D.11 Mbole to the Mongo Group (C.60); his assignment of C.43-45 and Kango to the Northwestern Zone, he says, is made "most hesitantly" (loc. cit.). Such a choice would have been avoided had Cope noticed that the weak link in the chain C - northern/western D - eastern D - E was within Zone D. Interestingly, his final classification recommends the splitting of Zones E and L, which shows that he might have been convinced to split Zone D.

Our judgment with the benefit of hindsight should not be too severe. Of all the decisions Cope makes, this is the only one he marks as problematical on his map (p. 232). His final word on the choice is as follows:

There are Sub-Bantu and transitional languages on both sides of the boundary, which may or may not be satisfactory-- only a specialist is in a position to express an opinion.

Though we cannot demonstrate that the languages of the present Zone D (Guthrie's original zone, minus the languages of Zone J) belong more with Zone C than Zone J (and thus should be placed in Cope's Northwestern Area), on a lower level, it is hard to accept such a cleavage between the Bua Group and the languages of Zone D.

2.3.1.4. Other sources

Two other sources follow Guthrie for the most part but are of interest for the number of varieties they cite.

In the Ngombe group, Voegelin and Voegelin (1977:64-69) list not only Benge-Bati, Boa, and Angba (Ngelima), but also "Yewa" (which is almost certainly an erroneous reading for Yewu) and Kango. "Apagibeti" (our Pagibete) is mentioned as a Bati dialect, with alternate name Gezon. Boa and Yewu are both also called "Napagibetini" (as well as Bali and Bango!). Bwela is listed under the name 'Doko, and Binza and Genja are mentioned as Ngombe dialects.

Voegelin and Voegelin list Liko (our Lika) as a "Lega-Kalanga language", along with Bali, Komo, and others. As in Bastin, Bali is characterised as "southeastern Bua".

The Linguistic Atlas of Zaïre (Kadima et al. 1983) lists the following languages as belonging to Zone C: leángbá (lebeo, kingelima), apáákibété, lebáátí, lebéngé, lebóá (lebóale) [four dialects], and lególó; as well as ebango, libinza (libenza, kibenza) [code 310 in Haut-Zaïre, distinguished from another libinza, code 309 in Equateur], embujá (kimbudza), and lingombe [six dialects, including libwela and ligenza]. Also classified as a Zone C language is libaalí, presumably because of others' equation of Bua and Bali. However, lilikó (kilíka) is listed as belonging to Zone D, as is kekómo (kikúmu).

2.3.1.5. A synthesis

The table compares these names with our perception of the relationships.

Table 3: Some Bantu tongues of northern Zaire

<u>Variety</u>	<u>Ethnologue</u>	<u>Guthrie/Bastin</u>	<u>V&V</u>	<u>ALZ</u>
ANOTHER...				
Libinza	LIZ	C.31...	(NW067?)	309
(not to be confused with Benza/Genza)				
BUDZA GROUP				
Budza	BJA	C.37	NW028	329
Babango	BBM	(misclassified?)		307
NGOMBE GROUP				
Ngombe	NGC	C.41	NW100	334
- Bwela, Doko?	BWL	C.42	NW035	yes
Benza/Genza	LGZ	--	(NW100)	310
BUA GROUP				
Pagibete	PAG	--	(NW013)	302
Bua cluster	BWW	:	:	:
Benge-Bati		C.43	NW013	:
Benge	yes			308
Bati	yes			304
Legolo?				314
Bua east		C.44		311
(see next table)				
Kango	KTY	--	NW048	---
Ngelɛma	AGH	C.45	NW002	301
OTHERS...				
Lika	LIK	--	NE051	417
Bali	BCP	D.21	NE002	303
Komo	KMW	D.23/37	NE047	412
(all actual or potential "Lega-Kalanga" languages)				
"Dialect"	<u>Ethnologue</u>	<u>Guthrie/Bastin</u>	<u>V&V</u>	<u>ALZ</u>
Bua east	(part of BWW)	C.44	:	311
(as distinct from Benge-Bati)				
Bambesa Zone			NW016	
Buta Zone				
Gbe				yes
(in Mobati coll.)				
Nganzolo				yes
(own collectivity)				
Yew	yes		NW146	
Ngingita	yes			
Balisi				yes
(Yew in three colls., including Barisi-Mongengita-Bakango)				
Banalia Zone				
Kiba, of Kole				yes

In other words, for Bua, the least fine distinction is in Ethnologue (Benge-Bati considered part of Bua), the next finest is Guthrie's (that two-way division), followed by Voegelin & Voegelin (three-way division: Benge-Bati vs. Bua vs. Yew), followed by the Linguistic Atlas of Zaïre (four-way division: Benge vs. Bati vs. Legolo vs. Bua). However, Ethnologue has separate entries for Pagibete (unlike anyone else but ALZ among the sources we cite) and Kango (unlike anyone else but V&V).

2.3.2. The evidence of the present survey

2.3.2.1. Lexicostatistical analysis

A computer database was prepared using the WordSurv program (see Wimbish 1989), entering new lexical data in the three Pagibete varieties, Benge and Bati, and the Kiba, Yew (Buta), and Boa-le (Bambesa) dialects of Bua; data collected in Kango-Uélé, Benza and Ngombe during the survey, and Budza afterward; data in Babango, Ngombe, and Libale

from Fultz and Morgan's data notebooks; and data in a number of other eastern Zaïre Bantu languages, including Komo, Lika, and Bali, from the Eastern Zaïre Group files. There were 300 records in the database, including a record for most of the glosses on all of the previously elicited lists. Since the lists varied among 120, 150, 200, 250, and 280 items in length, some of the similarity calculations are based on many comparisons while others are based on relatively fewer. A fuller treatment of the lexicostatistical analysis may be found in Appendix H.

Lexicostatistical analysis of these data indicates that Babango (spoken in the Basoko Zone of Haut-Zaïre) is most closely related to Budza. Benza, or Genza, spoken in parts of the Bumba Zone (among the Budza) and of the Aketi Zone (between the Bua and the Bati), is most closely related to Ngombe. Each of these four languages is at least 50% similar to the others, and between 35% and 50% similar to Lingala. Libale, spoken along the Dua river near the Omveda dialect of Pagibete, is quite similar to Ngombe.

Budza, Babango, Ngombe, and Benza show approximately 40% similar vocabulary with Pagibete, Bua, and Ngelema. Their similarity to several other languages of eastern Zaïre, including Lika, Bali, and Komo, is also between 35% and 40%. It would seem to be more appropriate to associate Budza-Babango and Ngombe-Benza and to dissociate them from Pagibete-Bati-Bua-Ngelema, rather than retaining the classification of the Bua bloc as part of the "Ngombe group".

Each of the eight Bua bloc dialects in which 300-item word lists were compiled (three Pagibete, Bengé, Bati, and three Bua) is at least 80% similar to the others. They are all between 70% and 80% similar to Kango, as spoken along the Uélé river in the Bambesa Zone, and between 60% and 65% similar to Ngelema. Similarity to Lika is between 55% and 60%. (Oddly, however, Ngelema and Lika show less than 50% similarity.) The languages represented by these eleven lists (Pagibete, Bua (cluster), Kango, Ngelema and possibly Lika) may for convenience be called the "Bua group" (a replacement for the eastern part of the former Ngombe group).

Bali and Komo are between 40% and 50% similar to each of the languages of the "Bua group" so defined. Bali is over 50% similar to Lika. It could be argued that the "Bua group" should therefore be considered to be east-central Zairean Bantu languages. In other words, as far as lexicon is concerned, there is as much reason to group Pagibete and Bua with some of the "D" languages as with the closest of the "C" languages.

Heretofore, as part of the "Ngombe group", the Bua bloc languages have been counted as languages of the "North Western Area" of Bantu (comprised of Guthrie's Zones A, B, and C). This is one of six areas posited by Guthrie; another is the "North Eastern Area", comprised of his Zones D, E, and F (which are now called Zones D, J, E, and F). Thus, the decision to assign Bua to the Ngombe group or to a new group associated with, say, Komo or Bali, has high-level classificatory ramifications, beyond the identification of separate languages, or the joining of languages into "groups" or even into "zones". In fact, as far as Ethnologue (Grimes 1992a) is concerned, the ramifications are even greater, since there the contrast between "Northwest" and the other five areas (comprised of twelve other zones) is treated as the primary division of the Bantu languages.

In fairness, however, it should be pointed out that Guthrie proposed Zone D to fill in the gaps between Zones C, L, E, and F. "There are reasons for not placing any of these groups in the neighbouring zones, but few, apart from geographical contiguity, for making a zone out of them." (Guthrie 1948:40) Furthermore, of Zone C he says: "On the west and south the limits of the zone are well defined, but the eastern boundary is somewhat arbitrary, although the languages just over this boundary are sufficiently different to justify their being placed in another zone." (op. cit., p. 36)

2.3.2.2. Other linguistic features

Guthrie remarks that some of the words which occur in Zone D languages "appear to be related to those in languages to the west rather than [to those in languages] to the east." (op. cit.:40) What we have noticed, however, goes beyond that: Komo, Bali, Lika, etc. (in Zone D) have a higher proportion of similar vocabulary with Pagibete and Bua (supposedly in Zone C) than do certain Zone C languages.

We should remember that lexicon is only one of several areas in which linguistic similarity is reflected. It has been observed, however, that except in the case of pidginised languages (in which the lexicon of one language is overlaid on the structure of another, often quite different, language), lexical similarity is usually an indicator of similarity in other areas. Thus, although computer-assisted adaptation of texts between related languages (CARLA) does not depend on the languages' having a high proportion of similar vocabulary, it has been found that when there is strong lexical similarity, CARLA is generally feasible (Bill Mann, personal communication).

Still, a few comments can be offered on the classificatory role of phonological and morphological features observable from the data collected.

- According to Guthrie 1948, Zone C languages have a seven-vowel system, as do all the Zone D languages except those of the Rwanda-Rundi group, which have five vowels. These five-vowel "D" languages have since been reassigned to Zone J. According to Guthrie 1971, Bangubangu also has a five-vowel system; however, Bastin (1978:142, 147) recommends transferring this language from Zone D to Zone L. In D.30, Guthrie 1971 only mentions the number of vowels in Nyali: supposedly five.

However, recent research has led SIL to believe that at least seven languages in the "northern Bantu borderland" (corresponding to the Bira-Huku and northern Lega-Kalanga groups of Zone D), historically seven-vowel languages, now have developed nine-vowel systems. These include Lika and Bali; the others are Budu, Nyali, Vanuma, Ndaka and Mbo. (Of these, Guthrie only mentions Bali, as a seven-vowel language, and Nyali, as a five-vowel language with contrastive vowel length.) Four others, Komo, Bhele, Bera and Amba, are known to have a seven-vowel system. Another, Bila, has an interesting hybrid system: there are seven contrasting vowels in noun roots but nine vowels with cross-height harmony in the verb (Kutsch Lojenga 1994).

The phonology of Kaiku, which is lexically most similar to Bhele and Bila, has not been studied. There is disagreement concerning the number of vowels in Lengola. Guthrie placed it in the Enya group (apparently on geographic grounds as much as anything), which would imply that it has seven vowels (Guthrie 1971:42). However, a word list gathered in Lengola in 1990 by Bettina Gottschlich, of the Budu translation project, distinguishes nine vowels, and Stappers 1971 only distinguishes five.

Since in the Bua Bloc languages and Ngelema we find what appears to be a seven-vowel system, we do not have phonological evidence for grouping them with Lika and Bali. Budza, Ngombe, and other Bantu languages of Equateur, including Lingala, have seven vowels; on the other hand, so does Komo.

- Another characteristic of Zone C languages, says Guthrie (1948:37), is a lack of contrast between /k/ and /g/ except as the first consonant following a nasal. In Zone D languages, on the other hand, these phones are usually contrastive. (Later, Guthrie mentions that the phonemic contrast of voiced and voiceless is neutralised after a nasal in groups 40 and 60. These are now classed as Zone J, except for Nyanga, which retains the contrast, according to Kahombo 1969 (see for example, pp. 68, 74).)

Our data in speech varieties of Equateur, other than Pagibete--Lingala, Budza, Ngombe and Benza (supplemented by the data of Fultz and Morgan in Babango and Libale)-- support this characterisation of Zone C languages. Generally speaking, Lingala words with (unprenasalised) /g/ are rare, and for our corpus the same holds true for the other languages. Even when other languages have a cognate form with (unprenasalised) /k/, these languages generally have /g/ instead. Guthrie 1971 (pp. 39f) seems to say something different: for many of the Zone C languages, k/g in the proto-form are realised as (zero)/k. Perhaps since his attention was focused on "Common Bantu" (proto-forms with wide applicability to all of Bantu), his phonetic descriptions do not take regional vocabulary into account.

Although all the obstruents (p, t, k, kp, b, d, g, gb, s, z, j) may be prenasalised in Lingala (the double stops more rarely), many speakers do not pronounce the nasal before voiceless consonants, including /k/. As a matter of fact, in some languages of Zaïre, prenasalised voiceless consonants are rare, and possibly nonexistent in some.

The tendency not to pronounce the nasal before voiceless obstruents is also reflected in the lexical data for the other languages of Equateur; for example, the word for 'hoe' ("nkongo" in Lingala) is pronounced "kongo" in Budza and Ngombe (and in western Pagibete). On the other hand, in eastern Pagibete, all the Bua cluster dialects surveyed, as well as Ngelema and Benza (of Haut-Zaïre), the old Bangala word "gita" is used. Observe the word-initial /g/.

There is ample evidence of a /k/-/g/ contrast in Pagibete, Bua (east), Kango, Ngelema, Lika, Komo, and Bali, in a variety of environments (prenasalised, root initial, noun root medial and verb root final). The contrast between /k/ and /g/ in Bengé and Bati is neutralised in some environments (but not after a nasal). There are consistent correspondences /ng/-/ng/ and /g/-/k/ between Pagibete and other Bantu languages of Equateur, such as Ngombe and Budza. There are no cases of /ŋk/ (prenasalised /k/) in the corpus for the Bua bloc languages or for any language of Equateur other than Lingala. There are, however, a few cases of a syllabic /ŋ/ before a root beginning with /k/ in Pagibete and Bengé-Bati.

Since unprenasalised /g/ seems to be almost as rare in Budza, Ngombe, etc. as in Lingala, but more common in the Bua Bloc varieties (apart from Bengé-Bati), as it is in many of the Bantu languages of Haut-Zaïre and Maniema, we have some phonological evidence for an affinity of the Bua bloc to languages of eastern Zaïre.

- Guthrie states that in Zone D languages, the infinitive prefix is generally ko- (class 15); Zone C languages have a variety of prefixes, but "none appears to have... ko-" (Guthrie 1948:38). Lingala, however, is but one of several apparent exceptions, and this is noted in Guthrie 1971:40. (Other classes cited for Zone C infinitives are: 5, 7, 9, 11, 14; some with suffixes.)

For selected data analysed for this survey report, we find that in the Mongala area of Equateur, the infinitive prefix is 6o-, which may indicate that infinitives are in class 14 (data in Ngombe and Benza, in contradiction to Guthrie's pointing to class 9 for Ngombe; Budza has the prefix wo-, where 6/w is a regular sound correspondence, and the suffix -ke). Southern Lega-Kalanga languages have ku- (data in Zimba and Lega-Shabunda; Lega-Mwenga has ?u-, where k/? is a regular sound correspondence). The Budu Group languages and Lika, and also apparently Enya and Mituku, have kV-, where the vowel is determined by rules of vowel harmony. This prefix may point to class 15 or class 12 (but not class 5, as Guthrie claims for Nyali). Bali has u-, and Bembe, Nyanga, and Lengola apparently have (r)i- (class 5; Guthrie 1971:43, Kahombo 1969:75, Stappers 1971:290).

The Bua bloc languages, plus Ngelema and Kango, have a zero infinitive prefix, as do the languages of the Komo group (which could mean class 9 or, in Sub-Bantu, classless). These languages are thus distinguished from most of the other languages of Zones C and D.

- According to Guthrie, the class 7 prefix is almost always e- in Zone C languages, rather than ke-. This is true for Lingala, Budza, and Ngombe. It is also true for Pagibete, the Bua cluster and Ngelema. Similarly, within their nine-vowel systems, Bali and Budu have the equivalent prefix, I- (where the capital letter refers to the high front (unrounded) vowel with the same ATR value as the root). Lengola, however many vowels it has, belongs with the languages already cited. Komo, of course, has no class 7, and for Lika, see below.

In contrast, Mituku has ke- (vowel?), Enya has ce-, and Nyanga, Zimba, and Lega-Shabunda have ki- (Lega-Mwenga and Bembe "?i-", where there is a regular correspondence between Shabunda "k" and "?" in these tongues). Interestingly, on this point, not only do the Bua Group languages belong with Zone C, so do Bali, the Budu Group languages, and possibly Lengola. Since the Komo Group languages are Sub-Bantu, they could belong with Zone C or with the Lega Group.

- Guthrie (1948:40) comments on the unusual genders in some of the northern D languages, such as Bali. In his later work, he gives more examples (though actually fewer for Bali). Of Bua, he says there are "several genders with skewed prefixes" (1971:40). He lists unusual "genders" for some of the (other) C languages also, including Ngombe (loc. cit.), though it is Bobangi which takes the prize, with nine class pairings listed (p. 39). Several unusual genders are cited for Nyali (p. 43), many of which are also found in Budu, although on the face of things, it is redundant to postulate both 5a "li-" (plural of 5 "i-")

and 8 "di-" (plural of 7 "i-"). Since we are inferring the class system only from singular-plural matchups on the word lists, it is possible that we are overlooking unusual genders in Bua (then again, if it is Sub-Bantu, we may not be overlooking much).

One curious feature of the Bua bloc languages is the existence of class suffixes (in addition to usual class prefixes) in some of the noun classes, namely the singular class 7 (and in Pagibete, possibly another, either 15 or 14) and the plural class 4 (and often also classes 2 and 8). This phenomenon is not found in Lingala, Budza, Ngombe, Benza (Zone C languages), Kango, Ngelema (considered members of the "Bua group"), Bali, Komo (showing nearly 50% lexical similarity with the Bua bloc), or any other of several Zone D languages for which the Eastern Zaïre Group has lexical data-- except Lika.

Although one may only guess at noun classes using only singular and plural data from a standard word list (even one as long as that used for the present survey), it appears that there are class suffixes for at least two, and possibly as many as four, Lika noun classes. The class assignment of which we can be most certain is class 15 (kU- -kO), which can have plurals in classes 6 (ma-/mo-) or 10 (zero prefix). The next most likely is class 7 (sI- -sO), with plurals in class 8 (6I-) and possibly another class. (The capital letters stand for the vowel of the appropriate height and frontness/roundedness with the same ATR as the root.)

The seven possible "class seven" nouns are cited below. On the strength of the first three, we postulate class eight. The other four could be classes 7 (sI- -sO)/ 10 (ø- -tO), or some other classes, such as 12/13 or 19/13. The corresponding prefixes in Proto-Bantu are ka- (12), to- (13) and pi- (19) (Hinnebusch 1989:466). According to Stappers (1971:262), Lengola has words in classes 12/13 with prefixes si-/to-. It is possible, however, that his "class 12" should have been analysed as "class 19", since this class, marked by the prefix "si-", is apparently attested in some of the Zone D languages. For example, Masumbukowa-Busungu (1979:62, 106) shows that in Lega-Shabunda, class 13 (tu-) is the plural both of class 12 (ka-) and class 19 (si-).

Probable 7/8

'neck'	síngásò	ǂíngò	prob. cognate to kǂngó?
		(Lingala, Bua Group), Bali síngó, Swahili shingo	
NOTE:	9/10 in Lingala and Swahili, analysed as 1a/2 in Bua Group		
'thigh'	sǐbésù	ǂǐbǐ	cognate to Lingala èbèlò,
	Pagibete èbékè, and Bali sǐbélù.		

NOTE: class 7/8 in Lingala, analysed as 7/8 in Bua group;
Komo (mostly classless) kǐbǐ has k- (cf. Bua èbǐ)

'year'	sǎángásù	bǎngà	(apparent cognates:
	Komo ǎngá, Bua-east sǎngá, Pagibete èkyàngákè,		
	Lingala & Ngelema èlángà, Ngombe cluster (e)yangà)		
NOTE:	class 7/8 in Lingala, Ngombe, Pagibete; 1a/2 in Bua?		

Possibly 7/10 (?) (or 19/13 or 12/13)

'thorax'	sǐkúbá:sò	kǐbátò	(possibly cognate to 'chest':
	Lengola ègumbà (class 7/8), Swahili kifua (7/8),		
	Enya nǔkúǂà (cl. 9/10), Lega ǂùkúǂà (14/6?))		
'skin'	(s)ǐbòkúsò	bòkútò	[no known cognates]
'feather'	sǐsàsásù	tàtátù	[no known cognates]
'tree'	(s)ǐbùkúsò	bùkútò	[no known cognates]

There is room for doubt concerning the apparent suffixes. There is a morphophonemic change [s] → [t] in the root for 'feather'; this could mean that there is no "-tO" suffix at all, an argument for placing the plural in class 10. But if the "-tO" is not a suffix, perhaps the "-sO" is not a suffix either, in which case the singular may not be class seven. The existence of a "-tO" suffix would be a strong argument that the plural is class 13; it is a matter of conjecture whether the singular ("sI- -sO") is better placed in class 7 (avoiding the need to posit another class), 12 (expected singular for class 13, rather than or alongside class 19) or 19 (phonetically more similar to sǐ- than to ka-).

That only Lika and the Bua bloc languages of the C/D languages surveyed have class suffixes may be meaningful for classification and raises the questions: Did these languages historically have class suffixes, or did the suffixes evolve? If the former, why only these, and why did Ngelema lose them? If the latter (which is more likely), did Lika and the Bua bloc languages innovate separately, or did the suffixes appear while there was still contact between their speakers?

- Guthrie also refers to differences in noun phrase and verbal constructions which may distinguish Zones C and D. However, the present survey does not provide adequate grammatical data to apply them to the question of classification.
- Of course, Guthrie's own characterisation of some languages as Sub-Bantu, though beyond the scope of the present survey, points to a link between the Bua bloc and languages to the east not shared with any of the languages to the west. (Guthrie 1971: "Boa has concord of class 1/2 only" (p. 40); similarly for Amba, and "Komo has no class agreements" (p. 42).)
- Another area for research in northern Zaïre Bantu languages is whether there are separate roots for the substantives associated with certain verbs. The Bua bloc languages seem to have separate roots more often than other languages. To avoid comparing nouns with verbs in lexical similarity analysis, it is good to elicit both, e.g. "to dance a dance". Komo, Budza and Lingala seem most likely to have the same root in both the noun and the verb, i.e. for the noun to be derived from the verb. From the partial data at hand, we observe the following:

<u>gloss pair</u>	<u>same root</u>	<u>different roots</u>
dream, to dream	Lingala, Swahili, Bali, Budu	Budza, Bua bloc, Lika, Komo, Lega
fear, to fear	Budza, Lika, Bali	Ling., Ngombe/Benza, Swah., Bua bloc, Komo, Budu
laugh, to laugh	Lingala, Budza, Swahili, Komo	Benza, Bua Group, Lika, Lengola, Enya
journey, to walk	Budza, Komo, Budu	Lingala, Swahili, Bua bloc, Lika, Bali
dance, to dance	Lingala, Budza, Komo, Lika, Bali, Enya?	Bua bloc, Bangala, others

(usually, when a language is not mentioned for a given gloss pair, it is a sign of missing data)

However, there are cases in which even the Bua bloc data have derived nouns: weeping (to weep), song (to sing), game (to play).

One final remark on Van Bulck and Hackett's famous classification by speech formula: Apart from "Pagibete" (which isn't even a complete formula), none of the related terms cited in the report of the Survey of the Northern Bantu Borderland was offered in the course of the present survey. In fact, one Bengé man whom we interviewed said that "Napaki bete" was the way he said 'I say that...'—NOT "Napagi sene" or "Napaki sene" (as one would expect from SNBBL). Speculation is more entertaining than helpful, but it may be that the "sene" of Van Bulck and Hackett's formula is the Zande greeting, once reportedly used as a greeting throughout the Bangala-speaking area.

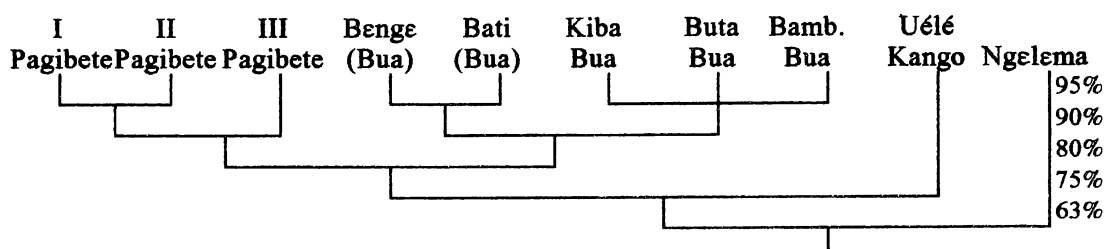
2.3.2.3. Summary

The present survey confirms the existence of a Bua Group of Bantu languages, consisting of Pagibete, Bua (including Bengé-Bati), Ngelema, and at least some "Kango dialects" (where it is possible that not all the baKango—"fishers"—even speak the same language). Not included in the Bua Group of languages, contrary to one might conclude from alternate names in the literature, are Bali, Bango, or Benza. However, Benza may be an ethnonym among the baNgelema, used for ethnic Benza who have adopted the Ngelema language.

The Bua Group is lexically most similar to Lika (especially), Bali and Komo. (The other members of the Komo Group were not in the database used for analysis.) The Bua Group

has more similar vocabulary in common with the closest Bantu languages of Equateur than with Lengola, Búdu, or the Lega Group.

Within the Bua Group, the lexical similarity relationships may be diagrammed as shown in the Figure. The ten lists are distinct, but those representing two of the Pagibete dialects are extremely similar, as are those for Bengé and Bati. The three "Bua east" lists are also more than 95% lexically similar. The first nine lists are, on average, 63% similar to the Ngelema list and 57% similar to the Lika list, yet the Ngelema and Lika lists are less than 50% similar.



I: Mongwapele II: Omveda

This supports the subdivisions of the Bua Group, including Bua Bloc (similar at 80%), "Pagibete language" and Bua cluster (each similar at 90%), "Bua east" (similar at over 95%), and "Pagibete west" and "Benge-Bati".

Such divisions are also confirmed by phonetic correspondences, e.g. probable devoicing of root-medial intervocalic consonants in "Benge-Bati".

The Bua Bloc languages are seven-vowel languages, as are most of the languages of both Zones C and D. Evidently, some of the border Bantu languages of Haut-Zaïre have developed a nine-vowel system, but none of these are in the Bua Group. A study of vowel correspondences in cognate words (Pagibete and Bua vs. Lika, Bali, and, in the case of verbs, Bila) would be of particular interest.

There is a contrast between /k/ and /g/ in the Bua Bloc languages; this contrast is neutralised in Benge-Bati in some environments. Such a contrast is more usual in eastern Zaïre (Zone D) than in Equateur (Zone C). In this respect, then, the Bua Group has more affinity with the east than the west.

The noun-class system of the Bua Bloc languages appears to be broadly similar to Zone C languages, although our research did not extend to the entire concord system. There do not seem to be unusual noun class pairings (Guthrie's "genders") as are evidenced in some of the northern Zone D languages, nor does there appear to be a class 13 (with associated classes 12 and/or 19), as in some of the other Zone D languages.

However, we observe noun class suffixes, a characteristic shared with Lika in Zone D, and Guthrie characterises Bua as a Sub-Bantu, of which all the others (apart from Bangala) are placed in Zone D. It would be valuable to know whether any other Zairean Bantu languages have noun class suffixes (other than for infinitives, already observed in Budza). Further description of "Sub-Bantu" would be helpful, including a study of the presence or absence of a system of concords in the Bua Group languages. In particular, there seem to be two degrees of Sub-Bantu: in some tongues, plurals are the same as singular or are "ba-" plus singular; others have a full set of "independent" prefixes but supposedly lack the full set of agreements on adjectives, indirect object prefixes, associative constructions, etc.

2.4. Linguistic Findings

2.4.1. Phonetic

The languages under study all have about the same phonological inventory: all have seven vowels, for example, and all have two implosive stops (bilabial and alveolar) and voiced and voiceless double stops.

Nasalisation of vowels appears not to be phonemic; the Bua dialects (particularly "Bua East") have a nasalised vowel in certain class prefixes, corresponding usually to a syllabic nasal in Pagibete (and sometimes Benge-Bati).

As far as consonants are concerned, the word lists show regular phonetic correspondences in three or four areas:

- a neutralisation of contrast between voiced and voiceless phones in some environments in Benge-Bati;
- the apparent dropping of root-medial or root-final [k] in many of the Bua bloc dialects;
- the realisation in the easternmost Pagibete dialect of prenasalised [g] as the velar nasal; and
- a possible correspondence between a palatalised plosive in western Pagibete and a fricative elsewhere.

To all appearances, the phonemic inventory of all the Pagibete and Bua dialects include the full series of voiced and voiceless obstruents /b, d, g, gb, z/ and /p, t, k, kp, s/. However, it would seem that in Benge and Bati, there is a neutralisation of contrast between voiced and voiceless phones in some environments (root-medial for noun roots and adjectives, root-final for CVC verb roots). As in all the speech varieties under study, there are prenasalised voiced obstruents in Benge and Bati. In addition, they have voiced obstruents root-initial and in loan words (for example, "ndobani" 'fishhook'). It is interesting that this pattern of neutralisation is found only in these two varieties, and neither in Pagibete to the west nor in the (other) Bua dialects to the east. We can offer no explanation for this phenomenon.

The best-attested case of the relation between voiced and voiceless obstruents is that of [p] and [b]. As seen in the table, Benge and Bati have [p] root-medial in two-syllable noun roots and root-final in CVC verb roots for which the cognates in Pagibete and Bua-east have [b]. Prenasalised /b/ is realised [mb] in all the speech varieties under study, including Benge and Bati.

Table 4a: Correspondence of bilabial plosives

<u>gloss</u>	<u>most BWW & PAG</u>	<u>Benge-Bati</u>
navel	lekubu	lekupu
to laugh	-tebe	-tepe
wound	(l)e-poa-le	(same)
to exit	-pup-	(same)
to dance	-bine	(same) (see also 'stone')
tree	-mbaso-	(same)
moon	tembe	(same)
to sing	-lembe [Bua]	(same)

Although there are fewer data on which to base the judgement, it would seem that the same phenomenon of distribution applies to the corresponding alveolar, velar, and labiovelar stops and alveolar fricatives.

Table 4b: Correspondence of other obstruents

<u>gloss</u>	<u>most BWW & PAG</u>	<u>Benge-Bati</u>
joy	sode [Bua]	sote
younger sibling	tadi	tatu(?)
ear	-tu-	(same)
night	ɔiti	(same)
to dream	-lotoge	-lotoke
friend	dame	(same)
tail	-(n)kondo-	(same)
long ago	endele	(same)
to tie	-kand-	(same)

stone \	lebogo	leboko	
plantain /	(minimal tone pair)		
to say	-pag- [Bua]	-pak-	
to urinate	-nog- [Bua]	-nyok-	
bone	-(n)kue	(same)	
neck	kingo	(same)	(see also 'hen')
seed	(l)ekopo	(same)	
mountain	gangala [Bua]	(same)	
to weep	-game	(same)	
elephant	mbongo	(same)	
to request	-kong-	(same)	
knee	e-digba-	e-dikpa-	(only example)
chest	-(n)kpoto-	(same)	
to dig	-kpoe	(same)	(see also 'to sweep')
axe	gbono	(same)	(see also 'rat')
red	-gbam-	(same)	
house	ngbaale (Bua only)	(same)	
to sweep	-kpaze	-kpase	
(to awaken)	-zuzu-	-zusu-	
	(inferred from phrase list)		
news	-(n)sambo-	(same)	
left (side)	-mes-	(same)	
to burn	-sumbe	(same)	
to wash	-sos-	(same)	
drool	(ba)zangai	(ba)zangali	
hot	-zong-	(same)	
body	nzoe	(same)	
good	enza	(same)	

There are a number of words which have a root-medial [k] in the western dialects of Pagibete (Mongwapele and Omveda) and in Kango, but no [k] in the other Pagibete dialect or in the Bua cluster. These are of two kinds: between like vowels (in a noun root) and otherwise (between unlike vowels in a noun root, or verb-root final). If the (eastern) Pagibete/Kango [k] is between like vowels, these vowels coalesce in the other Bua bloc dialects. Of the six best examples, four have [ɔ̀kò] in Pagibete east corresponding to [ɔ̃] in the other dialects. Note that not only do the vowels coalesce, leaving a single-length vowel, but so do the high and low tones, leaving a falling tone. There is no vowel or tone coalescence in the other cases.

Table 5: Correspondences [k] versus no [k]

<u>gloss</u>	<u>PAG-west</u>	<u>PAG-east and BWW</u>
mouth	nòkó	ènô [Benge-Bati not cognate]
arm	òbókò	òbô; èbô
	(possibly Pagibete and Bua have different noun classes)	
fear	bókò	bô
	(others with o: old man, snake)	
child	míkí	mî (some dials) mí (others)
visitor	mbike	mbie (PAG, B-B); embie (BWW-east)
rat	gbuke	gbue
dream	nzaki	nzai
to shut	-dike	-diε
to send	-teke	-teε
to throw	-make	-maε
to vomit	-loke	-loε
to bite	-koke	-kwe [note apparent ATR diff.]
to open	-wuke	-(w)uwe
to pour	-duke	-du(y)ε

Normally, the Ndundu-Sana dialect of Pagibete (Dialect III) has the velar nasal without a stop when the other Bua bloc dialects have a prenasalised velar stop [ŋg].

In addition to these major correspondences, there are a few anomalies which may point to rarer correspondences.

There are three instances of [gy] in the western Pagibete dialects ('to play', 'to do', 'to fight'). In the first two cases, the Bua cluster had a cognate with corresponding [z]. Similarly, the man from Ndundu-Sana (speaker of "Pagabete III") did not give a word for 'to do', but in the other two cases his datum had [z]. There are three instances of [ky] in the western Pagibete dialects but for one of these ('to push') there are no cognates in the other dialects. The others ('dry season'/'year' and 'long'; see the Table) suggest a partial parallel to the [gy]-[z] correspondence already discussed.

Table 6: Correspondences with Omveda [ky]

	<u>'year'</u>	<u>'long'</u>
Mongwapele (Pagibete I)	--[no datum]--	e-kyaka
Omveda (Pagibete II)	ekyangake	e-kyaka
Ndundu-Sana (Pagibete III)	taŋa	e-taa
Benge, Bati	sanga	e-saa
Kiba	sange	e-saa
Yew, Bua proper	sanga	e-kia/e-kyā

Note the characteristic [ŋ] (not [ŋg]) in the Ndundu-Sana word for 'year', in accordance with the observation already made.

An unexplained (irregular) phonetic difference among the dialects of Bua-east is observed in the word for 'knife'. A group at Buta, asked to give examples of dialect differences, said that in Kiba and Boa-le (outside the Buta Zone), one says "fòká" and that in Yew (in the Buta Zone), the word is pronounced "pòká". According to the actual wordlist data, however, in Yew one hears [f] and in Kiba [p]. Possibly there is free variation. The lexical corpus includes a dozen or so words for which all the Bua bloc dialects have [p]; for another, 'to dry', all have [f]. The only other instance of [f] in the corpus is in a word used in four Bua dialects (but apparently not Kiba) for 'to swell'. It seems reasonable to conclude that /f/ is a phoneme in the Bua bloc languages, but one which occurs in relatively few words.

Although we are not counting Ngelema as a Bua bloc language, one striking correspondence in cognate words is worth mentioning. Ngelema often has [h] when the Bua bloc languages have [s]. This occurs in several environments: root initial in verbs and nouns (after both zero and nonzero noun prefixes), root medial in nouns and root final in verbs (and possibly in the causative verb extension "-is-", which would be "-ih-"). There is [s] in some of the Ngelema data. Of six words noted, two are not clearly cognate to Lingala or Bua ('feather', 'grass'), three are cognate to, or borrowed from, Lingala ('fear (noun)', 'seven', and 'to do'), and one is cognate to both Lingala and the Bua bloc ("son-" 'to sew'). In all six cases, the [s] appears to be root-initial.

2.4.2. Lexical

Since we have already discussed the lexicostatistical aspects of the data, we shall limit our comments here to isoglosses.

One of the supplementary words we elicited was 'tortoise'. In Bua-east, the most frequent response was "àngilé", and in the Buta area (but not in the Bambesa and Banalia Zones--Boa proper and Kiba), also "èkùndá". All the Benge and Bati speakers interviewed said that their word is "èkùndá", and that "àngilé" is a Lingala word. The Pagibete speaker from Ndundu-Sana (Bumba Zone) also gave "èkùndá". The other Pagibete speakers owned neither "àngilé" (which they also thought was Lingala) nor "èkùndá", insisting on a third word, "kóbá"; this is also a Lingala word, according to at least two Lingala dictionaries!

In short, the word preferred in Bua-east, "àngilé", and the word used in the two western Pagibete dialects, "kóbá", are both also Lingala words. (This is not to rule out the possibility that one or both is also a home-grown word.) A third word, "èkùndá", is used in Benge, Bati, and Pagibete III. Of course, this particular lexical item is not very important

in Bible translation; in fact, Lev. 11:29 may be the only place the word is needed (though if there is a shortage of words for reptiles, more than one of these words may be needed there). Of course, if the tortoise has a symbolic meaning in traditional culture, it may be used elsewhere, e.g. in Luke 13:32.

For 'cooking (palm) oil', Bua-east has "lèkàú" and "èśú" (synonyms). Bengé and Bati have neither of these, only "ḡgwá". Pagibete has neither "lèkàú" nor "ḡgwá", only "nsú(kú)", which is cognate to the Bua-east word "èśú". (The Bali word is also cognate: "ḡsúkú".)

Here are a few instances of the different ways isoglosses split the Bua bloc languages:

Table 7: Bua Bloc Isoglosses

Three Pagibete lists vs. five Bua lists:	
'chief'	ḡgba(ye) vs. kumu
'manioc greens (mpondu)'	(e)ḡgunza vs. ḡbanda
'hill'	m̄wambe vs. ḡḡḡḡḡḡ
'to go'	-ḡen- vs. -ḡi-
'to return'	-pan- vs. -si-
Pagibete east vs. PAG-III vs. Bua cluster:	
'God'	Nyombo vs. Nwemolo vs. Kunzi
Two Pagibete lists vs. PAG-III and Bua:	
'field'	tiko vs. matiya
'big'	-dingi vs. -si(i) [but see below]
'to say'	-kp̄et- vs. -pag- [B-B: -pak-]
'to pull'	-wot- vs. -ko-
Pagibete and Bengé-Bati vs. Bua-east:	
'meat'	nsoni vs. esui
'knife'	sone vs. foka/poka
'to come'	-bi- vs. -ḡo-
Pagibete and Bua east vs. Bengé-Bati:	
'mouth'	noko/eno vs. ḡḡake

Some items show more than a three-way split, e.g. 'manioc root', 'straight', and 'to follow'.

Some of those interviewed gave only loan words for the numbers between five and ten. Others gave the traditional expressions, which involve arithmetic. The forms given by the Pagibete-Omveda speaker differed from the ones given by the Pagibete-Ndundu-Sana speaker and the Bua-east speakers, however, as follows:

	<u>six</u>	<u>seven</u>	<u>eight</u>	<u>nine</u>
PAG-Omveda	3+3	4+3	4+4	4+5
PAG-III & BWW-east	5+1	5+2	5+3	5+4

Examples of synonymous words found only in some dialects:

- 'foot' Bengé and Bati have "s̄det̄eli" in addition to "et̄indi"
- 'person' Bengé and Bati have "ndumbe" in addition to "nto"
- 'fog' Buta and Kiba have "ndumba" as well as "ntutu"
- 'big' Bengé and Bati have both words (see above)

Examples of glosses for which there was a local word in some dialects and not in others:

- 'canoe' two Pagibete dialects use the Lingala word, Pagibete III has its own word, all Bua cluster dialects have another word.
- 'fish hook' the whole Bua cluster use the Bangala word, but the Pagibete have a local word "ḡwedēs̄".
- 'hoe' most, if not all, dialects have the same word for a small hoe ("ḡḡosa"), but two Pagibete dialects use the Lingala word "kóngó" for a big hoe, while Pagibete III and the Bua cluster use "ḡítà" (or "ḡétà"), which is used in Bangala but was probably a Bua word first.
- 'river' Pagibete dialects have "dùà" (which is the name of the river that crosses/borders their territory) and Bengé-Bati have "ḡḡbàngí". The other Bua dialects (Bua-east) and Nḡelema seem not to have a

specialised word, only 'big water' or 'small water'. All, however, have "(l)ègbúdí", which may be glossed 'pool'.
 'tomorrow' two Pagibete dialects (and Ngelema) have a word distinct from the word for yesterday. Pagibete III and the Bua cluster dialects tend to use a compound, sometimes involving the word also used for 'yesterday' (cf. Lingala, which uses the same word for 'yesterday' and 'tomorrow').
 'price' all the Pagibete dialects have "tálò", which may be a Lingala loan or may be the original Pagibete word. The whole Bua cluster have "(è)ntíò".

2.4.3. Semantic

Of course, very little can be said about the comparative range of meanings of words in a set of languages on the basis of a standard lexical survey. It is worthwhile, however, to comment on findings of this nature noted in the course of the Bua bloc survey.

The form (l)èkópò 'seed (for planting)' was attested in all the Bua and Pagibete dialects. It is also used in Ngelema and Kango. The gloss 'fruit' was not elicited in these last two, but was lèmbùá in all the Bua dialects. There is apparently no separate word for 'fruit' in Pagibete, however, and the cognate word èmbùká means 'testicle' (in at least the Ndundu-Sana dialect). This could cause embarrassment should the a text in Bua Scripture containing the word (apparently the most obvious word for 'fruit' in Bua) be read in a church meeting in Pagibete-land.

The complexes 'leg, foot, sole' and 'arm, forearm, hand, palm' are notoriously difficult to elicit reliably, especially when working through Lingala or Swahili, trade languages in which 'leg' and 'foot' are signaled by the same word, as are 'arm' and 'hand'. However, the following seems to hold:

Table 8a: 'Leg' and 'Foot' in Bua group

Lingala	Pagibete-west	PAG-III&B-B	Bua-east	Ngelema
lòkòlò [both]	ògókò 'leg'	(no cogn.)	(no cogn.)	ègòlò 'leg'
(no cognate)	ètíndí 'foot'	ètíndí 'leg'	ètíndí [both]	ètíndí 'foot'
(no cognate)	(no cognate)	èdètèlí 'foot'	(no cognate)	(no cognate)

The differences do not necessarily hinder communication.

Some of the three words for 'house' attested across the Bua chain have different meanings from place to place.

Table 8b: 'House' in Bua bloc

PAG-Mongw.	PAG-Omv.	PAG III	Benge-Bati	BWW-Kiba	other BWW
èbèmbé	èbèmbé	(no data)	(no cogn.)	lèbèmbé	lèbèmbé='hut'
				(Kiba: "pejorative")	
(no data)	ètókè	ètòè	ètòè='natte'	(no data)	ètòè='natte'
	(ètókè: not at Butu)				
(no data)	(no data)	(no data)	ngbáálè	ngbáálè	ngbáálè

There may be a semantic difference in the area touched on by the gloss 'affair' (abstract 'thing', Lingala "likambo"): The old Bangala word, also used in Ngelema, is "kùlá". Most of our data sources did not offer this word (when queried, one Pagibete speaker said it was not used in his dialect). Someone from the Dingila area did say "kùá", and the Kiba speaker (interviewed last) said "kwē", which is probably cognate. At Buta, we were told that "kùá" means 'to die' (this cannot be tested, since they gave a conjugated form when we elicited the verb, but 'death' is "kwá").

The word offered by all the Pagibete and Benge-Bati speakers was "mbàngé" (three syllables in Pagibete, two in Benge and Bati). The Kiba speaker said this meant 'something which brings about a court trial' and the group interviewed in Dingila said "ngbàngé" meant 'judgement' (Lingala "ngwanga"). Apparently, they agreed with the person who had already given "kùá" as the word for 'affair' in their dialect. The source of the Buta dialect data preferred "mbàngé" (syllabic m), having glossed "kùá" as 'to die'.

2.4.4. Morphological

Based on a corpus of over 120 nouns in each of the eight dialects in focus (most of them in both singular and plural forms), we offer the following preliminary sketch of the noun classes in the Bua bloc languages:

Table 9: Noun class affixes in Bua bloc

Probable Class	leBoa-le	Yew	Kiba	Benge-Bati	PAG-III	Omveda	Mongwap.
1	é	é-	ě-	N-	N-	N-	N-
	(when the N Rt starts with an oral or prenasalised cons.)						
1	è-N	è-N	è-N	ø-N	ø-N	ø-N	ø-N
	(when the noun root starts with a nasal consonant)						
1a	ø-	ø-	ø-	ø-	ø-	ø-	ø-
2	βà-	βà-	βà-	βà-	βà-	βà-	βà-
or	βà- -βà	βà- -βà	βà- -βà	βà- -βà	βà- -βà	βà- -βà	βà- -βà
3	ě-	ě-	ě-	N-/ě-	N-	N-	N-
	(when the N Rt starts with an oral or prenasalised cons.)						
3	è-N	è-N	è-N	ø-N	ø-N	ø-N	ø-N
	(when the noun root starts with a nasal consonant)						
4	mò- -mè	mò- -mè	mě- -mè	mă- -mè	N- -mè	N- -mè	N- -mè
	(when the noun root starts with an oral consonant)						
4	mò- -mè	mò- -mè	mè- -mè	mà- -mè	N- -mè	N- -mè	N- -mè
	(when the N Rt starts with a nasal or prenasalised cons.)						
5	lè-	lè-	lè-	lè-	è- ?	è- -ko ?	è- -ko ?
6	mà-	mà-	mà-	mà-	mà-	mà-	mà-
7	è-	è-	è-...e	è-	è-	è- -kè	è- -kè
	(when the noun root ends with "a", "ε", or "e")						
7	è-	è-	è-	è-	è- -è	è- -kè	è- -kè
	(when the noun root ends with "i")						
7	è- -è	è- -è	è- -è	è- -è	è- -è	è- -kè	è- -kè
	(when the noun root does not end with a front vowel)						
8	βè-	βè-	βè-...a	βè-	βè-	βè-	βè-
or	βè- -βè	βè- -βè	βè- -βè	βè- -βè	βè- -βè	βè- -βè	βè- -βè
9	ø-	ø-	ø-	ø-	ø-	ø-	ø-
10	ø-	ø-	ø-	ø-	ø-	ø-	ø-
11??	lè-	lè-	lè-	lè-	è-	è-	lè- (?)
14?	βò-	βò-	βò-	βò-	ò-	ò-	ò-
15??	ø-	ø-	ø-	ø-	ø-	ø-	ø-

Key: ỹ- = nasalised vowel (v= e or a)

N = homorganic nasal with low tone; ø = zero

- Pairings (singular/plural): 1/2, 1a/2, 3/4, 5/6, 7/8, 9/10, 11/6; No plural: "14" (one datum), "15" (infinitive—may actually be 9)
- Sometimes Kiba N- or ø-N in class 1
- Benge-Bati class 3 sometimes has ě- prefix (like Bua-east) and sometimes N- prefix (like Pagibete)
- For some pairs, the root ends with "e" or "ε" in the singular (classes 1, 1a, 3, 7) but with "a" in the plural (classes 2, 4, 8)—but, for some of the pairs, there are exceptions in some of the dialects!
- Corpus is too small to postulate rule on when classes 2 and 8 have a suffix
- Kiba class 7 tends to end with "e" or "ε" rather than "a".
- The Pagibete data with "o- -ko" class affixes (corresponding to "le-", class 5 in Bua) may actually be class 15: the plural is class 6 in any case....

The realisation of the class 1 and class 3 prefix as a nasalised vowel has been observed in other languages besides Bua, notably Bali. Of course, when the first phone of the root is a nasal, it is more difficult to judge whether the class 1, 3, or 4 prefix is nasalised.

It is interesting that the "k" is retained in the class suffix for Pagibete (east) class 7 (and possible class 15), but not in the prefix. It is not surprising that the class 7 suffix in the other speech varieties has no "k" since this is a consistent phonetic difference. That there are class suffixes at all is interesting, as has already been noted.

There is no separate phonetic explanation for the fact that the class 5 prefix in Pagibete is "e-" and not "le-".

The personal subject prefixes on verbs are similar across the chain, as demonstrated by the three phrase lists elicited in the Buta dialect of Bua, in Bengé, and in Pagibete-Omveda.

		Bua-Buta		(present)		Bengé		Pagibete	
1 sg	1 pl	na-	ta-	na-	ata-	na-	ta-	na-	te-
2 sg	2 pl	o-	ma-	o-	ama-	o-	ma-	o-	ma-
3 sg	3 pl	a-	ba-	aya-	baba-	a-	ba-	a-	ba-

Our data do not permit us to comment further with any confidence on comparative morphology.

2.4.5. Syntactic

The grammatical data gathered in the course of the survey (presented in Appendix G) will only support a few tentative conclusions. Generally, however, it would appear that Pagibete, Bengé-Bati, and the Buta dialect of Bua are all SVO languages. (I refer to Bengé-Bati since a Bati speaker, an observer at the "phrase list" interview with the Bengé speaker, said that his answers would have been the same.)

Pagibete differs from the other two varieties in several ways. The Pagibete respondent consistently used a double negative (triple, if one counts the utterance-initial "No,"). He placed the negative word "ka" at the end of a negative imperative sentence, unlike the others; in addition, his negative imperative construction contained a morpheme "se" not evidenced elsewhere. He also used a different word to introduce the second clause of a contrafactual sentence than did the Bengé and Bua-east respondents.

There was evidence of an apparent copula in the Bua bloc languages. In Pagibete it was "aliki" (singular, class 9?) and "bamiki" (plural, class 2). Bengé had singular "ali" or "ale" and plural "bami", which may be only the phonetic difference already described. Bua had "ali" in one case, but only singular "a" and plural "mi" on the final page.

Bua seems to have an allomorph of "give" (used in the imperatives "give me" and "give us") not evidenced in the other two varieties. The subordinator "that" (the "bete" of "Apagibete") has three syllables in the Buta dialect ("betene") in contrast to Pagibete and Bengé-Bati.

The three lists evidenced three different strategies for compound sentences, of which our list had two kinds: conditional ("If this happens, then do this") and eventual ("When this happens, then this will happen").

It appears that all three groups have an anaphoric marker ("yo" in Pagibete, "ya" in the other two). All had the same word for the demonstrative 'this' ("mo") but there is no clear evidence on their word(s) for 'that'.

2.5. Degree of Comprehension within the Bua Bloc

2.5.1. As stated by the interviewees

The Bati speakers we interviewed in Aketi mentioned three Bua bloc varieties: Bati, Bengé, and Bua; the last includes Yewu, Nganzolo, and Kiba. They said they understood Bengé better than Bua and Bua better than Ngelema (in part because the Ngelema "put Benza in their speech").

More specifically, the interviewees said that there are no problems of comprehension among the Bati people in different places. Though Benge is almost the same as Bati (they supposedly are interintelligible without need for the interlocutors to slow down or simplify), according to one interviewee, the Bati people would have trouble understanding some Benge proverbs. There is some need for Bati and Bua interlocutors to change the way they speak (or explain some things in Lingala) in order to understand each other. Bati children do not immediately understand Bua (meaning "Bua east").

The Benge interviewees, all originally from Likati or Bondo but now residents of Aketi, also expressed the difficulty of understanding Bua (east). One of them had apparently been to Buta and had thought the baBua understood him better than he them, but that not all the baBua could understand him. Another Benge man, whom we interviewed later at Bunduki, said that even a little child from Likati could understand a Bati person from Bunduki right away. He said that a speaker of Benge and a Bua person from Buta might have to have recourse to Lingala when they could not understand each other's dialects.

Another Benge speaker interviewed at Aketi on our return trip said that due to contact, the Benge and the Bua understand each other in spite of vocabulary differences.

A group interviewed at Malingwia, between Buta and Bambesa (in the Yew area, northeastern Buta Zone), claimed to understand not only Bua-east but also Bati, Benge, and liKango with no difficulty. They admitted having difficulty understanding Ngelema. Upon hearing the recorded texts, they were able to identify the first two as Benge and Bati; they said that the third one (representing Pagibete of Ngakpo) was "lèGòlò", a variety of Benge spoken near Muma or Ngaye (western Aketi Zone—Boguru/Bogboma *groupements*), which they admitted having more difficulty understanding. In fact, they said they understood Ngelema better than leGolo. In other words, they supposed that they could understand the speech of the baNgelema (whom they didn't even list among the Bua peoples) better than that of the Pagibete people.

The group interviewed at Dingila first reported understanding Bati (and Kiba) better than leGolo (and Kango), and leGolo more than Benge. After hearing the taped stories, they admitted that they'd have to change the ranking: they understood Benge better than Bati, and Bati better than the speech form used in the other text (i.e., Pagibete, which someone had also identified with leGolo).

Let us turn now to the Pagibete. According to the group interviewed at Ngakpo, there is 100% understanding among the speakers of Pagibete in the Businga Zone, both in the Mongwapele *groupement* on the Businga-Lisala road and in the Ngakpo area. They claimed that they could generally understand the speech of the Ndundu-Sana *groupement* (Fultz & Morgan's Dialect III, or at least the southern part of it); however, they would not always understand everything in rapid speech.

This same group said that they would understand Bua only with difficulty. Two of the participants said that they had met a Bua speaker in Businga once and though they had to listen carefully, they could understand him in part. Their impression (and his) was that he could understand them well. A man from Ndundu-Sana who was present at the interview added that he can understand even rapid Bua, and a moBua would understand him also. Two days later, we learned that his grandparents came from Bua country. He did understand all four taped texts (in Bua-Buta, Benge, Bati, and Ngakpo Pagibete); however, an indeterminable part of this comprehension may have been acquired.

2.5.2. Results of Comprehension Testing

We recorded texts for testing dialect intelligibility in four dialects: Bua, Bati, Benge, and Pagibete. Literal translations of the texts appear in Appendix D. The texts were tested in the following locations: Mondongbo (Pagibete), Ngakpo (Pagibete), Bunduki (Bati), Aketi (Benge), Malingwia (Bua), and Dingila (Bua). We found that speakers of Bua, Bati, and Benge could consistently reproduce the texts in each of those three dialects. They were able to supply most details when pressed, only occasionally missing a word here and there. (Bua speakers encountered slightly more difficulty with Bati and Benge than vice versa, see chart below). They were also able to identify which dialect was being spoken.

However, in testing when the Pagibete tape was played for these three dialects, the people had significantly more difficulty in re-telling the stories. They could usually give the basic story-line, but important events in the story were often missed. They could not identify the dialect that was being spoken.

Likewise, the Pagibetes in general had difficulty re-telling the stories from the other three dialects. Some of the younger men even would give up telling the story-line out of frustration. The following chart gives a summary of the findings for the four dialects. Complete results from all of our test points are given in Appendix E.

Table 10 Comprehension of Texts between Bua Dialects

		Language of Text			
Test Location		BUA	BATI	BENGE	PAGIBETE
		GOOD	GOOD*	GOOD*	SOME
	BUA	GOOD	GOOD*	GOOD*	SOME
	BATI	GOOD	GOOD	GOOD	SOME
	BENGE	GOOD	GOOD	GOOD	SOME
	PAGIBETE	SOME	SOME	SOME	GOOD

* —indicates a "qualified" good. The subjects understood the text well, with one or two minor difficulties.

The purpose of RTT is to ascertain the degree of inherent intelligibility between dialects, which means to what extent something said in one dialect can be understood by a speaker of another dialect based only on the degree of linguistic similarity of the two dialects. In practical experience, this is hard to measure accurately whenever there is a history of contact between speakers of the two speech forms. In such cases, there is usually a sociolinguistic overlay (such factors as the hearer's attitude to the other speech form and the relative status of the two speech forms in the society), and a greater or lesser degree of passive or active learning of one another's speech.

While there is no evidence of deeply-felt negative or positive attitudes among speakers of Bua bloc dialects to the other dialects, or of differences of status, we were concerned about the effect that acquired comprehension would have on our results. For, as Blair states (1990:74), Recorded Text Tests actually measure comprehension, rather than intelligibility.

Our terminology differs from Blair's at this point. We find it useful to refer to "immediate" and "acquired" comprehension. "Immediate" comprehension is based on inherent intelligibility. "Acquired" comprehension is based on contact factors and includes both active language learning and passive acquisition of the ability to understand. Blair uses the expression "acquired intelligibility". This term is problematic in that it is people who acquire the ability to understand more of the language than is inherently intelligible to speakers of their dialect.

The difference is semantic, as is illustrated by our preferred restatement of his explanation (p. 25) that "a community is homogeneous with regard to inherent intelligibility; it is usually heterogeneous with regard to acquired intelligibility". "A community is homogeneous with respect to immediate comprehension of another community's speech; it is usually heterogeneous in how much they actually understand of that speech, due to differing degrees of acquired comprehension."

However the phenomenon is phrased, the researcher needs to be aware of how the results may not directly reflect the degree of inherent intelligibility. Blair (1990:25) points out that if the standard deviation of the (conventional) RTT scores on a given text in a given community is large, the researcher can be pretty sure that acquired as well as immediate comprehension has been measured. Although such a criterion cannot be applied in the RA RTT (in which a group is tested and the results are qualitative), as we conducted the interviews, we looked for similar variability in individual comprehension.

We found that people who had travelled widely, or many times, were often more able to retell the stories accurately than those who had not. This was particularly the case in the Pagibete region. Several Pagibete men who had been to Haut-Zaïre were able to recount the (non-eGezo) stories in detail, while younger men who had not been out of the region had great difficulty. Thus we considered it important to test a variety of people, and we made a point of finding those who had not travelled much. If the ability of people to retell

the stories in a particular village varied widely, we concluded that those who were able to tell the stories well could do so because of active or passive learning of the dialect in question. We thus considered the lower "scores" (lesser apparent comprehension) to be a more accurate reflection of the actual inherent intelligibility of the dialect in question in the community.

The groups of Bua, Bengé, and Bati dialect speakers interviewed were able accurately to identify each other's dialects, usually after hearing less than fifteen seconds of the reference tape. This led the survey team to believe that some of the observed comprehension had been acquired through contact. There is, in fact, much interaction between speakers of these dialects, especially in Buta. Nevertheless, we are sure that the difference in observed comprehension (GOOD versus SOME in Table 10) is not only due to contact factors but also in part to greater inherent intelligibility. The word lists showed a greater degree of lexical similarity among the Bua, Bati and Bengé dialects than between any of these and Pagibete, a pattern of similarity which appears also to apply on the phonetic, morphological, and syntactic levels.

It should also be remembered that not only does geographical and social proximity tend to promote acquired comprehension, it also helps prevent linguistic divergence and can actually cause linguistic convergence. That is, people of different speech varieties who talk to each other not only understand in spite of differences of expression, they are less likely to develop new differences and may even adopt new similarities.

It was our experience that the RA RTT is not able to distinguish well between dialects when there is a high degree of immediate comprehension. For this reason, intelligibility testing does not help very much in the choice of reference dialect. This may not be a flaw of the rapid appraisal approach only, but may be true of recorded text tests in general for Zaïre (cf. Girard 1991).

2.6. Sociolinguistic Findings

2.6.1. Language Use

Since Reeder and Stoothoff investigated language use in the Ngakpo area, our research was limited to the dialects spoken in Haut-Zaïre. The groups interviewed in Buta and in Aketi agreed that almost all the (baBua and boBati) people in their areas between the ages of six and sixty years know Lingala. Similarly, both groups said that a few city-born people would not know their ancestral language, but all the baBua and boBati who live in "the village" do know the mother tongue.

The CECCA/16 church in Malingwia said that although there were formerly adult literacy classes in Bangala, these have ceased due to lack of books. (Apparently the bookshop in Ibambi still has Swahili books in stock but not books in Bangala.) At Buta, we were told that Christian Education includes literacy classes, mainly in Lingala but also some in French, and many people participate, but that teachers were not always paid. According to those we interviewed at CBZN/Aketi, there are no literacy classes in the Bati area.

The highest-ranking CECCA/16 pastor at Buta (not himself a Bua speaker) remarked that the Bua people love to use their language. If the leaders at a meeting are Bua, the baBua can "run away" with the meeting; sometimes the non-Bua must "protest" so that the meeting can be conducted in Bangala again. Because we did not conduct individual interviews, we cannot elaborate on the domains in which Bangala/Lingala and Bua are used.

2.6.2. Language Attitudes

The Bati speakers interviewed in Aketi considered the Bati of the Bongi and Bodongola collectivities to be both the purest and the prettiest kind of Bati. All six of them agreed that it would be good to use that kind of Bati for a written standard form of their language. They imagined that if there were Bua scriptures, they would be interested in attending literacy classes in order to learn to read them.

In an interview the next day, some Bengé speakers confirmed that Bati and Bengé were very much alike. One said there were only a few words which were different. A second group of boBati interviewed in Bunduki were of much the same opinion, adding that because of occasional vocabulary differences, children would not immediately understand.

Probably, then, the Bati and the Bengé would use the same literature; if one group could and would read written materials in Bua (east), so would the other. If separate literature is needed for one or both, one body of literature would probably suffice for both.

Those who live in the Buta and Bambesa Zones, probably feeling secure that the reference dialect would be chosen from among their (Bua-east) speech forms, said that they would like to see all the dialects included in a project. For practical reasons, Buta would be a natural choice for placement of the project centre.

There was a range of opinions concerning the value of what had already been published in the Kiba dialect. The Malingwia pastor may also have expressed the point of view of others when he said that the greatest flaw in the "Wooni" tract is that it is in kiKiba, and he would prefer to see literature which includes all the dialects.

Those we interviewed in Ngakpo claimed that theirs was the purest Pagibete speech. This is Fultz and Morgan's Dialect II, the "Principal" dialect, Momveda. They said that Mongwapele residents (Dialect I speakers) put Ngombe words in their speech and those at Butu (Yakoma Zone) put Ngbandi words in theirs. At Ngakpo, however, they don't put in anything foreign ("except Lingala", remarked Douglas Boone at this stage).

There was a feeling that the dialects don't differ very much. The group interviewed at Mondongbo declined to identify a "center" of Pagibete. (However, the Preacher there said, "Pagibete is spoken from here eastward", effectively omitting Mongwapele! His reply may have been coloured by having just heard three stories on tape told in Bua-Buta, Bengé and Bati.)

The Ngakpo group said that though a six-year-old from Ngakpo would not understand someone from Ndundu-Sana (Dialect III) well, these are "one-month dialects". (Mr Boone's notes say that a Dialect III child would "catch on" to Dialect II within a month; Mr Olson's, that in less than a month a Dialect II child would get used to speaking Dialect III.)

2.7. Institutions

We asked about schools and health centres only at Buta. There we were told that there are many primary schools in the Bua area, but that many parents were having difficulty paying the school fees. Most of the teachers, they said, are baBua. Not every *groupement* has a health centre and not all health centres have adequate medicine.

3.0 ANALYSIS

3.1. Classification

The present survey generally supports the classification of Van Bulck and Hackett (1956) concerning the Bantu languages of what is now the Nord-Ubangi, Mongala, Bas-Uélé and northern Tshopo regions. Ngombe and Benza/Genza seem to belong with Budza and Babango rather than with Pagibete and Bua. Ngelema, by the definition of this report not part of the "Bua bloc" (since it was not in focus in this survey), would appear nevertheless to be part of the "Bua group", based on the degree of lexical similarity shown with the Bua bloc languages. At least some forms of "Kango", which may be an umbrella term for an undetermined number of riverine dialects, belong in the "Bua group" as well, for the same reason.

We cannot confidently attach Lika and Bali to this group or any other. A tantalising possibility is that if we had more data in the "Extreme North Group" languages (Van Bulck and Hackett 1956:74-75), the relationship between the Bua group and Lika would become more clear. Not only is this suggested by the fact that Van Bulck and Hackett call Lika and Bali "Extreme North Group Transition Languages", but also by the statement that several of them (e.g. Kari, Nyanga-li, Gbati-ri, Ngbee) had "vestigial suffixes" and by the lexical similarity between Lika and the few Gbati-ri data one of the present surveyors (Boone) was able to collect near Niangara in Sept 1993.

As McCord and McKee note (1986:1), Van Bulck and Hackett seem to have been "splitters" rather than "joiners" in the assignment of language and dialect names. McCord and McKee, of course, had to limit their study to the Bambesa and Buta Zones, where the uniformity of speech led them to conclude that there are two minimally different dialects of

kiBua. Certainly, the six-way division of the Bua bloc offered in SNBBL is misleading since some varieties are much more similar than others. We see no reason not to retain the Ethnologue division into three languages, namely, "Pagibete", "Bwa" and "Ngelima". If it is later discovered that Bengé-Bati (or even Pagibete-east) needs separate literature, new Ethnologue entries can be added.

As "joiners", we have not distinguished Bua dialects within the Buta Zone (e.g. Yew, Ngingita). We were not able to confirm the status of other varieties, such as leGolo.

We believe that Guthrie's organisation of the Bantu languages of north-central Zaïre into "Groups" (Guthrie 1971:40, 42-43; Bastin 1978:140-142), is misleading. Without much more study of the dialects of the "Bangî-Ntomba Group", we cannot critique the placement of Budza and Ngombe in different groups. However, his "Ngombe Group", "Lega-Kalanga Group" and "Bira-Huku Group" each need to be split.

Generally, we endorse Bryan's Bali, Bira and Nyali Groups, though in the case of the Bali group (Lika and Bali), this is more a matter of convenience than conviction. We prefer to rename the Bira and Nyali groups, calling them the Komo and Budu groups. Komo and Budu are spoken by larger numbers of people than Bera and Nyali, and are also better documented (more fully described). Believing that Ngombe and Bua do not belong together (at the "group" level), we propose a separate Bua Group.

(There are two other reasons not to use the name "Bira Group". The term "Bira" has been used to refer to or include various combinations of Bera, Bila, Amba, and a so-called "Western Bira" near Kisangani, cf. Bryan 1959:89ff. Also, there is a markedly lower lexical similarity between Bira and Amba, on the one hand, and Bila, Kaiku, Bhele, and Komo, on the other.)

The alternate names given for Bali (that is, S.E. Bua and Bango) are quite unhelpful. Bua and Bali are in fact separate languages, though not necessarily belonging to different "zones". Bango refers not to a Bua Group dialect but to a Budza-like speech form (possibly a Budza dialect) in the Basoko Zone. Based on present knowledge, SIL would accept the assignments Bryan makes to her Lega Group, except for Bangubangu and probably Holoholo. However, we would not be surprised if Enya, Mituku, and especially Lengola were eventually better removed from the (Lega) group. Thus, Guthrie's "Lega-Kalanga" has been heavily reworked: losing members to the Komo (Bira) Group to the north and the Luba Group to the south and gaining members from some or all of D.10 to the west and D.40-50 to the east.

Further research is needed before a definitive assignment of the Bua Group to a larger classificatory entity is possible. Perhaps the present situation should be retained-- an attachment to Guthrie's Zone C languages (Northwestern Bantu) with a resulting disassociation from the languages of eastern Zaïre. It is more likely, however, that another solution will be necessary, in order to show the affinities between the Bua Group and languages both to the east and to the west. This could take the form of a revised classificatory system, taking more Groups into consideration.

For an account of relationships within the Bua Group, see section 2.3.2.3.

3.2. Comprehension

As mentioned above, we found in our study of comprehension that there was a clear line of demarcation between Pagibete, on one hand, and the dialects to the east on the other. This line interestingly coincided with the political boundary between Equateur and Haut-Zaïre. Pagibete speakers showed lower comprehension of the eastern dialects than the comprehension found between the several eastern dialects.

According to Stalder (see Appendix C), if comprehension is found to be good, and language attitudes are positive, there is a high probability that the use of a common translation is possible. However, if the comprehension level is only "some", while language attitudes are positive (such as the situation between Pagibete and the eastern dialects), then there is a LOW probability that a common translation could be used.

Given these criteria, and in the absence of conclusive language attitude information, we can conclude that intercomprehension between Bua, Bengé, and Bati is at a sufficient enough

level so that there is a high probability that these speech varieties can share common literature. At the same time, we also conclude that there is a low probability that Pagibete could share common literature with the dialects to the east.

3.3. Relationship to Existing Language Projects

3.3.1. Academic affairs

It seems likely that projects in Bua bloc languages (in the near future, Bua and Pagibete; if needed later, in other "dialects") could derive benefits from exchanges with projects in other Bantu languages in Zaïre in which SIL or another group (such as Pioneer Bible Translators) is working.

In eastern Zaïre, this means particularly the existing projects in the Bali and Komo languages, and another to be undertaken in kiLika beginning in late 1995. Such exchanges could include participating in a text adaptation (CARLA) network established between some or all of these other languages, conferring on language (and cultural?) analysis, and coordinating orthography decisions. A wider CARLA network may be feasible, incorporating existing Scriptures in, for example, the Ngombe language. The possibility of text adaptation from Lobala to Pagibete has also been raised. (According to the Guthrie classification, Lobala is in the Bangi-Ntomba group, C.30.)

Apart from Lobala, all of the existing SIL language projects in the Equateur region are in non-Bantu Niger-Congo languages, that is, languages of the Adamawa-Ubangian family (Ngbaka-Minagende, Ngbandi, Mbandja, and Mono, see Boyd 1989:192-193 and Barreteau and Moñino 1978:198). In addition, because of high bilingualism rates in Lingala (Fultz and Morgan 1986), there are no plans for new projects in any of the Bantu languages of the Ubangi and Mongala Sub-Regions, such as Budza and liGenza. This could reduce the academic support available for Pagibete from the Congo-Western Zaïre Group.

Teams working in Bua bloc languages might also do well to confer on cultural analysis and orthographic decisions with any personnel working in neighbouring though unrelated or less closely related languages such as Ngbandi, Zande, Barambo, Budza, or Ngombe. At present, however, there are no translation projects in these languages. Although there are Scriptures in Ngbandi and Zande, there is need for literacy work. Barambo is a priority language for allocation by SIL in eastern Zaïre.

3.3.2. Logistics

Once the Ngakpo airstrip is cut, there will be easy air access to a Pagibete centre. Until then, Ngakpo is over twenty kilometers by trail (via bicycle, motorcycle, or on foot) from the nearest semi-practicable road (from which it is also separated by a wide river which must be crossed by dugout canoe). No language project is planned for the immediate vicinity, though SIL is seeking a literacy team for Ngbandi. Radio contact already exists with all the church stations in CECU and CEUM, including the sites of all the other SIL projects in Equateur and of the SIL administrative centre in Gemena.

There can be no doubt that the best approach to work in Pagibete is through CEUM sponsorship and SIL administration from the Congo-Western Zaïre Group. CEUM is the only ECZ denomination with work in the Pagibete area, and has already expressed interest in a translation project there. Furthermore, CWZG has worked closely with CEUM and CECU (Communauté Evangélique du Christ en Ubangi) for many years, drawing mother-tongue translators from these communities and organizing literacy programs in conjunction with the work of the two churches. Finally, the CEUM, CECU, and CWZG all have their administrative offices in Gemena.

It is likely that initial work in Bua will begin in Buta, a hard two-day drive from the nearest SIL project and from the SIL support centre in Isiro. There is a possibility of future SIL work in the Barambo language, but teams in Bua and Barambo would still be separated by at least 250 km, rendering frequent face-to-face interaction virtually impossible. There is, however, an airfield at Buta, with regular commercial air traffic (direct connections to Isiro and to Kisangani and from these to Bunia and to Kinshasa).

There are also possibilities for mission aviation support. The MAF programme at Nebobongo, south of Isiro, which already serves several language projects and the SIL

centre in Isiro regularly, can also serve Buta and/or Malingwia. (Projects served by MAF-Nebo include Mangbetu, Mayogo, Budu, and potentially Lika, with partial service to Bali and Komo.) In addition, an MAF-Nyankunde aircraft, the Caravan, passes very close to Buta on its approximately monthly run between Nyankunde and Karawa. The Buta airfield can easily accommodate a DC-3, but the present flight paths of AIM-Air's DC-3, based in Nairobi, bypass Buta. If permissions are received for international traffic in and out of Buta, or even in and out of Ango, missionaries in Buta may be able to enjoy passenger and freight service via DC-3.

(See the map in Appendix A.)

3.4. Reference Dialect

It seems best to initiate two projects, with later evaluation of the need for additional projects. Accordingly we comment on reference dialects for a Pagibete project and for a Bua project.

The factors to be considered are linguistic, sociolinguistic (language use and attitudes) and non-linguistic. Many of the language use and non-linguistic factors (e.g. socioeconomic importance of a dialect, social ties such as marriage alliances) do not apply. Besides, the present survey did not explore all of the potentially important factors. To a certain degree, the question of reference dialect may be moot since both among the Pagibete and among the Bua, the church appears not to have raised the question.

3.4.1. Pagibete

Of three recognised dialects of Pagibete, there can be no question that "Dialect III" is the most linguistically divergent. This is shown by our data in the areas of phonetics and lexicon, and is also reflected in a comment during the group interview that dialects I and II are only slightly different.

It is reasonable to accept the testimony of those who said that there is no difficulty of understanding among the Pagibete speakers of the Businga Zone (Dialects I and II), and that their understanding of the other Pagibete dialect(s) is imperfect. The one interviewee from a non-Businga dialect (Ndundu-Sana) may be atypical, but he seemed to understand the speech of Ngakpo. (Children from his home area reportedly would not understand so well, but could acquire the ability in a short time.)

The only language use factor which applies to the question of Pagibete reference dialect is the existence of several books prepared in Gemena in the 1980's by Pastor Mbangiye, who comes from Ngakpo. The quality and acceptability of these for the people of Ngakpo was amply demonstrated during our visit.

As far as attitudinal factors are concerned, there is a feeling that the Ngakpo dialect is more "pure" than the Pagibete spoken in the Mongwapele *groupement* or outside the Businga Zone. Fultz and Morgan (1986:10) say the same thing in other words: the Ngakpo dialect is the "principal" one in part because of its relative isolation from other ethnic groups. This dialect may also score highest in "reputed understanding" (as distinct from "declared understanding", reported above).

Fultz and Morgan's other criteria (loc. cit.) for naming the Ngakpo dialect the "principal" dialect are non-linguistic, specifically geographic and demographic. The Bomongo, Bonzwambe, and Bozame *groupements* enjoy a central position between Mongwapele on the west and the Yakoma/Bumba Zone dialect (Dialect III) on the east. According to Fultz and Morgan, they also represent a larger territory. While we have no way to verify this, it appears that Dialect III is spoken in seven *groupements* (three listed in Fultz and Morgan), compared to Dialect II's three. On the other hand, the Ngakpo territory is forested, with great distances between some of the villages, and could easily cover a greater area.

Fultz and Morgan claim that Dialect II has more speakers than the other dialects (not necessarily counting Bua!). Their own population figures contradict this and since they accidentally counted the Babale among the "principal" dialect speakers, Dialect II could represent only 25-30% of the total Pagibete population (about 6,000 people). By any interpretation of the data, there are more speakers of Dialect III (estimates vary between 7,300 and 13,000). More recent census data from all three zonal offices should be used to make the final judgement.

An important criterion not used by Fultz and Morgan is the presence of institutions, such as health centres, schools, church centres, etc. Since Ngakpo has an airstrip, a major dispensary, a primary school, two Catholic chapels and a Protestant church centre (According to Reeder and Stoothoff's map), its dialect is well-positioned. However, in the absence of comparable information for the other Pagibete areas, it is not possible to choose a "most-favoured dialect".

The choice of Ngakpo for project location is natural because of its relative isolation (while potentially accessible by air) and its importance in the CEUM church (reflected by the existence of books in the dialect and the presence of a major dispensary). The "Ngakpo dialect" is also advantageously placed (geographically central): we were able to talk to speakers not only of the local dialect but even of the other dialects the weekend we were in Ngakpo. It may also be linguistically central, although our survey did not reveal any shared features between Dialects II and III which were not also shared with Dialect I.

On the other hand, the central dialect may not have as many speakers as the eastern dialect. Taking the two Businga Zone dialects together, however, their speakers may outnumber those of Dialect III. Surface area of the dialect area may not be an important issue, and complete data on dialect intelligibility are not available.

The Ngakpo dialect (Omveda, dialect II) is the most likely choice for reference dialect. More up-to-date census information and careful intelligibility testing could conceivably put dialect III into greater prominence. The strategy for serving all the Pagibete people must be agreed upon by both SIL and CEUM, and probably this precludes the choice of dialect III.

3.4.2. Bua

Generally speaking, the linguistic data support a basic unity of the Bua dialects at a certain level (called the Bua cluster in this report) and a natural division into Bua east and Benge-Bati at another. This is supported by a comparative study of phonetics, lexicon, and morphology on the five word lists, and by analysis of the two phrase lists in the Buta dialect and in Benge. In some respects, the eastern Pagibete dialect shows similarities to the Bua cluster. In some other respects, the Kiba dialect is dissimilar to the Bua dialects spoken in the Bambesa and Buta Zones. Linguistic factors alone do not allow us to choose a reference dialect for all the Bua cluster dialects, nor to choose between Benge and Bati for a "Benge-Bati" standard, nor to choose between Bambesa and Buta dialects for a "Bua east" standard.

There are no clear results on either "declared" or "measured" understanding. Different interviewees from the same dialect often gave different impressions of how well they could understand those who spoke other dialects. The RTT results may be interpreted to suggest that people in the western zones (Benge and Bati speakers) can understand the eastern dialects a little better than vice versa. However, it is not possible to be emphatic on the point.

Except for two hymns in the CBZN songbook, which are apparently in the Bati dialect, the only printed material in Bua represents the Kiba dialect. This consists of Scripture tracts. There are a number of spiritual songs in Bua (transmitted orally); some of these may be the ones Past. Moziki wrote about thirty years ago, based on Scripture. We do not know whether these are adapted into the local dialect wherever they are sung, or even if they cross dialect boundaries.

Unfortunately, we are lacking in information on dialect attitudes in the Bua cluster. Both at Buta and at Malingwia, people said that the Yew dialect is best understood throughout Bua-land (this is the dialect spoken at Malingwia). It is possible that "Bua proper", being the dialect with the name "leBoa-le", might be considered the most pure. "Purity" can also mean freedom from outside influence, which might also favour Bua proper or Yew. (Benge, Bati, Kiba, and any Buta Zone varieties other than Yew have many non-Bua neighbours.) It is a matter of conjecture which varieties might be "prestige dialect(s)".

There are evidently more Bua speakers in each of the Buta and Bambesa Zones than in the Aketi, Banalia, and Bondo Zones together (though we don't have census figures from the

latter two zones). Certainly, there are many more speakers of "Bua east" than of Benge and Bati.

The Buta Zone is physically central to the Bua area, and shares a border with all of the other four zones in which Bua cluster dialects are spoken. Buta town is the politico-administrative centre for the whole area and is located at the intersection of the road to Kisangani and the road between Bumba and Isiro (as well as the presently non-functioning railroad covering the same ground).

Strictly speaking, Buta is on the edge of the Bua language area, but the sole neighbouring non-Bua collectivity only borders the northwest quadrant of town. Other towns and missions on the main east-west road are Dingila/Tobola and Bambesa (on the eastern and northern edge of the language area), Malingwia (in a Yew collectivity), Aketi (outside the language area) and Bunduki (on the west edge of the language area). Likati, a Benge centre, is on the road through Bondo to the Central African Republic. Kole is on the road through Banalia to Kisangani.

Buta is the seat of CBZN/14 and is a regional centre for CECCA/16 (though the district seat is Malingwia). Buta is also a Roman Catholic diocese (bishop's seat). CNCA and CADZ each have several chapels in Buta as well. Buta also has a teacher training college (Institut Supérieur Pédagogique). The Bua area is so large that no market, secondary school, or hospital could be expected to serve the whole area.

Another non-linguistic criterion, anthropological factors, may apply. It appears that the residents of the "Mobati" collectivity of the Buta Zone are ethnic Bati who now speak the Buta dialect. If there is in fact a single "Buta dialect", it would be the dialect of several clans, including Yew, Ngengita and Monganzulu (all of which are named in collectivities). To refer to the "Bambesa dialect" is to claim a single dialect for as many as five clans (suggested, again, by collectivity names).

Apparently, the Bua community has not addressed the subject of reference dialect. Of those interviewed, those who had given the most thought to Bua translation were CECCA/16 leaders in Buta and Malingwia. Their ideal was apparently a union translation so that all dialects could be included; such a plan, especially if efforts are made to accommodate the Bati and Benge, may not be feasible due to dialect differences. Buta is the best choice for project centre in order for all the interested church bodies to be involved in the project; an inter-denominational committee, to meet in Buta, has already been formed. Were the project only aimed at the people of the Buta and Bambesa Zones, a CECCA/16 project centred at Malingwia would make more sense.

On geographical grounds, the choice of Buta is the best. Buta or points further east could also be chosen on linguistic and demographic grounds. Kiba, in which Past. Masini has already done translation, is only spoken in the CNCA area and if the baKiba are willing and able to use written materials with a wider appeal, it would be better not to encourage separate development of the various dialects. It remains to be seen whether there will be agreement on use of the "Buta dialect" as reference dialect, or whether the team concludes that more than one translation will be needed.

3.5. Contacts

There are many educated people who speak the Bua bloc speech varieties. We are especially aware of several who speak the proposed reference dialects (Pagibete-Omveda and the "Buta dialect" of Bua). A few of these are listed in Appendix F.

At least two men have prepared translated Scripture in Bua bloc dialects and had their work printed. These are Past. Mbangiye, working in Pagibete, and Past. Masini, working in the Kiba dialect of Bua. They may be contacted through CEUM in Gemena and CNCA in Kisangani, respectively.

Past. Moziki of CBBU (now CBZN) in Buta, who died in 1972, reportedly translated almost two hundred Scripture songs in Bua. His manuscripts no longer survive, though his assistant, Temeni Aliaba, committed several of the songs to memory. Two hymns in Bua translated by Past. Moziki were included in the CBBU hymnal/songbook "Lembo na Sasaipi" (the title is in old Bangala and ironically means "new songs" or "current/modern

hymns"). These may have been in the Bati dialect (or maybe the pastor simply came from the Bobati collectivity of the Buta Zone).

Mr Kozapebo Charles, from the Yeu clan (Buta Zone), prepared a Bua-French "dictionary" in 1982. Apparently, he also noted dialect differences in Kiba and Bati. It has not been printed. The size of the work may be inferred from the fact that he wrote it in a 48-page exercise book. A graduate of the National University of Zaïre and an English speaker, he was very keen to help us.

A Bua language translation committee was formed in December 1991 and has invited the participation not only of CECCA/16, CBZN, and CNCA, as recommended by Tim Girard after his survey, but also of other church bodies, including CADZ/12.

Since CEUM is the principal protestant church with work in the Pagibete area, a Pagibete language project team, working under CEUM sponsorship, would need only to keep the Roman Catholic church and the collectivity and *groupement* chiefs informed of their progress.

4.0 CONCLUSIONS

It appears from our results that a combined Pagibete-Bua language project would not be feasible, both on linguistic and logistical grounds. Although the survey affirmed the reality of a "Bua bloc", the bloc divides clearly into the "Pagibete cluster" (three dialects) and the "Bua cluster" (in which we list five dialects).

Lexical dissimilarity is a clear example of the division. Each of the Pagibete lists was at least 90% to the other two. Each of the five Bua lists was at least 90% similar to each of the others. However, only the easternmost Pagibete dialect was more than 80% similar to Bua. Furthermore, in comprehension testing, the Pagibete speakers had difficulty understanding Bua, Bati and Bengé, and vice versa.

Logistically, the Pagibete people are separated from their Bua neighbors by a political boundary (Pagibete people in Equateur, others in Haut-Zaïre), and ecclesiastical boundaries (CEUM vs. other communities). In addition, especially at this time period, travel between the two regions is very difficult (we were unable to reach the Pagibete dialect area III because of the roads). Access to the Pagibete region is much more feasible from the west, while access to the Bua region is better (but not excellent) from the east.

Our research confirms the previous recommendations that projects be based in Ngakpo and Buta. The extent of the Bua bloc is now better documented: it may be said to consist of the three Pagibete dialects identified by the 1985 survey, the two very similar Bua dialects identified by the 1986 survey, plus Bengé, Bati, and Kiba, but not Bango or Benza.

During the course of the survey, contacts were made with most of the churches concerned by potential translation projects in Bua bloc languages. It was encouraging to find that the Buta area churches had indeed organised a Bua translation oversight committee; it was challenging to hear at Malingwia that they had been waiting twenty-two years for help with Bua translation.

Proficiency testing was not done in Bangala or Lingala.

5.0 RECOMMENDATIONS

Based on our findings, we want to reconfirm the recommendation from Fultz and Morgan (1986) for a Bible translation and literacy project in Dialect II (Ngakpo) of the Pagibete language. This translation should meet the needs of at least those Pagibete speakers in Dialect areas I and II, plus the Pagibete speakers resident in Businga (a sum total of over 12,000). The translation team should be administered by the Congo-Western Zaïre Group.

In addition, additional survey work needs to be conducted in the Dialect III area in order to determine more precisely the distinctiveness of the two dialects. We suggest that such a survey be carried out after a small body of literature has been produced in Dialect II, in order that it could be used for testing during that survey.

We also recommend that advisors be sent to Buta to help the Bua people set up a project of linguistic analysis, Bible translation and literacy. Such an assignment will depend on the

conviction of the Eastern Zaïre Group that SIL can provide all the necessary support for such a project, especially since it would represent a major geographical expansion of the work of SIL in eastern Zaïre. The SIL advisors should work closely with the existing Bua (project oversight) committee to develop a project team (working committee). The project team should include people from a variety of church bodies and with a variety of professional backgrounds.

The representation of the different dialects is a more complex matter. On the one hand, too much diversity may lead to deadlocks if a "union" translation is not feasible. However, the oversight committee may not readily accept the idea of a single reference dialect, which would probably call for a preponderance of speakers of that dialect on the working committee (for linguistic and stylistic work, though not necessarily for exegesis).

It is our impression that it would be better not to cherish the goal of a "union" translation if the result does not reflect natural kiBua. This could mean risking non-acceptance of the translation by speakers of dialects other than the reference dialect. We hope this can be avoided by having all the church bodies included in the decision-making. If possible, Past. Masini should be involved, as CNCA's leader in Bua translation (in an outlying dialect), even if only by correspondence. If funding is found to underwrite travel, it would be worth arranging air passage for Past. Masini and perhaps others who would not otherwise be able to travel to Buta.

It seems likely that at least some of the Kango people will be able to read Bua literature. The Bua team would be the natural choice to conduct survey among the Kango, perhaps with help from the Survey Department. At present, there are not even any reliable population figures for Kango. A complete Kango survey may not be feasible if, as has been suggested, there are different varieties on the Bima, Api, and Bomokandi rivers as well as at several points along the Uélé river, potentially along practically its entire length. (Van Bulek and Hackett 1956:80-81 and map, who suggest that Kango may be spoken further downstream than Bondo; Baxter and Butt 1956:map, who show a Kango speaking area along the Kibali upstream from Dungu.)

Finally, we propose the following revisions for the Ethnologue entries BWA, KANGO, and PAGABETE:

BWA (BOA, BUA, BOUA, LIBUA, LIBWALI, KIBUA, KIBWA, BATI, BAATI, BENGE-BATI) [BWW] 200,000 (1994 SIL). Haut-Zaïre Region, Bambesa, Buta, Banalia, Aketi, and Bondo Zones. Buta is considered the center. The majority, living in the Buta and Bambesa Zones, speak very similar dialects. The Kiba dialect is spoken in the Banalia Zone and Benge and Bati in the Aketi and Bondo Zones. Niger-Congo, Atlantic-Congo, Volta-Congo, Benue-Congo, Bantoid, Southern, Broad Bantu, Narrow Bantu, Northwest, C, Ngombe (C.50). Dialects: LEBOA-LE, YEWU, KIBA, BENGE, BATI. Bangingita, Bagunzulu, and Bokipa are clan names. First three dialects are only 90% similar to Benge and Bati. Approx. 80%-85% lexical similarity with Pagabete, 65%-80% with Kango, 60-65% with Ngelima, 55-60% with Lika, 47% with Komo, 43% with Bali, 37% with Ngombe, 35% with Budza. Dissimilarity from Ngombe suggests need for a new classification. Some noun classes have suffixes in addition to the usual prefixes. Different from Bua of Chad. Bible portions 1938.

KANGO (LIKANGO) [KTY] Haut-Zaïre Region, Bas-Uélé Subregion, along the banks of the Uélé River and its affluents. Niger-Congo, Atlantic-Congo, Volta-Congo, Benue-Congo, Bantoid, Southern, Broad Bantu, Narrow Bantu, Northwest, C, Ngombe (C.50). Said to be a pidginized language, Kango may be a catch-all term for a variety of dialects spoken by fishers (who are called "Bakango"). Lexical similarity for one dialect was 75-80% with Bwa, 70-75% with Pagabete, 66% with Ngelima, and 60% with Lika; but Bomokandi dialect has only about 60-65% with Bwa and Pagabete, 68% with Lika, and 58% with Ngelima. Both dialects about 50% similar to Bali and Komo. Different from Dikango, a pygmy language spoken among the Bali. Survey needed.

PAGABETE (APAGIBETE, APAGIBETI, APAKIBETI, APAKABETI, EGEZO) [PAG] 25,000 (1985 SIL); about 4,500 Mongbapele, 6,000 Momveda. Equateur Region, Businga, Yakoma, and Bumba Zones. Mongbapele is along the road south of Businga, Momveda in the area around Ngakpo on the north side of the Dua River, across from Gumba. Residents of Butu in the Yakoma Zone also speak Momveda dialect; a third dialect is spoken to the south of Butu and at Ndundu-Sana in the northern Bumba Zone. Niger-Congo, Atlantic-Congo, Volta-Congo, Benue-Congo, Bantoid, Southern, Broad Bantu, Narrow Bantu, Northwest, C, Ngombe (C.50). Dialects: MOMVEDA, MONGBAPELE, NDUNDUSANA (GEZON, EGEZON, EGEZO, EGEJO). First two dialects quite similar, about 90% lexical similarity to the third dialect. Similarity to Bwa: about 80% for Momveda and Mongbapele, about 85% for Ndundu-Sana. 60-65% lexical similarity with Ngelima, 45% with Komo, 43% with Bali, 35-40% with Ngombe and Budza, 30% with Lingala. Dissimilarity from Ngombe suggests need for a new classification. "Pagabete" comes from the speech-initial formula "Apagi bete" = "He says that..." Bwa dialects have similar formulas. Some noun classes have suffixes in addition to the usual prefixes. Vigorous language use. Most know Lingala; in one study, more than half of those tested in Ngakpo were at level 3 or above. Some also speak Ngbandi.

Several things are retained by necessity. Alternate names which appear in the literature are never omitted. The notation "Different from Bua of Chad" could be helpful to someone. The lexical similarity figures are an update of figures printed in the twelfth edition (Grimes 1992a) and may be omitted at the editor's discretion. Clearly, the figures cited at Lika, Bali, Komo, etc. will also need to be updated for the thirteenth edition. The classification ("Northwest, C, Ngombe (C.50)") is unchanged, although it would be better to change all the Zone C group numbers: Ngando and Ngundi of Rep. of Central Africa from C.10 and C.20 to C.00 and C.10, respectively, Mboshi of Congo from C.30 to C.20, Bangi-Ntomba from C.40 to C.30, Ngombe from C.50 to C.40, Kele from C.60 to C.50, Mongo from C.70 to C.60, Tetela from C.80 to C.70, and Bushong from C.90 to C.80.

The Bible Translation Needs Bulletin entries for Bwa should also be revised:

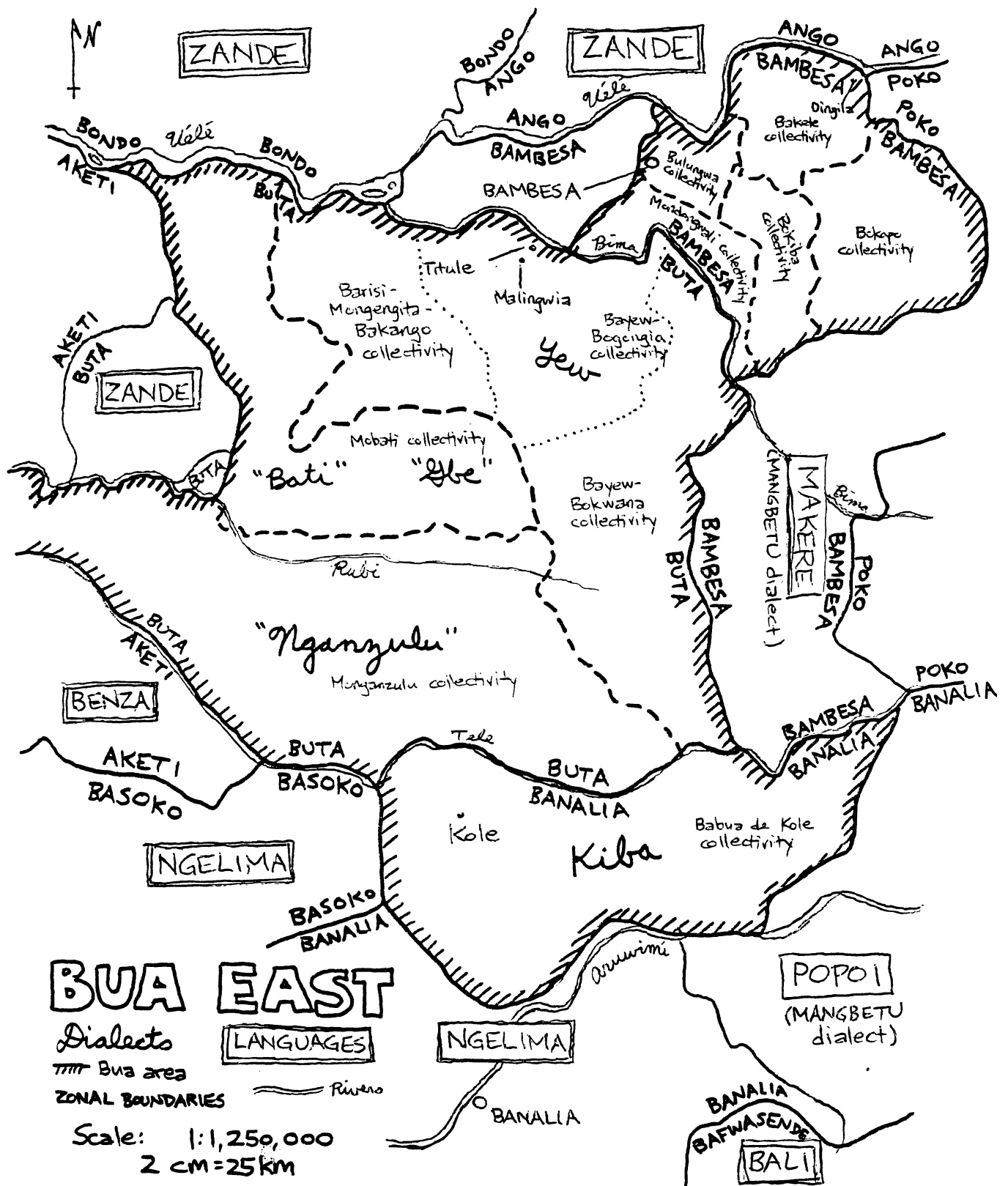
BWA [BWW] Bible portions 1938. Definite translation need. CECCA 16, CBZN/MBN, CNCA, CADZ, RC.

The former CEHZ, founded by UFM, is now Communauté Nations de Christ an Afrique. The former CBBU is now Communauté Baptiste au Zaïre du Nord. There are no longer any WEC or UFM missionaries in the Bwa area. (If, however, the mission must always accompany the Zairean church body, these would be CECCA 16/WEC, CNCA/UFM, CADZ/AOG.)

BIBLIOGRAPHY

- Barreteau, Daniel, ed. 1978. *Inventaire des études linguistiques sur les pays d'Afrique noire d'expression française et sur Madagascar*. Paris: Conseil International de la Langue Française.
- and Yves Moñino. 1978. Les langues oubanguiennes. In Barreteau (ed.), pp. 195-208.
- Bastin, Yvonne. 1978. Les langues bantoues. In Barreteau (ed.), pp. 123-185.
- Baxter, P.T.W. and Audrey Butt. 1953. *The Azande, and Related Peoples of the Anglo-Egyptian Sudan and Belgian Congo*. London: International African Institute.
- Bendor-Samuel, John, ed. 1989. *The Niger-Congo languages*. Lanham, Maryland: University Press of America.
- Bergman, Ted G. 1991. Rapid appraisal of languages. Notes on Scripture in use and language programs 28.3-11. Dallas: Summer Institute of Linguistics.
- Blair, Frank. 1990. Survey on a shoestring: a manual for small-scale language survey. Dallas: Summer Institute of Linguistics. (esp. Chapter 3)
- Boone, Douglas W. 1989. Toward an African area standard wordlist. Manuscript presented at the International Language Assessment Conference, Horsleys Green, England, pp. 217-228.
- Boyd, Raymond. 1989. Adamawa-Ubangi. In Bendor-Samuel (ed.), pp. 178-215.
- Bryan, M. A. 1959. *The Bantu Languages of Africa*. London: Oxford University Press.
- Casad, Eugene. 1974. *Dialect intelligibility testing*. Norman, OK: Summer Institute of Linguistics.
- Cole, Desmond T. 1959. Doke's classification of Bantu languages. *African Studies* 18(4), pp. 197-213.
- Cope, A. T. 1971. A consolidated classification of the Bantu languages. *African Studies* 30(3-4), pp. 213-236.
- Fultz, Jim and David Morgan. 1985. *Enquête dialectale de l'Ubangi et de la Mongala, région de l'Equateur: Troisième Partie*. Gemena, Zaïre: Association Wycliffe pour la traduction de la Bible.
- Girard, Tim. 1991. Bwa survey. Unpublished ms. Nairobi, Kenya: Summer Institute of Linguistics.
- Grimes, Barbara F. (ed.). 1992a. *Ethnologue: Languages of the World*, 12th edition. Dallas: Summer Institute of Linguistics.
- (ed.). 1992b. *Bible Translation Needs: Bulletin 2*. Dallas: Wycliffe Bible Translators, International.
- Guthrie, Malcolm. 1948. *The classification of the Bantu languages*. London: Oxford University Press.
- . 1971. *Comparative Bantu. Part 1, Vol. 2*. Farnborough, England: Gregg International Publishers.
- Hinnebusch, Thomas J. 1989. Bantu. In Bendor-Samuel (ed.), pp. 450-473.
- Kadima Kamuleta et al. 1983. *Atlas Linguistique du Zaïre, Inventaire préliminaire*. Agence de Coopération Culturelle et Technique, Equipe Nationale Zaïroise.
- Kahombo Mateene. 1969. *Essai de Grammaire Générative et Transformationnelle de la Langue Nyanga*. Doctoral thesis, Sorbonne. (Reprinted by Presses Universitaires du Zaïre, 1980.)
- Keller, Chuck (with John Bendor-Samuel?). 1977. *Preliminary Assessment of Translation Needs in Zaïre*. Unpublished ms.
- Kutsch Lojenga, Constance. 1994. *The Vowel System of Bila, A Bantu language of the Northern Bantu Borderland in Zaïre*. Paper presented at a Leiden Colloquium. Unpublished ms. (9 pp.)
- Masumbuko-wa-Busungu. 1979. *Esquisse Grammaticale du Lega: Phonologie et Morphologie*. Travail de fin d'études, Institut Supérieur Pédagogique, Bukavu.
- Mathis, Ed and Sheryl Mathis. 1990. *Report on Conversation with Pastor Masini*. Unpublished ms. (2 pp.)
- McCord, Michael and Robert McKee. 1986. *Kibua survey report (draft)*. Unpublished ms. (4 pp. plus map). Nairobi: Summer Institute of Linguistics.
- Morgan, David. 1989. *Testing comprehension of scriptures in a bilingual situation*. Manuscript presented at the International Language Assessment Conference, Horsleys Green, England, pp. 419-428.
- Phillips, John. 1992. *Developing a sentence repetition test to study Lingala bilingualism*. Brazzaville: Summer Institute of Linguistics.

- Radloff, Carla F. 1991. Sentence repetition testing for studies of community bilingualism. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
- Reeder, JeDene and Sharon Stoothoff. 1994. Report on February 1994 Pagabete survey, Momveda dialect. Unpublished ms. (4 pp. plus bibliography, map and appendix). Gemena: Bureau de Traduction de la Bible.
- Stalder, Jürg. 1993. Rapid Appraisal: Strategy and methodology as applied in Cameroon. Manuscript presented at the second International Language Assessment Conference, Horsleys Green, England.
- Stalder, Jürg. To appear. Recorded Text Testing in the context of Rapid Appraisal. Unpublished ms.
- Stappers, Leo. 1971. Esquisse de la langue lengola. In: *Africana Linguistica* V. (Annales, Série in-8°, Sciences Humaines no 72.) Tervuren, Belgium: Musée Royal de l'Afrique Centrale.
- Thomas, Elaine. 1979. Report on visit to Kisangani, Jan 23-29, 1979. Unpublished ms. (2 pp. plus map)
- Van Bulck, G. and Peter Hackett. 1956. Report of the Eastern Team. In *Linguistic Survey of the Northern Bantu Borderland*, pp. 63-122. London: Oxford University Press for the International African Institute.
- Voegelin, Charles F. and Florence M. R. Voegelin. 1977. *Classification of the World's Languages*. New York: Elsevier for the Indiana University Foundation.
- Wimbish, John S. 1989. WORDSURV: A Program for Analyzing Language Survey Word Lists. Dallas: Summer Institute of Linguistics. (Occasional Publications in Academic Computing 13)



Appendix B: Sociolinguistic Questions for Group Interviews

Chief participants:

Village or dialect in focus:

Date and place of interview:

NOTA BENE: [LM] = Langue maternelle = lokóta na yó/bínó = *mother tongue*
 lokóta = langue = *language*, monoko = dialecte = *dialect*
 [GE] = Groupe ethnique = libóta na yó/bínó = *ethnic group*

1. (GLOSSONYMS, ETHNONYMS)

- 1.1. Na lokóta na bínó, bobiángaka lokóta na yó ndéngé níni?
In your own language, what do you call your language?
- 1.2. Batu mosúsu bobiángaka lokóta na bínó na nkómbó níni?
What name do other people have for your language?
- 1.3. Na lokóta na bínó, bobiángaka bínó mpenzá ndéngé níni (libóta na bínó)?
In your language, what do you call the name of your ethnic group?
Ndéngé níni bobiángaka motu mókó (kati na libóta na bínó)?
What do you call one person (in your ethnic group)?
- 1.4. Bakonzi na letá bobiángaka lokóta na bínó na ndéngé níni?
What do government officials call your language?

2. (DIALECTS)

(Nsimá na kozua "liste" na minoko na [LM])

(After getting a list of dialects)

- 2.1. Na káti na minoko óyo totángákí, monoko níni ekéséní bobélé moké mpenzá na monoko na yo?
Amongst the dialects that we listed, which ones differ just slightly from your dialect?
- 2.2. Mpé, nsimá na yangó, monoko níni mosúsu ekéséní moké lokóla?
(kotuna mbala lokóla minoko bazalí, kínó kokuma na óyo na mosíká koleka)
After that, what other dialect differs slightly? (Ask for each dialect up to the one that differs the most).
[i.e. first two questions are to rank the dialects in similarity to hometown]
- 2.3. Banáni basololaka [LM] na sôló? Mpó na níni bolobí boyé?
Who speaks the true [LM]? Why?
- 2.4. Monoko níni na [LM] ezalí kitóko koleka minoko (bamonoko) mosúsu?
Which dialect of [LM] is the best/prettiest of all dialects?
- 2.5. Mbóka mosíká ezalí, óyo batu yangó basololaka ndéngé mókó lokóla bínó?
Are there villages far away who speak the same language as you?

3. (HISTORY)

- 3.1. Batu [GE] baútákí wápi? *Where do the [GE] people come from?*
- 3.2. Mpó na níni balongwákí esíká yangó? *Why did they leave that place?*
- 3.3. Bakúmákí áwa ntángo níni? *When did they come here?*
- 3.4. Nkómbó na bankóko yangó ezalákí banáni? *What were the names of your ancestors?*

4. (CHURCH LIFE)

- 4.1. Bakompanyí (ba-communautés) níni basálaka mosálá kati na mokili na batu [GE]?
< tó, lingómbá níni azalí áwa? >
What churches are working in this region?
- 4.2. Lokóta níni bosololaka na yangó mbala míngi kati na losámbo mpo na:
What language do you use in church to:

a. kosakola? <i>preach?</i>	d. kosámbela? <i>pray?</i>
b. konjémba? <i>sing?</i>	e. kopésa mayébísi? <i>make announcements?</i>
c. kotánga Maloba na Nzámbe? <i>read the Bible?</i>	f. kotánga liturujia? <i>recite liturgy?</i>

(NB. C'est courant de dire «ndáko na Nzámbe» au lieu de «losámbo»)

- 4.3. Ntángo basakolaka na lingála, motu azalí kobálisa/kobóngola yangó na [LM]?
When the preaching is in Lingala, does someone translate into [LM]?
- 4.4. Okosala níni mpó na kozua Biblia na lokóta na bínó mpenzá?
What will you do in order to receive the Bible in your own language?

5. (YOUTH)

- 5.1. a. Bána na mbóka óyo, babandaka koyékola lokóta níni úta bomwána mpenzá?
What language do children in this village learn first?
b. Nsima na yangó, bayékolaka lokóta níni na míbalé, mpé ndéngé níni?
Next, what language do they learn second? How?
c. Ndé kolanda, bayékolaka lokóta níni na mísáto, mpé ndéngé níni?
Then, what language do they learn third, and how?
- 5.2. Bilengé (Ba-palangá) na mbóka óyo basololaka [LM] ndéngé ekokání mpenzá?
Do the youth in this village speak [LM] properly?

6. (INSTITUTIONS)

- 6.1. Kelási na libosó [écoles primaires] ezalí wápi (ba-mbóka níni)?
In which villages are there primary schools?
- 6.2. Kelási na nsima [écoles secondaires] ezalí wápi (mbóka níni)?
In what village is the secondary school?
- 6.3. Bána míngi bakendaka na kelási? [Sókó tê,] mpo na níni?
Do many children go to school? If not, why?
- 6.4. Bána babáli bazalí míngi koleka bána básí, tó bazalí penepene esíkámókó?
Are there more boys than girls, or are there the same number of each?
- 6.5. Balákisi basololaka lokóta níni kati na kelási?
What language do the teachers use in class?
- 6.6. Bisíká na mônganga [dispensaires] ezalí wápi?
In which villages are there dispensaries?

7. (INTERMARRIAGE)

8. (MIGRATIONS)

9. (DIALECT RELATIONS)

*** ASK THESE QUESTIONS CONCERNING THE "CENTRES" OF EACH OTHER LANGUAGE OR DIALECT FOR WHICH INTELLIGIBILITY IS IN QUESTION ***

- 9.1. Ntángo okeí kati na mbóka (X), bokokí koyóka malámu lokóta kúná?
When you go to the village, (X), can you understand the language there?
- 9.2. Sókó bosololi lokóta na bínó kúná, bakoyánola na lokóta níni?
If you speak your language there, in what language will they respond?
- 9.3. Bolingí kosolola lokóta níni kati na mbóka yangó?
What language do you prefer to speak in that village?
- 9.4. [Sókó ezalí [LM],] Bobálisi/bobóngólí maloba na yó mpó ete bangó báyóka yó, tó bangó babálisi/babóngólí maloba na bangó, tó batu nyóso basololaka lokóla bamésání?
[If [LM],] Do you change your language so that they can understand you, or do they change their language, or does everyone speak as they are used to speaking?
- 9.5. Mwána moké na áwa óyo azalí na mbúla motóbá, akokí koyóka malámu batu na mbóka (X)? *Can a six-year-old child from here understand the people from the village (X)?*
- 9.6. [Sókó boyé tê,] Eséngélí yé ázala na mbúla bóni mpo na koyóka malámu monoko na bangó? *[If not,] how old does he have to be in order to understand their language?*

10. (MEDIA)

11. (ATTITUDES)

12. (HOMOGENEITY)

- 12.1. Awa na mbóka na bínó, batu bazalí óyo bayébi kosolola [LM] tē?
Are there people in this village who do not know how to speak [LM]?
- 12.2. [Sókó boye,] Basololaka na lokóta/nkóta níni? *[If so,] What language do they speak?*
- 12.3. Bamekí koyékola [LM] (lokóta na bínó)? *Do they try to learn [LM]?*
- 12.4. Motu azali kati na mbóka na bínó, óyo ayébi [LM] káka? Bazalí náni?
Are there people in this village who know only [LM]? Who?
- 12.5. [Sókó boye tē,] Bozalí na batu óyo basololaka lingála tē? Náni?
[If not,] Are there people here who do not speak Lingala? Who?
- 12.6. Batu na [GE] bazalí óyo bayébi/basololaka [LM] lisúsu tē? Bazalí náni?
Are there people of [GE] who no longer speak [LM]? Who?

13. (CHURCH STATISTICS)

14. (LANGUAGE DEVELOPMENT)

- 14.1. Osílí komóná búku ekomámí na [LM]? *Have you seen a book written in [LM]?*
[Sókó boye,] Okókákí kotángá yangó? *[If so,] Could you read it?*
Batu bakokí kozua búku yangó lisúsu leló? *Is that book still available?*
- 14.2. Osílí koyóka nsango na elóko mókó ekomámí na ntína na batu barambo tó lokóta na bangó? (Ou, «osílí koyóka nsango na motu óyo ekomákí sókó búku, sókó travail de fin d'études likoló na batu barambo?»)
Have you heard of anyone who has written a book or a "travail de fin d'études" on your ethnic group or language?
- 14.3. Kelási ezalí mpo na mikóló (batu óyo bakólí) mpó ete báyékola kotángá?
Are there literacy classes for adults?
Náni abandísákí yangó? *Who began them?*
Batu míngi balandákí kelási yangó? *Are there many people who take that class?*

Former questions for dialectology:

- 2.1. Tángélá ngái mbóka mosúsu óyo batu na mbóka yangó basololaka ndéngé mókó na bínó... *Name for me other villages where the people speak the same language as you.*
- 2.2. Tángélá ngái mbóka mosúsu óyo monoko na batu yangó ezalí mwa godi...
Name for me other villages where the language is somewhat different.
- 2.3. Tángélá ngái mbóka mosúsu óyo monoko na batu yangó ekésení (ezalí godi) kasi boyókí maloba nyóso...
Name for me other villages where the language is different, but you understand it well.
- 2.4. Tángélá ngái mbóka mosúsu óyo monoko na batu yangó ekésení míngi, mpé bokokí koyóka maloba nyóso tē.
Name for me other villages where the language is very different, and you don't understand it well.
- 2.5. Mbóka mosúsu níni bazalí na lokóta godi mpenzá?
What villages have a completely different language?

Batu [GE] nyoso basololaka mpenzá ndéngé mókó, tó bokeseni (bogodi) ezalí?

Do all the people in this ethnic group speak the same, or are there differences?

(Sókí bokeseni ezalí,) Batu wáná óyo basololaka ndéngé mosúsu bazalí wápi?

(If there are differences,) Where are those who speak differently?

Osílí kosolola na motu óyo aútí kúna?

Have you spoken with someone who came from there?

(Sókí boye,) Okokákí koyóka makambo nyóso alobákí?

(If so,) Could you understand all that he said?

Ntángó osolólákí na motu yangó, osolólákí na ndéngé mókó osololaka na motu óyo

afándaka áwa, tó obúngólákí ndéngé na kosolola mpó motu yangó áyóka yó?

When you spoke with that man, did you speak in the same manner that you speak with people here, or did you vary your speech so that he could understand you?

Appendix C: Description of the Rapid Appraisal RTT

The following modified Recorded Text Test procedure is drawn from Stalder (to appear) [Ed.: I have edited freely, so that while the outline conforms to Stalder's description, the wording is often different]. In keeping with the spirit of the Rapid Appraisal Approach, the RA RTT is intended to provide a general impression of the sociolinguistic situation in an area without recourse to involved and time-consuming procedures.

Like the conventional Recorded Text Test for dialect intelligibility testing described by Casad (1974), the first step consists of recording and transcribing 3-4 minute stories in several dialects. The recordings are then played in different dialect areas in order to determine the amount of inter-dialect communication (an indirect test of inherent intelligibility).

The RA RTT differs from the "Casad" method at this point. Unlike conventional RTT scoring, communication is evaluated in a qualitative manner. Also, testing is not done individually, but in a group, chosen to include men and women of a variety of ages.

More specifically, one administers the test as follows:

- Choose one story as reference tape [if several have been recorded in the target dialect].
- Let them [the group at the test point] hear the entire story.
- Replay the story bit by bit. After each section, let people retell the section [NOTE: in the trade language, in our case, Lingala or Bangala]. Make sure that the group as a whole is involved. Do not hesitate to interact and to ask for details if necessary.
- If it is not possible to rate the level of comprehension as mentioned below, use the other story(ies).

The amount of comprehension [NOTE: Stalder uses the term "intercomprehension"] is rated using a three-tiered scale:

1. No comprehension. The group is not able to repeat even the general story line.
2. Partial comprehension. By retelling the different sections, people invent and add to the story. If asked, they are not able to answer details.
3. Good comprehension, i.e. the story is retold accurately and the people are able to give details.

Usually the group being tested consists substantially of the same people who have participated in the group interview which is central to the RA method. It is interesting to observe attitudes during the RA RTT and to compare actual understanding with their responses to the interview questions having to do with intercomprehension.

Stalder suggests the following levels for decision-making, which correspond to the levels of comprehension:

1. Use of a common translation is impossible.
2. If the attitude is positive, there is a low probability for the use of the translation in the test language.
3. If the attitude is positive, there is a high probability for the use of the translation in the test language.

Appendix D: Recorded Texts used in RA RTT

Literal Translations of Texts from 1994 Bua Bloc Survey, used as Rapid Appraisal Recorded Text Tests (Words in brackets are implied)

Bua, recorded in Buta

That which I recount comes from an old man, an old soldier, whose name is Kabui, that he told when we were in church. He lived in a region where there were many thieves. In that region, there was a man who raised a goat. As he knew that there were many thieves in that region, he didn't want that anyone would steal his goat. Since the goat was still young, he wanted to go everywhere that [the man] went. Even if he went to draw water, he was accompanied by the goat; he would attach the goat to a tree; he would draw water, then he would get his goat to return. Even when he went to the field, he was always accompanied by this goat. When he arrived at the field he would attach the goat to a tree, and then cut his field.

During the tree felling time, since he was used to going with the goat, he brought his son. This time, the goat was attached a bit farther away, so that a tree would not fall on it. During this time, when he went in to cut the trees, he would send his son to go see the goat. When his son went, he found the goat still attached.

Now the thieves wanted to steal the goat. They came silently, since the owner could not see [the goat] because of the trees. They cut the goat's throat. They took the head and attached it to a tree. They took the body. And then they disappeared.

Afterward, when [the man] sent his son to go see the goat, he found the face of the goat looking at him. (But the body of the goat was no longer there, only the head). They continued to fell the trees, and at the end, they came and found only the head. They came in order to detach the goat, but it was only the head. The father began to scold his son. "Since I began to send you to see the goat, the goat was no longer here, only the head. The goat was already stolen."

His son explained, "Since I saw the head, I thought that the goat was still there." Now the goat was already stolen.

They went home unhappy. This is an example. Many Christians in the church are like that. Satan has already cut them at the throat, and taken the body and the heart with him. Only the face remains in the church. All the time, he comes only with the face for showing to the church, but he personally is no longer there. He goes home, he comes back, but he is no longer there. He shows only his face, he shows only his face, then he returns home.

That's the moral of the story that the man told. I think that that's the end.

Benge (Likati), recorded in Aketi

My name is Ngalika Zoka. One day, I left the house and went to the fields. I began to weed the peanut field. As I was weeding the peanut field, the rain began to fall. Since it began to rain, I left and went in the field shelter. The rain continued for a long time, so I could not return to town.

The night began to arrive. I took my clothes and attached them to my bicycle rack; including my identity card. I left the field and was on the main road. The night arrived. When I came to the road, and my clothes fell off the bicycle.

I arrived in town, and I asked my wife, "Come look at the bicycle to see if the clothes are still there."

She said, "They aren't there."

I cried, "It's death! What will I do?" After two days of looking, I didn't find them.

One of my friends came from the fields during the night. He had a flashlight in his hand. He shined the flashlight on the ground, and he saw the clothes. He didn't "consider" the

clothes, so he didn't take them. The clothes had fallen to the side; the papers had fallen further on. He saw the identity card, he took it, and he put it in his shirt pocket.

When he looked at the papers, he saw my name on the work card. Since he knew how to read, he knew that this man worked at the Catholic mission. The man went with the identification papers to the Priests. They told him where my house was. This friend came with the papers, he found me, and he gave them to me.

That friend was a good Christian. He helped me well. If all Christians were like him, if they all followed his example, it would be great. That's the end.

Bati (Bogbasa), recorded in Aketi

Me, a Bati, I am speaking. My name is Magindo. I have one story. There once was a papa. Before, we would dig big holes. Animals would fall in, and we would eat them. We didn't know about guns.

One day, this man left to go to the forest. He went to the forest to look for animals that had died (in the traps) to bring back to eat. He was walking, walking, walking. A friend of his had made a hole. He fell into it. As he was walking he didn't know that he would fall. There was no means to get out of it. No one else came to let down a branch to help him get out.

At that time, his friend (not the one who had dug the hole) was passing by on the road. He heard cries, he heard cries, he heard cries. "Why is that man crying? I will go see what is happening there." He went, he arrived there, and he found his friend who had fallen in a hole that had been dug by someone else. His friend asked, "What's going on there?"

"Me, I came to look for game. I pierced the hole that my friend had dug, and I fell in." He [the third person] got his things. He put them down. He took his machete, and he cut some bamboo in order to save the other. But the other said, "Before I leave here, go tell my wife that she should get a chicken; I'll eat it, then I'll leave. She should come, give me (the chicken), I'll eat it, then I'll leave."

Then he went, he got his wife, and said, "I found your husband fallen in a hole. He said that before he would leave, you should cook a chicken. Come with it, he will eat it, then he will leave."

The wife chased the chicken, then she caught it. She prepared the chicken for cooking for going there. The rain was threatening, the rain was threatening. The rain began a little bit. Then it became strong. The rain filled the hole. The man drowned. The hunter died in this way.

Thus in our days, people are serving God. They say to you and me that we should accept the Lord. We should listen to only that which God says. Because if we listen to the Word of God, we will be saved. But, you and me, we say that before we accept the Lord, let me go prepare alcohol, let me go finish mourning (the death of) my father, let me smoke cigarettes. Then I will come accept God. When you are there, death will come to kill you in your impurity. You lack the kingdom that God has prepared for me and you. The hole that we talk about is that road.

Pagibete, recorded in Ngakpo

My name is Londo Nvungbo. My father is a hunter. He called me to accompany him in the forest to look for animals. We arrived at the camp. In the morning, we left and we saw the tracks of a wild boar. We began to pursue it. We went and the wild boars entered into a thicket. We began to walk on our knees. We walked, we walked, we saw a wild boar in its bed. Others were farther on, eating. (My father) came closer to it and had a little gun. He shot it in the head. Seeing that it was already dead, I arrived behind (the pig). On arriving behind him, the wild boar began to "jump around." The father began to cry that he (the kid) should bring the big gun. I said, "Papa, move away!" I fired and hit the buttocks of the animal.

He stood up to pursue me, and I jumped up and perched on a tree branch. On seeing this, papa cried, "What is this animal?" On hearing the voice, the animal chased papa and papa, in his turn, perched in a tree, and both of them began to tremble. As the wild boar began to throw himself here and there, we were still in the trees. He left blood, and left.

That shows that it is important to seek the kingdom of heaven. Because if we look for that which is on earth, it does not have its reward. At that hour there, if we had been dead, it would have been for the problem of the earth, and there would be no reward. At the end of the world, we will have the reward with our Father in heaven. That's all.

Appendix E: Comprehension testing results

(1)

Language group: Pagibete (Ngakpo dialect)

Location: Mondongbo

Date: June 30, 1994

Bua: Some comprehension

Benge: Good comprehension

Bati: Good/some comprehension

Comments: We tested three or four men in this village. The CEUM pastor pretty much dominated the testing. He had traveled widely.

(2)

Language group: Pagibete (Ngakpo dialect)

Location: Ngakpo

Date: July 2, 1994

Bua: Some

Benge: Some

Bati: Some

Comments: Men over 30 who had traveled had good comprehension. They could re-tell the stories in detail. However, younger men had much difficulty even giving the story-line.

(3)

Language group: Pagibete

Location: Ngakpo (man from Bauma)

Date: July 4, 1994

Bua: Good

Benge: Good

Bati: Good

Comments: This elderly man from Bauma had good comprehension. He said that the language is called "Pagibete" and not "eGezo". He claimed to understand Bua well, even fast speech. He also said that boBua could understand him.

(4)

Language group: Bati

Location: Bunduki

Date: July 8, 1994

Bua: Good

Benge: Good

Pagibete: Some

Comments: They had a few difficulties with the Bua story. As a group, they were able to reconstruct the Pagibete story with details, but not individually. They identified where the Pagibete story came from.

(5)

Language group: Benge

Location: Aketi

Date: July 10, 1994

Bua: Good

Bati: did not test (many Bati in town, so we assumed good comprehension)

Pagibete: Some

Comments: They thought that the Pagibete text was Ngɛɛma!

(6)

Language group: Yeu

Location: Malengwia

Date: July 13, 1994

Bɛngɛ: Good

Bati: Good

Pagibete: Some

Comments: They understood Bati the best. They thought that Pagibete was a variety of Bɛngɛ. Thought that Pagibete was more difficult than Ngɛɛma!

(7)

Language group: Bua

Location: Dingila

Date: July 14, 1994

Bɛngɛ: Good

Bati: Good

Pagibete: Some

Comments: They understood Bati the best. They actually could give more details for the Pagibete story than any other test group could!

Appendix F: Some educated speakers of Bua bloc dialects

This list is neither complete, representative, nor even very suggestive. With time, we expect that very many more names will be added. We mention these few names so that they are not overlooked in the future.

Bua		
<u>Name</u>	<u>Affiliation</u>	<u>Residence</u>
APONZA	<i>Coordinateur des Ecoles Protestantes</i> (à Yew)	Yew
BANALE	Head pastor, Malingwia District, CECCA (graduate of Isiro Bible Institute)	Malingwia
KOZAPEBO Charles	Clerk at subregional office (has studied at Univ. of Lumbambashi. a Yew. Speaks English well)	B.P. 1, Buta
MBOKPA Alphonse	Pastor, head of Buta Parish, CECCA (studied at Ibambi Bible School; speaks "Buta dialect")	Buta
MOKPUTU	Pastor, CBZN (a Bati from Aketi Zone)	Buta
TEMENI Aliaba Alphonse	<i>Surveillant</i> , CBZN	Gbobo (Buta)

At least four *travaux de fin d'études* were done on some aspect of the Bua language at the Institut Supérieur Pédagogique - Buta between 1988 and 1991. There were doubtless several more done in previous years and possibly some done at other ISP's or universities.

Pagibete		
<u>Name</u>	<u>Affiliation</u>	<u>Residence</u>
BONZA	(University graduate)	Belgium
BOWA Malikési	<i>Pasteur Responsable</i> (CEUM)	Mombassa (south of Bumba)
ELANGA	Doctor	Belgium
ELOMO Manzimbi	Nurse	Ngakpo
KPOBE	Pastor (inactive)	Businga
LOKAME	?	Ngakpo
MAKPANDA	?	Ngakpo
MBANGIYE	<i>Pasteur Responsable</i> (CEUM)	Abuzi
Epapola		
Justin MBOYAYE	son of Pastor Mbangiye	?
NGIATE Mawa	Headmaster (<i>préfet</i>)	Mongili
NZONGO	<i>Aumônier, Institut Biblique</i>	Gbado-Gboketsa
NZONGO Zéphyrin	School director (a Catholic, <i>professeur responsable, Institut Kokomedj</i> , his training is: <i>cycle long, littéraire (option Humanités)</i>)	Bombwa
YANGBA Mbo	School director(?)	Businga

All the above are speakers of the Omveda dialect. There are four other graduates of the Univ. of Kinshasa, still resident at Kinshasa: KIMENGO, LIAMA, MAGONGO, SOMBIKA Aloïs

Appendix G: Grammatical data—Statements, questions and answers, and other phrases in Bua bloc dialects

These data were elicited in June-July 1994 by Ken Olson and Douglas Boone as part of the Bwa bloc survey. The three lists represent:

- Pagabete (as spoken in the Ngakpo area) by DWB in Mondongbo on 4.VII.94.
- Benge (closely related to Bati) by DWB in Aketi on 9.VII.94.
DWB had Pagabete data on hand.
- Bua ("Buta dialect") by KSO in Buta on 17.VI.94.

It was hoped that these data would shed light on grammatical similarities and differences between the various speech forms. Vocabulary differences (nouns and verbs) are less in focus than how sentences are formed. Some of the grammatical elements under consideration are verb tenses, negative constructions, demonstratives and anaphoric markers, and possessives.

1. He ate (mpondu) (yesterday).
Il mangea (du mpondu) (hier).
Yé alíákí (mpondú) (lóbí).
Pagabete Yí ayí yáka. [Yí apí ɔ̃guⁿza (yoko).]
Benge Yí àní (gbàⁿdá) (yíò).
Bwa-Buta Yí àyáyí tò.

COMMENTS: Pagabete cites different form for "no object".
I checked Pagabete "ayí" and he said it was correct.

2. The man ate (mpondu) (yesterday).
L'homme mangea (du mpondu) (hier).
Mobáli alíákí (mpondú) (lóbí).
Pagabete Nòkó apí yaka. [Nòkó apí ɔ̃guⁿza.]
Benge Mòló àní (gbàⁿdá) (yíò).
Bwa-Buta Mòló (mò) àyáyí tò.

COMMENTS: Buta "mo" = demonstrative? (cf. #11) Otherwise all same as #1.

3. He is eating (mpondu).
Il est en train de manger (du mpondu).
Yé azalí kolía (mpondú).
Pagabete Yí (ndó) gaka ayá ya. [Yí (ndó) gaka apá ɔ̃guⁿza.]
Benge Yí àpálê (gbàⁿdá). OR Yí páná ɲă (gbàⁿdá).
Bwa-Buta àpá yâ yâ.

COMMENTS: As in ##1-2, first two have p in verb but Buta has y.
Pagabete word order is unexplained (cf PARADIGM on last pg.)
Benge first option may be "àpá èlè"
Two or three different strategies for present progressive?

4. He will eat (mpondu) (later).
Il mangera (du mpondu) (plus tard).
Yé akolía (mpondú) (nsima).
Pagabete Yí ayá lo. [Yí apá lo ɔ̃guⁿza.]
Benge Yí àpá (gbàⁿdá) (kátìⁿdà).
Bwa-Buta Yí àyâ yâ tò.

COMMENTS: Benge speaker said that "Yi apá lo" means 'Ye álía leló'.
(on #4) ('Let him eat today.')

5. Who will eat (mpondu)? The man will eat (mpondu).
 Qui mangera (du mpondu)? L'homme mangera (du mpondu).
 Níni akolía (mpondú)? Mobáli ndé akolía (mpondú).
 Pagabete Kǎné mò lò ǎná Ƴguⁿza? Nòkó Ƴdò lò ǎná Ƴguⁿzà.
 Benge Kǎné mǎná (gbàⁿdá)? Mòló ná ǎná (gbàⁿdá).
 Bwa-Buta Kǎné mǎyâ yâ? Mòló àyá yà tò.

COMMENTS: Significant change of word order in Pagabete? (cf. #4)
 Division of Pag. "mòlò" is speculative; poss. "mo lò" (cf. #11)
 [perhaps 'Who is this who will eat? The man is the one...']
 Bwa-Buta not elicited with emphatic (Lingala 'ndé').

6. Did he eat (mpondu)? No, he did not eat (mpondu).
 Est-ce qu'il a mangé (du mpondu)? Non, il n'a pas mangé (du mpondu).
 Yé alíákí (mpondú)? Tê, yé alíákí (mpondú) tê.
 Pagabete Yí ǎní áká Ƴguⁿzà? Ká, yí kǎní àká Ƴguⁿzà kà.
 Benge Yí ǎní tò (gbàⁿdá (mò))? Ká, yí kǎní tò (gbàⁿdá (mò)).
 Bwa-Buta Yí àyáyí tó? Yí kàyáyí lê.

COMMENTS: Compare question with statement (#1). Benge has demonstrative?
 Buta speaker omitted initial 'No' in answer. Pag. triple neg.!

7. Is he eating meat? No, he is not eating meat.
 Est-ce qu'il mange de la viande? Non, il ne mange pas de la viande.
 Azalí kolía mosuni? Tê, yé azalí kolía mosuni tê.
 Pagabete àyágà nǎmé? Ká, yí kàyégà nǎmé kà.
 Benge àyéìlè nǎmé? (Yí) kàyèìlè nǎmé.
 Bwa-Buta Yí àyá tò nǎmé? Yí kàyé nǎmé.

COMMENTS: Pagabete speaker may have used the habitual tense (cf. #10)

8. What is he eating? He is eating fish.
 Qu'est-ce qu'il mange? Il mange du poisson.
 Yé azalí kolía níni? Azalí kolía mbísi.
 Pagabete Yí àyágá kí? ỳyágà bàsí.
 Benge Yí àyèlè té? àyéìlé Ƴsí.
 Bwa-Buta Yí àyá ètè? óyálé Ƴbàsí.

COMMENTS: Word list 'what': èké; tê; èsínà (subject position?) (resp.)
 Pag. and Buta subj. prefix o- in response. (due to pl. obj.?)

9. If (when) he eats meat, tell me.
 Si (quand) il mange de la viande, dis-moi.
 Sókó yé aléí mosuní, lobela ngái.
 Pagabete Gwété yí ǎná nǎmé, èkpélégé èmè.
 Benge àyí àyàlè nǎmé, èpáké mè.
 Bwa-Buta álŷátè áyéyámé, èpágí námè.

COMMENTS: Word list 'say': -kpèté, -pàké, -pâgè (resp.) [cf. "Apagibete"]
 Pagabete vb. may be 'tell unto' (cf. Lingala 'lobela')
 Missing syllable in Buta 'he eats meat'? ("ayeya nǎmé")
 Apparently three different strategies for the conditional

10. If he had eaten meat, that would have been good (but he didn't).
 S'il avait mangé de la v, cela aurait été bon (mais il ne l'a pas fait).
 Sókó yé alíákí mosuni, mbele ezalákí malámu (kási alíákí yangó tê).
 Pagabete èlì áyákà yí àyá pámé, yíbà àlìkí éⁿzà.
 Benge álì àyé pámé, áðì àl' éⁿzà.
 Bwa-Buta àlYá bètè àyí pámé, ábì àb' éⁿzà.

COMMENTS: Possibly Benge and Bwa share "aði" which Pagabete does not use.

11. This man (here) eats meat (every day).
 Cet homme (ici) mange de la viande (tous les jours).
 Mobáli óyo (áwa) alíaka mosuni (mikolo nyóso).
 Pagabete Nòkó mó àyágà pí pámé (móné màázù).
 Benge Mòló mò nàgíwè àyákàkà pámé (bàbìtí yásò).
 Bwa-Buta Mòló mô àyá tò pámé.

COMMENTS: Buta shows no distinction between pres. prog. and habitual
 (note no 'every day' in that datum either)
 Meaning of Benge "nagiwe"? (maybe 'here')
 Strategy for 'days': Pagabete 'suns' and Benge 'nights'
 All have "mo" for demonstrative (tones?) (see also last pg.)

- | | |
|----------------------------|------------------------------------|
| 12. Don't eat meat! | 13. Eat fish! |
| Ne mange pas de la viande! | Mange du poisson! |
| Olía mosuni tê! | Olía mbísi! |
| Pagabete Kòyé sè pámé kà! | Pagabete óyè ⁿ dé bàsí! |
| Benge Ká kòyé pámé! | Benge ná òsí! |
| Bwa-Buta Ká kòyé pámé! | Bwa-Buta óyá bàsí! |

COMMENTS: Pagabete prefers negative at end, others at beginning
 Meaning of Pagabete "se" in negative commands?
 Free variation in positive command? ("oye" vs "na", cf. #14)

- | | |
|------------------------------------|----------------------------------|
| 14. Eat this fish (here)! | 15. Don't eat that fish (there)! |
| Mange ce poisson-ci! | Ne mange pas ce poisson-là! |
| Olía mbísi óyo (áwa)! | Olía mbísi wáná tê! |
| Pagabete òyé òsí mò! OR ná òsí mò! | Pagabete Kòné sé òsí mòkó kà! |
| (pl. òyé bàsí bà!) | |
| Benge ná òsí mò! | Benge Ká kòné òsí mòná! |
| Bwa-Buta òyá bàsí bò! | Bwa-Buta Ká kòyé bàsí bò! |

COMMENTS: Unfortunately, the Lingala could be singular or plural.
 Similar treatment of 'this' but maybe not for 'that'

16. He says that the meat is spoiled. Don't eat that meat (already spoken of).
 Il dit que la viande est pourrie. Ne mange pas ce viande (dont on vient
 Yé alobí ete mosuni ebebí. Olía mosuni yangó tê. [de parler).
 Pagabete Yí àkpétí bétè pámé yò à^mbà^ugákà. Kòyé sè pámé yò kà.
 Benge Yí àpàkí bétè pámé àpòitò. Ká kòyé pámé yà.
 Bwa-Buta Yí àpàgí bétèⁿé pámé àpòitò. Kákòyé.

COMMENTS: Cf. names: Bwa "Napagibetini", Bati "Napagisene", & Apagibeti (on #16) Benge-Bati have "k" instead of "g", Pag. "-kpet-" for "-pag-" Anaphoric marker? Maybe Pagabete "yo", Benge "ya" (Bwa 'it'?) Second Bwa sentence lacks direct object

17. If the meat were spoiled, that would be bad (but it isn't).
 Si la viande était pourrie, cela serait mauvais (mais elle ne l'est pas).
 Sókó mosuni ebebákí, mbele óyo ezalákí mabé (kási ebebí tē).
 Pagabete èlí áyákà pàmé è^mbà^ŋgá, yíbà yó àlikí épè.
 Benge àlí pètè pàmé ápò, àbí àlí 'nè.
 Bwa-Buta àliyètè pàmé yóðð àpòitò, ábì àlí 'nè.

COMMENTS: Meaning of Bwa-Buta yóðð (demonstrative?)
 Tentative analysis of this and #10:
 Pagabete: èlí áyákà..., yíbà àlikí....
 Benge: ali (pete)...., àbí àli....
 Bwa-Buta: àli (bete?)...., ábì àli....
 "àlikí, àlí" may be copula (see last page)
 [high tone here would be explained by elision with adj.]

18. Someone was sleeping. A dog barked. That person (I mentioned) woke up.
 Quelqu'un était en train de dormir. Un chien a aboyé.
 Cette personne (déjà cité) s'est réveillée.
 Motu mókó azalákí kolála. Mbwa angángákí. Motu yangó alamúkákí.
 Pagabete ntó mótí àliàgiàkà kò àlálà.
 Ndé ^ŋvé àgbókà. Ndé ntó mô àkólè.
 Benge mótó mótí àlí àlálà. ^ŋve àgbóì. Mótó yà ázùsùà.
 Bwa-Buta mótó mótí àlí àlálà. ^ŋvé àgbóì. Mótí yà mô àzùz^wátù.

COMMENTS: All data were elicited in Lingala; thus mótí 'one'.
 "k" in àgbókà vs. no "k" in àgbóì is regular phonetic corresp.
 Part of first Pag. sentence is unexplained. (wà^ŋgé 'lie down'?)
 Demonstratives: cf. #11 ("mo" 'this'), #16 ("ya" anaphor).

19. If (when) a dog barks, the man will wake up.
 Si (quand) un chien aboie, l'homme se réveillera.
 Sókó mbwa angángí, mobáli akolamuka.
 Pagabete k^wété ^ŋvé àgbókà, nòkó mò àkólè ló.
 Benge ^ŋvé àgbóì, mòló ázùsùà.
 Bwa-Buta àlià bétèné ^ŋvé àgbòè, mòló ázùzùà.

COMMENTS: Three diff. treatments of 'If (when)': Pagabete "k^wete",
 Benge (empty), Bwa apparently 'if that' (cf. ##17, 16).
 Benge source: yes, "azusua" is both past and future.

- | | |
|--|-----------------|
| 20. That person (there) has not yet awakened. | 21. Wake up! |
| Cette personne (là) ne s'est pas encore réveillée. | Réveille-toi! |
| Motu wâná alamúkí naino tē. | Lamúká! |
| Pagabete ntó mòkó kà kúlílí kà. | Pagabete Kólé! |
| Benge Mótó mò ká zùsúli. | Benge Zùsúá! |
| Bwa-Buta Mótó mô ká zùzùá lílè. | Bwa-Buta Zùzùá! |

COMMENTS: Inconsistent transcription of Pag. 'wake up'; still double neg.
 Demonstrative: cf. #15 (Pag. mòkó, Benge mònà); but "mo"='this'
 All have "-li" in a 'not yet' situation.

22. What do you want me to do? Give me water.
 Que veux-tu que je fasse? Donne-moi de l'eau.
 Olingí násála níni? Pésá ngái mái.
 Pagabete òwé òbá bé nэгè èké? èpá èmè líbá.
 Benge obái bètè nòzé gònò? èpá mè líbà.
 (OR) obái bètè názè tè?
 Bwa-Buta òbá té názè tè? ègó líbà.

COMMENTS: Word division is uncertain: E.g., acc. to a group at Dingila, "bétònò" = 'what', but "bètè nò-zé" makes more sense here. Basic word order is '(thou) want that(?) I-do what'. Verb has -e/ε suffix in all three (contr. #23) [= subjunctive?]. Compare bé, bètè, té with subordinator (#16) bétè, bétènè. Word list data for 'love', 'to do', and 'what': báê, g^yé, èké (Pagabete); bâ, zâ, tê (Benge); bâ, zê(, èsínà) (Bwa) Bwa "égó" is unexplained: special form for 'give me'?

Give him water.	Give us water.	Give them water.
Donne-lui de l'eau.	Donne-nous de l'eau.	Donne-leur de l'eau.
Pésá yé mái.	Pésá bísó mái.	Pésá bangó mái.
Pagabete mpá yí líbá.	Tépá búsú líbá.	òpá bú líbá.
Benge èpá yí líbá.	èpá búso líbá.	òpá bú líbá.
Bwa-Buta è ^m pá í líbá.	égó búsù líbà.	ópá bú líbà.

COMMENTS: Pagabete: obj. prefix on verb? Bwa: lp obj. "égó" cf. above Bwa intonation on direct object 'water'?

23. What are you doing? I am eating (mpondu).
 Qu'est-ce que tu fais? Je mange (du mpondu).
 Ozalí kosála níni? Nazalí kolía (mpondú).
 Pagabete (Owe) óg^yàgá 'ké? (émé) pànágàkà ùgùⁿzà.
 Benge òzàlè tè? Nàpélé (gbàⁿdá).
 Bwa-Buta (wèá) òzá tè? àná yâ yâ.

COMMENTS: Compare ##3, 8. Some consistent diffs. Pag. vs. Benge vs. Bwa. Nevertheless, all have the word order (ProN) + V + "what". Benge present (here and ##3, 7, 8) may contain copula "ali"

24. What did I do (yesterday)? You woke up, and (then) you ate (mpondu).
 Qu'est-ce que j'ai fait (hier)?
 Tu t'es réveillé, et (ensuite) tu as mangé (du mpondu).
 Nasálákí níni (lóbí)? Olamúkákí, mpé olíákí (mpondu).
 Pagabete ((Yókó) émé) nègí 'ké? òkúlí (yókó), ndé òpá ùgùⁿzà.
 Benge Nàzí tè (íyò)? òzúsú òpá gbàⁿdà.
 Bwa-Buta (Yò) nàzí tè èè? òzúzú òpá gbàⁿdà.

COMMENTS: Sequential action (cf. #18): Pag. "ndé", others (empty?)

ADDITIONAL PARADIGMS -- PARADIGMES SUPPLEMENTAIRES -- MASOLO MOSUSU

you are eating; we are eating; you (pl.) are eating; they are eating
 tu mange; nous mangeons; vous mangez; ils mangent [en train de]
 ozalí kolía; tozalí kolía; bozalí kolía; bazalí kolía

Pagabete S òpágàkà 0; S tènágàkà 0; S màpágàkà 0; S òápágàkà 0

Benge òpálè 0; tápálè 0; màpálè 0; bápálè 0

Bwa-Buta óyà yà; átàyà yā; ámayā yā; (bú) bâbàyā yā

(THE REST OF THE PARADIGM IS FOUND IN RESPONSE 23 AND STATEMENT 3)

I ate (mpondu); we ate (mpondu); you ate (mpondu); they ate (mpondu)

je mangeai; nous mangeâmes; vous mangeâtes; ils mangèrent

naliákí (mpondu); toliákí (mpondu); boliákí (mpondu); baliákí (mpondu)

Pagabete nàpí yàkà 0; tèní yàkà 0; màpí yàkà 0; òápí yàkà 0

Benge nàpí tò 0; tápí tò 0; màpí tò 0; bápí tò 0

Bwa-Buta nàyáyí tò; tàyáyí tò; --; (bú) bàyáyí tò

(MORE OF THE PARADIGM IS FOUND IN STATEMENT 1; RESPONSE 24 NOT PARALLEL)

COMMENTS: Resp. 24 was supposed to be parallel.. but isn't! (sequential?)

2p pl. not gotten in Bwa (3p pl. instead, as often happens)

my father; thy father; his/her father; our f.; your f.; their f.
 mon père; ton père; son père; notre père; votre père; leur père
 tatá na ngái; tatá na yó; tatá na yé; t. na bísó; t. na bínó; t. na bangó

Pagabete èbámè; èbákò; èbàkè; èbàsù; èbànù; èbàðù

Benge àbómù; àbá ; àbé ; àbàsù; àbànù; àbàðù

Bwa-Buta àbámè; àbá à; àbé è; àbàsù; àbànù; àbàðù

COMMENTS: Differences in 2p and 3p sg. may be phonetic only (k vs no k)

this is a house

c'est une maison

óyo ezalí ndáko

Pagabete yé àlikí èbè^mbè

Benge yé àlé ðgbáálè

Bwa-Buta èyébè à ðgbélè

these are houses

ce sont des m'sons

óyo ezalí bandáko

yé àlikí màbè^mbè

yé àlé ðáðgbáálè

èyébè à ðáðgbálè

this is my child

c'est mon enfant

óyo ezalí mwána na ngái

Pagabete mú àlikí míkí mòðgámè

Benge mú àlé mànàmè (?)

Bwa-Buta môbé à mí mòðgámè

these are my children

ce sont mes enfants

óyo ezalí bána na ngái

ðà ðàlikí ðàmíkí ðàðgámè

ðà (ðálé) ðà mí ðòðgámè

bôbé bàbà mí ðòðgámè

COMMENTS: Again, the Pagabete intervocalic "k" with no "k" on other lists
 Pagabete and Benge seem to be more similar than either with Bwa
 Not intended to showcase demonstr., but cf. ##2, 11, 14-15, 18
 Apparent copula "ali"—cf. ##10, 17, 18.

Appendix H: Details of statistical analysis

The following matrices are based more or less on the elicitation list used by Mary Allen McMaster for her doctoral dissertation. Most of the data, however, are not hers, since SIL had word lists in all of the languages she studied except "Ngenda", and these word lists generally included all but seven of her glosses (namely, 'liver', 'new', 'left', 'right', 'straight', 'to swim', and 'to extinguish').

In the first matrix, similarity percentages appear below the diagonal, and a measure of range of sampling error (called VARIANCE by the WordSurv program), appears above the diagonal.

Lingala	6.6	6.6	7.6	7.9	7.6	6.6	6.5	7.5	7.5	6.5	7.4	6.4	7.7	5.5	5.4
66 Bangala	6.8	7.8	7.9	7.9	6.9	6.8	7.8	7.8	6.7	7.5	6.6	7.5	5.7	5.6	
59 48 Budza	7.2	7.4	7.4	6.4	6.4	7.4	7.4	6.4	7.2	6.5	7.5	5.5	5.2		
44 42 63 Ngombe	6.8	8.7	7.4	7.4	8.4	8.3	7.4	8.3	7.4	8.4	6.3	5.9			
46 40 63 Benza	9.1	7.7	7.7	8.7	8.7	7.6	8.7	7.6	8.7	6.8	6.1				
29 31 31 33 34 Ngenda	8.0	8.0	9.1	9.1	7.9	8.9	8.0	8.2	7.3	5.8					
40 48 43 44 47 49 Bati	3.3	5.6	7.0	6.3	7.3	6.3	7.1	6.0	4.7						
39 48 45 46 49 49 93 Bua	4.8	6.9	6.2	7.2	6.3	7.1	6.0	4.7							
41 49 46 47 48 53 83 88 Kango	7.7	7.1	8.3	7.3	8.1	6.7	5.7								
39 44 43 42 45 43 67 68 71 Ngelema	7.4	8.3	7.4	8.2	6.8	5.5									
37 38 42 46 44 41 61 63 65 55 Lika	7.1	6.4	7.1	6.1	5.0										
40 38 44 45 45 42 51 53 55 51 60 Bali	7.3	7.9	7.1	5.5											
37 38 47 51 52 42 61 60 60 54 59 52 Komo	7.3	6.3	4.8												
40 33 45 41 41 25 32 33 34 33 32 31 37 Mituku	7.2	6.5													
21 21 23 24 27 27 30 29 28 29 32 35 35 34 Budu	4.8														
30 31 28 27 28 20 21 21 24 22 24 23 22 36 21 Swahili															

The percentage figures are calculated by dividing the TALLY (shown below the diagonal in the next matrix) by the N(TOTAL) (shown above the diagonal). That is, the degree of similarity is amount of similar vocabulary as a proportion of all items compared. The total number of items counted for a list is shown to the right.

Lingala	85	92	90	85	90	91	92	90	89	91	92	93	86	90	93	total
56 Bangala	88	85	82	86	87	88	87	85	87	88	88	88	82	85	88	88
54 42 Budza	96	89	97	97	98	96	95	98	99	98	98	92	95	96	99	99
40 36 60 Ngombe	90	95	95	96	95	95	96	97	96	93	97	94	98			98
39 33 56 73 Benza	87	89	90	89	88	89	89	90	86	89	89	90				90
26 27 30 31 30 Ngenda	97	98	96	95	98	100	96	91	94	95	100					100
36 42 42 42 42 48 Bati	100	96	96	98	99	97	92	94	96	100						100
36 42 44 44 44 48 93 Bua	97	97	99	100	98	92	95	97	101							101
37 43 44 45 43 51 80 85 Kango	94	96	98	96	92	94	94	99								99
35 37 41 40 40 41 64 66 67 Ngelema	96	97	95	90	94	94	98									98
34 33 41 44 39 40 60 62 62 53 Lika	100	97	91	95	96	100										100
37 33 44 44 40 42 50 53 54 49 60 Bali	98	93	96	97	102											102
34 33 46 49 47 40 59 59 58 51 57 51 Komo	92	95	97	99												99
34 27 41 38 35 23 29 30 31 30 29 34 Mituku	92	90	94													94
19 18 22 23 24 25 28 28 26 27 30 34 33 31 Budu	94															97
28 27 27 25 25 19 20 20 23 21 23 22 21 32 20 Swahili	98															98

Since we had no other "Ngenda" data, the similarity figures shown between Ngenda and each of the other languages listed are the most reliable we can have. This set is "maximal" for Ngenda. The 49% and 53% figures (Ngenda compared with Bati, Bua and Kango) are significantly higher than the 29-34% figures (Ngenda compared with languages of Equateur). The 41-43% figures (Ngenda compared with Ngelema, Lika, Bali, and Komo) are significantly higher than the 20-27% figures (Ngenda compared with Mituku, Budu and Swahili).

Even for a set of only 100 data, the affinity of Budza, Ngombe, and Benza (potential "Budza Group"), and of Bati, Bua, and Kango ("Bua Group", without Ngelema), is already evident. Ngelema, Lika, Bali, and Komo display a moderate similarity to these last three.

(optional discussion of the items not counted)

Based on up to 103 data, with between one and fifteen missing for each list.

Missing data:

(Ngenda datum counted for only one of two related words)
 leg (Ngenda datum counted as 'foot')
 animal (Ngenda datum counted as 'meat')
 seed (Ngenda datum counted as 'fruit')
 fear (Ngenda datum counted as verb) <- somehow counted too
 (gap in McMaster's appendix)
 sand: Mituku
 new: Ngombe, Benza, Mituku, Budu
 left: Ngombe, Benza, Mituku, Budu
 right: Ngombe, Benza, Mituku, Budu
 straight: Ngombe, Benza, Mituku, Budu
 (compound word)
 feather: Bangala, Kango
 tree bark: Lingala, Bangala, Benza, Swahili
 tree root: Bangala
 sun: Benza
 to bite: Bangala
 foot: Komo
 left: Lingala, Budza, Kango, Komo
 right: Lingala, Budza, Ngelema, Komo, Swahili
 to swim: Ngombe, Bati, Lika, Budu
 to extinguish: Lingala, Bangala
 (doublet) --> though sometimes the other word isn't on the list!!
 seed (= fruit): Bangala
 fat (= oil): Lingala, Bangala, Swahili ['oil' not on list]
 ashes (=dust): Lingala [though 'dust' is not on the list]
 star (= fire): Lingala
 cloud (=smoke): Bangala
 earth (= sand): Budza
 many (= big): Ngombe, Benza, Ngelema, Lika, Budu
 red (= blood): Mituku
 sleep (gerund): Bangala [though the verb is not on the list]
 fear (gerund): Bangala, Budza, Lika, Bali
 foot (= leg): Lingala, Bangala, Benza, Swahili
 meat (=animal): Bangala, Benza, Kango, Mituku
 hand (= arm): Lingala, Bangala, Budza, Komo, Swahili;
 Benza, Bati, Bua, Lika
 (semantic shift)
 cloud (fog): Bati, Bua
 to fear (fear): Ngelema
 (loan word excluded)
 cloud: Ngelema
 left: Bangala
 right: Bangala
 (no data yet)
 horn: Ngelema
 fat: Budu
 year: Lingala
 louse: Bangala
 fruit (seed): Benza, Kango
 to fear: Ngelema
 liver: Mituku
 to swim: Mituku

The next set of matrices is based on the elicitation list used by Fultz and Morgan in 1985-86. As before, similarity percentages are found below the diagonal, and probable bounds of error above the diagonal. Note that the range of error is slightly less than for the 100-item list; this is a result of basing the calculations on a longer elicitation list (more glosses).

Lingala	5.4	5.3	6.7	6.7	5.9	6.1	5.1	5.1	5.1	5.1	5.1	5.1	5.1	5.7	4.9
70 Bangala	5.8	7.3	7.2	6.5	6.5	5.8	5.8	5.8	5.8	5.8	5.8	5.7	6.3	5.7	
54 45 Budza	4.2	6.6	6.1	5.8	5.3	5.3	5.3	5.3	5.3	5.3	5.2	5.9	5.2		
48 43 89 Babango	8.4	7.7	7.5	6.5	6.5	6.4	6.5	6.5	6.5	6.4	7.2	6.5			
48 41 57 53 Libale	6.6	7.0	6.6	6.6	6.5	6.5	6.4	6.6	6.5	7.3	6.5				
37 35 57 54 75 Ngombe	5.9	6.0	6.0	5.9	5.9	6.9	6.0	5.9	6.7	5.9					
44 38 65 59 70 77 Benza	6.0	6.0	6.1	6.0	6.0	6.1	6.0	6.7	6.0						
35 42 42 37 40 38 39 Pag-Mongl	1.1	3.6	4.2	4.2	4.0	5.2	6.1	5.4							
35 42 44 38 41 39 41 99 Pag-Omv	3.3	4.2	4.1	4.0	5.1	6.0	5.4								
33 42 40 34 36 36 40 87 89 Pag-east	3.8	3.7	3.6	5.3	6.1	5.4									
33 43 40 36 36 37 40 81 81 85 Benge	1.1	2.9	5.2	6.0	5.4										
34 44 41 36 36 38 41 81 82 86 99 Bati	2.5	5.2	6.0	5.4											
35 44 43 37 39 40 44 83 83 87 92 94 liBua	5.2	6.1	5.4												
32 36 38 35 38 36 38 64 65 61 62 63 64 Lika	6.1	5.4													
32 33 39 34 35 35 36 45 45 44 45 48 54 Bali	6.1														
30 35 38 37 39 37 39 51 51 49 53 54 54 50 47 Komo															

This is the maximal count for Babango, Libale, and (for verbs) Ngombe.

Lingala was the language of elicitation for Fultz & Morgan data.

(Old) Bangala was added by Douglas Boone for comparison (higher similarity to Bua bloc languages).

All others based on new data (1994-95).

Benge, Bati, Lika, Bali, and Komo were not studied by Fultz & Morgan.

In the next matrix, the raw data are presented as before: TALLY (numerator) below the diagonal, N(TOTAL) (denominator) below the diagonal, and total data counted for each list in the right margin.

Lingala	119	143	138	141	139	140	141	142	140	141	142	141	140	141	142	145
83 Bangala	119	115	116	115	116	118	119	118	118	119	119	118	119	117		120
77 54 Budza	141	141	141	141	144	144	142	143	144	143	144	144	144	144		147
66 49 125 Babango	137	136	138	139	139	137	138	139	138	139	139	139	139			141
68 47 80 72 Libale	138	138	139	140	138	139	140	139	138	139	140					143
52 40 80 73 104 Ngombe	137	139	140	138	139	140	139	138	139	139						143
61 44 92 82 97 105 Benza	140	140	138	139	140	139	138	138	140							142
49 49 60 51 56 53 55 Pag-Mongl	145	143	144	145	144	142	142	142								145
49 50 63 53 58 54 57 144 Pag-Omv	144	145	146	145	142	143	142									146
46 49 57 47 49 50 55 125 128 Pag-east	144	144	143	140	141	140										144
47 51 57 49 50 51 56 116 117 123 Benge	145	144	141	142	141											145
48 52 59 50 51 53 58 118 119 124 143 Bati	145	142	143	142												146
49 52 61 51 54 56 61 119 120 125 133 136 liBua	142	142	141													145
45 42 55 49 53 49 53 91 93 85 87 90 91 Lika	143	141														144
45 39 56 47 49 49 49 64 65 63 63 65 68 77 Bali	141															145
43 41 55 52 54 51 55 72 73 69 75 76 76 71 66 Komo																145

Nominally based on the 150 items on the Fultz & Morgan elicitation list, excluding 'to learn', 'to create', 'healer', 'spirit' and verb conjugations and using both of 'head/body hair'.

These data support the grouping of the lists as follows:

- Lingala and Bangala 70% similar, generally < 50% similar to any other list
- Budza, Babango, Libale, Ngombe, and Benza more than 50% similar, and about 40% similar to all the other lists (Babango is about 90% similar to Budza; Libale and Benza about 75% similar to Ngombe; these two clusters are about 55% similar.)
- Two Pagibete dialects are nearly identical, and so are Benge and Bati.
- Six lists join at 80% to form the "Bua bloc"; Lika is 60% similar to these.

The very same groupings are suggested by the RATIO of DEGREES of PHONETIC DIFFERENCE matrix also produced by WordSurv. This is a measure of phonetic (dis)similarity, calculated only for similar words, and represents the number of significantly different phones or tones out of 100 comparable phones or tones. The smaller the figure, the more similar the lists.

Lingala	LIN-BXG (9)													
9 Bangala - old														
14 22 Budza	BJA-BBM (10)													
17 21 10 Babango	five plus LIN (21)													
17 21 21 22 Libale														
14 21 16 17 9 Ngombe	NGC plus two (12)													
18 24 19 19 12 12 Benza	remarkable cleavage between													
28 25 29 31 26 25 26 Pag-Mong	first seven and Bua bloc													
28 25 29 31 26 25 27 3 Pag-Omv														
29 26 29 32 26 23 26 10 9 Pag-east	Bua bloc (14)													
26 23 26 30 23 20 21 13 13 9 Bengc														
27 23 26 30 23 21 22 13 14 9 1 Bati														
29 23 27 29 26 22 25 14 14 11 9 8 liBua														
28 21 26 27 26 24 27 20 20 23 20 21 21 Lika														
26 20 28 28 27 24 26 28 28 23 25 26 25 21 Bali														
24 20 26 28 26 24 23 23 23 23 23 23 26 23 22 Komo														

Some of the data in the database were collected using an elicitation list developed in 1990, consisting of 150 words. The next section shows the similarity figures for nineteen lists based on only those 150 glosses. First, PERCENTAGES below the diagonal, and VARIANCE above the diagonal.

Lingala	5.4	5.4	5.1	5.1	5.1	6.6	4.9	5.7	5.1	5.8	5.0	4.9	5.0	5.9	6.1	5.7	4.3	4.3
64 Bangala	5.6	5.6	5.6	5.6	7.0	5.4	6.1	5.5	6.0	5.2	5.2	5.2	6.1	6.4	5.8	4.4	4.9	
52 45 Budza	5.3	5.3	5.3	6.9	5.1	5.9	5.3	5.8	5.0	4.9	5.0	5.9	6.0	5.7	4.6	4.2		
33 41 44 Pagibete	3.9	3.8	6.6	4.9	6.0	5.3	5.4	4.7	4.5	4.8	5.3	5.5	5.7	4.9	3.8			
32 43 41 84 Bengc	2.3	6.6	5.1	6.0	5.3	5.5	4.8	4.5	4.8	5.4	5.6	5.6	4.6	3.8				
33 44 43 85 95 liBua	6.5	5.0	6.0	5.3	5.5	4.8	4.6	4.9	5.5	5.7	5.7	4.7	3.8					
34 36 43 63 62 65 Kango	6.3	7.7	6.9	6.7	6.0	5.8	6.1	6.8	7.2	6.9	5.8	5.0						
31 36 39 69 65 67 70 Lika	5.8	5.3	5.5	4.6	4.5	4.8	5.3	5.6	5.6	4.9	3.9							
33 34 40 55 51 54 56 60 Bali	6.1	6.1	5.3	5.1	5.4	5.8	6.2	6.3	5.6	4.4								
32 37 42 56 58 58 50 52 50 Komo	5.9	5.0	5.0	5.0	5.8	6.1	6.1	5.0	4.1									
35 33 35 28 29 29 24 30 28 37 Zimba	5.9	5.9	5.9	6.8	7.1	6.7	5.4	5.2										
33 31 34 27 28 29 26 27 28 34 54 Bembe	3.7	5.0	5.9	6.1	5.8	4.6	4.4											
31 30 32 23 24 25 23 24 24 32 52 85 Lega-Mw	4.7	6.0	6.2	5.8	4.6	4.5												
33 32 35 28 28 30 27 29 28 33 50 63 71 Lega-Sh	5.9	6.1	5.9	4.7	4.4													
37 33 40 26 27 28 25 26 24 35 51 52 51 57 Mituku	6.5	6.8	5.4	5.1														
37 39 38 27 28 29 28 28 26 38 51 52 51 56 69 Genya	7.1	5.7	5.4															
31 28 32 32 30 31 26 31 30 43 37 35 36 38 42 47 Lengola	6.1	4.9																
20 19 25 29 25 26 23 31 31 33 28 26 25 27 27 29 43 Budu	3.6																	
31 40 27 20 20 20 21 23 21 24 37 34 35 34 34 39 28 18 Swahili																		

Bangala = old Bangala; Pagibete = Omveda; Kango = Bomokandi; Budu = Ibambi

This is maximal for Kango, Mituku, and Zimba (data from 1990).

Bembe, Lega-Mwenga, Lega-Shabunda, Genya, Lengola and Budu data were collected using a longer list of which this one was a subset.

Data in all other languages were collected in 1994-5 with a much longer list.

Next, TALLY below the diagonal, N(TOTAL) above the diagonal:

```

Lingala 129 143 140 140 140 130 144 142 140 142 146 144 144 142 132 138 143 147
83 Bangala 129 127 127 128 119 131 129 125 130 131 130 130 127 121 126 129 132
74 58 Budza 143 143 143 131 149 146 142 145 149 147 147 145 138 142 146 146
46 52 63 Pagibete 147 147 134 145 146 142 144 147 146 145 143 135 141 143 143
45 55 59 124 Benga 147 134 145 145 142 142 146 145 144 142 134 141 143 143
46 56 62 125 139 liBua 134 146 145 142 143 146 145 144 142 134 141 143 143
44 43 56 85 83 87 Kango 133 133 131 130 134 133 132 130 124 129 133 133
44 47 58 100 94 98 93 Lika 148 144 146 150 148 148 146 138 143 148 147
47 44 59 80 74 79 74 89 Bali 142 145 149 148 147 145 136 141 146 145
45 46 60 79 82 83 65 75 71 Komo 141 145 143 144 142 133 141 143 143
49 43 51 41 41 41 31 44 40 52 Zimba 150 149 149 147 136 142 144 145
48 41 50 40 41 42 35 41 41 49 81 Bembe 152 152 150 140 145 148 149
44 39 47 34 35 36 30 36 36 46 78 129 Lega-Mw 151 148 138 143 146 147
47 41 51 40 40 43 36 43 41 48 75 96 107 Lega-Sh 149 139 144 146 147
52 42 58 37 39 40 33 38 35 49 75 78 76 85 Mituku 138 142 144 145
49 47 53 37 37 39 35 38 36 50 69 73 70 78 95 Genya 134 136 135
43 35 45 45 42 44 33 45 42 60 53 51 51 55 60 63 Lengola 141 141
28 25 37 41 36 37 31 46 45 47 41 39 37 39 39 40 60 Budu 146
46 53 39 28 28 28 28 34 30 35 53 51 51 50 49 53 40 26 Swahili

```

The total number of data considered for each list is shown below:

```

LIN BXG BJA PAG BGE BWV KTY LIK BCP KMW ZMB BMB LGM LEA ZMQ GEY LEJ BUU SWC
147 132 150 149 148 148 136 152 151 146 150 154 152 152 150 141 145 150 150

```

The RATIO of DEGREES of PHONETIC DIFFERENCE matrix for these lists follows.

```

Lingala          3 truly Equateurian          three notable
10 Bangala                               clusters ( < 20 DD)
14 19 Budza /
27 24 28 Pagibete      3 Bua bloc
26 21 26 12 Benga
28 22 27 11 8 liBua /
23 18 25 25 22 24 Kango
25 22 26 21 21 23 18 Lika
30 23 31 27 24 24 27 21 Bali
24 21 26 21 21 24 22 22 21 Komo
19 20 25 28 28 30 21 23 27 25 Zimba /      BMB & LGM
17 23 27 32 31 32 29 33 33 26 23 Bembe
21 22 30 35 32 32 30 35 33 30 25 13 Lega-Mw /
22 23 32 30 30 31 26 30 28 29 22 22 24 Lega-Sh
21 22 24 30 26 28 21 22 29 26 19 25 26 22 Mituku
27 31 32 32 31 33 31 35 33 28 29 26 28 26 23 Genya
23 27 27 29 28 31 30 29 30 22 29 29 30 27 27 30 Lengola
28 31 31 36 34 36 35 25 30 30 30 33 36 34 30 36 29 Budu
20 13 26 28 28 29 24 24 26 25 20 27 21 22 22 28 29 30 Swahili

```

Throughout the 1980's, a different standard elicitation list was used, consisting of approximately 200 words. Once again, 19 lists are compared.

PERCENTAGE and VARIANCE:

Lingala	4.6	4.6	4.4	4.4	4.4	4.2	4.9	5.0	4.5	4.3	4.3	4.3	4.4	5.1	4.9	3.6	3.7	3.7
68 Bangala	4.9	4.9	5.0	5.0	4.7	5.2	5.4	4.8	4.8	4.5	4.5	4.6	5.4	5.0	3.7	3.7	4.2	
49 43 Budza	4.6	4.5	4.6	4.4	5.0	5.1	4.6	4.5	4.3	4.2	4.3	5.0	4.7	3.8	3.9	3.5		
33 39 42 Pagibete	3.6	3.6	4.5	5.2	5.4	4.8	4.7	4.0	3.8	4.0	4.7	4.6	3.9	3.9	3.2			
32 43 40 81 Benga	2.0	4.6	5.2	5.3	4.7	4.7	4.0	3.9	4.1	4.6	4.6	3.9	3.9	3.2				
34 43 42 81 95 liBua	4.5	5.3	5.4	4.8	4.7	4.0	3.9	4.2	4.7	4.7	4.0	4.0	3.2					
29 36 35 63 59 62 Lika	5.1	5.4	4.7	4.6	4.0	4.0	4.1	4.6	4.7	4.2	4.3	3.3						
32 32 38 51 47 50 57 Bali	5.5	5.3	5.3	4.5	4.4	4.6	5.3	5.4	4.8	4.8	3.7							
31 33 34 47 43 45 51 71 Beeke	5.4	5.5	4.4	4.4	4.5	5.4	5.5	5.2	5.2	3.5								
33 33 40 49 49 50 48 56 54 diKango	4.3	4.2	4.3	4.3	5.1	5.0	4.4	4.4	3.6									
31 34 39 51 52 53 47 50 52 70 Komo	4.3	4.3	4.3	5.2	5.2	4.3	4.3	3.4										
30 28 32 24 25 25 25 25 21 28 30 Bembe	3.7	4.5	5.3	4.9	3.8	3.9	3.8											
31 28 30 21 22 22 25 24 21 29 30 79 Lega-Mw	4.3	5.4	5.0	3.9	3.9	3.8												
32 30 32 25 26 29 28 27 22 30 31 57 68 Lega-Sh	5.4	5.1	3.9	4.0	3.8													
34 33 33 25 24 26 24 26 25 31 36 45 47 53 Genya	6.1	4.7	4.7	4.5														
30 27 28 28 26 27 28 29 29 32 40 32 34 37 44 Lengola	5.2	5.2	4.1															
18 17 21 24 22 23 30 30 37 30 30 22 23 22 25 39 Budu-Ib	2.3	2.9																
19 17 22 25 22 23 31 31 37 31 31 23 23 24 26 40 93 Budu-Wa	2.9																	
30 38 24 17 19 19 21 20 16 24 22 31 31 32 33 26 15 15 Swahili																		

At present, this is maximal for Bembe, Lega-Mwenga, Lega-Shabunda, Genya, Lengola, and Budu, although it does not exhaust the database for most of these languages since additional data have been elicited or found in the literature. All other data were elicited in 1994-95 using longer lists.

Note that diKango, which is spoken by a pygmy group living among the Bali, is different from the speech of the Bakango (fishers) far to the north.

Based on what was the standard eastern Zaïre elicitation list in the 1980's.

TALLY below the diagonal, N(TOTAL) above the diagonal:

Lingala	168	192	187	188	187	191	190	177	179	187	191	190	189	179	188	188	188	196
114 Bangala	167	163	164	164	168	167	160	157	161	166	167	165	158	166	166	167	170	
95 71 Budza	190	192	190	196	195	182	184	190	193	192	192	185	191	191	190	194		
61 63 80 Pagibete	194	194	190	192	179	182	186	190	189	189	181	188	186	187	189			
61 70 77 158 Benga	194	191	193	181	183	188	190	190	189	181	190	187	187	190				
63 71 79 158 184 liBua	191	191	180	181	186	189	188	188	180	188	186	186	189					
56 60 69 119 112 118 Lika	195	184	184	190	197	195	194	185	193	194	193	195						
60 54 74 97 91 96 112 Bali	186	187	189	194	195	193	184	191	192	193	193							
55 52 61 84 78 81 94 132 Beeke	177	177	181	181	180	171	181	182	183	180								
59 52 73 90 90 91 88 105 96 diKango	186	185	185	184	174	184	182	182	182									
58 55 74 94 98 98 89 94 92 131 Komo	190	189	188	178	190	187	186	190										
58 46 61 46 47 48 49 49 38 51 57 Bembe	199	199	184	193	192	193	195											
58 47 58 40 41 42 48 47 38 54 56 157 Lega-Mw	197	182	192	191	192	194												
61 50 62 48 50 54 54 53 40 56 59 114 134 Lega-Sh	183	191	190	191	192													
60 52 61 46 44 46 44 47 43 54 64 82 85 97 Genya	181	181	181	182														
56 44 53 52 49 51 55 55 53 59 76 62 66 71 79 Lengola	189	188	191															
33 29 40 45 42 42 59 58 67 55 57 42 43 42 46 73 Budu-Ib	196	191																
35 28 41 46 42 42 59 60 67 56 57 44 45 46 47 76 183 Budu-Wa	191																	
59 64 46 33 36 36 40 38 28 43 41 60 60 62 60 50 29 29 Swahili																		

Totals:

LIN BXG BJA PAG BGE BWW LIK BCP BKF diK KMW BMB LGM LEA GEY LEJ B-I B-W SWC
197 171 200 196 197 194 191 201 187 189 195 202 201 200 190 197 197 198 200

The RATIO of DEGREES of PHONETIC DIFFERENCES matrix:

Lingala	LIN and BXG																								
9	Bangala	/																							
13	17	Budza	/																						
26	22	26	Pagibete	Bua bloc																					
25	20	24	12	Benge																					
26	20	25	12	8	liBua	/																			
25	21	25	21	21	22	Lika																			
27	21	28	27	25	24	21	Bali	(BCP and BKF)																	
29	25	30	28	28	29	22	17	Beeke																	
24	20	26	23	24	24	22	24	29	diKango	(DiK and KMW)															
24	21	25	21	22	24	22	22	22	12	Komo															
18	22	27	31	31	32	34	31	35	29	28	Bembe	(BMB and LGM)													
21	20	27	32	30	30	34	31	34	29	30	14	Lega-Mw													
22	22	30	29	29	29	31	26	32	28	30	22	23	Lega-Sh												
25	30	30	32	31	32	34	31	32	31	29	25	27	26	Genya											
23	25	26	28	29	30	30	29	28	24	23	29	29	26	28	Lengola	/									
26	29	30	34	32	33	24	28	24	34	30	32	36	34	35	28	Budu-Ib	BUU								
33	31	36	38	37	36	27	28	24	37	33	37	40	38	39	30	8	Budu-Wa	/							
20	14	24	27	27	28	24	24	29	21	25	26	20	22	28	28	29	31	Swahili							

In 1993, a new standard elicitation list was developed, with 85% overlap with the previous 200-item list, and including about 70% of the glosses from the list used by Fultz and Morgan. Once again, we compare 19 speech varieties.

PERCENTAGES below the diagonal, VARIANCE above the diagonal:

Lingala	4.4	4.5	5.3	5.0	4.3	4.3	4.3	5.0	5.0	4.1	4.7	4.9	4.3	4.3	4.3	3.5	3.6	3.6							
69	Bangala	4.7	5.6	5.3	4.7	4.8	4.8	5.5	5.5	4.5	5.0	5.2	4.6	4.6	4.5	3.6	3.6	4.0							
50	43	Budza	5.3	5.1	4.4	4.3	4.3	5.0	5.1	4.2	4.8	4.9	4.5	4.4	4.2	3.7	3.7	3.4							
39	39	59	Ngombe	5.1	5.3	5.2	5.3	6.0	6.1	5.1	5.9	5.9	5.5	5.4	5.1	4.3	4.3	3.9							
39	37	56	78	Benza	5.0	4.9	5.0	5.6	5.8	4.8	5.5	5.5	5.1	5.1	4.9	4.2	4.3	3.6							
33	41	39	39	38	Pagibete	3.4	3.4	4.5	5.1	4.4	5.0	5.2	4.6	4.5	3.9	3.9	2.9								
33	44	36	36	37	82	Benge	2.3	4.4	5.0	4.4	5.0	5.3	4.6	4.5	4.0	3.8	3.8	3.0							
34	45	37	39	40	82	93	liBua	4.1	4.9	4.4	5.1	5.2	4.6	4.5	4.1	3.8	3.8	3.0							
36	47	38	38	36	73	75	80	Kango	5.7	5.0	5.8	6.0	5.2	5.2	4.6	4.3	4.3	3.7							
34	41	38	39	40	60	63	66	65	Ngelema	5.2	5.9	6.0	5.3	5.3	4.6	4.4	4.4	3.6							
29	35	32	34	32	61	58	61	61	51	Lika	5.0	5.2	4.6	4.5	3.9	4.0	4.1	3.1							
31	31	35	37	35	47	44	48	50	47	55	Bali	5.3	5.1	5.1	4.5	4.7	4.7	3.5							
30	33	32	31	30	46	43	45	48	43	49	71	Beeke	5.3	5.3	4.2	5.0	5.0	3.3							
32	34	39	41	40	49	49	51	53	49	46	57	55	diKango	4.1	4.2	4.2	4.2	3.4							
32	35	38	40	39	50	52	54	55	49	49	52	53	72	Komo	4.2	4.2	4.2	3.3							
33	32	32	33	32	24	26	28	27	25	25	26	20	28	29	Lega-Sh	3.8	3.9	3.8							
18	17	21	19	21	24	23	22	22	23	28	31	36	30	29	22	Budu-Ib	2.3	2.9							
20	17	22	20	22	25	23	23	22	23	29	32	36	31	30	24	93	Budu-Wa	2.9							
29	35	23	21	19	16	17	17	20	18	19	19	15	22	21	33	15	15	Swahili							

Bangala = old Bangala; Pagibete = Omveda; Kango = Uélé; diKango = pygmy lang.

Lists using the 1993 Standard elicitation list (200 items)

Lingala, Bangala, Swahili are trade languages, not elicited.

1994 Ngombe list did not include verbs, shortfall made up from Fultz & Morgan

Komo, Lega-Shabunda, and Budu (x2) lists predate 1993.

TALLY below the diagonal, N(TOTAL) above the diagonal:

Lingala	181	201	176	198	200	199	199	195	189	201	200	188	191	197	194	198	199	205
124 Bangala	181	161	178	180	179	180	176	169	182	181	174	173	176	175	179	181	184	
100 77 Budza	180	201	204	204	203	200	194	207	206	194	197	201	199	202	202	204		
69 62 107 Ngombe	176	180	178	178	176	171	180	180	169	171	176	178	178	179	179			
78 65 113 138 Benza	202	201	201	197	193	203	200	189	191	196	193	199	199	201				
66 73 80 71 76 Pagibete	208	208	204	197	205	206	194	197	200	198	200	202	203					
65 79 74 64 74 171 Benga	207	203	195	205	205	194	196	200	197	200	201	202						
68 81 76 70 81 171 193 liBua	202	195	205	204	194	195	199	196	199	200	202							
70 82 75 67 71 148 152 162 Kango	191	201	202	189	193	196	195	198	199	199								
64 69 74 67 77 119 123 128 125 Ngelema	195	194	182	186	189	188	190	191	192									
59 63 67 62 65 126 119 126 123 99 Lika	206	195	197	202	198	204	204	204										
62 56 73 66 69 96 90 98 100 92 113 Bali	196	198	200	199	202	204	203											
57 57 62 53 56 89 83 88 91 78 95 139 Beeke	189	189	187	192	194	191												
62 58 76 70 76 96 97 99 102 92 91 112 104 diKango	198	191	194	195	194													
63 61 76 70 76 100 104 107 107 93 98 104 101 143 Komo	196	199	199	200														
64 56 63 58 62 48 51 54 53 47 49 51 38 53 57 Lega-Sh	196	197	197															
36 30 42 34 42 48 45 44 44 43 58 62 69 58 58 43 Budu-Ib	205	201																
39 30 44 36 43 50 46 45 43 43 60 66 69 60 59 48 191 Budu-Wa	202																	
59 64 46 37 39 32 35 35 39 34 39 39 29 42 41 65 30 30 Swahili																		

Totals:

LIN BXG BJA NGC Bnz PAG BGE BWW KTY AGH LIK BCP Bke diK KMW LEA B-I B-W SWC
205 184 209 184 205 210 209 208 206 198 211 210 197 200 205 203 206 208 208

This set is maximal for (Ngombe,) Benza, Uélé Kango, Ngelema, and Beeke.
(Ngombe) since it is at once incomplete (no verbs in 1994) and supplemented by
1985-6 data from Fultz and Morgan.

Thanks to McMaster's list, there are a few other data for all but Beeke.

The standard list has 200 items, but I often elicited some or all of:

'hand' (with 'arm')	'bad' (with 'good')
'wing' (with 'feather')	'today' (with 'yesterday' & 'tomorrow')
'oil' (with 'fat')	'to fear' (with the noun 'fear')
'firewood' (with 'fire')	'a laugh' (with the verb)
'fog' (with 'cloud')	'a smell' (with 'to smell')
'lake' (with 'river')	'rapidly' (with 'to run')

and I usually could isolate the infinitive, causative, and reciprocal affixes and the final
vowel of the infinitive, for as many as 216 data.

RATIO of DEGREES of PHONETIC DIFFERENCE matrix:

Lingala BXG is like LIN
9 Bangala / these five moderately similar
13 18 Budza
14 20 18 Ngombe
19 22 20 13 Benza /
27 23 28 25 26 Pagibete
25 20 26 21 22 13 Benga group including
27 21 26 23 25 13 8 liBua the Bua bloc
20 16 21 20 22 19 16 16 Kango
25 20 27 23 24 23 18 16 17 Ngelema
27 21 26 25 28 20 21 22 19 25 Lika /?
27 21 28 27 27 28 25 25 24 40 20 Bali BKF is like BCP
29 25 30 27 28 29 28 29 28 48 20 17 Beeke /
24 22 26 27 26 23 24 25 19 24 21 24 28 diKango diK is like KMW
24 22 26 25 25 21 22 24 18 37 22 23 22 12 Komo /
20 21 30 26 27 30 29 29 25 27 31 26 31 28 30 Lega-Sh / BUU=BUU
26 30 31 31 28 34 32 33 30 31 23 27 23 33 30 34 Budu-Ib
33 32 36 34 30 37 36 35 33 34 27 28 24 36 34 37 8 Budu-Wa /
21 15 26 28 29 30 30 31 24 27 25 24 29 24 26 22 29 31 Swahili

Swahili dissimilar to all but Bangala (moderately like Lingala, Lega-Shabunda)

Finally, using the entire database, we are necessarily limited to data elicited during or since the Bua bloc survey (plus the three trade languages). This time, the similarity figures were calculated to three decimal places, and the decimal point was accordingly removed from the VARIANCE figures.

PERMILLAGES below the diagonal, VARIANCE above the diagonal for entire database (300 records)—three decimal places:

Lingala	40	39	37	37	36	36	36	37	37	37	35	40	37	31
676 Bangala	43	42	42	42	42	42	43	43	42	43	41	44	40	36
481 403 Budza		38	38	38	37	38	39	38	38	38	36	41	37	28
328 378 381PAG-Mongw	11	24	30	30	30	30	30	30	30	39	44	39	25	
313 367 378 978 PAG-Omv	23	30	30	31	31	30	38	44	40	24				
291 367 341 891 899 PAG-east	28	28	28	28	27	39	44	39	24					
299 392 346 808 809 848 Benga	8	23	22	22	39	43	39	25						
304 398 354 810 812 854 989 Bati	24	21	21	39	44	39	25							
308 391 369 815+ 805+ 853 899 903 Kiba	15	13	39	45	40	25								
301 387 357 813 804 842 906 917 959 Yew	8	39	44	40	25									
315+ 401 360 822 816 858 910 924 966 986 liBua	38	44	40	25										
272 318 306 572 591 558 551 563 592 569 581 Lika	44	39	26											
275- 280 325- 432 445- 435+ 420 428 454 448 456 522 Bali	44	29												
297 303 335- 458 466 444 472 475+ 482 473 487 460 459 Komo	28													
277 344 201 147 141 134 151 152 153 148 152 172 161 191 Swahili														

The + and - is noted after "permillages" ending in 5 so that one knows which way to round to two significant figures (i.e. percentages).

As before, these are calculated using TALLY (below diag.) and N(TOTAL) (above):

Lingala	219	264	262	259	258	261	260	250	259	260	261	262	259	267
148 Bangala	216	217	215	215	217	216	215	217	217	217	217	218	211	221
127 87 Budza	265	262	261	266	263	255	263	264	268	268	263	264		
86 82 101 PAG-Mongw	276	275	276	274	265	273	275	271	273	262	265			
81 79 99 270 PAG-Omv	276	277	276	267	275	277	269	272	262	263				
75 79 89 245 248 PAG-east	277	274	265	273	275	267	271	261	262					
78 85 92 223 224 235 Benga	277	268	277	279	272	274	265	265						
79 86 93 222 224 234 274 Bati	267	276	276	268	271	263	264							
77 84 94 216 215 226 241 241 Kiba	268	268	260	262	255	255								
78 84 94 222 221 230 251 253 257 Yew	277	269	270	262	263									
82 87 95 226 226 236 254 255 259 273 liBua	272	272	263	264										
71 69 82 155 159 149 150 151 154 153 158 Lika	276	265	267											
72 61 87 118 121 118 115 116 119 121 124 144 Bali	266	267												
77 64 88 120 122 116 125 125 123 124 128 122 122 Komo	262													
74 76 53 39 37 35 40 40 39 39 40 46 43 50 Swahili														

Totals:

LIN	BXG	BJA	P-1	P-2	P-3	BGE	Bati	Kiba	Yew	Bamb	LIK	BCP	KMW	SWC
271	224	276	281	279	278	282	278	270	278	280	284	284	274	276

- (5) The more data one collects, the greater the likelihood of having data which are comparable to those collected by another researcher in another speech variety using a different elicitation list.
-

The bottom line on lexical similarity, if one discounts all percentages less than 50%, is something like this:

Lingala	Ngenda	Bwa bloc (see permillages for internal facts)					
68% Bangala	(basically alone)	75%	Uélé	Kango			
		65%	65%	Ngelɛma			
Budza		60%	??	??	Bomokandi Kango		
90% Babango		57%	60%	50%	65-70%	Lika	
55% 55% Libale						52%	Bali
55% 55% 75% Ngombe						70%	Beeke
65%? 55% 75% 75% Benza					50%??	55%	55% diKango
					50%??	55%	55% 70% Komo

even Bali, supposedly a "Lega-Kalanga" language, has little in common with:

Bembe	
80%	Lega-Mwenga
57%	70% Lega-Shabunda
	53% Genya

There would seem to be a "maNgala Group", "a Budza/Ngombe Group" (which may actually be two groups), a "Bua Group", a "Komo Group" (including diKango as well as Bhele, Kaiku, Bila, Amba, and Bera), a "Budu Group" (not shown), and a "Bembe-Lega Group". Lika and Bali are not firmly attached to any groups.