# The numeral system of Proto-Niger-Congo

## A step-by-step reconstruction

Konstantin Pozdniakov

language science press

Series information: ./langsci/seriesinfo/nc-info.tex not found!

# The numeral system of Proto-Niger-Congo

## A step-by-step reconstruction

Konstantin Pozdniakov

Konstantin Pozdniakov. 2018. *The numeral system of Proto-Niger-Congo*: *A step-by-step reconstruction* (Niger-Congo Comparative Studies 2). Berlin: Language Science Press.

Ирине Поздняковой

# Contents

Contents

Contents

# Preface

## 1 Introduction

### 1.1 Niger-Congo: the state of research and the prospects for reconstruction

It is quite predictable that the title of this book may be met with skepticism by specialists in the comparative-historical studies of African languages. The first question that may arise is whether a Niger-Congo (NC) reconstruction is achievable at all, considered that the reconstruction of proto-languages underlying particular families and their branches has not been completed (or even properly started, as is the case for some groups and branches of NC). Before we turn to the structure of the book, let us try to answer this fundamental question. To do so, it seems reasonable to very briefly outline the present state of affairs in NC comparative studies.

First, it should be noted that presently there is no general scientific discipline such as "NC comparative studies". Instead, there are individual researchers who work on particular families, groups, sub-groups or branches of NC. Among these, comparative-historical Bantu studies has flourished the most. However, the Bantu languages comprise only a branch of the Southern Bantoid languages that (together with Northern Bantoid) go back to Proto-Bantoid. Hence Bantu is merely one of 16-17 Bantoid branches, as can be gleaned from the chart below.

The progress of comparative-historical studies of the Bantoid languages has been less impressive than that of Bantu studies. Proto-Bantoid, as well as a number of other proto-languages, goes back to the Proto-Eastern-Benue-Congo. In turn, the latter (along with Proto-Western-Benue-Congo and possibly some other languages that do not belong to these two major groups of Benue-Congo) goes back to Proto-Benue-Congo (BC). Hence, the Bantoid branch is merely one of 14-15 branches of Benue-Congo, as demonstrated by the chart below (Table 2).

The traditional reconstruction of Proto-BC based on regular correspondences between the proto-languages underlying the separate branches listed in table 0.2 has developed rapidly in recent years. However (and I hope that my colleagues will take no offence at this statement), despite numerous brilliant studies dealing

Table 1: Bantoid languages
This book does not investigate the genealogical classification of Niger-Congo as a whole, nor of the individual families of this macro-family. The schemes presented here take into account the most well-known classifications (sometimes with small deviations due to the specific purposes of our study). The scheme of Bantoid languages given here is based mainly on the classification in https://mpi-lingweb.shh.mpg.de/numeral/Niger-Congo-Benue-Congo.htm . It generally reproduces the John Watters' classification (1989: 401) with some deviations, which are not considered here.

| Northern Bantoid: | Dakoid | Mambiloid | Fam | Tiba (Fà) |
|---|---|---|---|---|
| Southern Bantoid: | **Bantu** | Beboid | Yemne-Kimbi | Ekoid |
| | Jarawan | Mamfe | Mbam | Mbe |
| | Ndemli | Tikar | Tivoid | Wide Grassfields |

Table 2: Benue-Congo languages
Inventory of Benue-Congo groups is given mainly by Williamson 1989a: 266-269. The main difference in Table 0.2 is that Jukunoid is separated from Platoid, which allows us to better compare the forms of numerals of these groups, as well as the fact that Lufu has been added to isolated languages. The division of the BC into the Western and Eastern branches does not always reflect the genealogical characteristics of languages.

| *Western BC | *Eastern BC |
|---|---|
| Nupoid | Kainji |
| Defoid | Platoid |
| Edoid | Cross |
| Igboid | Jukunoid |
| Idomoid | **Bantoid** |
| Isolated BC: Oko, Akpes, Ikaan, Lufu ||

with the subject, this is still a relatively 'young' science.

Finally, in addition to Proto-BC there are probably more than ten proto-languages underlying other language families that together comprise the Niger-Congo macro-family (see Table 3).

Table 3: Niger-Congo languages
The grouping of 12 families of NC into 5 geographical zones is convenient for technical purposes of generalization of data. So, it means nothing else. As for a genealogical tree of NC languages, as of today there are insufficient grounds for creating one, in my opinion.

| Atlantic | Mande | Dogon | Ubangi | Kordofan |
|----------|-------|-------|--------|----------|
| Atlantic | Mande | Gur | Ubangi | Adamawa |
| Mel | Kru | Kwa | Ijo | BC |

Most of the works presently available in NC comparative studies do not reach beyond this point. Exceptions are rare, and examples of the comparative-historical approach to the NC reconstruction are few. Moreover, the most significant works of this kind (e.g. those of Westermann 1927, Greenberg 1966, Sebeok 1971, etc.) are not that recent and usually date to the middle of the 20th century. Comparative studies of the African macro-families had a jump start but nearly had come to little by the end of the 20th century (important works such as Bendor-Samuel 1989 including Williamson 1988; 1989b are few in this period).

So, what happened?

By the 1990s, our knowledge in the field of African languages had begun to grow exponentially. Hundreds of new language descriptions had been published, and the few dozen experts working in NC comparative linguistics were simply unable to digest this avalanche of new information.

The main problem in the 1960s was that we knew too little. From the 1980s on, we have faced the opposite problem: we know "too much". Not only do scholars not have enough time to absorb new results, sometimes they do not even have enough time to acquaint themselves with those results. During the last four decades, amidst this dialogue between linguistic knowledge and language data, African linguists have remained in listening mode. But I am convinced that the time has come for linguists to say something new again. Unlike even ten years ago, today we are well equipped to do so.

Firstly, we have really exceptional databases. The best one is the RefLex database elaborated by Guillaume Segerer (Segerer & Flavier). It contains more than one million words from African languages (2017), and each entry contains a link to a

PDF file of the corresponding source page. It provides a huge range of information and is maximally user-friendly to comparative linguists: it can be solicited for establishing regular phonetic correspondences, for reconstruction and for ranking reflexes as well as for various kinds of statistical data analysis. This new database is being constantly updated.

A big database is something much more than just a huge amount of data. When a database reaches certain degree of plenitude with respect to the main families and branches of the NC macro-family, it opens up prospects for both working with the distribution of words that do exist and with the distribution of **gaps** in postulated cognates. The distribution of filled cells and lacunes is a powerful tool allowing 1) identification of important innovations, 2) targeted searches for unusual phonetic reflexes, 3) detection of diachronic semantic changes and 4) refinement of genealogical classification.

In my opinion, the opportunity to rely on both the apparent cognates as well as on the missing reflexes of reconstructed prototypes in particular languages dramatically changes the approach to the reconstruction itself.

The following case may serve as an illustration to this statement. Suppose we need to assess one of Greenberg's proposals, e.g. a Niger-Congo root meaning 'hill'. Among the reflexes quoted by Greenberg for this root are: "(2) Busa *kpi* 'mountain', Kweni *kpi*; (4) Gã *kpɔ*; Gwa *ogba* 'mountain'; (5) Nungu *agbɔ*, Ninzam (Ninzo) *igbu*. Kordofanian: (2) Tagoi *(c)ibe*." (Greenberg 1966: 155). The phonetic correspondences underlying the comparison of these forms will not be discussed here (we will just assume that they are valid), for the main problem is elsewhere. A reader with no access to a representative lexical database on the NC languages is always uncertain about a number of key issues, including:

1. whether the root in question is widely attested in the families and groups for which the author postulates the reflexes?

2. whether the root is present in other NC families and groups and how widely it is attested in them?

3. are there any other roots possibly interpretable as NC terms for 'hill'?

The RefLex database establishes that:

1. there are plenty of forms phonetically similar to those of Greenberg (cf. e.g. Boko (in the same sub-group as Busa) *kpii* 'mountain', Gwari (Nupoid, BC) *ōpé* 'hill, mountain', etc), but the postulated root is at best only marginally attested in the families where Greenberg finds it.

2. The root is absent in other branches and families (even if the proposed phonetic correspondences are approached most liberally), although, if wished, its "reflexes" can be found in any of the NC families, cf. e.g. Ibani (Ijo) *kpókpó* 'hill', etc.

3. Most importantly, several other roots with the meaning 'hill, mountain' are distinguishable in the NC languages. All of them (unlike the one proposed by Greenberg) are valid candidates for the reconstruction of the NC prototype. One of these roots is presented in the chart below (0.4) (one could mention some other roots nearby):

Table 4: *ʈʊnd* 'hill, mountain' in Niger-Congo

| | | tɔ́rɔ́ | | |
|---|---|---|---|---|
| *tʊnd | *tinti, *ton | | | |
| tul- ? | tōḍō | tu? | tɔ́ndɔ́ | tʊ̀ndà |

The exact correspondence between Proto-Bantu (*tʊ̀nda`, zones HJKPMNRS > ( ?) *dɔ́ndʊ̀, zones EGHJKLMNRS), Ijo (Ibani *tʊ́ndʊ́*) and Atlantic languages (Atlantic Bak: Manjak *ntʊnda*, Atlantic North: Basari *e-tɔ́nd*, Bapen *ɛ-tʌnd*, Laala *tundə*, Fula *tulde*, Wolof *tund*) is reason enough to postulate the root *tʊnd* 'hill, mountain' at the Proto-NC level, especially since these languages have apparently been out of direct contact[1]. In addition, the absence of this root in Gur-Ubangi-Adamawa may prove to be a shared innovation in these languages.

Using the databases, the focus of our research could be redirected toward the basic meaning of the lexemes (rather than on the occasional phonetic similarities between the forms). This approach may help in answering the following question: if a Proto-NC term for 'mountain, hill' existed, how did it sound? The answer would probably be as follows: this word could sound like *tʊnd*, *kong/ keng* or *kudu* ('hill, rock, stone'), but not like *dima* (PB *dìmà*, zone EGJ), *mut* (Proto-Jukunoid *muT*) or *pi* (PB *pìdì*, zone KLMN).

Upon arriving at these unconventional "results", one could bring them to the attention of specialists in particular NC languages and branches for further eval-

---

[1]We shall repeat that nearby there are some other candidates for 'mountain' in NC, which we do not treat here.

uation. Without such professional evaluation there can be no hope for success. Moreover, in recent years it has become evident that this evaluation needs to be collaborative (i.e. made by dozens of specialists working together) for the simple reason that today no specialist can be proficient in the languages of more than one or a maximum of two NC families. Hence, it is important that these specialists are asked questions they can answer, so ideally the approach outlined above should be applied to every family within Niger-Congo. For example, according to the etymological database of the Atlantic languages (Pozdniakov-Segerer 3700 cognates, 2017) only *tʊnd and *thəng are potentially interpretable as the terms for 'hill, mountain' in Proto-Atlantic.

Initially I thought of numerals as of an ideal group of terms to test this approach. On the one hand, the core group of numerals must have existed in Niger-Congo. On the other hand, they represent a relatively compact lexical-semantic group with minimum potential for semantic shifts. My initial question seemed simple: what is the most probable Proto-Niger-Congo root for 'two'? The term for 'two' (being the only numeral on the Swadesh list) is generally recognized as one of the most persistent numerals. Why not try reconstructing it on the basis of the NC evidence? It appeared, however, that such a reconstruction is beset with difficulties, so what was originally intended as an article turned into this very book. The structure of the book is described in the section below. As I hope to demonstrate, this structure is conditioned by specific issues encountered in the course of the reconstruction of NC numerals.

## Sources and the monograph structure

### Sources

Numeral terms included in the majority of lexical sources hold a privileged position. The information pertaining to the Niger-Congo numerals is more than extensive, it is nearly exhaustive. In addition to the above-mentioned RefLex database by Segerer-Flavier which contains over 17,000 entries marked as "numeral" (state April 2017)) a number of other databases with expansive coverage of the Niger-Congo languages are available. One of them is the "Numeral Systems of the World's Languages" database created by Eugene S. L. Chan and edited by Bernard Comrie (Chan) The data regarding the number systems of about 4,300 languages (with hundreds of the Niger-Congo languages among them) is incorporated into it. Two or even three sources (often unique) are accessible for some of the languages via this neatly organized and user-friendly database. Another

universal database that provides numerical data is "Numerals 1 to 10 in over 5000 languages" by Rosenfelder. It was consulted to a somewhat lesser extent because it only includes evidence pertaining to the first ten numerals, for which a simplified transcription is used. Finally, a number of unpublished databases that incorporate the evidence of specific Niger-Congo families and groups were consulted, e.g. the etymological databases of Atlantic (Pozdniakov-Segerer) and Mande (Valentin Vydrin).

As a result, a total of 2,200 sources for Niger-Congo languages were used in this study. This raises the issue of references, since it is impossible to provide a complete list of sources for every NC language. The language index at the end of this book lists the nearly 1,000 languages cited. For these 1,000 languages, the main sources I used are indicated in Appendix 5. The index of sources in Appendix 5 is structured according to the NC main families in alphabetical order.

For each language, I provide not only the source(s) that can be found in the bibliography, but also the name of every contributor in Chan's database [Chan]. The list of contributors is many pages long, but their names should be known, even if their data are unpublished. This is the least I can do to express my sincere gratitude to each of them.

## Monograph structure

Noun class affixes are present in numerical terms in the majority of the Niger-Congo languages. At the same time, many forms that are considered primary at the synchronic level have frozen noun class affixes that are no longer productive. In such cases it is extremely difficult to distinguish the etymological root within a numerical term. Without it, however, both the comparison and reconstruction of roots is impossible. This is why the first chapter of this book is devoted to the study of various uses of noun class markers in numeral terms.

The second chapter deals with the alignment by analogy in numeral systems. As in other languages, numerals represent a lexical-semantic group that is especially subject to alignment by analogy due to its closed structure, where words are associated in a paradigm. A textbook example is the term for 'nine', with Indo-European *$\mathbf{n}$- irregularly reflected in Proto-Balto-Slavic as $\mathbf{d}$- (Russian *dev'at'* '9' instead of the expected *\*nev'at'*) by analogy with the term for 'ten' (Russian *des'at'* '10'). This yielded a minimum pair *dev'at'* ~ *des'at'* that forms a "class of the upper numerals" within the first ten. Adjacent numerals may be alined with each other in the NC languages by a similar formal marker. Thus, no satisfactory etymology can be suggested for the forms attested in Mumuye (Adamawa; *ziti* '2' ~ *ta:ti* '3' ~ *dɛ̃:tì* '4') without the analysis of alignment by analogy. The issues

pertaining to both detection and analysis of such alignments are addressed in Chapter 2.

Chapter 3 offers a step-by-step reconstruction of number systems of the proto-languages underlying each of the twelve major NC families, on the basis of the step-by-step-reconstruction of numerals within each family. The term "reconstruction" related to numerals throughout this book calls for a definition. As mentioned above, the use of this term has been questioned, mainly because systems of regular phonetic correspondences between the languages within NC families remain unknown. This is why Kay Williamson opted for the term *pseudo-reconstructions* (marked with # instead of *): "Reconstructions proposed by their authors as based on regular sound correspondences are preceded by an asterisk. Pseudo-reconstructions based on a quick inspection of a cognate set without working out sound correspondences are proceded by a #" (Williamson 1989a: 251). In his numerous online publications Roger Blench uses # as well, but his terminology is different: he prefers the more neutral term of *quasi-reconstructions*. Modern comparative studies of the NC languages is a relatively young science, so the opposition between "real" and "pseudo-/quasi-" reconstructions seems irrelevant to me at this stage. The more so that nearly all of our reconstructions (maybe with the exception of Bantu and some other branches) should be marked with #, including the large proportion of reconstructions allegedly based on the evidence of historical phonetics. On the other hand, I think that many colleagues would agree with the following statement: although we do not know the regular phonetic correspondences between the languages that belong to different NC families, there is hardly any doubt that the NC root for 'three' sounded something like *tat*.

Throughout this book the term "step-by-step reconstruction of number systems" (e.g in the Atlantic family) is used in reference to the method that includes the following steps:

1. While comparing the forms of numerical terms attested in the languages under study, their most likely prototypes were established within both of the Atlantic groups, i.e. Northern (Proto-Tenda, Proto-Jaad-Biafada, Proto-Fula-Sereer, Proto-Wolof, Proto-Cangin, Proto-Nalu-Baga Fore-Baga Mboteni) and Bak (Proto-Joola-Bayot, Proto-Manjak-Mankanya-Pepel, Proto-Balant, Proto-Bijogo).

2. On the basis of these prototypes, the most likely forms of Proto-Northern Atlantic and Proto-Bak Atlantic numerals were suggested.

3. On the basis of these more ancient forms, the most plausible reconstruction of Proto-Atlantic numerals was offered.

Chapter 4 deals with the reconstruction of the Proto-Niger-Congo numeral system on the basis of the step-by-step-reconstructions offered in Chapter 3 for each of the twelve major families and a handful of isolates. The reconstruction described in Chapter 4 inspired the analysis of the distribution of reflexes of the NC proto-forms within each of the twelve families (as well as within the isolates) in order to establish:

1) the most archaic NC families / groups / branches (i.e. those that preserve the inventory of Proto-NC forms most fully);

2) NC families / groups / branches that are the most distant from Proto-Niger-Congo in what pertains to the reflection of numerals.

The results of this analysis are presented in Chapter 5.

To illustrate the logic of the complex structure of the monograph, let us consider one example.

In Chapter 3, along with other NC families, the numerals of the Atlantic languages are analyzed (section 3.12). Atlantic languages are divided into two main groups – North Atlantic (section 3.12.1) and Bak Atlantic (section 3.12.2).

In Sections 3.12.1.1.–3.12.1.7, systems of numerals are considered consecutively in the seven main subgroups of the North Atlantic languages. In particular, in §??, numerals in the Jaad-Biafada subgroup are considered and it is established that in these languages, for the numeral '10', the form *-po* is attested. In the final section of 3.12.1, namely in §?? the forms of numerals in the seven northern subgroups are compared, and in particular it is concluded that for Proto-Northern Atlantic, the most probable reconstruction for the numeral '10' is the reconstruction of *\*pok.*

In Sections 3.12.2.1-3.12.2.4, the numeral systems in each of the four subgroups of the second Atlantic group, namely Bak, are discussed consecutively. The final section concerning the Bak group (3.12.2.5) concludes that the only candidate for reconstructing '10' in the Proto-Bak (in addition to the possible model 10 = 5 * 2) is the root *\*-taaj.*

In the final paragraph of section 3.12, namely in 3.12.3, the systems of the North Atlantic languages and the Bak Atlantic languages are compared. This paragraph concludes that the comparative evidence points to the total absence of common roots present in both groups. The only exception to this is the root *\*tɔk / \*tVk* 'five'. Accordingly, it is concluded that it is impossible to reconstruct the Proto-Atlantic root for the numeral '10' without the Niger-Congo context.

In Chapter 4, reconstructions for each family are compared. Accordingly, Chapter 4 has a different structure. If in Chapter 3 each of the sections is devoted to a particular family of languages (in particular, §?? is devoted to the Atlantic languages), then in Chapter 4 each section is devoted to the prospects for the reconstruction of each Niger-Congo numeral. So, in §?? all intermediate reconstructions for the numeral '10' are considered. It turns out, in particular, that the form *-taaj* reconstructed for '10' in the Proto-Bak does not find parallels in other Niger-Congo branches. In contrast, the root *pok '10', reconstructed for the North Atlantic languages, can be related to the roots reconstructed for the vast majority of Niger-Congo families (it seems to be missing only in Ijo, Dogon and Kordofanian). Based on the NC comparison, the root for '10' is reconstructed as *pu / *fu.

Chapter 5 traces the history of the numerals of Niger-Congo, reconstructed in Chapter 4, in each individual family of languages. Accordingly, each section, as in Chapter 3, is devoted to one of the NC families. So, §?? is devoted to the Atlantic languages. In particular, it is concluded that in the North Atlantic languages the term for '10' has been preserved in three sub-groups (Wolof *fukk, Proto-Tenda *pɔxw, Proto-Jaad-Biafada *po). In the other subgroups it is replaced with isolated innovations. The forms of the Bak languages are also innovated.

So, the basic logic of the chosen structure of the book is as follows: we will consistently move from reconstructions in individual families (Chapter 3) to the reconstruction of each Niger-Congo numeral (Chapter 4) and to the interpretation of each individual family in the Niger-Congo context (Chapter 5). We will take into account the provisions formulated in the preliminary chapters concerning noun classes in numerals (Chapter 1) and changes by analogy in systems of numerals (Chapter 2).

# Acknowledgments

Today the greatest benefit to being a researcher is the opportunity to directly contact leading specialists in the comparative studies of African languages. Even the best database does not ensure the proper interpretation of the results achieved by other scholars. In the course of my work on this monograph I have benefited from the help of many colleagues, whose comments and suggestions I greatly appreciate. My particular thanks go to Guillaume Segerer (Atlantic languages and RefLex database), Valentin Vydrin (Mande languages), Raymond Boyd (Adamawa languages), Larry Hyman (Bantu languages and Benue-Congo in general), Mark Van de Velde (Bantu languages), Marie-Paule Ferry (Tenda languages), Pascal Boyeldieu (Bua languages and Laal), Marion Cheucle (Bantu A.80), Denis Creissels (Balant), Sylvie Voisin-Nouguier (Buy), Ekaterina Golovko (Baga Fore), Odette Ambouroue (Orungu) and many others. It is a great pleasure for me to thank you all!

My special gratitude is addressed to colleagues who read the first version of the manuscript of the book and made a number of valuable critical remarks. These are members of the editorial board of this series of Niger-Congo Comparative Studies: Valentin Vydrin, Larry Hyman, John Watters and Guillaume Segerer. I tried as much as possible to take their remarks into account. Naturally, all responsibility for inevitable mistakes and shortcomings lies with me.

# Abbreviations

**Language groups and proto-languages**:
  BC – Benue-Congo
  GD – Ga-Dangme
  GTM – Ghana–Togo Mountain
  Juk. – Jukunoid
  NC – Niger-Congo languages
  PB – Proto-Bantu
  PLC – Proto-Lower Cross
  PP – Proto-Platoid
  PTB – Proto-Potou-Tano-Bantu
  PUC – Proto-Upper Cross
  SE – South-Eastern Mande
  SWM – South-Western Mande
  **Others**:
  CL – noun class
  CL.SG. – noun class of singular
  CL.PL. –noun class of plural
  CM – noun class marker
  dial. – dialect
  PL. – plural
  redupl. – reduplicated
  SG. – singular.

# 1 Noun classes in the Niger-Congo numeral systems

In most NC languages, the numeral stems are combined with noun class markers. More often we are dealing with the dependent markers of noun classes (in particular, in the numeral '1', as well as in the numerals '2'-'5') in those languages where there is an agreement between numerals and nouns. But class markers appear in many languages, even without any agreement. For example, when counting, numerals are often used in a nominal function and include obligatory markers of noun classes. In this case, numerals as nouns and, on the other hand, numerals as proper numerals can have different class markers (and different roots). Thus, in Likile (Bantu C) *li-yɔɔ* 'ten' (Cl5), *mo-túkú / mi-* 'dozen' (CL3 / CL4) (**Carrington1977**).

In many languages, nominal classes in numerals are easily recognized. In other languages, as a result of phonetic processes at the junction of CM and numeral stem and/or as a result of changes by analogy in the paradigm of numerals, it might be difficult to determine which noun class is included in the numeral, although we can distinguish a lexical root. Thus, in Lulamoji (Bantu J) in some derived numerals (*mm-kágá* '60' < *mu-káagá* '6' and *mm-sáánvu* '70' < *mu-sáánvu* '7'), an obscure CM **mm-** is observed (Larry Hyman, p.c.). It is not homorganic, so we can not treat it as CL10. Meanwhile, in the majority of other languages within this group, it is clearly CL10 which is observed in these forms: cf. for example, in Gwere *ṇka: ga* '60', *ṇsanvú* '70', cf. *lù-kúmì* '1000' / *ṇkúmì, βìβírì* '2000' (clearly CL11 / CL10)[1]. Such cases are not sufficiently dramatic for reconstruction.

However, in a number of languages in synchrony we do not have sufficient criteria to decide whether we are dealing with the root of a numeral or with combinations of a root with an archaic noun class marker. In other words, we cannot isolate the root, and therefore we cannot compare it with the roots of other languages. E.g. we posess no formal proof that the Kobiana (Atlantic) term

---

[1]The irregular allomorph of CL.10 may have arisen as a result of a change by analogy with the basic numeral '6' and '7': N homorganic (CL.10) in these derived forms > **mm-** by analogy with **mu-** (CL.3).

*sana* 'four' is composed of **sa**- being a class prefix adduced to the lexical base (-**na**). This base is only distinguishable by means of external comparison, although this method alone is admittedly insufficient, since the Kobiana term may as well be interpreted as an innovation (*sana* '4').

In more complicated cases, it should be assumed that a noun class affix replaced one of the segments of the stem, thus becoming an integral part. The Wolof (Atlantic) numerals provide a good example of this phenomenon. The following numerical terms are attested in Wolof at the synchronic level: *ñaar* '2', *ñett* '3', *ñeent* '4'. Normally the noun class affixes are not included in the lexical base in Wolof, so synchronically we do not have to interpret the first consonant of Wolof numerals as a prefix. At the same time, there are a number of important arguments in favor of the presence of the frozen prefix *Ñ- in the Wolof numerals. First, these are the only numerals that agree in the Ñ class, being one of the two plural noun classes preserved in Wolof (cf. *fukk* 'ten' which agrees in the singular noun class B). Secondly, the forms *yaar* '2' and *yett* '3' (with the initial consonant being identical to the other plural noun class - **Y**) which agree in the **Y** class have been preserved in some Wolof dialects. Finally, as we hope to demonstrate below, the unification of numerals by class in Niger-Congo languages is characteristic of terms covering the sequence from 'two' to 'four'. Thus, in the diachronic perspective, the consonants in question should be viewed as characteristic of class markers rather than stem segments. However, if this assumption is correct, we are forced to conclude that these markers have been integrated into the stem, having replaced the original initial consonants of the terms in question, the more so that VC-roots are uncommon in Wolof (numerical roots most probably had CVC structure, see **PozdniakovRobert2015**: 615-616). This means that the Wolof terms are of little significance for the reconstruction of the terms for '2-4' in Proto-Atlantic.

Most of the issues (theoretical ones included) that have complicated our reconstruction while studying noun classes in the families and branches of Niger-Congo pertain to the relationship of noun classes and numerals at the synchronic level. These problems are often left aside in the grammatical descriptions and do not attract sufficient attention from linguists. I am not aware of any work which discusses them systematically. Meanwhile, I am sure this question is worthy of attentive study because it reveals additional characteristics of noun class systems.

The first five numerals in Niger-Congo usually agree with nouns, for example in Sereer: *o-koor o-leng* 'one man', *a-koy a-leng* 'one monkey', *Ø-naak Ø-leng* 'one cow'. In some languages and branches of the macro-family, the inventory of numerals that show agreement is reduced.

As noted, the noun class marker may appear in numerals in some contexts which are not related to the agreement.

1. For instance, for counting, the majority of languages include a class marker (CM); moreover, different numerals may have different affixes. For example, in Biafada for the numerals '1', '6-7' the class **N** is used, for '2-4' the class **bi**-, **gə** – for '5', **Ø** – for '8-9', **ba** – for '10'.

A lot of languages use CM in numerals starting from '6' and higher, that is in the numerals that do not show agreement in class, and not only in counting. For example, in Manjak *ngə-bʊs ngə̀-təb* 'two dogs' (agreement), *ngə-bʊs ò-ntaja* 'ten dogs' (lack of agreement, numeral '10' with CL **ò-** is used in an independent form).

The choice of the noun class for numerals in the two aforementioned contexts (in counting forms, and in numerals with no agreement) represent a very interesting case which I will outline hereinafter.

2. The interaction between noun classes and numerals cannot be limited to the aforementioned contexts. Noun classes emerge as well in derived numerals. The three main cases will be highlighted as follows.

Firstly, in the majority of Niger-Congo languages (and, apparently, even in Proto-Niger-Congo) the numeral '8' was formed from '4' by the reduplication of the first syllable of the original root *CL-*na(h)i* '4' > *CL-*na-na(h)i* '8'. Often the noun class marker of '4' and '8' coincides, but sometimes they do not. A question therefore arises: which factors define the choice of a noun class in a derived numeral?

Secondly, the Niger-Congo languages use compound numerals extensively, as do the majority of languages in the world. For example, the numeral '40' is formed following the model '40' = '4*10' (in many Bantu languages, for instance) or '40' = '20*2' (in the majority of Atlantic languages). The latter model is based on finger-counting, when two hands and two feet give a sum of 20. The numeral '20' goes back to the lexeme 'chief' or 'man'. In these languages the numeral '15' is often formed following the model 'two hands and one foot'. This model is well known and is discussed in the literature. However, the question of the choice of noun class in the first and second formative of these compound numerals was often left aside. Meanwhile, this question needs more clarification. The following questions will be discussed in the present study.

In a compound numeral, for example, '20' = '10*2', the class marker is often absent in the second formative. For example, in Bomwali (Bantu, A80) we have: *Ø-kamɔ* '10' (CL9)[2] , *ɓe-ɓa* '2' (CL2), *mɔ-kamɔ Ø-ɓa* '20'. In this type of language,

---

[2]For a reader who is not aware of the tradition of Bantu linguistics, it is necessary to explain that

we have additional causes to discuss derivative words rather than syntagms.

In a compound numeral, both formatives include class markers, for example, '20' = 'CL-10*CL-2'. The CM can be different or the same in the two formatives: Pinji (B30) **n**-*dzìmà* **dí**-*bàlè* '20' (10*2), Nsong (B80) **ma**-*kwǐm* **m**-*ɔːl* '20' (10*2). In the latter case, a particular type of *agreement* can be observed, that is, the second formative agrees in class with the first formative.

If in a compound numeral both formatives include class markers, as in '20' = 'CL-10*CL-2' then theoretically we can expect that the noun class of the first formative will coincide with the class of the independent numeral '10'. This strategy is very rare. One of the unique examples comes from Moghamo (Grassfields) **ì**-*γùm-bē* '20' (*ì-γùm* '10', *í-bē* '2'). In the majority of cases the noun classes of the two formatives do not coincide. For instance, in the same branch of Benue-Congo (Grassfields): Laimbue **mì̀**-*γím-bò̀* '20' (*í-γím* '10', *bò̀* '2'), the number '10' changes its class, being part of the first formative of the numeral '20'. The interpretation of this strategy in Niger-Congo languages will be given later. The same problem arises with the second formative. Very rarely does its class coincide with the noun class of the initial numeral (in the present case we deal with the numeral '2'). In the majority of cases it differs. The cause is, as it was already mentioned above, that the second formative agrees with the first one. For example, in the same group of languages (Grassfields): Mundani **è**-*γɛm* **ye**-*be* '20' (*è-γɛm* '10', **be**-*be* '2'). In some languages, noun classes of simple and compound forms differ even if agreement is absent.

3. Finally, the strategy of forming numerals only by the change of the noun class and with no changes in the lexical root represents a real parade of paradigmatic values of noun classes in numerals. This strategy was systematically developed in one zone of Bantu languages, that is zone J (although it can be encountered sporadically in some other Niger-Congo languages). For example, in Chiga (Bantu J): **ì**-*βìɾí* '2' > **à**ː-*βìɾí* '20' ; **mù**-*kâːgà* '6' > **ŋ**̩-*kâːgà* '60', **mù**-*nâːnà* '8' > **kì**-*nâːnà* '80'.

It is interesting that the same language combines all three strategies. Thus, in Chiga:

1. The numeral '8' is formed by reduplication of '4': *í-nà* '4' > *mù-nâ̂ː-nà* '8'

---

in Bantu languages there is a stable inventory of noun classes, each having a fixed number. The ongoing numeration of Bantu was found useful for the study of noun classes in Niger-Congo in general, where the numeration of classes of non-Bantu languages represents a concrete etymological hypothesis. If a scholar assigns the number '6' to the class -**ɗam** of Fula (Atlantic language), it means that etymologically it should be related to the class **\*ma** (CL 6N) of Proto-Bantu.

(and we can observe the variation of noun classes 5 (**í**-) and 3 (**mù**-);

2. The numeral '200' is formed by a word-combination, but not by the combination of '100' and '2' as we would expect. Instead, it is formed by the combination of '10' and '2': *βì-kùmì βì-βírì* '200' (*ì-kúmì* '10', *ì-βìrí* '2'). Thus, '200' (CL.PL) is a plural form of '10' and '2' (CL.SG). Furthermore, the second formative agrees in noun class with the first.

3. The numeral '20' is formed from '10' by changing the noun class exclusively: *à:-βìrí* '20' (*ì-βìrí* '2'), and by the use of a different noun class, different from the one we find in '200', that is CL.PL **à:**-.

## 1.1 Noun classes in the counting forms of numerals

In some Niger-Congo languages, numerals do not have noun class markers in the counting form, but the number of these languages is very low. In the Atlantic family the only language with this feature is Balant. In the majority of Niger-Congo languages while naming a numeral (for example, in counting) noun class markers are used. These markers may be the same for all numerals, but this is a rare case. More often, for the numerals 1-10 there are three to four different markers (furthermore, special class markers may be used for the numerals '20', '100', '200' and others).

A fragment of the Tetela (C80) numeral system is presented below (Table 1.1.):

Table 1.1: Tetela numerals

| 1 | ó-tɔy | 9 | di-vwá |
|---|---|---|---|
| 2 | ha-énde | 10 | dí-kumi |
| 3 | ha-sátu | 20 | á-kumi á-ende |
| 4 | a-néy | 90 | á-kumi di-vwá |
| 5 | a-tánu | 100 | lo-kámá |
| 6 | a-samále | 200 | n-kámá y-éndé |
| 7 | e-sambɛ́ɛ́lé | 1000 | ki-nùnu (yínŋa) |
| 8 | e-náánéyi | 2000 | ø-nunu p-énde |

We see here a variety of classes as well as plenty of mini-clusters (note the noun class switch that occurs when a number becomes a part of a compound term; this phenomenon is characteristic of the Niger-Congo languages). The terms

for 'one' (**ó**- class), 'hundred' (**lo**-) and 'thousand' (**ki**-) appear to be isolated on account of their noun class. At the same time, the following groups of terms are distinguishable: '2-3' (**ha**-), '4-6/20' (**a**-, «/» refers to the grouping of non-adjacent numerals), '7-8' (**e**-), and '9-10' (**di**-). It should be noted, however, that even in such systems some numerals can be used without noun class markers ('2000').

Three issues need to be mentioned here.

The noun class markers are easily distinguishable in Tetela. However, for the majority of the NC languages (especially the non-Bantu ones) this is not the case. The criteria that would allow for distinguishing between the markers and the segments of stems are often lacking, which means that we have no idea which stem in a language under study is to be used for comparative purposes. The situation is even more grave in those numerous cases where an additional class marker is added to a numeral which contains an archaic class marker integrated in a stem.

The mechanism underlying the grouping of numerals into the mini-clusters (by including them in a common noun class) remains virtually unexplored, although it is certainly worthy of investigation and thorough consideration from the theoretical point of view. What was the motivation behind the use of the class marker **ha**- with the Tetela terms for 'two' and 'three', while in case of 'nine' and 'ten' the class marker **di**- was preferred in this language? The answer to this question is probably not to be sought within the semantics of a given noun class. On closer examination, the choice of a noun class in such distributions is often unmotivated by anything other than the need to formally distinguish a group of numerals (as opposed to other groups). In this respect, this mechanism is very similar to the alignment by analogy as applied to numerals in many languages. This strategy (implying an irregular alteration of a part of a lexical stem) can be compared to a radical surgery, which is never an easy option. Languages with noun classes have less traumatic means to achieve the same result, e.g. by using different noun class markers to distinguish between the groups of numerals. This elaborate marking technique is widely attested in the Niger-Congo languages. The grouping of numerals is typologically interesting as well: some of the groups are fairly common whereas some are quite rare. Moreover, it is probable that these groups were formed independently in different languages: a situation where a pair of closely related languages exhibit radically different grouping and vice versa is not uncommon.

Some numerals are not normally subject to grouping and tend to be marked with a specific noun class, thus standing in opposition to the rest of the numerical terms. The use of this specific class is especially frequent with the terms for 'one',

'hundred' and 'thousand', cf. e.g. specific noun classes observable in the Tetela terms for 'one' (**ó-** ) and 'hundred' (**lo-**).

Let's look at the distribution of numerals in noun classes for the languages where this information is available. This observation will be made on a selection of 254 Benue-Congo languages (among these, 166 are Bantu languages, evenly distributed by zones). Our sampling comprises languages that are known to employ noun classes on the numerical terms used in counting.

### 1.1.1 The specific marking of numerals

As mentioned above, specific noun classes are used with the terms for 'one' and 'ten' especially often: 174 languages out of 254 mark the numeral '1' in a distinguished way, and 151 languages mark the numeral '10' separately.

Examples of systems with the term for 'one' being in opposition to the rest of the numerals (marked with a different noun class)[3] are provided below (Table 1.2).

Table 1.2: Specific noun classes in '1'

| Branch | Language | '1' | '2' | '3' | '4' | '5' | '10' |
|--------|----------|-----|-----|-----|-----|-----|------|
| J30 | Nyole | **ndala** | ebiri | edatu | ené | etaanu | ehúmi njereere |
| Defoid | Ede Ica | **ɔkɔ̃** | eɟi | ɛta | ɛ̃ɛ̃ | ɛwu | ɛya |
| Defoid | Ede (dial.) | **ɔ̀kɛ̃** | mɛ́ɗʒì | mɛ́ta | mɛ́hɛ̃ | mɛ́hú | mɛ́wá |
| Defoid | Ifè | **ènɛ / ɔ̀kɔ̃̀** | méèdzì | méɛta | méɛrɛ̃ | méɛrú | maá |
| Mbe | Mbe | **ómè** | bɛ́pʷâl | bɛ́sá | bɛ́ñî | bɛ́tʃân | bɛ́fwɔ̂r |
| Mbam | Nomaande | **ɔmɔtɛ́** | béfendí | batátɔ́ | bényíse | batáánɔ́ | bɔ́ɔ́háta |
| Mbam | Tuotomb | **ɔ́mɔ̀** | péfáⁿd | pédààt | pínìs | pétàn | pʷówàt |
| Mbam | Tuki | **umwê:sií** | mówá | mótátó | mwé:né | motá:nó | mwábɔ́tɔ́ |
| Mbam | Yambeta | **ímùʔ** | mɔ́bààn | mɔ́dáád | múnìʔ | mɔ́táàn | mɔ́wád |
| Mbam | Nubaca | **pòmóhò** | mʷăntʃì | mùtát | mùɲíhì | mùtâ:n | mʷapʷat |
| Mbam | Yangben | **pùmòm** | mándɛ̀ | matát | ménì | mátàn | mát |
| Mbam | Numaala | **bùmʷòm** | mâ:ndɛ̀ | mádádɔ̀ | ménî | mátʰán | mátʰ |
| Mamfe | Denya | **gɛ́mâ** | ópéá | ólɛ́ | ónì | ótà | ófíà |

Examples of one other strategy (the term for 'ten' being a noun remains in opposition to the rest of the numerals by means of a noun class) are given in Table 1.3.

Another strategy with the terms covering the sequence from 'two' to 'nine' being opposed to the terms for 'one' and 'ten' is characteristic of the languages

---

[3]Considering the fact that numerals '2-9' belong to the same noun class, the numerals '6-9' are not included in Tables 1.2-1.5.

Table 1.3: Specific noun classes in '10'

| Branch | Language | '1' | '2' | '3' | '4' | '5' | 10 |
|---|---|---|---|---|---|---|---|
| S30 | Kgalagadi | (bʊ):ŋwɪ | (bʊ)bɪrí | (bʊ)rá:rʊ | (bʊ):nɛ | (bʊ)tʰá:nʊ | **lɪʃʊ́:mɪ** |
| S10 | Kalanga | (ku)ŋómpèlá | (kù)bìlí | (kù)tàtú | (kù)nnà | (kù)ʃánù | **gùmí** |
| Cross-River | Bete-Bendi | ìkèn | ìfè | ìkíé | ìnè | ìdíɔŋ | **lèhʷó** |
| Mbam | Nugunu | gímmue | gáandɛ | gádadɔ | génni | gátáanɔ | **sɛ́ɔdɔ** |
| Idomoid | Eloyi | ńgwònzé | ńgwòpó | ńgwòlá | ńgwòndó | ńgwolɔ́ | **úwó** |
| Jukunoid | Akum | ájì | afã̀ | ata | aɲì | acóŋ | **īkùr(ù)** |
| Platoid | Tyap (Kataf) | anyuŋ | afeaŋ | atat | anaai | afwuon | **swak** |

represented in Table 1.4.

Table 1.4: Common noun classes for '2'-'9'

| Branch | Language | '1' | '2' | '3' | '4' | '5' | '10' |
|---|---|---|---|---|---|---|---|
| Cross-River | Ebughu | **sìŋ** | ìbà | ìtɛ́ | ìnìàŋ | ìtîŋ | **lùgò** |
| Cross-River | Oro | **ki** | íbà | íté | ínîaŋ | ítiŋ | **lughu** |
| Cross-River | Usakade | **tʃɛ̀n** | m̀bà | ǹtá | ǹnìɔŋ | ǹtʃôn | **nùòp** |
| Cross-River | Leggbo | **wɔ̀ni** | àfɔŋ | àttan | ànnaŋ | àzen | **dzɔ** |
| Platoid | Ayu | **ɪdɪ** | afah | ataar | anaɲaʃ | atugen | **iʃog** |
| Grassfields | Mundani | **yea-mɔʔ** | bebe | betat | bekpì | betã̀ã̀ | **èɣɛm** |
| Igboid | Ekpeye | **ŋìnɛ́** | ɓɨ̀bɔ́ | ɓɨ́tɔ́ | ɓɨ́nɔ̀ | ɓíse᷄ | **ɗì** |
| Tivoid | Ipulo | **émɔ̀** | víàl | vétàt | véɲì | vétàn | **épɔ́:t** |
| Isimbi | Isimbi | **kēnə̄** | mə̄rākpə̄ | mākə̄lə̄ | mōnī | mātə̄nə̀ | **būɣù** |
| A40 | Bankon | **(i)yǎ** | (bi)ɓá | (bi)íyâ | (bi)nân | (bi)tán | **iɓŏm** |
| A80 | Bekwil | **wát / ŋgɔ́t** | e-ɓá | e-lêl | e-nâ | e-tên | **kǎm** |
| A80 | Koonzime | **gwár** | bìbá | bìlêl | bìnâ | bìtên | **kám** |
| B20 | Kélé | **nwúntù** | bàbá | bàlál(è) | bànáyì | bàtán | **dyúm(ù)** |
| B20 | Ntumbede | **íwótó** | bɔ́bà | bɔ́rárè | bɔ́náyè | bɔ́tánè | **dʒómè** |
| J20 | Jita | **kamʷi** | βiβiri | βisatu | βina | βitanu | **ɛkumi** |
| K20 | Mbunda | cimo | vivali | vitatu | viwana | vitanu | **likumi** |
| M20 | Ndali | **kamukene** | fi-ŵiri | fi-tatʊ | fi-na | fi-hano | **kaloⁿgo** |
| N30 | Nyanja | **cimɔ́dzi** | (zi)βíri | (zi)tátu | (zi)nái | (zi)sanu | **kʰúmi** |
| N20 | Tumbuka | **ka-môza** | tu-ŵîri | tu-tâtu | tu-nâyi | tu-nkʰonde | **kʰûmi** |
| P20 | Makonde | **iímo** | mbiíli | nnaátu | n̄tʃe:ʃɛ | mwaánu | **likuúmi / kuúmi** |

At the same time, the terms for 'one' and 'ten' can form a group opposed (by means of a noun class) to the rest of the numerals (Table 1.5).

With the exception of the terms for 'one' and 'ten', a specific marking of numerals by means of a noun class is rarely attested. A specific noun class (different from noun classes in other numerals) was found in only 6 languages for the numeral '3', and in only 7 for the numeral '4'. It should be noted, however, that a specific marker is often employed for the terms within the sequence from 'six'

Table 1.5: Common noun classes for '1' and '10'

| Branch | Language | **'1'** | '2' | '3' | '4' | '5' | **'10'** |
|--------|----------|---------|-----|-----|-----|-----|----------|
| Platoid | Ayu | **ɪ-dɪ** | a-fah | a-taar | a-naŋaʃ | a-tugen | **i-ʃog** |
| Tivoid | Ipulo | **é-mɔ̀** | v-íàl | vé-tàt | vé-ɲì | vé-tàn | **é-pɔ́:t** |
| Bantu-A40 | Bankon | **(i)yǎ** | (bi)ɓá | (bi)íyâ | (bi)nân | (bi)tán | **i-ɓǒm** |
| Bantu-M20 | Ndali | **ka-mukene** | fi-ŵiri | fi-tatʊ | fi-na | fi-hano | **ka-loⁿgo** |

to 'nine', e.g. the term for 'nine' bears a specific noun class marker in the 151 languages under study.

## 1.1.2 The grouping of numerals by noun class

Adjacent numerals are more often grouped by their noun classes. Among different numeral grouping types, several are diffused across all main branches of Benue-Congo. I will list 15 of the more frequent groupings of numerals and illustrate each of them with an example. These groupings are reported in Table 1.6.

Even limiting Table 1.6 to 15 groupings demonstrates the fact that some numerals (for example, '2') are grouped by noun class more often than other numerals (for example, '8'). By analyzing the whole table of groupings (reported in Appendix A-B), the following observations can be made regarding each numeral.

**Numeral '1'.** Groupings of the numeral '1' are relatively rare: the majority of languages, obviously, prefer to oppose '1' to all other numerals. In case it is grouped with other numerals, the most frequent grouping is within the first five ('1-5') or six ('1-6') numerals. In the analyzed database there are four languages which differentiate the first two numerals '1-2'. For instance, Ngoreme (Bantu-E10): *e-mʷe* '1', *e-beɾe* '2', but *i-satɔ* '3', in Gitonga (S60) *mwéyò* '1', *mbìlì* '2', but *dzì-ná* '4'.

**Numeral '2'.** The numeral '2' reveals the maximum predisposition to groupings. The most frequent are: '2-5' and '2-6'. The grouping '2-4' is significantly less frequent but remains present in the majority of Bantu zones and in other groups of Benue-Congo languages.

**Numeral '3'.** '3' is often found in groupings but is very rarely opposed by noun class to '2'. However, some very interesting examples exist. For example, Mbuun (Bantu-B80): *umwɛ́s* '1', *byĕl* '2', *í-tár* '3', *í-na* '4', *í-tân* '5'. It is worth mentioning that grouping of '3-8' and '3-10' were not encountered in any of the languages examined.

**Numerals '4' and '5'.** The only frequent grouping involving '4' is '2-4' (except

groupings that include four numerals or more) and for '5' it is '2-5' or '2-5/10'. The grouping '5-9' was encountered only in five languages and the grouping '5-10' and '5-8' (in combination '5-8/10' – only in one language. The lack of a frequent grouping of '5-9' can seem even more strange because in many languages numerals '6-9' are based on 5 (moreover, this type of derivational model can be reconstructed for Proto-Bantu and, perhaps, for Proto-Benue- Congo, with the sole exception of the numeral '8' which was apparently formed from '4'). Another unexpected case is the lack of grouping for '5/10', that is the lack of a specific class for '5' and '10', considering the fact that in many languages '10' is formed from '5'. This model was encountered only in one dialect of Eggon: *ò-tnó* '5', and *ó-kpo* '10', while in other numerals the noun class is not marked (I am not aware whether the different tone on the prefix indicates a different noun class).

**Numeral '6'.** A high number of groupings of '6-9' is natural. In many languages it becomes '6-8' because of the specific derivation of the number '9'. In contrast, groupings '6-10' are very rare.

**Numeral '7'.** It is worth mentioning the frequent grouping of '7-8' (21 languages). We are dealing not with one concrete class in Benue-Congo but rather a similar way of marking the numerals '7' and '8'. In the three examples reported in Table 1.3 the presumably common CL7 (Cilungu **tʃí**-, Sakata **ke**-, Xhosa **si**-) was found, in other languages a number of different classes can be encountered (Table 1.7).

**Numerals '8', '9', '10'.** The same charactetistic is typical for the frequent groupings of '8-9' and '9-10', shown in Tables 1.8-9.

1. The first column contains a stable grouping of numerals illustrated by an example. The second column indicates the number of languages which have this grouping (out of 254 languages under consideration). The rows in the table are organized in decreasing order. The third column lists all the groupings based on the noun class for a concrete language. Groupings of the adjacent numerals are indicated by a hyphen. Groupings of non adjacent numerals are indicated by a slash. Thus, the formula in the third column of the last row can be interpreted as follows: in Eleme there are three groupings of numerals – '2-3' (class **ɔ̀**-), '4-6' and '9' (class **ɛ̀**-), and also '7-8' and '10' (class **à**-).

Table 1.6: The most frequent groupings of numerals based on noun classes in Benue-Congo languages

| Grouping | Number of languages | Entire grouping | BC branch | Language | '1' | '2' | '3' | '4' | '5' | '6' | '7' | '8' | '9' | '10' |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2-5 | 58 | 1,2-5,6,7-8,9,10 | Bantu-F10 | Cilungu | tʃóónjá | ví-ílí | ví-tátú | ví-nì | ví-sáánò | mú-táándá | tʃí-nímbálí | tʃí-náání | fúúndúmbálí | i-kúmi |
| 2-6 | 42 | 1,2-6,7-8,9,10 | Bantu-C40 | Sakata | némo | í-pé | i-sar | i-ni | i-tsir | i-sonj | ke-ʃo | ké-né | leva | jõ |
| 2-4 | 24 | 1,2-4,5,6,7,8,9,10 | Bantu-C50 | Pagibete | moti | e-ɓale | e-salo | e-kwanane | ɓunoti | motoɓa | sambo | mwambe | libwa | zomi |
| 2-9 | 22 | 1,2-9,10 | Grassfields | Mundani | yea-mɔʔ | be-be | be-tat | be-kpi | be-tãã | be-ntúa | be-sããmbe | be-fãã | be-bɔʔa | è-yem |
| 7-8 | 21 | 1,2-6,7,8,9-10 | Bantu-S40 | Xhosa | ɲè | m-bini | n-tátʰu | ˀn-nè ? | n-tˡãnù | n-tˡándátʰu | sí-ˡlʰɛɲl̀ɛ | sí-bɔ̀zɔ̃ | lí-tʰɔ́ɓá | lí-ʃúmi |
| 6-9 | 20 | 1,5,6-9,10 | Ekoid | Nde-Ndele | n-dʒi | m-ba | n-sa | n-ne | n-dɔn | a-sighasa | a-simma | a-neghane | a-sima-wobo | wobo |
| 9-10 | 16 | 1,2-5,6,7,8,9-10 | Platoid | Lijili | lɔ̀ | à-bɛ̀ | à-tʃɛ́ | a-nàrɔ̀ | à-sɔ̀ | mì-nzí | mú-tá | rúnɔ̀ | zà-tʃɛ́ | zà-bɛ̀ |
| 1-6 | 15 | 1-6,7,8,9,10 | Bantu-E10 | Simbiti | ka-mʷe | ka-ɓere | ka-tatɔ | ka-nne | ka-taanɔ | ka-saⁿsaβa | mu-huⁿgate | mɔ-naane | keⁿda | i-komi |
| 6-8 | 15 | 1,5,6-8,9,10 | Bantu-F30 | Nilamba | ka-mwe | ka-beli | ka-tatu | ka-nee | ka-láno | mu-tandatu | mup-ungate | mu-naana | kyenda | kyumi |
| 2-10 | 14 | 1,2-10 | Mbe | Mbe | ómè | bé-pʷál | bé-sá | bé-ñí | bé-tʃán | bé-sésár | bé-tánèbépʷál | bé-ñíbɛ́ñí | bé-tánèbɛ́ñí | bé-fwɔ̀r |
| 1-5 | 14 | 1-5,6,7,8,9,10 | Grassfields | Ghomala | yɔ̀-mũ̀ʔ | yɔ̀-pwɔ̀ | yɔ̀-tá | yá-pɬúà | yɔ̀-tɔ̃ | ntɔ̀kɔ́ | sɔmbwɔ̀ɔ | hɔm | vɛ̀ʔɪ̀ | yám |
| 8-9 | 12 | 1/7,2-6,8-9,10 | Bantu-H30 | Ngongo | m-wisi | b-wol | bé-tat | be-wan | bé-tan | be-saman | ns-ambwadi | ke-nan | ke-bva | è-kwom |
| 1-10 | 9 | 1-10 | Defoid | Ayere | ì-kã̀ | ì-dʒì | ì-tã | ĩ̀-jẽ̀ | ĩ̀-tṹ | ì-fã̀ | ì-dʒʷ̃ì | ì-rõ | ĩ̀-dã̀ | ì-gʷ̃ã |
| 7-9 | 9 | 1,2-5,6/10,7-9 | Idomoid | Alago | ó-je | é-pà | é-ta | é-nè | é-hɔ | ì-hírí | à-hapà | à-hatá | à-hánè | ì-gʷ̃õ |
| 2-3 | 9 | 1,2-3,4-6/9,7-8/10 | Cross-River | Eleme | ǹ-nɛ | ɔ̀-berɛ | ɔ̀-taa | è-táale | è-wò | è-ʔɔ̀rɔ̀ | à-ʔàràbà | à-ʔaataa | è-siraʔò | à-ʔò |

11

Table 1.7: Groupings of '7'-'8' by noun classes

| Branch | Language | '6' | **'7'** | **'8'** | '9' |
|---|---|---|---|---|---|
| Bantu-B70 | Teke-Tyee | bísémɛnɛ | **n-tsaama** | **m-pwɔ́mɔ** | Owá |
| Bantu-C80 | Tetela | asamále | **e-sambɛ́ɛ́lé** | **e-náánéyi** | Divwá |
| Bantu-J30 | Nyore | bisasaba | **mu-safu** | **mu-nane** | Sienda |
| Platoid | Yeskwa | èncí | **tò-nvà** | **tó-ndát** | tyúôrá |
| Cross-River | Eleme | ɛ̀ʔɔ̀rɔ̀ | **à-ʔàràbà** | **à-ʔaataa** | èsiraʔò |

Table 1.8: Groupings of '8'-'9' by noun classes

| Branch | Language | '7' | **'8'** | **'9'** | '10' |
|---|---|---|---|---|---|
| Bantu-B10 | Myene | ò-ɾwáɣénô | **è-nánáyì** | **è-nóɣòmì** | ì-ɣómí |
| Bantu-B20 | Sake | bì-tánènɛ̀bìbá | **rì-mwâmbì** | **rì-bvùwɔ́** | dʒúmù |
| Bantu-B80 | Mpiin | n-sámwêːn | **bí-nán** | **bí-vwa** | kub |
| Bantu-H10 | Kikongo | sàmbúwàlì | **í-nànà** | **í-vùwà** | kúmì |
| Bantu-B80 | Songo | n-sambwar | **ki-nan** | **ki-va** | kwim |
| Bantu-J40 | Nande | eri-rínda | **omú-nani** | **omw-énda** | erí-kúmi |
| Bantu-J50 | Tembo | βi-rínda | **mú-nanɛ** | **mw-ɛnda** | ɛ́-kumi |
| Grassfields | Ngomba | sambá | **yé-né-fom** | **ye-ne-pfúʔú** | ne-gɯ́m |

Table 1.9: Groupings of '9'-'10' by noun classes

| Branch | Language | '8' | **'9'** | **'10'** |
|---|---|---|---|---|
| Bantu-B70 | Teke-Tyee | m-pwɔ́mɔ | **o-wá** | **o-kwúúmu** |
| Bantu -C40 | Budza | mo-nánáye | **li-bwá** | **ly-ómo** |
| Bantu -C80 | Tetela | e-náánéyi | **di-vwá** | **dí-kumi** |
| Bantu -G60 | Hehe | m-nane | **nyi-gonza** | **nyi-chumi** |
| Bantu -J60 | Rundi | umu-naáni | **i-tʃeénda** | **i-tʃúmi** |
| Platoid | Lijili | rúnɔ̹́ | **zà-tʃɛ̹́** | **zà-bɛ̹** |

## 1.2 Noun classes in derived (reduplicated) numerals

Reduplication is widely attested as a means of constructing numerical compounds in NC. This is especially applicable to the pattern '8 = 4 redupl.' which, as we hope to demonstrate below, can be reconstructed at the Proto-Niger-Congo level. Another common pattern (attested, however, with a somewhat lesser degree of frequency) is '6 = 3 redupl.'. Three main strategies pertaining to the use of the noun classes are employed within this derivation scenario:

1. Reduplicated terms preserve the class marker of the source-term in both segments, cf. e.g. Ndoe (Ekoid) *be-ra* '3' > *be-ra-ba-ra* '6', *be-ne* '4' > *be-ne be-ne* '8'; in Alege (Cross-River) *é-cɛ* '3' > *é-ce-e-ce* '6'.

2. The original class marker is preserved in only the first segment of the reduplicated form, and omitted in the second: Okpamheri (Edoid) *e-sa* '3' > *e-sa-Ø-sa* '6', *e-ni* '4' > *e-ni-Ø-ni* '8'.

3. Finally, the class marker of the first segment of the reduplicated form is different from that of its source-form: Kwa (Ekoid) *e-sa* '3' > *a-sa-ka-su* '6', *i-ni* '4' > *a-ni-ka-ni* '8'.

The number of these strategies is reduced to two in cases where a derived term is non-separable (e.g. derived by partial reduplication). In such cases, the class marker of the source-term can be either employed (Kikuyu *i-tatu* '3' > *i-tatatu* '6'), or not (Vinza *ka-ne* '4' > *mu-nane* '8').

We might expect that while forming '8' from '4', the singular class of the latter would be switched to the plural class of the former. In Bantu languages, however, this is not the case. Apparently already in Proto-Bantu we should reconstruct the derivational model **\*ì-nàì** '4' (CL.SG.5) > **\*mʊ̈-nànàì** '8' (CL.SG.3). However, from an etymological point of view, the class **mu-** represents the reflex of the class 6B.PL and not a reflex of the class 3.SG in Niger-Congo. This question raises an additional and very important topic which cannot be examined in the present study (the arguments in favor of class 6B.PL **mu** in Proto-Niger-Congo can be found in **Pozdniakov2013**).

**Bantu languages**. The following presents partial data on the numeral system in Myene (B10)[4] (Table 1.10).

First of all, it is interesting to highlight a variety of noun classes in the left column of the table and their uniformity in the right one. In the numerals from

---

[4]Thanks to Odette Ambouroué for some clarifications and for a profiatable discussion on noun classes in Myene.

Table 1.10: Myene numerals

| | | | |
|---|---|---|---|
| 1. | *N-mɔ̀rì (> mɔ̀rì) | | |
| 2. | *N-bànì (> mbànì) | 20. | à-ɣóm á-mbánì (10*2) |
| 3. | *N- ɾáɾò (> tʃáɾó) | 30. | à-ɣóm á-ɾáɾò |
| 4. | *N-náyì (> náyì) | 40. | à-ɣóm á-náyì |
| 5. | ò-tání | 50. | à-ɣóm á-tánì |
| 6. | ò-ɾówá | 60. | à-ɣóm ó-ɾówà |
| 7. | ò-ɾwá-ɣé-nómò (6+1) | 70. | à-ɣóm ó-ɾwá-ɣénô |
| 8. | è-ná-náyì  (2*4) | 80. | à-ɣóm é-ná-náyì |
| 9. | è-nó-ɣòmì (10−1) | 90. | à-ɣóm é-nó-ɣòmì |
| 10. | ì-ɣómí | 100. | *N-kámá. |
| | | 200. | kámá mbání |

1 to 10, the system includes four different singular noun classes: N- (CL9) – '1-4', ò- (CL3) – '5-7' (the numeral '7' is formed as '6+1', where *nómò* means «the only one, the same»), è- (CL7) – '8-9' (the numeral '8' is a reduplicated form of '4', the numeral '9' is formed as '9 = 10 − 1') and finally, ì- (CL5) – '10'. A homorganic nasal can be quite reliably reconstructed in '1-4', sometimes appealing to indirect characteristics. For example, in *tʃáɾó* '3' the nasal is absent but in Myene **tʃ-** is not a reflex of *t*. In this language *t- > r-*, as can also be seen in the second formative of '30'. The initial **tʃ-** can be traced back to *N-r-*.

In numerals of dozens only CL6 **à-** is used, which is one of the plural classes (with a collective meaning). An interesting detail: in '20' – '50' the second formative agrees with the first one in noun class (**á-**), and in '60' – '90' there is no agreement (the second formative maintains noun classes which mark the units as in independent forms; its high tone is due to the high tone in the preceding root *ɣóm*).

Non-derived numeral '100' belongs, as '1', to the singular class CL9. Does the second formative of '200' agree with the first one? It is impossible to say, because the noun classes of both formatives coincide when used singularly.

Finally, it is possible to formulate the principle of derivation with reference to the noun classes: the numeral '10', being a formative of numerals '20' – '90', maintains its meaning but changes the singular noun class to a plural noun class following the most standard SG ~ PL correlation in the language. For CL.SG.5 (**ì-** in Myene) which is expressed through *ì-ɣómí* '10', the standard correlate is CL.PL.6 (**à-**). Concerning the second correlate (units), it agrees with the first one (dozens)

in the numerals that even in independent use show agreement with nouns (in Bantu numerals '1-5' show agreement with nouns). For this reason in numerals '20'–'50' units from '2' to '5' agree with '10' in its plural form and in '60'–'90' second formatives '6'-'9' do not show agreement.

If we confront the numeric characteristics of simple and derived forms, the formation of numerals in Myene can be represented by SG > PL-PL and numerals '60' – '90' by SG > PL-SG.

This system is quite typical for Bantu languages, although the variation is considerable. The main variations are illustrated in the Table (1.11), including languages only from the zone J:

Table 1.11: Number patterns in derived numerals

| SG > SG-PL | 10 > 200 | CL5 > 5-8 | Hema | 10 | ikumi | 200 | ikumi bibiri |
|---|---|---|---|---|---|---|---|
| SG > SG-PL | 1000 > 2000 | CL11 > 11-8 | Hema | 1000 | rukumi | 2000 | rukumi bibiri |
| SG > PL-PL | 2 > 20 | CL5 > 6-6 | Gundu | 3 | ìsátɔ̌ | 30 | makumi gasatʊ |
| SG > PL-PL | 100 > 200 | CL5 > 6-6 | Shi | 100 | igana | 200 | magána abiri |
| SG > PL-PL | 10 > 200 | CL5 > 8-8 | Chiga | 10 | ìkúmì | 200 | βìkùmì βiβírì |
| SG > PL-PL | 100 > 200 | CL7 > 8-8 | Ganda | 100 | tʃikúmì | 200 | bìkúmì bìbírì |
| SG > PL-PL | 1000 > 2000 | CL7 > 8-8 | Shi | 1000 | cihumbi | 2000 | bihumbí bibiri |
| SG > PL-PL | 1000 > 2000 | CL11 > 10-8 | Ganda | 1000 | lùkúmì | 2000 | ŋkùmí bìbírì |
| SG > PL-SG | 8 > 80 | CL3 > 6-3 | Shi | 8 | múnaani | 80 | mákumi galí múnaani |
| SG > PL-SG | 9 > 90 | CL3 > 6-3 | Shi | 9 | múénda | 90 | mákumi galí múénda |
| SG > PL-SG | 1000 > 2000 | CL11 > 10-5 | Soga | 1000 | lùkúmì | 2000 | ŋ́kùmí ìβírì |
| PL > PL-PL | 2 > 20 | CL8 > 6-6 | Shi | 2 | bibiri | 20 | mákumi abiri |
| PL > PL-PL | 3 > 30 | CL8 > 6-6 | Shi | 3 | biʃarhu | 30 | mákumi aʃarhu |
| PL > PL-PL | 4 > 40 | CL8 > 6-6 | Shi | 4 | bíni | 40 | mákumi ani |
| PL > PL-PL | 5 > 50 | CL8 > 6-6 | Shi | 5 | birhaanu | 50 | mákumi arhaanu |

The Hema example demonstrates that the pluralization of the class for the formation of derived numerals is not mandatory (at least, for hundreds and thousands), although it unconditionally dominates in the languages of this group (Shi, Chiga, Ganda, Soga). If the simple numeral is already marked for plural class (there are examples demonstrating this), the first formative of the derived numeral appears with a new plural class (for example, in Shi). In the combination SG > PL-PL the plural classes in a composed derived numeral can be different (Ganda, derivation '1000' > '2000').

While forming a word combination from one word, the number of possible combinations of singular and plural classes amounts to eight. As shown in the table, only four of these combinations are actually encountered. No languages show combinations SG > SG-SG, PL > SG-SG, PL > SG-PL, PL > PL-SG This distri-

bution demonstrates how pluralization is used for the formation of numerals of higher rank. This strategy can be systematically found in other branches of Niger-Congo.

**Atlantic languages**. In order to be able to compare the principles of derivation of numerals in Bantu and in Atlantic languages systematically, we need to first formulate at least three main differences between these systems.

First of all, it is important to highlight that the system of Bantu is decimal, which is not typical for other branches of Niger-Congo, nor for other branches of Benue-Congo. The overwhelming majority of Altantic languages are '20'-based and not decimal. In these languages, accordingly, '40 = 20*2' (and often '100 = 20*5') and very rarely '40 = 10*4'.

Secondly, in Atlantic languages the numerals '6-9' are systematically formed following the model '5' + '1, 2, 3, 4'. This model does not permit the change of noun classes for the numerals '6-7' and/or '7-9'. The numerals '6-9' maintain all the characteristics of '5' (first formative) and '1-4' (second formative).

Thirdly, contrary to Bantu, the majority of forms of '5' are formed from the lexeme 'hand', maintaining the noun class of this lexeme. In Proto-Bantu 'hand' and 'five' are reconstructed as different roots.

The sum of the abovementioned factors explains the fact that noun classes in the numerals '6-9' are of no concern to the present study. Nonetheless, as will be further demonstrated, the main principle of interaction between noun classes and numbers in the numeral system of Atlantic languages is similar to that of Bantu.

Apparently, derived numerals were already formed following the model '40 = 20*2', '60 = 20*3', '80 = 20*4' in Proto-Atlantic. Different strategies of agreement are partially shown in the table (Table 1.12, (only the most simple cases were reported):

Table 1.12: Atlantic languages: noun classes in the derived numerals

|  | '20' | CL | '40' | CL-CL | '2' | CL |
|---|---|---|---|---|---|---|
| Bijogo | o-joko ('person') | SG | ya-joko ya-n-som | PL-PL | n-som | PL |
| Banjal | 'ə-vːi ('chief') | SG | 'u-vːi ɣuː- βɐ | PL-PL | 'suː-βɐ | PL |
| Kasa | ə-yiː ('chief') | SG | ku-yi: ku-l̹uβə | PL-PL | 'si-l̹uβə | PL |
| Bayot (Sénégal) | 'ə-yi ('chief') | SG | 'ku-yi kʊ-ɪɾɪgːə | PL-PL | 'ɪɾɪgːə | PL |
| Bayot (Guinea Bissau) | ga-bamɔgol ('person') | SG | gʊ-mɔgol-gʊ-ɾɪgˑga | PL-PL | tɪgˑga |  |
| Kwaatay | butuman | SG | ba-k-an ba-ka-suba | PL-PL | ku-suba | PL |
| Nyun Gunyamolo | buruhur | SG | ɟamaŋ ɪ-nakk | PL-PL | ha-nakk | PL |
| Karon | ə-wi | SG | ə-wi e-supək | SG-SG | su-supək | PL |

As demonstrated in Table 1.12, the majority of Atlantic languages within the Bak branch (Bijogo, Banjal, Kasa, Bayot) show that in the numeral '40' ('60', '80') the units '2' ('3', '4') agree in general according to a plural class and not according to the class of the numeral '20'. The same principle is characteristic for the languages of Benue-Congo. In all four abovementioned languages, the formation of '40' is based on the agreement in number as for animated nouns cl1.sg – cl2.pl (this is very clear especially knowing the etymology of the numeral '20').

Pluralization as a form of derivation is used when the form of the numeral '20' is not transparent (Kwaatay *butuman* '20', unclear etymology, Nyun Gunyamolo *buruhur* '20' (possibly from «price + man»); in the numeral '40' lexemes are used with the meaning 'people'). In some languages (Karon) the agreement is based on the singular class of the numeral '20' and not on its plural correlate.

In Atlantic languages that, like Bantu, systematically follow the decimal system, the pluralization of the class permits the formation of new numerals (more often as word combinations) (Table 1.13):

Table 1.13: Agreement in numerals derived from '10'

|         | SG<br>'10' | PL<br>'40' | SG, PL<br>'4' |
|---------|-----------|-----------|-----------|
| Basari  | ɛ-pəxw    | ɔ-fəxw ɔ-nɐx | ɓə-nɐx    |
| Sua     | Ø-tɛŋi    | i-tɛŋi i-naŋ | b-nan     |

In such cases agreement of the formatives can be observed, that is the same noun class is used for dozens and units. In the languages where '20' is formed from '10' (10*2), the units more often do not show agreement:

- Mankanya *i-ɲên* '10' (literally: «hands»), *i-ɲêŋ ŋí-tèp* '20' (*ŋí-tɛp* '2'), *i-ɲêŋ ŋɨ-bakɨr* '40' (*ŋɨ-bakɨr* '4');

- Jaad *pa-ppo* '10', *pa-ppo ma-ae* '20' (*ma-ae* '2'), *pa-ppo ma-nne* '40' (*ma-nne* '4'),

- Palor *dɐːŋkɛh* '10', *dɐːŋkɛh kɐ-nɐk* '20' (*kɐ-nɐk* '2'), *dɐːŋkɛh niːkiːs* '40' (*niːkiːs* '4').

Even in the following case the use of a plural class for units is possible: Baga Fore *ɛ-tɛlɛ* '10', *ɛ-tɛlɛ mɛn-di* '20' (*ʃi-di* '2'), *ɛ-tɛlɛ mɛ-nɛŋ* '40' (*ʃi-nɛŋ* '4').

Finally, in order to complete the description, hybrid composed forms will be reported, that is when '40' can be traced the root '20' and not '10' but in units where '4' is used and not '2'. This means that in '20' – '90' the root '10' is used, which is different from the main root:

- Nalu *tɛ bɪ-lɛ* '10' (literally: «two hands», *bɪ-lɛ* '2'), *alafaŋ bi-lɛ* '20', *alafaŋ bi:-na:ŋ* '40' (*bi:-na:ŋ* '4');

- Pepel *o-diseŋene* '10', *ŋ-taim pugus* '20' (*ŋ-pugus* '2'), *ŋ-taim ŋ-uakr* '40' (*ŋ-uakr* '4');

- Limba *kɔɔ-hi* '10', *kɔ-ntʰɔ ka-aye* '20' (*ka-aye* '2'), *kɔ-ntʰɔ ka-naŋ* '40' (*ka-naŋ* '4').

In spite of plurality of strategies, the modern systems of agreement of units in the dozens reflect a significant distinction that is characteristic of the two main branches of Atlantic languages – Northern and Bak. Apparently, the proto-languages of the Bak group maintained the principle of agreement which was typical for Proto-Niger-Congo, that is, the agreement of units following the plural correlate of '10' or '20'. This principle was lost in the system of the Northern branch, where it can be encountered in only one of the Tenda languages, Basari. It is also present in Nyun Gunyamolo, but in this language, as it is highlighted by different scholars, the numeral '20' (and probably the whole agreement model) is borrowed from Joola (Bak).

The model of agreement in '200'/ '2000' works in a similar way, as shown in Table (1.14):

Table 1.14: Agreement in '200' and '2000'

|  | Language | '100' | '200' | '1000' | '2000' | '2' |
|---|---|---|---|---|---|---|
| 1 | Balant | geme | g-geme g-sibi | wili mbooda ('1') | g-wili g-sibi | -sibi |
| 2 | Bayot | ɛ-tɛmel | ɪ-tɛmel i-rig·ga | ɛ-ʊlɪ | ɪ-ʊlɪ–i-rig·ga | tɪg·ga |
| 3 | Banjal | ɛ'-kɛmɛ | sɪ'-kɛmɛ 'su:-βɐ | 'e-uli | 's-u:li 'su:-βɐ | 'su:-βɐ |
| 4 | Kwaatay | temer | si-temer sú-suba | ẽ-ñjune | sú-ñjune sú-suba | kú-suba |
| 5 | Baga Fore | bɔ ben ('1') | ʃu-bɔ ʃi-di | tɛngbeŋ ben ('1') | ʃi-tɛngbeŋ ʃi-di | ʃi-di |
| 6 | Nalu | m-laak | a-laak bi-lɛ | m-ɲaak | a-ɲaak bi-lɛ | bi-lɛ |
| 7 | Basari | kɛmɛ | ɔ-kɛmɛ ɔ-ki | wəli | ɔ-wəli ɔ-ki | ɓə-ki |
| 8 | Konyagi | keme | wɐ-keme wɐ-ki | wəli | wɐ-wəli wɐ-hi | wɐ-hi |

As observed for dozens, the agreement in '200' and '2000' can be systematically observed only in the languages of the Bak group (languages 1-5 in Table 1.14).

In the Northern group this agreement is found only in Basari (7). Even in Konyagi, the fact of agreement is not clear because in this language the CM of '2' in '200' and '2000' coincides with the CM of CL2 in independent use (for the same reason it is not clear whether we encounter agreement in Baga Foré (5). Moreverover, there is no agreement in Nalu (6), a language of the same branch.

In the majority of languages, the noun classes of '200' and '2000' systematically differ from the noun classes of units and dozens. This is typical for Niger-Congo, perhaps because in '100'/'200' and '1000'/'2000' we are often dealing with borrowings.

**Mel languages**. The present analysis will be limited to the data from one Mel language, that is Temne (Kərata dialect) collected by David Odden (Table 1.15):

Table 1.15: Noun classes in Temne numerals

| 1 | p-ín | | | | |
|---|---|---|---|---|---|
| 2 | pɨ-rʌ́ŋ | 20 | kɨ-gbá | | |
| 3 | pɨ-sas | 30 | kɨ-gbá 'tɔ́-fɔ́t (20+10) | | |
| 4 | pa-nlɛ | 40 | tɨ-gbá tɨ́'-ríŋ (20*2) | | |
| 5 | tamáṭ 5 (*ta-tam-at) | 50 | = 20*2+10 | | |
| 6 | du-k-ín (X+1) | 60 | tɨ-gbá tɨ́-sas (20*3) | | |
| 7 | dɛ-ríŋ (X+2) | 70 | = 20*3+10 | | |
| 8 | dɛ-sas (X+3) | 80 | tɨ-gbá tâ-nlɛ (20*4 ) | | |
| 9 | dɛ-ŋanlɛ (X+4) | 90 | = 20*4+10 | | |
| 10 | tɔ-fɔ́t (< * ta-fu-at) | 100 | k-ɛmɛ́ k-ín | 200 | t-ɛmɛ́ tɨ́'-ríŋ |
| | | 1000 | ʌ-wúl 'ŋ-ín | 2000 | ɛ-wúl jɛ-ríŋ |

The numerals '1-4' in counting forms belong to CL.SG **pV-**. The numeral '5' can be traced back to the form with positive meaning of definiteness (*ta-tam-at*) – as well as 10 (< *ta-fu-at*), initially having the structure CV-CVC-VC, where CV- and -VC are allomorphs of the noun class in a definite form and CVC is the root (**Pozdniakov1993**)[5]. For us, it is important that the numerals in '5' and '10' can be reconstructed with CL.SG **ta-**. The non-derived numeral '20' can be traced to CL.SG, and in particular **kə-**. The numerals '40' – '90' are formed with the change of the noun class in the first formative to CL.PL **tə-**. Furthermore, the second formative agrees with the first one in noun class and consequently is also

---

[5]It is clear that '5' and 'hand' have assonance in the languages of the group. Due to space limitations, it is impossible to explain the complicated emergence of this assonance. Let's also leave aside details on the first formative in the numerals '6-9'.

included in the class **tə-**. That is to say, this is the same derivational model as in Bantu and in Atlantic languages. This model emerges as well in the formation of '100' and '200'. In the borrowed form *kɛmɛ* '100' the initial root consonant can be interpreted as a singular CM (the same noun class as in '20'). That means that '200' is used as its plural correlate and the original root consonant gives us **t-**. Finally, the correlation of '1000' ~ '2000' can be interpreted as correlation in number but with a new pair of classes: CL.SG ʌ- ~ CL.PL **ɛ-**.

**Gur languages**. An example of an interesting system from the Ditammari language (Oti-Volta) follows (Table 1.16):

Table 1.16: Ditammari: agreement in the derived numerals

| SG | PL–PL | SG |
|---|---|---|
| tɛ-pii-tɛ '10' | si-pi-si-dɛ '20' | dɛɛ-ni '2' |
| | si-pi-si-tâadi '30' | tâadi '3' |
| | si-pi-si-wɛi '90' | n-wɛi '9' |
| di-tu-si-di '100' | yɛ-tu-si-ɛ yɛ-déɛ '200' | dɛɛ-ni '2' |
| di-yɔɔ-di '1000' | yɛ-yɔɔ-d-ɛ yɛ-dɛ̀ '2000' | |

In this example we can see the correlation of number classes in derivatives and «agreement» between the parts of syntagm in '200' and '2000' using different structures of class markers (prefixes, suffixes, confixes, or the lack of marker).

Similar formation strategies of derived forms can be found in another language from the Gurma group (Oti-Volta), Miyobe (Table 1.17):

Table 1.17: Miyobe: noun classes in derived numerals

| SG | PL, SG-PL, PL–PL | SG |
|---|---|---|
| **kɛ**-fi '10' | **ɑ**-fɛɛ-rɛ́ '20' | -tɛ́ '2' |
| | **ɑ**-fɛɛ-nɑ '40' | **n**-nɑ '4' |
| **pí**-lɛ '100' | **pí**-lɛ-**pí**-lɛ **mɛ**-tɛ́ '200' | |
| **kú**-kotokú '1000' | **á**-kotokú **ɑ**-tɛ́ '2000' | |

In '20' (10*2) and in '2000' (1000*2) a plural correlate CL.SG **kV-** (CL.PL **á-**) is used. In '2000' the numeral '2' agrees in noun class with '1000' (the root is formed from the word with the meaning 'sack'). In '200' the reduplication of '100' and a special class marker (CL.PL **mɛ**) for the formative '2' are used.

Another language from Gurma group Ntcham follows the same standard model (Table 1.18):

Table 1.18: Ntcham: noun classes in derived numerals

|  | SG |  | PL–PL |  | SG |
|---|---|---|---|---|---|
| 20 | ḿ-mùǹkú | 40 | ì-mùǹkú ì-lí | 2 | ǹ-lí |
| 100 | di-làátàà-l | 200 | kú-làáfaa-u |  |  |
| 1000 | Ø-kùtùkú | 2000 | Ø-kùtùkú-bì bì-lí | 2 | ǹ-lí |

The numeral '200' is formed from '100' by changing from the singular class to the plural one.

The existence of similar strategies for use of plural class markers for the formation of numerals of higher rank in different areas of Niger-Congo (Benue-Congo, Atlantic languages, Mel languages and Gur languages) permits us to presume that similar principles of interaction between noun classes and numbers were typical for the system of Niger-Congo as well. There are no traces of derivative pluralization in Kru and Ijo languages, but they can surely be found in Kwa languages. I did not manage to find similar strategies in the Adamawa and Ubangi languages, nonetheless traces can be found in Kordofanian languages.

Here is an example from Koalib, a Kordofanian language (Table 1.19):

Table 1.19: Koalib example

|  | SG |  | PL–PL |  | SG |
|---|---|---|---|---|---|
| 20. | t-úɽì | 40. | r-úɽì r-ìɽèn | 2. | -iɽɐn |
|  |  | 2000. | á-lep (< arab ) w-ìɽèn | 200. | mîɐ kw-ìɽèn |

A prefix for the plural class is used for the formation of the numeral 40. The formative '2' in '40' agrees with the formative '20' in the noun class. In '200' the prefix of singular class CL1 is used, which includes animated nouns and borrowings. In '2000', in the formative '2' is used for the prefix **w-**, a standard agreeement marker for vocalic noun classes.

Traces of pluralization of noun classes as a means of derivation in numerals can be found in Moro and Acheron (both are Kordofanian languages).

This distribution gives us sufficient grounds to assume that derivation for the formation of dozens in Niger-Congo was similarly established in Proto-Niger-Congo.

## 1.3 Noun class as a tool for the formation of numerals

Finally, there is one (perhaps the most interesting) strategy for formation for derived numerals. It consists exclusively of changing the noun class for the formation of a derived form. The system from Efik is partially reported below (Table 1.20):

Table 1.20: Efik example

| 2 | í-bá | 40 | à-bà |
|---|------|----|------|
| 3 | í-tá | 60 | à-tá |
| 4 | í-náŋ | 80 | à-nàŋ |

In Efik, as in the majority of Niger-Congo languages, a stable correlation in number CL5.SG ~ CL6.PL can be found: in Efik reflexes of these classes are accordingly **í-** ~ **à-**. A simple change of singular class to plural (with no compound forms and no reduplication) is enough to form '40' from '2', '60' from '3' and '80' from '4'. Apparently, this system uses '20' as its primary base.

The formation of new numerals by a change in noun class can be encountered in some languages of Benue-Congo, including Bantu (Table 1.21):

Table 1.21: Benue-Congo examples

| Bantu-B80 | Tiene (Tende) | 4 | i-níì | 40 | mu-níì |
|-----------|---------------|---|-------|----|--------|
| Bantu-C40 | Sengele | 4 | í-nɛi | 40 | mo-nɛi |
| Bantu-C90 | Ndengese | 4 | i-nej | 40 | bo-neji |
| Grassfields | Limbum | 4 | Ø-kjè: | 40 | m̩-kjè: |
| Edoid | Degema | 2 | i-βə́ | 40 | ʊ-βá |

This technique is mostly used in Bantu languages within the zone J. The data reported in Table 1.18 does not necessarily signify that the conceptual base for derivation is the pluralization of original forms. In Tiene, Sengele, and Ndengese, derived numerals, as well as base numerals, belong to singular noun classes.

For example, for the languages J10 SG > SG is characteristic for four derivations which can be illustrated by Gundu language (Table 1.22).

Other derivations SG > SG can be found occasionally. Apparently, the forms *n-datu* '6' > *tʃí-ɾatu* '60' (CL9 > CL7) and *mú-nanɛ* '8' > *lú-nanɛ* '80' (CL3 > CL11) were encountered only in Tembo (J50). We can see that the choice of nominal classes

Table 1.22: Gundu number patterns in the derivations of numerals

|     | 8 > 80<br>cl3 > CL7 |     | 9 > 90<br>cl3 > CL7 |      | 10 > 100<br>cl5 > CL7 |      | 10 > 1000<br>cl5 > CL11 |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 8   | mɔ̀-ná:nèí | 9   | mʷè:-ⁿdá | 10   | í-kùmí | 10   | í-kùmí |
| 80  | ki-na:nei | 90  | kʲe:-ⁿda | 100  | ki-kumi | 1000 | ɾu-kumi |

differs in different languages, that is, it is not the symbolic semantics of nominal classes that is most important, but rather their paradigmatic modification.

In Bantu J10-J20 we find a triple derivation model cl5-*kumi* (or cl9-) '10' ~ cl7-*kumi* '100' ~ cl11-*kumi* '1000'. Thus in Hema, *i-kumi* '10' ~ *ki-kumi* '100' ~ *ru-kumi* '1000'.

This model can be found in Gur languages as well. In Nothern Nuni (Grusi group) dozens are formed exclusively by a change in noun class marker. The derivation from '20' to '50' is realized by the change of one singular class to another: *bì-là* '2' > *fīì-là* '20', *bì-twàà* '3' > *fīì-twàà* '30', *bì-nu* '5' > *fīì-nu* '50'. Formation of dozens by a change of class is encountered in some Senufo languages as well.

However, the derivational model sg > pl is much more active. In the Bantu zone J, six derivations are typical, illustrated by the following examples from Gwere (J10) (Table 1.23):

Table 1.23:  Gwere number patterns in the derivations of numerals

|     | 2 > 20<br>cl5 > cl6 |     | 3 > 30<br>cl5 > cl6 |     | 4 > 40<br>cl5 > cl6 |     | 5 > 50<br>cl5 > cl6 |     | 6 > 60<br>cl3 > cl10 |     | 7 > 70<br>cl3 > cl10 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 2  | ì-βírí | 3  | ì-sátú | 4  | ì:-ná | 5  | ì-tá:nú | 6  | mù-kâ:gá | 7  | mù-sá·ⁿvú |
| 20 | à:-βírì | 30 | à:-sátù | 40 | à:-nâ | 50 | à:-tâ:nù | 60 | ṇ-kâ:gà | 70 | ṇ-sá·ⁿvú |

For the numerals '20'–'50' cl6.pl is used, and for '60'–'70' cl10.pl is used. These classes demonstrate the correlation in number with the classes cl5.sg and cl3.sg respectively. In at least four languages in zone J, the model cl3.sg > cl10.pl was encountered for '9' > '90'. In Gwere and Tembo, the model cl5 > cl6 is used in derivation '2' > '20': Gwere *ì-βírí* '2' > *à:-βírì* '20'.

Only one language, and that is Tembo, systematically presents model pl > pl in the derivation cl8.pl > cl6.pl (Table 1.24):

This model is clearly secondary and was implemented as a result of re-interpretation,

Table 1.24: Tembo example

| 3 | βi-hátu | 4 | βí-nɛ | 5 | βi-tánɔ | 7 | βi-rínda |
|----|---------|----|-------|----|---------|----|----------|
| 30 | má-hátu | 40 | má-nɛ | 50 | ma-tánɔ | 70 | ma-línda |

atypical of zone J, of classes in numerals '2-5', '7' as plural classes opposed to '1'.

The fourth theoretically possible model, that is PL > SG, has never been encountered in any derivation which can be considered indirect evidence for the idea that the pluralization of numerals of higher rank is one of the key strategies for the formation of derived numerals, as was demonstrated. Nonetheless, this strategy does not explain everything.

In order to present this elegant mechanism of systematic use of noun classes in the derivation of numerals in greater detail, an example from derivation in Soga using the roots '10' and '2' will be schematically presented. The root meaning '10' matches in Soga with six different class markers, and the root meaning '2' matches with three of them, as shown in Figure 1.1.
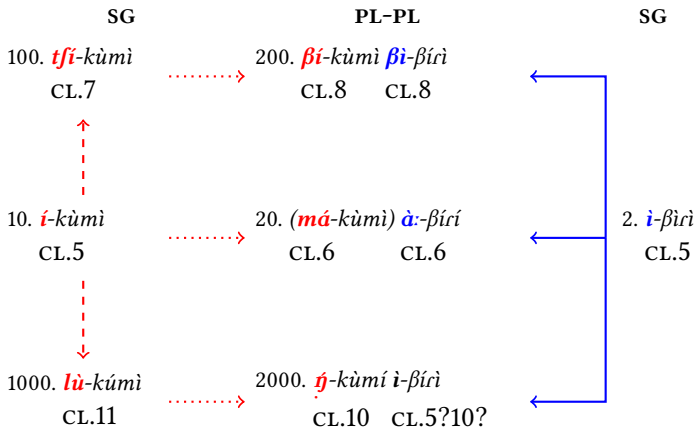


Figure 1.1: Soga numerals: derivations by noun classes

In the Soga language the root **kumi** takes part in three forms with singular class and three forms with plural class (one is facultative). In the derivations including forms of different numerals it is visible that the most stable correlations in number are: CL5-CL6, CL7-CL8 and CL11-CL10. At the same time, the choice of CL7 and CL11 for the derivations (as shown in Figure 1.1) seems to be arbitrary. According to Larry Hyman (p.c.) in the dialect Lulamoji, the archaic form of the

numeral '1000' belongs to the the CL11 and not to the CL14 (Hyman: «*óBu-kumí* '1000', older usage»).

The root ***βiri*** does not take part in singular derivates but was found in three derivates where ***kumi*** is marked by plural class markers. The main derivate from *ì-βìrì* '2' can function separately outside of the word combination (*à:-βírí* '20'). In this case, the main correlation in number for the class 5 is used (CL5-CL6). The difference in the class markers CL6 **mɑ-** and **ɑ:-** (in some dialects **ga-**) is related to the characteristics of the paradigms of agreement markers. A question about the nature of *ì-βírì* in '2000' emerges. Does it belong to CL5 or is this an homonymous form of the agreement marker in CL10? These questions are very hard to answer because we are dealing with derivational forms of class markers (often homonymous) and we cannot check on the context of agreement in order to test it.

In fact, the number of classes in numerals (both singular and plural) can be even bigger. In Soga, a singular form of '8' *mù-ná:-nà* (CL3) is always formed from the numeral '4' *í-nà* (CL5). In Mpumpong (Bantu, A80), the system of numerals includes four different plural noun classes, that is CL8 for units - *tên nɛ̀ i-nâ* '9' (5+4), CL6 – for dozens – ***mɛ̀-kàm mɛ̀-mbá*** '20'(10*2), CL4 for hundreds – ***mì-tsɛ̆t mì-mbá*** '200' (100*2), and CL2 for thousands – ***ò-tɔ́sìn ò-bá*** '2000' (1000*2).

The model of formation that was masterly developed by Soga has major relevance not only for the history of numerals in Niger-Congo, but for the theoretical analysis of the semantics of noun classes as well. The signifier of morphemes in noun class paradigms has a multilayer structure. This structure presumes that the semantics of each class can be defined through the paradigm at the intersection of four parameters: classificational, paradigmatic, syntagmatic and modal (for a more detailed discussion see **Pozdniakov2003**). It is useless to discuss the classificational aspect of noun class semantics in Soga numerals as we do when classes for humans, trees or animals are taken in consideration. The paradigmatic aspect of the signifier of the signs is the most relevant because the primary role is given to the correlation of classes in number, while some other paradigmatic correlations remain important as well.

In conclusion, it should be noted that the noun class switch as a derivation mechanism is not limited to Benue-Congo and can be reconstructed at the Proto-Niger-Congo level in at least one case (see chapter 4).

# 2 Analogical changes in numerals

## 2.1 Issues pertaining to the detection of alignments by analogy

In addition to the grouping of numbers by noun class, a number of more radical strategies are used in the Niger-Congo languages. One of them is the formal alignment of numbers resulting from the diachronic alignment of forms by analogy. This strategy implies irregular phonetic changes in lexical stems. As a result, contiguous numerals in the Niger-Congo languages often have similar forms, that is they have common phonetic element(s).

Such cases are not easily distinguishable from phonetic similarities conditioned by morphological changes, when affixes that are no longer productive blend into lexical roots, for instance, or archaic noun class markers exist in the numerals. Thus, in Wolof, as shown in the introduction, phonetic similarities arise in the numerals '2'–'4' (*ñaar* '2', *ñett* '3', *ñeent* '4') as a result of inclusion of the noun class marker Ñ in the lexical roots.

Only specialists of a concrete language can distinguish between morphological "accidents" and phonetic analogical changes, but sometimes even synchronic competence may not be enough. Table 2.1 shows the first six numerals in five Adamawa languages.

Table 2.1: Adamawa examples

|     | Languages | '1'   | '2'   | '3'     | '4'    | '5'    | '6'          |
| --- | --------- | ----- | ----- | ------- | ------ | ------ | ------------ |
| (1) | Tunya     | sèlì  | ari   | ata     | ana    | aluni  | nano         |
| (2) | Vere      | muo   | ituko | tariko  | nariko | gbanara | baburo      |
| (3) | Mom Jango | muzoz | itez  | taz     | naz    | gbana  | babez        |
| (4) | Dirrim    | nuan  | bara  | tara    | nara   | tona   | tini         |
| (5) | Pere      | dɔ̀ɔ̄  | īrō   | tā̰ā̰rō  | nārō   | núnnō  | nóndɔ́ɔ̄ (5+1) |

In Tunya (1) it is clear that the initial **a-** in the numerals '2'-'5' etymologically has the nature of the noun class marker. In Vere (2) the final syllable **-ko** can

hardly be considered a noun class marker, but it is very likely that we are dealing with a morpheme and not with a phonetic alignment of numerals. In Mom Jango (3) the final -**z** in '1'-'4' and '6' is difficult to comment on; it is likely that this is an analogical change but its direction is not very clear. In Dirrim (4) *bara-tara-nara* is the case of analogical change and, considering the diachronic context, the numerals '2' and '4' were clustered together with '3'. In Pere, the final -o in '2'-'5' may represent an analogical alignment or a morpheme.

Let us exclude all the cases of integration of noun class markers into stems and consider all the other cases of phonetic (or hidden morphological) clustering in the systems of numerals in Niger-Congo. We will deal mainly with two questions:

1. In which branches of Niger-Congo do analogical alignments have a major role and in which they are practically absent? This question is of crucial importance for the step-by-step reconstruction of numerals in Niger-Congo.

2. Which numerals phonetically align together and which analogical groupings are rare? This question is important not only for the etymology of numerals but also for the typology of analogical changes in numerals.

The topic of the present chapter is not relevant to all the branches of Niger-Congo. For instance, in Bantu and Benue-Congo there is no systematic analogic phonetic alignment. But in some other branches it is impossible to discuss the etymology of numerals without considering this factor. In the twelve main branches of Niger-Congo the situation is as follows:

In the first three branches the minus does not mean that there is no phonetic alignment of numerals. Some examples from Benue-Congo languages are given in Table 2.3.

Each of these examples is interesting for the study of concrete languages, but these seem to be the only languages, among hundreds of BC languages, where analogical changes have been found; therefore, no systematic changes of this type for the BC family have been attested.

In Mel there is only one case which is of interest to us, that is the unification of the initial root consonant in Krim: *yi-**g**in* '2', *yi-**g**a* '3'. The direction of analogical alignment in this case is not clear. It is impossible to study this particular case here, because the discussion of possible hypothesis would require a separate publication. It is important to underline that in other Mel languages cases of phonetic alignment of numerals have not been attested.

There are virtually no unifications of this type in Kru, excluding the phonetic alignment of the initial consonant in '4'-'5', reported in Table 2.4.:

Table 2.2: Analogic alignment in NC numerals

|    | NC family     | Analogy in numerals |
|----|---------------|---------------------|
| 1  | Benue-Congo   | –                   |
| 2  | Mel           | –                   |
| 3  | Ijo           | –                   |
| 4  | Kru           | –?                  |
| 5  | Mande         | –?                  |
| 6  | Atlantic      | +                   |
| 7  | Kwa           | +                   |
| 8  | Adamawa       | +                   |
| 9  | Ubangi        | +                   |
| 10 | Gur           | +                   |
| 11 | Dogon         | +                   |
| 12 | Kordofanian   | +                   |

Table 2.3: BC examples of analogic alignments

| Language    | '1'    | '2'   | '3'    | '4'     | '5'     |
|-------------|--------|-------|--------|---------|---------|
| Gweno (E30) | -m**wi** | -**vi** | -tharu | -nya    | -thwanu |
| Tiv         | mòm'   | h**ar** | -t**ar** | -nyin   | -tan    |
| Mmen        | mɔʔ    | b**ege** | t**ege** | kaiko   | ta      |
| Bute        | mui    | bam   | tare**b** | nasi**b** | -gi     |
| Kila        | mwe    | han   | t**ar** | **n**a**r** | tien    |
| Mama        | moʔon  | mari  | tar**u** | la jin**u** | ton**u** |

Table 2.4: Kru alignments in '4'–'5'

| Language        | '1'  | '2' | '3' | '4'   | '5'  |
|-----------------|------|-----|-----|-------|------|
| Gbe             | do   | so  | ta  | hyi~  | hm   |
| Southern Grebo  | do   | so  | ta  | ha    | *hm  |
| Bassa           | doo  | so  | ta  | hiye  | hm̋  |

I will dare to assume (based on these data) that the initial consonant in '4' has undergone analogical change with the consonant in '5'. The final judgment should be done by specialists. In Ijo this type of alignment is absent.

## 2.2 Mande

There are no systematic analogical changes in the systems of numerals in Mande languages.[1] Some languages like Busa, San (South-Eastern branch) and Soninke (Western branch) present exceptional cases.

In Busa, we are probably dealing with the fossilized suffix **-hõ** which can be found inside the lexical roots of '3' and '4': *\*a-hõ* '3', *\*si-hõ* '4', i.e. the phonetic similarity can be explained morphologically.

In San, apparently, the regular reflex of the three different consonants of proto-language of South-Eastern Mande is **s-** (see 3.10 below). Finally, three of the contiguous numerals start with the same consonant: *so* '3', *si* '4', *soro* '5'.

Soninke represents a more complicated case, wherein the last vowel of each numeral is not distributed randomly (Table 2.5):

Table 2.5: Soninke

| | | | |
|---|---|---|---|
| 1 | ba(a)ne | 6 | tu(n)mu |
| 2 | filo | 7 | nieru |
| 3 | siko | 8 | segu |
| 4 | (i-)nakato | 9 | kabu |
| 5 | karago | 10 | ta(n)mu |

In '1' there is a particular vowel -**e**. "Minor" numerals ('2'-'5') have the final -**o**, and all the higher numerals ('6'-'10') – final -**u**. Following the reconstruction of Nazam Halaoui (**Halaouï1990**): *fill-a* '2' (active voice) /*fill-e* '2' (passive voice) > *fill-e-nu* (PL) '2' > *fill-o* (PL) '2'. In other words, in the numerals '2-5' the vowel -**o** is interpreted by Halaouï as a phonetically conditioned allomorph of the plural morpheme -**nu**. But in the numerals '6-10' another vowel was found, not -**o**, but -**u**. Nazam Halaouï explains this in the following way: irregular final vowel -**u** initially appeared in the numeral '6' as a consequence of progressive assimilation (*\*tunm-o > tunmu*), and then following the analogy this vowel appeared in

---

[1] I would like to thank Valentin Vydrin for a detailed discussion of the history of numerals in Mande languages.

numerals '7'-'10'. Halaoui's hypothesis is not plausible (it presupposes a doubtful phonetic change *$e$-**nu** > -**o** in the numerals '2'-'5'), neither is it the only one possible.

Valentin **Vydrin2006** shows that Soninke has two different plural suffixes, -**u**/-**o** and -**ni**/-**nu** (the allomorphs -**u** and -**o** are dialectal variants, the same is true for -**nu** and -**ni**). It is not quite clear, do we have the generic plural marker -**u** in all the numerals from '6' through '10', or whether it is the alternative plural marker -**nu** that appears in '6' and '10', while the generic plural -**u** appears in '7' through '9'. In any case, it is evident that in the right column of Table 2.5, the final -**u** is of morphological origin, rather than a result of an analogical change. The fact of the appearance of a plural marker in the numerals '6'-'10' by itself is noteworthy; these numerals should be interpreted as pluralia tanta. Interpretation of the final -**o** in '2'-'5' is much more problematic. There is a singular morpheme -**o** in Soninke, however, Vydrin's data do not clarify why it is -**o**, rather than -**e** or -**Ø**. Therefore, it can be conjectured that the final vowel of the numerals '2'-'5' result from analogical changes.

Now let us move to the branches where analogical changes are systematic. Even in these cases we will encounter different examples.

## 2.3 Atlantic

In Table 2.6, the data on the first five numerals in ten various Joola languages will be compared.

Table 2.6: Joola

| Joola | '1' | '2' | '3' | '4' | '5' |
|---|---|---|---|---|---|
| Joola Karon | ɔ-ɔnɔ(ɔ)l | supə**k** | hə**ə**ciil | paa**k**ʊl | sa**k** |
| Bayot | ɛ-ndon | i-rigə**ʔ** | i-fiigi**ʔ** | i-βei**ʔ** | ɔ-r̥ɔ**ʔ** |
| Joola Gusilay | ya-nɔ | su-ruba | si-fe**gir** | si-ba**gir** | fu-tok |
| Joola Banjal | a-nu | si-gaba | gu-fi**gir** | si-ba**gir** | fu-tok |
| Joola Fogny | yɛ-kon | si-gaba | si-fe**gir** | si-ba**kir** | fu-tok |
| Joola Mlomp | yɔ-nɔɔ**l** | si-sube**l** | si-heji**l** | si-baci**l** | ŋa-suwaŋ |
| Joola Kasa | ya-no(r) | si-lube | si-hej**i** | si-bak**i** | hu-tok |
| Joola Ejamat | a-yʊnka | ku-lube | si-heji | si-bacir | fu-tok |
| Joola Kerak | ya-nɔr | si-sube | si-heji | si-bacir | hu-tok |

In the last group, apparently, there is no reason for the establishing phonetic alignments. In the meantime, in the first two groups such alignments are evident. In the first group the velar consonant is spread, and in the second group, the liquid consonant; furthermore, the roots are mostly related. These are classical "symptoms" of analogical change. It is clear that it is useless to etymologize the numerals without an in-depth analysis of these alignments.

Joola languages form one of the four branches of the Bak group in Atlantic. In Bijogo, there are no analogical changes in numerals. In the other two branches, these changes of various types can be found, and such changes differ from the type of changes in Joola.

In Pepel (Manjak branch) in some sources the numerals '2' and '3' have a final -s, in other sources they have a final -ṭ and in Koelle (1963[1854]) the final consonants are different, which can correspond to the situation in proto-language (Table 2.7):

Table 2.7: Pepel

| '2'      | '3'      | Source          |
|----------|----------|-----------------|
| pugu**s** | ŋa-jen**s** | **Ndao2011**    |
| pugu**ṭ** | waa-jin**ṭ** | **Wilson2007**  |
| ge-pugu**s** | ga-ci**t** | Koelle 1963[1854] |

In the branch that is represented by isolated languages Balant (Senegal; according to the data from Creissels & Biaye 2015) for the numerals '2' and '3' the following forms exist (Table 2.8):

Table 2.8: Balant

| 2          | 3                      |
|------------|------------------------|
| CL-sìbí    | CL-hàbí ~ CL-yàbí      |
| sìɩbí      | yàabí                  |

Apparently the numeral '2' has undergone the analogical change following the numeral '3'. The sources on Balant Kentohe give different but also phonetically clustered forms: -*sebm* '2', -*abm* '3'.

It is important to underline that analogical changes in the three aforementioned branches of Bak languages are not historically related – these changes are of different origin. This means that for this group, the principle of phonetic alignments of numerals is characteristic, but different types of changes by analogy co-exist. A similar situation is also typical of Northern Atlantic languages, which show other types of phonetic alignments.

In Wolof, as previously mentioned, the alignment of the initial consonant in numerals '2'-'4' is of a morphological nature; these numerals maintain traces of the noun class prefix. Still, for native speakers these forms contain a similar phonetic marker that groups together the numerals for '2'-'4' and distinguishes them from other numerals.

In Sereer (Northern Atlantic), as in Joola (Bak Atlantic) the final velar can be clearly seen in the numerals '2'-'5': fi**k** '2', tadi**k** '3', nahi**k** '4', βeti**k** '5'. Here the clustering involves not only the final consonant but the precedent vowel as well, which creates an illusion of the existence of a specific morpheme ('suffix' -**ik**) used for marking the numerals '2'-'5'. As will be demonstrated later, this is a false intuition. In Sereer, for example, we deal with morphophonology and not with morphology. Moreover, the coincidence with Joola is not casual and reflects an important phonetic innovation which took place in Proto-Atlantic.

In Nyun (the branch Nyun-Buy, Northern Atlantic languages) form clustering occurs through the final velar -**k** as well: -**n**du**k** '1', -**n**a**k** '2', -re-**n**e**k** '4'. It is worth highlighting that the initial consonant of the aforementioned forms is also unified (**n**-).

The same isogloss can be encountered, although in its shorter version; in one of the five languages of the Cangin branch, that is in Palor, ka-na**k** '2, ke-je**k** '3'. For Cangin this alignment is definitely marginal, in all the languages of Cangin branch another analogical change is encountered: the initial consonant in the numerals '1'-'2' is unified, which is a rare phenomenon. In Proto-Cangin we have *ji- noʔ* '1', *ka-nak* '2' with the maintenance of the initial **n**- in all five languages (compare with the unifications in Nyun).

The final -**n** is the basis for phonetic alignment in Sua, though the affiliation to Atlantic languages has not been proven: sɔ**n** '1', m-ce**n** '2', b-rar '3', m-na**n** '4', sugu**n** '5'.

## 2.4 Kwa

54 out of the 111 sources for Kwa languages available in our database show a common initial consonant **n**- for the numerals '4' and '5'. For example, in Nzema:

*na* '4', *nu* '5'. In the other half of the sources forms with **n**- can be found for '4' and with initial **t**- for '5'; for example, in Gbe-Fon: *e-ne* '4', *a-ton* '5'. The latter forms correspond to Proto-Bantu numerals: *\*nài* '4', *\*táànò* '5'. The question then arises: where do the forms for '5' with initial **n**- come from?

Mary Esther Kropp Dakubu (**KroppDakubu2012**) includes the forms of the numeral '4' in the series of correspondences which go back to *\*n*- and reflect as **n**- in all of the main branches of the family except for Ga-Dangme (GD): Proto-Potou-Tano *\*-nã*, Tano *\*-nã*, GTM (Ghana–Togo Mountain) *\*-inâ*, Gbe *e-ne*. The author includes the numeral '5' in the series 15b where Akan and GD both have **n**-, in Gbe **t**-, and inside GTM are both **t**- and **n**- (Na-Togo). Mary Esther Kropp Dakubu suggests the following historical interpretation of these forms:

"The fact that GTM is reconstructed with *\*t*-, but its NA sub-group with *\*n*, suggests that the **n** of Akan and GD are also secondary, and that these forms are to be reconstructed as beginning in Kwa *\*t*" (ibid., p.24).

All the details of complex reconstruction will not be discussed here, but this shows that modern Kwa languages come from \*PTB (Proto-Potou-Tano-Bantu). It is worth underlining that the reported reconstruction does not explain why in some of the Kwa languages the numeral '5' with initial *\*t*- has changed to **n**-. Furthermore, she does not explain why this irregular change has happened in the aforementioned languages and not in the others.

The most natural answer to the first question is that in some languages, in the numeral '5' the initial consonant has undergone analogical change with the numeral '4'. As a result, the same consonant was formed in both numerals.

In order to answer the second question, it is necessary to observe the distribution of forms of '4' and '5' in different branches of Kwa, adding up in case of necessity forms for '3' and '2'. In order to extend the analysis of Mary Esther Kropp Dakubu, the Lagoon languages will be added to her database (Table 2.9).

Table 2.9: Akan

| Languages | '2' | '3' | '4' | '5' |
|---|---|---|---|---|
| Akan_Twi | abie-n | abie-sa | anan | anum |
| Ashanti | mie-nũ | mie-sã | enãn | enũm |
| Abron 1 | mie-nu | mie-sa | nain | num |
| Abron 2 | mie-nu**k** | mie-nza**k** | n-nai | n-num |

In all the Akan languages the alignment can be observed not only in '4'-'5' but (probably morphologically) also in numerals '2'-'3' (this phenomenon cannot be

found outside this cluster). Furthermore, one of the sources clearly indicates a final velar in Abron. Table 2.10 reports data on the main languages of Central Tano.

Table 2.10: Central Tano

| Language | '2' | '3' | '4' | '5' |
|---|---|---|---|---|
| Agni(Anyin) | ɲ́-ɲua | n-sa | n-na | n-nu |
| Baule | nɲo | sa | na | nũ |
| Nzema[2] | ɲ-ɲu | n-sa | n-na | n-nu |
| Anufo | ŋɲo | nza | na | nu |
| Baule (Baoulé)[3] | nɲon | san | nan | nun |
| Ahanta[4] | ayin | asan | anla | enlu |

Nearly identical forms are found in the other three branches of Tano (Table 2.11):

Table 2.11: Krobu-Ega, Western Tano, Tano Guang

| Branch | Language | '3' | '4' | '5' |
|---|---|---|---|---|
| Tano: Krobu-Ega | Krobu | n-sa | n-na | n-nu |
| Tano West | Abure | ŋ-ŋa | n-na | n-nu |
| Tano West | Eotile (Beti) | a-ha | a-ni | a-nu |
| Tano Guang | Dwang (Bekye)[5] | a-sa | a-na | a-nu |
| Tano Guang | Ginyanga | i-sa | i-na | i-noun |
| Tano Guang | Foodo | sa | naŋ | nu / nuŋ |
| Tano Guang | Larteh | sa | ne | nu |
| Tano Guang | Cherepon | i-sa | i-ne | i-ni |

---

[2]One of the sources on Nzema gives forms without an initial nasal: *sa* '3', *da* '4', *du* '5'. Let us note that even in this case the initial consonant is the same in the numerals '4' and '5'.

[3]In some sources Baule numerals '2'-'5' include also a final **-n**.

[4]Thus, in Ahanta the alignment of initial consonants for '4'-'5' is even more clear: **nl-**.

[5]The roots *-na* and *-nu* (for '4' and '5' respectively) can also be found in the Guang group in Awutu, Chumburung, Guang, Kplang, Krache, Nawuri, Nchumburu, Nkonya. For the subsequent exposition it is important that in all these languages the numeral '3' includes an initial **s-**.

Among the numerous Tano languages there is just one language in our database which does not have initial **n-** in '4' and '5'. This language is Ega, which is misleadingly put in the sub-group with Krobu; its attribution to Tano is also doubtful, according to the majority of specialists. The forms of these numerals provide one more argument against this grouping.

Some other languages display unification of the initial consonant in '4'-'5' outside of the Tano group.

As for Potou, forms with the initial **n-** in both '4' and '5': *ne-ni* '4', *ne-na* '5' were found only in Mbato, see Table 2.12.

Table 2.12: Potou

| Language | '3' | '4' | '5' |
|---|---|---|---|
| Mbato | ne-je | ne-ni | ne-na |
| Ebrie | bwa-dya | bwe-di | mwa-na |

Examples from Mbato permit us to reconstruct the unification of the initial consonant in '4'-'5' in Potou-Tano. Outside of Potou-Tano this unification, following Mary Esther Kropp Dakubu, was found only in some languages of Na-Togo (GTM). The numerals in the languages of this group are represented in Table 2.13.

Table 2.13: Na-Togo

| | Language | '3' | '4' | '5' |
|---|---|---|---|---|
| (1) | Anii | i-riu | i-naŋ | i-nuŋ |
| (2) | Logba | i-ta | i-na | i-nu |
| (3) | Selee | o-tie | o-na | o-no |
| (4) | Sekpele | cye | na | no |
| (5) | Lelemi | i-ti | i-ne | i-lo |
| (6) | Siwu (Akpafu) | i-te | i-na | i-ru |
| (7) | Adele | a-si | i-na | ton |

In languages (1-4) **n-** appears in '4'–'5' (Anii displays an utmost variant of alignment with the unification of the final consonant as well). In language (7) the most ancient proto-language initial **t-** is attested in '5', and this means that a reconstruction of *\*n-* in '5' for Proto-Na-Togo is problematic. Furthermore, in

languages (5-6) there is no alignment of the forms.

In other Kwa languages consonants in '4' and '5' differ. To be more precise, in Adjoukrou initial consonants are aligned but they are not nasals: jar '4', jen '5'.

All the other forms can be grouped into four main types:

1. the "basic" type, where, as in Bantu-Kwa, there is **n**- in '4' and **t**- in '5';

2. the type where '4' has initial **n**- while '5' shows a phonetic change of the initial consonant;

3. the type where '5' keeps **t**-, while '4' shows a phonetic deviation;

4. the most complicated type for the analogical interpretation which has **n**- only in '5' while '4' has a non-nasal initial consonant.

I will provide some examples followed by interpretations.

Type 1 is illustrated in (Table 2.14):

Table 2.14: n- '4', t- '5' (t- '3')

| Group | Language | '3' | '4' | '5' |
|---|---|---|---|---|
| Gbe | Aja | e-to | e-ne | a-to |
| Gbe | Ewe | e-to | e-ne | a-to |
| Gbe | Gen | e-to | e-ni | a-to |
| Gbe | Fon | a-to | e-ne | a-to |
| Gbe | Kotafon | a-to | e-ni | a-to |
| Gbe | Saxwe | a-to | i-ne | a-tu |
| Gbe | Xwla | a-to | e-ne | a-to |
| GTM | Kebu | ta | nia | to |
| Ga-Dangme | Dangme | e-to | e-ne | a-to |
| Ka-Togo | Akebu | ta | nie | tu |
| Ka-Togo | Ikposo-Uwi | i-la | i-na | i-tu |
| Na-Togo | Adele | a-si | i-na | ton |

It is clear that the basic etymological forms are represented extensively. They are not confined to Potou-Tano or the Lagoon languages but they can be found in four other branches of Kwa as well.

Type 2 is illustrated in (Table 2.15):

Table 2.15: n- '4', phonetic deviations in '5'

| Group | Language | '3' | '4' | '5' |
|-------|----------|-----|-----|-----|
| Ka-Togo | Avatime | o-ta | o-ne | o-cu |
| Ka-Togo | Tuwuli[6] | ɛ-lalɛ | ɛ-na | e-lo |
| Na-Togo | Lelemi | i-ti | i-ne | i-lo |
| Na-Togo | Siwu (Akpafu ) | it-e | i-na | i-ru |
| Lagoon | Avikam | a-za | a-na | a-ɲu |

Type 2, like Type 1, is not difficult to interpret. In the single languages the reflexes of the original consonant are maintained in '4', while in '5' *t- undergoes phonetic changes.

Type 3 is illustrated in (Table 2.16):

Table 2.16: t- '5', phonetic deviations in '4'

| Group | Language | '3' | '4' | '5' |
|-------|----------|-----|-----|-----|
| Ka-Togo | Igo (Ahlon) | i-ta | a-la | u-to |
| Ka-Togo | Nyangbo | e-tae | e-le | e-tie |

The proto-language consonant is maintained in only two languages in '5' (Ka-Togo and GTM) while the initial consonant in '4' undergoes regular phonetic change.

And finally, the most difficult type 4 is illustrated in (Table 2.17).

Here we see all the counter-examples against the hypothesis on the change *t- > n- in '5' as analogous to n- in '4'. The solution is to imagine that in certain languages belonging to different branches of Kwa (independently from each other), firstly, this analogical change occurred, the original *n-, which was the basis of the analogical change, but was then lost in the numeral '4'.

Finally, let us get back to the question raised above: why does analogical change in '5' take place in only some Kwa languages? Let us have a look at Table 2.18, where different initial root consonants in numerals '3'-'5' within different groups of Kwa are presented.

---

[6] **Harley2005** "With the exception of mɔa – 'one' and nviã – 'two', the citation forms of these numerals are derived using the expletive third person pronoun ke, which has become incorporated into the attributive numeral : ke ɛlalɛ '3' > kaalɛ, ke ɛna '4' > kɛna …".

Table 2.17: n- in '5' but not in '4'

| Group | Language | '3' | '4' | '5' |
|---|---|---|---|---|
| Potou | Ebrie | bwa-dya | bwe-di | mwa-na |
| Potou | Gã | e-tẽ | e-jwe | e-nũmo |
| Lagoon | Abé(Abbey) | a-ri | a-le | u-ni |
| Lagoon | Abiji | e ˈ-ti | a ˈ-la | e ˈ-ne |
| Nyo? | Ari (Abiji) | e-ti | a-la | e-ni |
| Central Tano | Ahanta | a-sa | a-la | e-nũ |
| Ga-Dangme | Dangme | e-te | e-ywi / e-wi | e-nuo |
| Lagoon | Alladian | a-o | a-zo | e-nri |
| Lagoon | Adioukrou | ɲa-hn | ya-r | ye-n |

Table 2.18: Kwa initial consonants in '3'-'5'

| Group Sub-Group | Bantu-Kwa | Tano Krobu | Tano Central Tano | Tano Akan | Tano Guang | Gbe Gbe | GD Gan-Dangme | GTM Ka-Togo |
|---|---|---|---|---|---|---|---|---|
| '3' | *t | s | s | s | s | t | t | t |
| '4' | *n | n | n | n | n | n | j/y | n |
| '5' | *t | n | n | n | n | t | t | t |

In the Kwa languages we see a clear tendency: in languages with the initial plosive *t- > fricative s-, the described analogical changes can be found. Where the plosive is maintained, this change is more difficult and can be found in only some of the languages (for example, some of the above-mentioned Na-Togo cases). In this case we have not *t- > n- '5', but *t- > s- > n. This observation can be interesting as a candidate for analogical changes – maybe, 'weak' consonants (for example, fricatives) can be more easily involved in analogical processes than 'strong' ones (plosives).

It is curious that this analogical isogloss can be found in a number of other branches of Niger-Congo, including Adamawa, Gur and Dogon (as well as Seenku from the Mande family).

## 2.5 Adamawa

In Adamawa the above-mentioned analogical change can be found in at least a dozen of languages (Table 2.19):

Table 2.19: Initial n- in '4'-'5' in Adamawa languages

| Language | '2' | '3' | '4' | '5' | '6' | '7' |
|---|---|---|---|---|---|---|
| Tula | rop | ta | na | nu | | |
| Kwa | negbe | ne mwan | ne nat | ne nu | | |
| Burak | rab | gbunuŋ | net | nob | | |
| Chamba | bara | te-ra- | nasa | tu-na- | | |
| Kolbila | inu | tonu | nereb | nunub | | |
| Bangunji | yob | **tar** | **nar** | **nu**ŋ | | |
| Yendang | ini | **tat** | **nat** | ghi-**n**an | | |
| Dadiya | yo | **tal** | **nal** | nu | | |
| Peere | i**ro** | ta**ro** | **naro** | **n**uno | | |
| Samba Leko | ki**ra**~ki**re** | tu**re** | **nara** | **n**unak | | |
| Gimme | idtige | tage | nage | nonige | nonge | nokidtige |

However, in Adamawa, analogies are much more widespread than in Kwa. For instance, in Gimme the numerals '2'-'7' share the same final syllable (morpheme?). In Chamba, only one similarity can be found for '4'-'5' and for '2'-'3' (the final syllable -**ra**). In Kolbila, the situation is quite similar to the one in Chamba ('2'-'3' share the same final syllable -**nu**) and in '4'-'5' both the initial **n**- and the final -**b** coincide.

Phonetic alignment follows more interesting models in Bangunji, Yendang, Dadiya, Peere and Samba Leko. In these languages, on the one hand, '4'-'5' are still grouped together (because of the initial consonant) and, on the other hand, ('2')-'3'-'4' are also grouped (because of the final syllable). The numerals with the meaning '4' have two simultaneously distinct features which mark two separate groupings. As a result, peculiar minimal pairs arise formed by contiguous numerals; for example, in Yendang: *tat – nat* '3'-'4', *nat – nan* '4'-'5'.

Another alignment of numerals (2), '3'-'4' takes place in Adamawa where there is no alignment in numerals '4'-'5'. Minimal pairs like in Dirrim *bara* '2' – *tara* '3' – *nara* '4' are a very widespread phenomenon for the languages within this family. Some examples are presented in the following table (Table 2.20):

This kind of assonance may seem insignificant, but I would like to underline once more that among hundreds of Benue-Congo languages, it is impossible to find any similar case.

Table 2.20: Adamawa analogical alignments in '3'-'4'

| Language | '1' | '2' | '3' | '4' |
|---|---|---|---|---|
| Vere (Mom Jango) | | ituko | tariko | nariko |
| Galke (Ndai) | | | ca-?a- | na?a |
| Dama | | | sa-i | nai |
| Mono | | | sai | nai |
| Mundang | | | sa-i | nai |
| Pam | | | sa-i | nai |
| Fali | | | tan | nan |
| Kam | | | car | nar |
| Bali | | | tat | nat |
| Kumba | | | sat | nat |
| Teme | | | tat | nat |
| Waka | | | tat | nat |
| Yendang | | | tat | nat |
| Wom | | ira | tara | nara |
| Taram | | bara | tara | nara |
| Fanya | | liru | taro | naro |
| Duupa | | ito | tato | nato |
| Kotopo | wate | i-to | tato | nato |
| Mom Jango | muzoz | itez | taz | naz |

## 2.6  Ubangi

Ives Moñino (1995) has reconstructed unified forms for '3'-'4' and partly for '5' in Proto-Gbaya. These forms resemble the above-mentioned "minimal pairs" in Adamawa. In Proto-Gbaya: *tar(a)* '3', *nar(a)* '4', *mor* '5' (notably, the numeral '5' coincides with the word 'hand'). In Ubangi-Sere, a different type of alignment can be found – the final **-o** in numerals '2'-'5' (in Ubangi-Zande – the final **-i**) (Table 2.21):

## 2.7  Gur

In some languages of the Gur family analogical changes in '4'-'5' can be found, as observed in Kwa and Adamawa (Table 2.21):

Table 2.21: Final vowel alignments in Ubangi

| Language | '2' | '3' | '4' | '5' |
|----------|------|--------|---------|----------|
| Ndogo | so | tao | nao | vo |
| Sere | so | tao | nao | vo |
| Tagbu | so | tao | nao | vuo |
| Pambia | a-vai | wa-tai | (h)avai | boinyaci |

Table 2.22: Gur initial n- in '4'-'5'

| Language | '2' | '3' | '4' | '5' |
|----------|-------|-------------|---------|-------|
| Baatonum | yiru | ita / yita | ne | nobu |
| Chala (dial.) | -la | -toro | -nara | -nuŋ |
| Buli | ba-yi | ba-ta | ba-nasi | ba-nu |
| Dagaara | ayi | ata | a-nar | a-nu |
| Delo | ala | atoro | a-nara | a-noŋ |
| Ditammari | deni | tati / tadi | na | numu |
| Nawdm | mrek | mtak | m-na | m-nu |
| Safaliba | ayik | atak | anaasi | anu |

Like in Chamba (Adamawa), some of the Gur languages have a common feature not only for '4'-'5' but also for '2'-'3'. For instance, in Nawdm and Safaliba, as can be deduced from Table 2.22, the numerals '2'-'3' have a final velar consonant. The final velar can be found in '2'-'3' in Hanga (*a-yik* '2', *a-tak* '3'), and in Dogose it is found in '2'-'5': *i-yok* '2', *i-sak* '3', *i-yị̆k* '4', *i-wak* '5'. Gudrun Miehe (Miehe et al. 2007: 157) shows in Khisa (Komono) the final -ʔ in '2'-'5': *ɲɔ́ɔ̀ʔ* '2', *sặaʔ* '3', *ɲéèʔ* '4', *ŋwáàʔ* '5'.

And finally I would like to report a rare case of strong alignment between the numerals '1' and '2' in Mbelime: *yɛ̃nde* '1', *yede* '2'.

## 2.8 Dogon

Assimilation of the initial consonant in '5' to the initial consonant **n-** in '4' (for example, Tommo So: *nay* '4', *no* '5') is characteristic of practically all the Dogon languages and should be reconstructed already for the Proto-Dogon. Other types

of unification cannot be found in this family.

## 2.9 Kordofanian

Phonetic / morphological alignments in this family are quite rare. In what follows, the most interesting cases are reported (Table 2.23):

Table 2.23: Kordofanian alignments

| Group | Languages | '1' | '2' | '3' | '4' | '5' |
|---|---|---|---|---|---|---|
| Kordofanian-Talodi | Tocho | pulu**k** | we-ra**k** | wa-ta**k** | | |
| Kordofanian-Talodi | Jomang | y-ílli**k** | y-ilra**k** | y-ida**k** | | |
| Kordofanian-Talodi | Nding | | -eta | t-ata**k** | -ibiɲi**k** | |
| Kordofanian-Talodi | Tegem | tléedi | **pad**erig | **pad**aig | | |
| Kordofanian-Katla | Katla | te:tá**k** | sek | | | |
| Kordofanian-Orig | Orig | | | | aru**m** | wura**m** |
| Kordofanian-Katla | Tima | | **eh**ek | **eh**oat | **eh**alam | |

In Talodi the final velar is present, similarly to other branches of Niger-Congo. Some cases of phonetic alignment can be found, though this alignment is reserved to singular languages rather than to the whole family.

In sum, the data examined in this chapter can be found in Appendix 3 where 50 different cases of probable analogical changes in Niger-Congo are highlighted. The Table in Appendix 3 permits the evaluation of the scale of analogical changes in the system of numerals in Niger-Congo in general.

It is worth mentioning that in the cases where numerals '6'-'10' are not derived, it is very unusual to find phonetic alignment in them (exceptional systems, such as that of Soninke, were previously discussed). For this reason, only the numerals '1'-'5' are included in Appendix C. Three main questions are to be answered concerning these numerals: 1) Which groupings of numerals are most typical for the Niger-Congo languages when we deal with analogical changes? 2) Which phonetic (or hidden morphological) means are used to produce the alignment of numerals? 3) Are there any reasons to consider that similar analogical changes in different branches of Niger-Congo can be diachronically related? Otherwise, can these materials be useful for the study of other isoglosses in Niger-Congo?

As demonstrated in Appendix 3, mostly contiguous numerals are aligned (see some rare examples above, for example in Nyun languages, where features for '1'–'2'/'4' are shared, but not for '3').

It is quite rare that '1' shares a submorphemic marker with the numeral '2', while for other contiguous numerals this is more common. Such rare examples are found in Ha (Bantu J) and in Mbelime (Gur). In both languages the forms of numerals '1' and '2' have minimal phonetic difference. As will be demonstrated in the following sections dealing with the etymology of numerals '1' and '2', the forms in Ha (*mbele* '1', *bhili* '2') are of great interest for the diachronic interpretation of numerals.

As can be seen in Appendix 3, the final phonemes have phonetic alignment much more often than the initial ones.

The appearance of the diachronically irregular initial **n**- in the numeral '5' as analogous to the regular form of the numeral '4' represents a common feature in different families of Niger-Congo: Potou-Tano (Kwa), Adamawa, Gur and Dogon. More attention should be paid to this phenomenon because it is unlikely that one analogical feature could appear in four different branches of Niger-Congo independently.

There are two remarkable cases in the alignment of final phonemes which are typical for several branches of Niger-Congo.

Firstly, there is the appearance of a final velar (-**k**) in the groupings of the numerals '2'-'5', '2'-'4', '2'-'3', '3'-'4' (in Kordofanian and Atlantic also '1'-'2'-('3')). This feature is typical for the Atlantic, Adamawa, Gur and Kordofanian groups (thus, one more common feature can be found for Adamawa-Gur). In Benue-Congo and Mande the reported examples are clearly marginal.

Secondly, similarly to the regular dental reflexes of the final consonant in the numeral '3' (*-**t(h)**), in '4' the final consonant undergoes an irregular change (non dental consonant becomes dental). This type of change is particularly characteristic for Atlantic, Adamawa and Gbaya (Ubangi), but it is also found in Kordofanian and in Benue-Congo, which do not have analogic changes as characteristic features.

The most common case is the appearance of the identical final vowel in some languages of different families (mostly in numerals '2'-'5'): Mama (Bantoid), Soninke (Mande), Peere (Adamawa) and Ndogo, Pambia (Ubangi).

All the reported cases should be taken into consideration for the process of etymologization of numerals, which will be done in the following chapter.

# 3 Step-by-step reconstruction of numerals in the branches of Niger-Congo

In this chapter we will try to create a step-by-step reconstruction of numeral systems for each separate family independent of the data from the other NC families. For each family we shall examine the range of basic numerals from '1' to '10' and then the numerals for '20', '100' and '1000'. We begin our overview with the largest family, Benue-Congo.

## 3.1 Benue-Congo

There is no Benue-Congo classification that is accepted by all scholars. As noted, the inventory of Benue-Congo groups mainly follows the classification of Kay Williamson (1989a: 266-269). We repeat here the scheme of BC given above, in the introduction:

Table 3.1: Benue-Congo languages

| *Western BC | *Eastern BC |
|:-----------:|:-----------:|
| Nupoid | Kainji |
| Defoid | Platoid |
| Edoid | Cross |
| Igboid | Jukunoid |
| Idomoid | **Bantoid** |
| Isolated BC: Oko, Akpes, Ikaan, Lufu | |

Let us begin our overview with the largest group of Bantoid languages.

### 3.1.1 The Bantoid languages (including Bantu)

The reconstruction of numerals in the Bantoid languages is based on 140 sources for the major branches of this family. What follows is the result of our step-by-step analysis of numeral systems in these languages.

#### 3.1.1.1 'One'

We shall collect the main forms for '1' in different branches of the Bantoid languages. The last column of Table 3.2. shows some isolated forms for '1' which seem to be innovations.

At first glance, the terms for '1' in the majority of the Bantoid languages appear to be quite homogeneous, their roots being traceable to either *\*moʔ* or *\*moi/mwi* of uncertain etymology. The misleading similarity of the Bantu roots *mòì, mòdì, mòtí* may be due to the merger of the noun class prefix **\*mò̀-** with the nominal base[1]. This hypothesis (developed in detail in **Vanhoudt1994**) has now found its way into the BLR (cf. BLR3 *sub mòdì* (NC): ʿ*plutôt mò̀-òdì: voir* **Vanhoudt1994** ʾ).

Among other common Bantu forms are *mócà* (zones KN), *mòtí* (ABCEGHKLRS) < *\*mò̀-òtì*, *mǒégá* (zones BH) (BLR3: «*mòì* + suffix»), and *mòì* (ABCDEFGJKLMRS). As will be shown below, the presence of a nasal prefix in the Bantoid numerals is suggested by the distribution of these forms in Benue-Congo. Those BC branches that have nasalless roots within the nominal classes 'one' and 'three' lack the terms for 'one' with a nasal consonant.

This interpretation, however, does not address two major issues, namely 1) whether the forms in question (e.g. *\* -òdì/ -oti/ -oʔi*[2]) consist of one or more roots and 2) whether the open back vowel belongs to the root.

A solution to the former problem may turn out to depend on how the latter is treated.

Within the context of Niger-Congo, it is conceivable that the Proto-Bantu *òdì* may go back to *\*ò-dì*, with *\****ò̀-*** being a marker of the NC noun class 1 (*\****ko-/ ʔo-*** according to my reconstruction). This hypothesis will receive a more detailed treatment in the next chapter. At this point, we will only note that it is quite problematic to explain the common reflexes of *\****-di***, *\****ti***, and *\****ʔ-*** in Bantu within this hypothesis. Moreover, the etymological relationship between these roots (disregarding *\*di* and *mɔ(m)* (Tivoid), *ó-mè* (Mbe), *ma* (Mamfe), etc.) would be much

---

[1]I agree with Larry Hyman who reacted to this point: "This would suggest that '1' was a noun; possible, just like '10', but note that '2'–'5' are not nouns!" (p.c.).

[2]Larry Hyman: "The glottal stop goes back to a velar in Grassfields; it could be either alveolar or velar in Tikar".( p.c.).

Table 3.2: Bantoid stems for '1'

| Branch | Language | '1' | '1' | '1' |
|---|---|---|---|---|
| **Northern** | | | | |
| Dakoid | Chamba-Daka | | | nòòní |
| *Mambiloid | | mwi | cin, jer | |
| Fam | | | | wuni[a] |
| Tiba (Fà) | | | à-kīn-á | |
| **Southern** | | | | |
| **\*Bantu[b]** | | **mòì/mòdì, mòtí** | | **p/m/b-ókó** |
| *Beboid | | mwi/mu | | baka, kpaŋ |
| *Yemne-Kimbi | | mwe | | |
| *Ekoid | | | ji(ŋ) / rəŋ? | yet?[c] |
| *Jarawan | | mɔʔ | | (ɗik) |
| *Mamfe | | mɔt / ma | | |
| *Mbam | | mwe/mùʔ | | |
| Mbe | Mbe | ó-mè | | |
| Ndemli | Ndemli | mɔ̀hɔ́ | | |
| Tikar | Tikar | mbɔʔ | | |
| *Tivoid | | mɔ(m) | | |
| *Esimbi | | | | nə |
| Wide Grassfields | Befang | moʔ | | |
| GF: Mbam-Nkam | Bamileke | moʔ | | cu |
| GF: Mbam-Nkam | Ngemba | mɔʔɔ | | |
| GF: Mbam-Nkam | Nkambe | moʔ(sír) | | |
| GF: Mbam-Nkam | Nun | moʔ | | |
| GF: Momo | | moʔ | | fiŋ |
| GF: Ring | | moʔ | | |

---

[a]The Fam and Tiba (Fà) forms are quoted according to **Blenchndc**) and **Boyd1999** respectively. The online version of Boyd (https://hal.archives-ouvertes.fr/hal-00323718v3) differs from the printed one.

[b]An asterisk (*) in the second column of the tables (here and below) means that in the corresponding line all the forms are reconstructed. However, with the exception of the Proto-Bantu line, which indicates real reconstructions in BLR3 (*), all other reconstructions are hypothetical (#) and reflect the most typical form / forms attested in a particular branch of Benue-Congo. Forms that may be related are grouped in tables within the columns. The last column of the tables shows isolated forms that are likely to be innovations.

[c]Concerning the form *yet* in Ekoid, I quote a precious remark of John Watters (p.c.): ''The actual root for Proto-Ekoid may be -**t** ~-**d**. The /aŋ/ in some Ekoid languages may be an accretion. The *yét* morphologically is /yé-t/ with the CV being a class agreement prefix, and -**t** being the root. So the -**t** may be closer to the Bantu *moti*. I'm not sure how *ó-mè* in Mbe figures in with the rest of Ekoid, but one possibility is that the -**mè** root derives from /me-t/. Ekoid needs further work''.

less transparent than that in case of *modi ~ moti* or even *-odi ~ -oti*.

The secondary PB form *\*ókó* (zones ABCHF) (BLR3: ''**Janssens1994**: alternance C1 p/m/b-ókó- protoforme secondaire, cf. 'seul'") is comparable to *\*baka* (Beboid: Fio *mbákâ ~ nbáhá*, Nchane (Mungong) *m⁴ba³ka⁴*). It should be noted that the above considerations allow us to explain the initial consonant (and the following back vowel) in these forms as noun class morphemes, too.

The Northern Bantoid *kin/cin* is remarkable and will be addressed later in this chapter.

The Bamileke *\*tʃu* (Fefe *ʃɯʔ*, Medumba a*ntʃʊʔ*, Nda'nda' *ŋtʃɔʔ*, etc.) is possibly related to the Bantu *\*tʊ* (BCDEGLP) 'alone, empty, vain'.

### 3.1.1.2 'Two' and 'Three'

Table 3.3: Bantoid stems for '2' and '3'

|  | **Language** | **'2'** | **'2'** | **'3'** |
|---|---|---|---|---|
| Northern | | | | |
| Dakoid | Chamba-Daka | | bààrá | tárā |
| *Mambiloid | | fee/ fal/hal | baa | taar |
| Fam | | | baale | tawnə |
| Tiba (Fà) | | | à-ɓə̄ə̄r-á | à-tár-á |
| Southern | | | | |
| **\*Bantu** | | | **bàdí / bìdí** | **tátò / cátò** |
| *Beboid | | fe | | tat, te |
| *Yemne-Kimbi | | fi(n) | | to |
| *Ekoid | | | ba(l) | sa/ ra |
| *Jarawan | | | ɓar | tat |
| *Mamfe | | pay / pea | | rat / lɛ |
| *Mbam | | fande? | bante? | tat |
| Mbe | Mbe | pʷâl | | sá |
| Ndemli | Ndemli | ifέ | | ítáá |
| Tikar | Tikar | | ɓî | lê |
| *Tivoid | | hal/ har/ vial | | tat |
| *Esimbi | | ra-kpə? | | kələ (< *lə?) |
| Wide Grassfields | Befang | fe | | táí |
| GF: Mbam-Nkam | Bamileke | pu/ pwe | bo/ bie | tat |
| GF: Mbam-Nkam | Ngemba | paa | baa / bəgə | tarə |
| GF: Mbam-Nkam | Nkambe | | baa | tar |
| GF: Mbam-Nkam | Nun | paa | baa | tɛt |
| GF: Momo | | | be | tat |
| GF: Ring | | | bo/ ba | tat |

Without exception, the reconstructed root for 'two' in all Bantoid branches has an initial labial consonant, either voiced (b-) or voiceless (p-/f-). A more precise reconstruction of the proto-form is beyond my cognizance. The forms cited above do not permit a conclusion with regard to the number of roots involved (one or two). When comparing the most commonly attested forms *pa/ fe and *baa, it

is necessary to keep in mind that at least the Proto-Bantu *bàdí/bìdí* could be a reflex of *di*. In the case of **ba-** the proto-form should be interpreted as a prefix of a plural noun class (possibly class 2).[3] The latter proposal finds support in the dialectal Proto-Bantu form *jòdè* (zones BH) (< *jò-dè*?). The main forms show the following zonal distribution: *bàdí* (ABCHKLR), *bìdí* (CDEFGJKLMNPS), *bídì* (?).

It was repeatedly stressed that the root for 'three' (*tat*) is one of the most stable in NC and in the Bantoid languages in particular. Phonetic variation within this root will be studied in Chapter 4.

---

[3]John Watters: "This analysis, if correct, could work also for most of Bantoid. So Ekoid would derive from **ba-** prefix and -**l** ~ -**d** ~ -**n** root. However, the /b/ may derive from /p/. Ekoid may derive from *-pal* and then you have the many other Bantoid languages with /p/" (p.c.).

### 3.1.1.3 'Four' and 'Five'

Table 3.4: Bantoid stems for '4' and '5'

|  |  | '4' | '4' | '5' | '5' |
|---|---|---|---|---|---|
| **Northern** |  |  |  |  |  |
| Dakoid | Chamba-Daka | nàà-sá |  | túùná |  |
| *Mambiloid |  | na(n) |  | tien/ tin/ con/ son | ngii? |
| Fam |  |  | daare | tʃwiine |  |
| Tiba (Fà) |  | à-nè-á |  | à-tǫ̀ǫ̀ŋ-á, tūùŋ |  |
| **Bantu** |  | **nàì / (nàí)** |  | **táànò / cáànò** |  |
| **Southern** |  |  |  |  |  |
| *Beboid |  | na, ne |  | ti(n) |  |
| *Yemne-Kimbi |  | ni |  |  | kpɔn |
| *Ekoid |  | ni |  |  | don/ lon[4] |
| *Jarawan |  | yi-ne? |  | towun/ twan |  |
| *Mamfe |  | n(w)i |  | ta(y) |  |
| *Mbam |  | ni(s) |  | taan |  |
| Mbe | Mbe | ñî |  | tʃân |  |
| Ndemli | Ndemli |  | itʃìjè | ítâŋ |  |
| Tikar | Tikar | ɲî |  | ʃæ̃ |  |
| *Tivoid |  | ɲi(n) |  | tan |  |
| *Esimbi |  | ɲi |  | tənə |  |
| Wide Grassfields | Befang |  | kʷà (kɥà) | ìtʲⁿ |  |
| GF: Mbam-Nkam | Bamileke |  | kwa/ kwo | tan |  |
| GF: Mbam-Nkam | Ngemba |  | kwa/ kya | taa(n) |  |
| GF: Mbam-Nkam | Nkambe |  | kwe/ kye | tan / ton |  |
| GF: Mbam-Nkam | Nun |  | kwa/ kpa | tan / tɛn |  |
| GF: Momo |  |  | kwe | tan |  |
| GF: Ring |  |  | kwi / kye / tsə | tan |  |

The well-known NC root *nai* 'four' is represented in all of the pertinent languages. The only exception is Grassfields, where it was replaced with the innovative *kwa/kya*. According to Roger Blench, Momo *-kpi* and Ring *kaìkò* as well as the Proto-Eastern Grassfields *-kùa* go back to the Proto-Benue-Congo #-kpà(ko) (**Blench2004**: #387). This root, however, is commonly found in Mbam-Nkam, i.e. in all Grassfields languages, and is barely attested outside this branch.

The root for 'five' is almost invariably *tan*. One possible exception is the Ekoid form, unless *don/ron/lon* (Ekajuk *nlɔn*, Ejagham *érôn*, Nkem-Nkum *írɔ̂n*) is a reflex of *tan*.

---

[4]John Watters: the Proto-Ekoid probably is *-ron (p.c.).

It should be noted that the Ndemli root *itʃìjè* may be related to *kwV* in the Grassfields languages. As we hope to demonstrate below, this is probably not a coincidence.

### 3.1.1.4 'Six'

The Grassfields languages show a common root *\*toʔo*. Outside Grassfields, it is attested only in Ndemli (just like the Grassfields root for 'five') and thus can hardly be reconstructed for Proto-Bantoid. However, we cannot exclude this, if PB *\*tʊ́ʊ́bá* '6' attested in zones ABCD is related to the Grassfields forms.

As in some other NC branches, three patterns that can be used to derive '6' from '3' are attested in the Bantoid languages (the following observations are even more relevant in the case of the patterns for 'eight' based on 'four'):

1. The change of a class prefix (or its addition): Ajumbu *tò* '3' > *kʲà-tò* '6'; this pattern is possibly attested in Tutomb (Mbam) *pɛ́-dààt* '3' > *pí-tʃín-dìt* '6', Elip *bʊ́-dáɖ* '3' > *bʊ́-thín-dàɖ* '6' (this pattern is marked '3PL' in the table above). To strengthen the etymology for 'six' in Tutomb, it should be noted that in Tunen (another Mbam language) that has *\*tat* '3' > *lal* (*bɛ́-lálɔ́*), the term for 'six' also contains [l]: *pɛ́-lɛ́ⁿdálɔ*.

2. The combination of 'three' and 'two': Lyive: *hʲâl* '2', *tàt* '3', *kɔ̀lɔ̀-kɔ̀-tàt* '6' (<'2*3'?).

3. The reduplication of 'three' (or the simple addition '3+3'): Ekajuk *n-ra* '3' >*n-ra-ke-ra* '6', Ejagham *é-sá* '3' > *è-sá-gà-sá* '6', Nkem-Nkum *i-ra* '3' > *i-ra-ra* '6', Mbe *bé-sá* '3' >*bè-sê-sár* '6', Tiv *ú-táŕ* '3' > *á-tér-á-táŕ* (this pattern is marked as '3+3' in the table above).

The Kenyang (Mamfe) form *bɛ́-tándât* '6' (cf. *bɛ́-rát* '3') deserves special discussion. This form is reminiscent of the common Bantu form *tándà* '6' attested in zones DGM. Its extended variant *tándàtʊ́* is found in EFGJS, while the GNS zones use the form *tántàtʊ́* which is even more interesting. Are the Bantu *tándà* forms cited above based on '3'? If so, *\*tat-tat > tatat* (*tántàtʊ́*) in the languages to which Dahl's law is applicable as well (> *tandat, tanda*).

In this case, the form *tʊ́ʊ́bá* (zones ABCD) that can be interpreted as '3*2': *\*tat-X-ba* may also be a derivative form.

If so, the aforementioned Bantu forms (as well as the Kenyang form) are probably not innovations. They may reflect a Proto-Bantoid model where 'six' is based

Table 3.5: Bantoid stems and patterns for '6'

| | | '6' | '6' | '6' | '6' |
|---|---|---|---|---|---|
| **Northern** | | | | | |
| Dakoid | Chamba-Daka | | | <5? | |
| *Mambiloid | | | | 5+1 | |
| Fam | | | | 5+1 | |
| Tiba (Fà) | | | | 5+1 | |
| **Southern** | | | | | |
| **\*Bantu** | | **tándà < 3redupl.?** | **tʊ́ʊ́bá** | | **càmb–, kaaga** |
| *Beboid | | | | | so |
| *Yemne-Kimbi | | 3PL? | | | |
| *Ekoid | | 3+3 | | | |
| *Jarawan | | | | 5+1 | |
| *Mamfe | | | | | kene? |
| *Mbam | | 3PL | | 5+1 | |
| Mbe | Mbe | 3+3 | | | |
| Ndemli | Ndemli | | tóhó | | |
| Tikar | Tikar | 3PL? | | | |
| *Tivoid | | 3redupl., 2*3? | | | |
| *Esimbi | | <3redupl.? | | | |
| Wide Grassfields | Befang | | ⁿdɔ̀fú | | |
| GF: Mbam-Nkam | Bamileke | | toɣo | | |
| GF: Mbam-Nkam | Ngemba | | toʔo | | |
| GF: Mbam-Nkam | Nkambe | | ntunfu | | |
| GF: Mbam-Nkam | Nun | | ntúwó/ tuʔo | | |
| GF: Momo | | | | | foɣ |
| GF: Ring | | | tufa | | |

on 'three'. It should be noted that a close parallel to the Kenyang form is attested in the Mbam branch: Nomaande *be-tíndétú* '6'.

In sum, it appears that the most probable word-formation pattern for 'six' in Proto-Bantoid is '3+3' or '3PL'.

### 3.1.1.5 'Seven'

The case of 'seven' seems pretty straightforward. In the majority of the Bantoid branches (including Bantu) the root is *\*samba/camba*. However, there is still a question whether this root is indeed primary: its Bantu reflex is strikingly similar to the root for 'six'. Here are some selected examples (Table 3.7):

It is noteworthy that the terms for 'six' and 'seven' show similarity not only in case of the root in question, but in case of other roots as well, e.g. J50: Fuliiru -*lindátù* '6'~ -*linda* '7', Shi *ńdarhu* '6'~ *ńda* '7'. This similarity is usually conditioned by one of the following factors:

- the terms for 'six' and 'seven' follow the patterns '10−4' and '10−3' respectively: Yeyi (Bantu R40) *vùndʒà ɛ́ nɛ́ɛ́* '6' ('10' 'break' '4 (fingers)'), *vùndʒà ɛ́ táâːtō* '7' ('10' 'break' '3 (fingers)'). This, however, is very rarely attested.

- the term for 'seven' is based on 'six' ('6+1'). This pattern is much more common (see Table 3.8).

- The similarity may also be due to the derivation of these terms from 'five' using '5+1' and '5+2' patterns, respectively (this is the most common case). It should be noted that there is another, much less transparent pattern for 'seven' ('X+2' or '5+X'). It is frequently attested not only in the Bantoid languages, but also in the Mande languages.

- Finally, we may be dealing with an alignment by analogy.

> maybe reference the relevant tables here

Staying within the Bantoid family, it is difficult to say which of these explanations should be applied in the present case. If it is alignment by analogy, we should reconstruct a Proto-Bantoid primary root *\*samba/camba* for 'seven' and then explain the many irregular shifts in the forms of 'six' (e.g. t > s) by analogy with this root (as shown above, the Proto-Bantu 'six' is based on 'three' (\*tat)). We may also be dealing with a derived proto-form *\*sam-ba/cam-ba* with the second element probably going back to 'two'.

Table 3.6: Bantoid stems and patterns for '7'

|  |  | '7' | '7' | '7' | '7' | '7' |
|---|---|---|---|---|---|---|
| **Northern** |  |  |  |  |  |  |
| Dakoid | Chamba-Daka |  |  |  |  | dùtím |
| *Mambiloid |  |  |  |  | 5+2 |  |
| Fam |  |  |  |  | 5+2 |  |
| Tiba (Fà) |  |  |  |  | 5+2 |  |
| **\*Bantu** |  | **càmbà-dì/càmbʊ̀-à-dì** | **6+1?** |  |  | **púngàtí** |
| **Southern** |  |  |  |  |  |  |
| *Beboid |  | fumba? | 6+1 | 4+3 |  |  |
| *Yemne-Kimbi |  |  |  | 4+3 |  |  |
| *Ekoid |  | sima? |  | 4+3? |  |  |
| *Jarawan |  |  |  |  | 5+2 |  |
| *Mamfe |  |  | 6+1 |  |  |  |
| *Mbam |  |  | 6+1 |  |  |  |
| Mbe | Mbe |  |  |  | 5+2 |  |
| Ndemli | Ndemli | sàᵐbá |  |  |  |  |
| Tikar | Tikar | ʃâmɓì |  |  |  |  |
| *Tivoid |  |  | '6+1 |  | 5+2 |  |
| *Esimbi |  |  |  |  | 5+2 |  |
| Wide Grassfields | Befang |  |  | 4+3 |  |  |
| GF: Mbam-Nkam | Bamileke | samba |  |  |  |  |
| GF: Mbam-Nkam | Ngemba | samba |  |  |  |  |
| GF: Mbam-Nkam | Nkambe | samba |  |  |  |  |
| GF: Mbam-Nkam | Nun | samba |  | 4+3 |  |  |
| GF: Momo |  | sambe |  |  |  |  |
| GF: Ring |  | samba |  |  |  |  |

Table 3.7: Similarities between '6' and '7' in Bantu

|  | '6' | '7' |
|---|---|---|
| PB | càmbànò (HL)/ cààmànò (ABCHLR)/ càmbombo (L) | càmbà-dì/càmbɵ̀-à-dì |
| A40 Bankon | bi-sámà | bi-sámbɔ̀k |
| A80 Kol | twáb | tábɛl |
| B20 Mbangwe | -syami | ntsaami |
| B60 Mbere | -syaami | ntsaami |
| B70 Teke-Tege | ósámìnì | ónsààmì |
| B80 Tiene | ísyam | nsam |
| C40 Sengele | ísama | ísambiálé |
| C90 Ndengese | isamo | isambé |

Table 3.8: Common stems for '6' and '7' in Bantu

|  | '6' | '7' |
|---|---|---|
| J50 Fuliiru | -lindátù | -linda |
| J50 Shi | ńdarhu | ńda |
| A80 Byep | tʷɔ́p | tʷɔ́p ɓɵ̀l (6+?) |
| C10 Yaka | βúè | βúè nà -mɔ̀tí (6+1) |
| D30 Budu | mèɗìà | mèɗìàníkà (lit: níkà 'to come') |
| M20 Malila | ʊ́mʊthaːⁿda | ʊ́mʊthaːⁿda na jěkʰa (6+1) |
| B10 Myene | òɾówá | òɾwáɣénô (6+1) |

Table 3.9: '6' and '7' from '5' in Bantu

|  | '6' | '7' |
|---|---|---|
| H10 Koongo | sàmbánù | sàmbú-wàlì (wálì '2') |
| K20 Nyemba | pàndù | pàndù vàlì (-vali '2') |
| K60 Mbala | sambanu | nsambwadi (mbadi '2') |
| L30 Luba-Katanga | isamba | isambaibindi (ibindi '2') |
| R10 Khumbi | epándú | epándúvalí (valí '2') |

### 3.1.1.6 'Eight'

Table 3.10: Bantoid stems and patterns for '8'

|  |  | '8' | '8' | '8' |
|---|---|---|---|---|
| **Northern** |  |  |  |  |
| Dakoid | Chamba-Daka |  |  | 7+1 |
| *Mambiloid |  |  |  | 5+3 |
| Fam |  |  |  | 5+3 |
| Tiba (Fà) |  |  |  | 5+3 |
| **Southern** |  |  |  |  |
| **\*Bantu** |  | **nainai(**4 redupl.**)/ nake** |  |  |
| *Beboid |  | ɲaŋ (<4?) |  |  |
| *Yemne-Kimbi |  | 4 redupl. |  |  |
| *Ekoid |  | 4+4 |  |  |
| *Jarawan |  |  |  | 5+3 |
| *Mamfe |  | 4PL |  |  |
| *Mbam |  | 4 redupl. |  |  |
| Mbe | Mbe | 4 redupl. |  |  |
| Ndemli | Ndemli |  | fɔ:mɔ́ |  |
| Tikar | Tikar |  |  |  |
| *Tivoid |  | 4 redupl. |  |  |
| *Esimbi |  | 4 redupl. |  |  |
| Wide Grassfields | Befang |  | éfómó |  |
| GF: Mbam-Nkam | Bamileke |  | fum/hum/fo? |  |
| GF: Mbam-Nkam | Ngemba |  | famə |  |
| GF: Mbam-Nkam | Nkambe |  | waami |  |
| GF: Mbam-Nkam | Nun |  | fame |  |
| GF: Momo |  |  | fami/ foŋ |  |
| GF: Ring |  |  | faamə |  |

Both Grassfields and Ndemli share the common primary root for 'nine' (*famV*). We have already seen this distribution, which only suggests that Ndemli belongs to the Grassfields branch (at least on the basis of their numeral systems). The majority of other branches point to the reconstruction of the term for 'eight' as based on 'four' (either by means of reduplication or by the noun class switch, or both).

### 3.1.1.7 'Nine'

Table 3.11: Bantoid stems and patterns for '9'

|  |  | '9' | '9' | '9' | '9' | '9' |
|---|---|---|---|---|---|---|
| **Northern** |  |  |  |  |  |  |
| Dakoid | Chamba-Daka |  |  |  |  | kúūm |
| *Mambiloid |  |  | 5+4 |  |  |  |
| Fam |  |  | 5+4 |  |  |  |
| Tiba (Fà) |  |  | 5+4 |  |  |  |
| **Southern** |  |  |  |  |  |  |
| **\*Bantu** |  | **bùá** | **5+4** |  | **10-1** | **kèndá/ jèndá** |
| *Beboid |  | bùkə? |  |  |  | fumḅɔ? |
| *Yemne-Kimbi |  |  | 5+4 |  |  |  |
| *Ekoid |  |  | 5+4 |  | 10-1 |  |
| *Jarawan |  |  | 5+4 |  |  |  |
| *Mamfe |  |  |  | 8+1 |  |  |
| *Mbam |  |  | 5+4 | 8+1 |  |  |
| Mbe | Mbe |  | 5+4 |  |  |  |
| Ndemli | Ndemli | bù?è |  |  |  |  |
| Tikar | Tikar |  | 5+4? |  |  |  |
| *Tivoid |  |  | 5+4 | 8+1 |  |  |
| *Esimbi |  |  | 5+4 |  |  |  |
| Wide Grassfields | Befang |  | 5+4 |  |  |  |
| GF: Mbam-Nkam | Bamileke | fu?u |  |  |  |  |
| GF: Mbam-Nkam | Ngemba | bu?u / pu?u |  |  |  |  |
| GF: Mbam-Nkam | Nkambe | bù?ὺ? búum? |  |  | 10-1? |  |
| GF: Mbam-Nkam | Nun | pu?u? |  |  |  | cipo? |
| GF: Momo |  | bok |  |  |  | ko? |
| GF: Ring |  |  |  |  | 10-1 |  |

It seems likely that there was a primary root for 'nine' in Proto-Bantoid. It can be tentatively reconstructed as *\*bukV*.[5] In Bantu, this root is found in the ABCDHL zones. The most common pattern '5+4' (as well as the less frequently attested '10–1') often develops independently in various languages. A marginal

---

[5]John Watters: ''Given the distribution of these forms for 'nine' I would conclude that Proto-Bantoid likely used 5+4 and that *\*bukV* was an innovation in the pre-Bantu era when Proto-Bantu had not yet separated from what became Grassfields and other closely located Bantoid groups''.

pattern '8+1', attested in Mamfe, Mbam and Tivoid is noteworthy. Because of its rarity, it is relevant for the genetic classification of the Bantu languages, since it is hard to imagine that this form developed independently in each of these branches. The last column of the table below lists bases that are exclusively found in a specific Bantoid branch.

### 3.1.1.8 'Ten'

At least two Bantoid roots (*\*fu* and *\*kum/ kam*) may be useful for our reconstruction purposes. Both of them are attested in no fewer than six of the Bantoid branches (note also the Chamba-Daka *kúūm* 'nine'). The Mambiloid languages show the greatest variety of roots.

It should be noted that a separate Proto-Bantoid form for 'ten' is not traceable in some of the pertinent languages. Despite this, it has been preserved as a part of the term for 'twenty', e.g. 'ten' is attested as *é-pɔ́:t* in Ipulo (Tivoid). This form is probably related to Tiv *púè/ púwè* and Lyive e*pùè* and may be attested in the Mbam branch as well (Nubaca *mwa-pwat* 'ten', etc.).

It is clear, however, that the Ipulo 'twenty' (*i-ham*) is derived from the Proto-Bantoid term for 'ten' by means of a noun class switch. The same can be applied to Bhele (D30): *mɔkɔ́* 'ten' but *e-kómi í-ɓalé* '20' (*í-ɓalé* 'two'). The root *kam* will be discussed below in connection to the terms for 'hundred'.

Table 3.12: Bantoid stems for '10'

|  |  | '10' | '10' | '10' | '10' | '10' |
|---|---|---|---|---|---|---|
| **Northern** |  |  |  |  |  |  |
| Dakoid | Chamba-Daka |  | kúūm-kɘ́rɘ́rɘ́ |  |  |  |
| *Mambiloid |  |  |  | cóŋ |  | job-, jer, jula ? fɛ́ŋ ? |
| Fam |  |  |  |  |  | kwoy |
| Tiba (Fà) |  |  |  |  |  | à-wɔ́ɔ́b-á |
| **Southern** |  |  |  |  |  |  |
| **\*Bantu** |  |  | **kɘ́mì/ kámá** |  |  | **dòngò** |
| *Beboid |  | jo-fi/ jo-fu |  |  |  |  |
| *Yemne-Kimbi |  | jo-fu |  | koŋ? |  |  |
| *Ekoid |  | fo |  |  |  | gol, wobo |
| *Jarawan |  |  |  |  |  | lum |
| *Mamfe |  | fia, bjo |  |  |  |  |
| *Mbam |  |  |  |  | p-wat/ b-wad |  |
| Mbe | Mbe | fwɔ̂r |  |  |  |  |
| Ndemli | Ndemli |  | dʒòm |  |  |  |
| Tikar | Tikar |  | wûm |  |  |  |
| *Tivoid |  | puɛ | *ham |  | pɔt |  |
| *Esimbi |  |  |  |  |  | bu ɣu? (< 9?) |
| Wide Grassfields | Befang |  | éɣúm |  |  |  |
| GF: Mbam-Nkam | Bamileke |  | ɣam |  |  |  |
| GF: Mbam-Nkam | Ngemba |  | ɣám |  |  |  |
| GF: Mbam-Nkam | Nkambe |  | ʔum |  |  | ri/ ru |
| GF: Mbam-Nkam | Nun |  | ɣom |  |  |  |
| GF: Momo |  |  | ɣum |  |  |  |
| GF: Ring |  |  | ɣɘm |  |  |  |

### 3.1.1.9 'Twenty'

It is not necessary to quote the forms for 'twenty', since in the majority of the Bantoid branches (including Bantu) this term is based on 'ten' and follows the pattern '10*2'. Some minor but peculiar variations should be noted here, but all of them are of little significance for our reconstruction. E.g. the term for 'twenty' often employs the plural noun class with the two components in agreement. At the same time, non-compound forms based on 'ten' or 'two' in the plural are also attested. For instance, in one of the Bafut dialects *báà* 'two', *tà-wûm / nì-wûm* 'ten' > *mì-wúm mí-mbáà* 'twenty', while *tà-ghûm* 'ten' ~ *mì-ghum* 'twenty' in another. At the same time, Limbum *bá:* 'two' ~ *m̩-bá:* 'twenty'. These patterns (especially the former) are common in the majority of the Bantu languages as well.

Primary roots for 'twenty' are rarely attested. They may go back to the lexical base 'man' (e.g. in D30 Komo *nkpá búi* 'twenty' = 'whole person'), 'head' (Suga (Mambiloid)) *ɓʉʉ bíb* 'twenty' < *ɓʉʉ* 'head') or some other lexical bases (e.g. Bantu A50: Bafia *ì-tín / mà-tín* 'twenty' < 'score')[6].

---

[6]John Watters: "The Bakor group of Ekoid attest something like *\*-tên* and Mbe has *-têl*. The other two Ekoid groups have a form *-rim* or *-sam*. I would reconstruct for Proto-Ekoid *\*-têl* or *\*-tên* which is like Bantu Bafia. They are a few hundred kilometers apart with many languages and a significant mountain range in between, so this is not borrowing'' (p.c.).

### 3.1.1.10 'Hundred' and 'thousand'

Table 3.13: Bantoid stems for '100'

| | | '100' | '100' | '100' | '100' | '100' | '100' |
|---|---|---|---|---|---|---|---|
| **Northern** | | | | | | | |
| Dakoid | Chamba-Daka | 20*5 | | | | | |
| *Mambiloid | | 20*5 | | | | | < fula |
| **Southern** | | | | | | | |
| **\*Bantu** | | | | | | kámá, gànà, tʋa, jànda | |
| *Beboid | | | | | gbi | | |
| *Yemne-Kimbi | | | | | gbi?ŋwe? | | |
| *Ekoid | | 20*5 | | | | | |
| *Jarawan | | | 10*10 | | | luru? | < Hausa |
| *Mamfe | | 20*5 | | | | | |
| *Mbam | | | | | | | < Engl. |
| Mbe | Mbe | 20 *5 | | | | | |
| Ndemli | Ndemli | | | | | m̀bókó | |
| Tikar | Tikar | | | | | nɗuʔ | |
| *Tivoid | | 20*5 | | | | | |
| *Esimbi | | | 10*10 | | | | < Engl |
| Wide Grassfields | Befang | | | | | bɔ̀mí ⁿdáŋgàŋ | |
| GF: Mbam-Nkam | Bamileke | | | | k(h)u | | |
| GF: Mbam-Nkam | Ngemba | | | | k(h)i/ kirə | | |
| GF: Mbam-Nkam | Nkambe | | | | ŋkʉ̀? | rdʒèe? | |
| GF: Mbam-Nkam | Nun | | | | ŋku | | |
| GF: Momo | | | | | ki, ko | | |
| GF: Ring | | | | | ɣí/ vi | ntu? | |

It appears that the term for 'hundred' cannot be reconstructed for Proto-Bantoid: in most of the branches the pattern employed is '20*5',[7] whereas in some of the branches the term is borrowed. Both Grassfields and Bantu show innovations. The Grassfields root may be tentatively reconstructed as *ku. Several roots are known for Bantu, their use being limited to certain zones: *kámá* ABCDHL, *gànà* DEFGJNPS, *tʋa* DL, *jànda* MNP. None of these roots is attested with this meaning elsewhere in the Bantoid languages, except for Bantu. The similarity of *kámá*

---

[7]John Watters: ''The distribution of this form is suggestive of an older vigesimal system for Bantoid rather than a decimal one. I would take the decimal ones as innovations'' (p.c.).

with the root reconstructed for 'ten' is noteworthy. Moreover, it is attested with the meaning 'thousand' in at least three of the Bantoid branches as the table below shows (Table 3.14):

Table 3.14: Bantoid stems for '1000'

|  |  | '1000' | '1000' |
|---|---|---|---|
| **Northern** | | | |
| Dakoid | Chamba-Daka | | 100*10 |
| *Mambiloid | | | ndúúŋ 'sack', < Fula |
| **Southern** | | | |
| **\*Bantu** | | | **nùnù, pɔ̀mbì, kɔ́tɔ̀** |
| *Beboid | | | cuku |
| *Yemne-Kimbi | | kam? | kia? |
| *Ekoid | | | 200*5? |
| *Jarawan | | | ? |
| *Mamfe | | | nka? |
| *Mbam | | | < Engl. |
| Mbe | Mbe | | 400*2+200 |
| Ndemli | Ndemli | | kòlí |
| Tikar | Tikar | ŋkæm | |
| *Tivoid | | | 20*10, engl. |
| *Esimbi | | | < engl |
| Wide Grassfields | Befang | | ítʃə́n ~ étʃə́n |
| GF: Mbam-Nkam | Bamileke | | tsa / sa? |
| GF: Mbam-Nkam | Ngemba | kamə? | tsuʔu? |
| GF: Mbam-Nkam | Nkambe | | cuki? |
| GF: Mbam-Nkam | Nun | | 100*10 |
| GF: Momo | | | < engl |
| GF: Ring | | kam | |

The root *kam* allows multiple interpretations. We will return to it after the evidence from other Benue-Congo branches has been examined.

The Proto-Bantoid numeral system can be reconstructed as follows (*) (Table 3.15):

Table 3.15: Proto-Bantoid numeral system(*)[8]

| 1 | m-o-ʔ, m-o-i, m-o-ti, mo-di | 7 | samba / camba (< *c/saN+2?) |
|---|---|---|---|
| 2 | pa/ fe, badi (*ba-di?) | 8 | na-nai (< 4 redupl.) |
| 3 | tat | 9 | bukV |
| 4 | nai | 10 | fu, kum/kam |
| 5 | tan | 20 | 10*2 |
| 6 | ta-ta(t) (< 3 redupl.?) | 100 | gbi? ki? 20*5? kam? |
| | | 1000 | ? |

According to Kay Williamson, the base for 'one' in Benue-Congo should be reconstructed as *#-kani*. The only form quoted in support of this hypothesis in her first article (Williamson 1989a: 255) is a supposed Bantoid reflex of the root in Tiba (*a-kina* '1'). Later (**Williamson1992**: 396) she adduced one more Bantoid form, a Southern Bantoid Esimbi term *keni* '1'. That Williamson gives too much weight to these two marginal Bantoid forms is evident from the fact that she reconstructs this base not only for Benue-Congo, but for Niger-Congo as well. This leads her to the idea (probably expressed in the latter work for the first time) that Niger-Congo originally roots had a triconsonantal structure, hence her reconstruction of the proto-form for 'one' as *\*\*-ˈkə'gəni*. This Niger-Congo etymology will be studied in detail below. At this point we will only note that the Esimbi form cited above is strikingly unusual for the Bantoid languages and was probably misinterpreted. The form *kēnə̄* '1' is indeed attested in some of the Esimbi sources (see Brad Koenig, https://mpi-lingweb.shh.mpg.de/numeral/Esimbi.htm). However, in other sources the form *ɔ-nə* is attested (Cristin Kalinowski in (Chan)), so the term for 'eleven' is *bùɣù nə-nə* (*bùɣù* '10'). In other words, the base for 'one' in Esimbi is *-ni/-nə̄* (!), while the first syllable should be interpreted as the noun class prefix, just as in other numerals (cf. the forms *mə̄rākpə̄* '2', *mōɲī* '4', *mātə̄nə̀* '5', etc. in Koenig).

As for Tiba, it is still not certain whether this language indeed belongs to the Bantoid group (cf. **Boyd1999**, where Tiba is considered an Adamawa language). The only Bantoid forms that could have been used by Williamson in support of her hypothesis are found in some of the Northern Mambiloid languages, cf. Twendi (Cambap) *tʃínī*, Mambila *tʃɛ́n* (with palatalization assumed). However, these forms are extremely marginal as well, so they cannot give ground for the

---

[8]My competence does not allow me to reconstruct the tones in the numeral Bantoid languages, especially in Benue-Congo.

proto-language reconstruction (in any case, not for Proto-Bantoid).

### 3.1.2 Benue-Congo (the Bantoid languages excluded)

After the numerals of the Bantoid languages, let's consider the numerals in each of the other groups within this vast family, namely Cross, Defoid, Edoid, Idomoid, Igboid, Jukunoid, Kainji, Platoid, Nupoid (Sections 3.1.2.1–3.1.2.9) and in some isolated BC languages – Ikaan, Akpes, Oko and Lufu (Sections 3.1.3.1–3.1.3.4). After this, we will generalize the results obtained in order to try to reconstruct the numerals of Proto-BC (§3.1.4).

#### 3.1.2.1 Cross

Let us consider the typical stems for numerals in the Cross languages.

Table 3.16: Cross stems for '1'

|  | '1' | '1' | '1' | '1' |
|---|---|---|---|---|
| **1. Bendi** | | | | |
| Bendi | ken | | -bɔ́ŋè? | |
| **2. Delta-Cross** | | | | |
| Upper | | ni (D[9]: *gʷá-nì) | wɔ̀n, guŋ? | mɔ́ɔ̀? |
| Central | | nin | | |
| Lower | sin/ cin, ki/ ge, kiet/ keed (D:*cèèd) | | | |
| Ogoni | zĩ̀ĩ̀ | nɛ(n) | | |

Let us dwell on this table, using it as an example for understanding the majority of the subsequent tables given in this book. Almost every table represents the synthesis of the primary data. We cannot publish all of these primary forms. Let's make an exception. In order to make clear to the reader on what basis the generalizations were made, we present in a separate Appendix (see Appendix D) all the forms available for the numerals '1' in the Cross languages, including intermediate Proto-Upper Cross and Proto-Lower Cross reconstructions, proposed by

---

[9]Here and below, index D introduces the reconstruction proposed by **Dimmendaal1978**.

**Dimmendaal1978** and **Connell1991**. From the Appendix D, it is clear that Connell accepts the Dimmentaal hypothesis, according to which in Upper Cross *ᵍwá-* is interpreted as a prefix, and the lexical stem is represented by *-ni*, attested also in Central Delta-Cross and Ogoni. Based on the 60 sources listed in Appendix D, in table 3.15 for the numeral '1', the root *ni(n)* is allocated. The table also identifies the second root for '1', also possibly represented in the three branches of their five. Connell reconstructs it as *cèèd*, but the data from various Lower Delta-Cross, as well as from Dendi, suggests that perhaps we are dealing with a palatalization of the velar before the front vowel: *ked / ket / kin > ced / cin* (unfortunately, for most groups of the Niger-Congo, including Cross, we do not have sufficient grounds for reconstructing the tones). Finally, the third root presented in Icheve *à-mɔɔ* is probably related to Bantu.

'Two' (Table 3.17)

Table 3.17: Cross stems for '2'.

|  | '2' | '2' | '2' |
|---|---|---|---|
| **1. Bendi** | | | |
| Bendi | | fe, ha? | |
| **2. Delta-Cross** | | | |
| Upper | | fa(n)/ poo (D:*ppán) | |
| Central | | | jal/ yal/ zal/ wal |
| Lower | bà (D:*íbà) | | |
| Ogoni | bàè/ bɛrɛ | | |

The roots *bae* and *po/pa* are noteworthy.

'Three' and 'Four' (Table 3.18)

The common Niger-Congo roots are attested for these numerals in all of the branches (*ta(t)/ ca(t)* and *na(n)* respectively).

'Five' (Table 3.19)

Two roots can be postulated for Cross, namely *tan* and its alternative, tentatively described as *gbo(k)*.

'Six' to 'Nine' (Table 3.20)

At this stage it seems reasonable to maintain the forms and patterns represented in the last line of the table.

'Ten', 'Twenty', and 'Hundred' (Table 3.21)

Table 3.18: Cross stems for '3' and '4'

|  | '3' | '3' | '4' | '4' |
|---|---|---|---|---|
| **1. Bendi** |  |  |  |  |
| Bendi | kie/ cia/ cat |  | ne |  |
| **2. Delta-Cross** |  |  |  |  |
| Upper | tat/ tan/ *sa/ , kia(t) (D: ttán ~ ttáD) | naan? | na (D: *nàŋì ~ này) |  |
| Central | sar/ rar |  | ɲa |  |
| Lower | tá (D:*ítá) |  | nàaŋ/ nìàŋ (D:*ìnìàŋ) |  |
| Ogoni | taa |  | nia | 3+1 |

Table 3.19: Cross stems for '5'

|  | '5' | '5' | '5' |
|---|---|---|---|
| **1. Bendi** |  |  |  |
| Bendi | taŋ |  | dʲoŋ |
| **2. Delta-Cross** |  |  |  |
| Upper | tɔ́ɔ́n/ tāɲ/ zen/ cen | gbo/ buo(k) |  |
| Central |  | oɣ/ wʊ? |  |
| Lower | tîŋ/ tin/ tion, go?(D:*ítíòn) |  |  |
| Ogoni | *rɛ̀ | ʔòò/ vòò/ wò/ *ʔa |  |

It should be noted that providing a detailed reconstruction for each of the Cross numerals lies beyond the scope of the present investigation, so there is probably no point in trying to establish which of the roots for 'ten' (*kpo* or *job* ) should be reconstructed in the Proto-Cross (especially impossible without external evidence).

The Cross languages are highly divergent in regard to numerals (an exception should be made for 'three' and 'four' which are remarkably stable in Cross, as well as in the other NC branches). At the same time, the forms cited above do not provide sufficient reason to suggest a closer relationship within any randomly selected pair of the Cross branches. Hence, it would be too daring to interpret

Table 3.20: Cross stems and patterns for '6'-'9'

|  | '6' | '6' | '6' | '7' | '8' | '9' |
|---|---|---|---|---|---|---|
| **1. Bendi** | | | | | | |
| Bendi | 5+1 | | | 5 + 2 | 5 + 3 | 5 + 4 |
| **2. Delta-Cross** | | | | | | |
| Upper | 5+1 | | ránē , 3+3 | 5+2, 4+3 | 4+4 | 10-1, 5+4 |
| Central | | di(n) | | ɗùal/ ɗuən | 4PL | súɣó |
| Lower | 5+1 | | | 5+2 | 5+3 | 5+4 |
| Ogoni | 5+1 | nìʔĩ̃? | ʔɔ̀rɔ̀? | 5+2 | 5+3 | 10-1, 5+4 |
| **CROSS** | **5+1** | **diʔ** | **3+3** | **5+2** | **4+4** | **10-1, 5+4** |

Table 3.21: Cross stems and patterns for '10', '20' and '100'

|  | '10' | '10' | '20' | '20' | '20' | '100' |
|---|---|---|---|---|---|---|
| **1. Bendi** | | | | | | |
| Bendi | kpu, hwo, fo | | ci/ si | | jam | 20*5 |
| **2. Delta-Cross** | | | | | | |
| Upper | | jo(b)/ zob/ jop (D:*jòb) | ti | lop, nip (D:*níb) | zol ... | 20*5 |
| Central | | ɗìoβ | | lisiíβ/ rusuβ | poɣ, 2PL | kùròn, 5*20, 80+20 |
| Lower | kɔp (D:*lùgòp) | duob/ duop, dugu/ lugu | | e-dip (D: *édíp) | | i-kie (D: *íkíɛ̀) |
| Ogoni | òb, ʔò | | | | tub/ cu | 5*20 |
| **CROSS** | **kpo** | **job** | **ti/ ci?** | **dip?** | | **20*5** |

the roots attested in both of these branches as shared innovations. Let us count the numbers of related numeral forms in different pairs of the Cross branches (Table 3.22):

Table 3.22: Number of related numerals in different pairs of the Cross branches

|          | Central | Lower | Ogoni | Upper |
|----------|---------|-------|-------|-------|
| Bendi    | **0**   | 4     | 4     | 5     |
| Central  |         | 2     | 2     | 4     |
| Lower    |         |       | 5     | 4     |
| Ogoni    |         |       |       | 4     |

This distribution is remarkable with regard to the total absence of shared forms (with the 'three' and 'four' excluded) between Bendi and Central Cross. Keeping this in mind, all of the established alternative roots and patterns can be reserved for a later discussion. At this point the following reconstruction of the Proto-Cross numerals can be suggested (Table 3.23):

Table 3.23: Numeral system of Proto-Cross (*)

| | | | |
|---|---|---|---|
| 1 | *kin/ cin, *ni(n), *gboŋ/ gwan | 7   | 5+2 |
| 2 | *bae, *po/pa                    | 8   | 4+4 |
| 3 | *ta(t)/ ca(t)                   | 9   | 10-1, 5+4 |
| 4 | *na(n)                          | 10  | *kpo/ kop, foʔ ʔoʔ *job |
| 5 | *tan, *gbo(k)                   | 20  | *ti/ ci ? dip ? |
| 6 | 5+1, diʔ, 3+3                    | 100 | 20*5 |

### 3.1.2.2 Defoid

The Defoid branch is relatively compact: it is composed of four languages including Yoruba and its dialects. Historical phonetics of these languages should be considered for a proper reconstruction of the Defoid numeral system, because most of the terms show great phonetic variety. E.g. for 'four' several forms are attested: *-nɛ* (Arigidi), *-jẽ̄* (Ayere), *-rin/ -hɛ̃/ -ɛ̃* (Yoruba), *-lɛ̀* (Igala). The main forms are given in the following table (Table 3.24), and their reconstruction will be discussed below.

Table 3.24: Defoid numerals

|  | Arigidi (dial.) | Ayere (dial.) | Yoruba | Igala | *Yoruba-Igala | *Proto-Defoid |
|---|---|---|---|---|---|---|
| 1 | kɛ̀ɛ́-ɲɛ̃ | ĩ̀-kã̄ | ē-ní, ɔ̀-kɔ̰̄ | é-ɲɛ́ / ŏ-kâ | *ɲɛ́ , ka(n) | *ɲɛ́ , ka(n) |
| 2 | kè-ji | ì-dʒì | è-jì | è-dʒì | *jì | *jì |
| 3 | ke-dà | ī-tā | ɛ̀-tā | ɛ̀-ta | *tā | *tā |
| 4 | ke-nɛ | ĩ̀-jẽ̄ | ɛ̀-rḭ̄ | ɛ̀-lɛ̀ | *lɛ(n) | *lɛ(n)/ ne, je |
| 5 | ké-ntɔ̀ | ĩ̀-tũ̄ | à-rṵ́ | ɛ̀-lú | *lú(n) | *lú(n) / tu(n) |
| 6 | ke-fà | ì-fà | ɛ̀-fà | ɛ̀-fà | *fà | *fà |
| 7 | ke-ɸi | ī-dʒʷī | è-jē | è-bʲe | *byē | *byē |
| 8 | ke-rò | ī-rō | ɛ̀-jɔ̄ | ɛ̀-dʒɔ | *jɔ̄ | *jo/ ro |
| 9 | ké-ndà | ĩ̀-dã̄ | è-sɔ́ | ɛ̀-lá | *sá(n) | *sá(n), dà |
| 10 | ké-è | ī-gʷá | ɛ̀-wá | ɛ̀-gʷá | *gwá | *gwá |
| 20 | u-gbɔrɔ̀ | ē-gbɔ̄lɔ̄ | ō-gṵ́ | ó-gʷṵ́ | *gwú(n) | *gwú(n)/ gbolo |
| 100 | 20*5 | 20*5 | 20*5 | 20*5 | 20*5 | 20*5 |

Following the Proto-Yoruba-Igala reconstruction (Pozdniakov, ms), the terms *lɛ(n)* '4', *lú(n)* '5' and *sá(n)* '9' are reconstructed on the basis of the following regular phonetic correspondences (Table 3.25):

These examples illustrate the phonetic correspondences coming from *l '(Table 3.26):

Table 3.25: Fragment of the Yoruba-Igala phonetic reconstruction

|       | Yoruba | Igala |
|-------|--------|-------|
| **\*l** | r      | l     |
| **\*r** | r      | d     |
| **\*d** | d/j    | d     |
| **\*n** | l/n    | n     |
| **\*s** | s      | l     |
| **\*ʃ** | s      | r     |
| **\*c** | ʃ      | c     |

Table 3.26: \*L-stems in Proto-Yoruba-Igala and their regular reflexes

| Meaning      | \*Yoruba-Igala | Yoruba   | Igala    |
|--------------|----------------|----------|----------|
| animal, meat | ɛ́lɔ̃           | ərɔ̃     | ɛ́la     |
| toad         | àkèlé          | àkèré    | àkèlé    |
| four         | èlĩ            | èrĩ      | èlɛ̀     |
| five         | èlṹ            | àrṹ      | èlu      |
| ant          | èlìlà          | èèrà     | èlìlà    |
| ashes        | élílú          | eérú     | élúlú    |
| feel         | gbɔ́ òlílù̃      | gbɔ́ òórù̃ | é-gbúlù  |
| star         | ìlàwɔ̀         | ìràwɔ̀   | ìlàwò    |
| small        | kékélé         | kékeré   | kékélé   |
| buy          | là             | rà       | é-là     |
| see          | lí             | rí       | é-lí     |
| plow         | lo             | roko     | é-lo     |
| body         | ɔ́la           | ara      | ɔ́la     |
| word         | ɔ̀là           | ɔ̀rɔ̀    | ɔ̀là     |
| sun          | ólìlù̀         | òòrù̃    | ólù      |
| sleep        | oólũ̃          | oorũ̃    | ólu      |
| neck         | ɔ́lù̃          | ɔrù̃     | ɔ́lɔ̀    |
| thirst       | òlùgbə̃         | òrùgbə̃  | òlùgbə   |
| ring         | ɔ́lù̃-ìka       | òrùka    | èlika    |
| run          | s�V̄lé          | sáré     | é-rúlé   |
| fat          | ùla            | ɔ̀rá     | ùlà      |
| seed         | úlú            | irú      | úlú      |

Yoruba [s] is correspondent to Igala [r] (< *ʃ) or [l] (< *s) in at least six examples, see Table 3.27 below:

Table 3.27: Reflexes of *ʃ and *s in Yoruba-Igala

| Meaning | *Yoruba-Igala | Yoruba | Igala |
|---|---|---|---|
| leg | έʃὲ | əsὲ | έrὲ |
| fruit | èʃo | èso | èro |
| block/ close | ʃé | sé | é-ré |
| launch | ʃɔ | sɔ | é-rɔ |
| nine | ὲsɔ̰́ | ὲsɔ̰́ | ὲlá |
| sleep | sṵ̀ | sṵ̀ | *é-lu- |

The reconstruction of the term for 'seven' (*byē) is based on the following correspondences (Table 3.28):

Table 3.28: One more fragment of the Yoruba-Igala regular correspondences

| | Yoruba | Igala |
|---|---|---|
| *by | j | by |
| *j | j | j |
| *b | b | b |

The reflexes of **\*by**- can be represented as follows (Table 3.29):

Table 3.29: Reflexes of *by in Yoruba-Igala

| Meaning | *Yoruba-Igala | Yoruba | Igala |
|---|---|---|---|
| **dog** | abyá | ajá | abyá |
| **blood** | ὲbyὲ | ὲjὲ | ὲbyὲ |
| **seven** | ebye | èje | ebye |

Finally, the terms *gwá '10' and *gwú(n) '20' are reconstructed in view of **\*gw** > Yoruba **w** (before [a]) / **g** (before [u]) ~ Igala **gw** (Table 3.30).

These correspondences are treated here in detail because they may be of special interest for the comparative study of the Defoid languages.

Table 3.30: Reflexes of *gw in Yoruba-Igala

| Meaning | *Yoruba-Igala | Yoruba | Igala |
|---------|---------------|--------|-------|
| **ten** | ègwá | ὲwá | ὲgwá |
| **beans** | ὲgwà | ὲwà | ὲgwà |
| **dig** | gwà | wà | é-gwà |
| **swim** | gwà | wὲ | é-gwà |
| **sweat** | (ò)úgwὺ | òógὺ | úgwù |
| **bone** | égwúgwú | egũgũ | ógwúgwú |
| **ascend** | gwὺ̰ | gὺ | é-tə-gwù |
| **war** | ógwũ | ogũ | ógwu |
| **twenty** | ōgwú | ōgú | ó-gwú |
| **vulture** | úgwṹnú | igúnugṵ́ | úgwúnú |

### 3.1.2.3 Edoid

The following reconstruction is based on nearly forty sources which represent twenty languages within this group. The reconstruction proposed by Elugbe was also considered.

Being no specialist in the comparative study of the Edoid languages (unlike Elugbe), I don't feel competent enough to criticize his ideas. Elugbe likely had his reasons for reconstructing the same consonant (**ch**-) in the terms for 'three', 'five', 'six' and 'seven'. Indeed, the comparison of data from the four Edoid branches confirms that the terms for 'three' and 'five' (but not for 'seven') have the same initial consonant. This is common for many of the NC branches (and probably for the Proto-NC as well).

In view of this, I would like to suggest a simplified reconstruction that is closer, in my opinion, to the actually attested forms (Table 3.31).

### 3.1.2.4 Idomoid

The roots attested in about ten of the Idomoid languages are represented in Table 3.32.

It should be noted that the data on the Yatye-Akpa branch (one of the two Idomoid branches) is systematically absent. The analysis is based on the Akweya languages only, so unexpected issues may arise.

Table 3.31: Edoid numeral systems and Proto-Edoid

|  | 1. Delta | 2. North-Central | 3. North-western | 4. South-western | Proto-Edoid (Elugbe) | *Proto-Edoid |
|---|---|---|---|---|---|---|
| 1 | βʊ | kpa, wo/ gwo | kpa | vʊ |  | **kpa, wo/gwo/vu** |
| 2 | βə/ βa | va | va | vɛ | i-və | **va/və** |
| 3 | saa | sa | sa | sa | ʊ-chaGʊ | **sa** |
| 4 | ni | ne | ni | ni | niə | **ni** |
| 5 | súwón /syònì | sen / ʃen | sie | soi/ siorin/ jorin | ii-chiNənhi | **sien/ su(w)on** |
| 6 | 3PL | 3+3? | 3+3 | 3PL? | chaN | **3PL, 3+3** |
| 7 | 5+2 | hiron/ hilon, 5+2 | sie/hi/rhi | ɣwrɛ́/ hre | i-chiə | **ghie?** |
| 8 | 4PL, 4 redupl | renren / lelen | nien | re(r)e | nhʊNanhʊ | **4 redupl**. |
| 9 | 10-1 | sin(rin), tili | 5+4 | rhi(r)i, zi | i-ciənhi | **cien/ sin** |
| 10 | gbeny/ gbei | gbe | gbe | kpe/ xwe | gbeNi | **gbe, kpe** |
| 20 | jow/yei | gie/je | gboro, ghe/ ze/ye | dhe/ ɟè/ ʒè | u-gheGi ~ u-ɟh | **gie/ jie** |
| 100 | 20*5 | 20*5 | 10PL | 20*5 |  | **20*5** |
| 1000 |  | ria / li, gbele | 500*2 | du, riorin |  | **du, ria/li** |

Table 3.32: Idomoid numerals 31

| 1 | nze/ je/ nye/ ye, kpokpoh?[a] | 7 | 5+2, renyi |
|---|---|---|---|
| 2 | pa, miyeh? | 8 | 5+3 |
| 3 | ta/ la | 9 | 5+4 |
| 4 | nè, ndo, he | 10 | gwo/ wo, jwo |
| 5 | do/ lo, ho, ro/rwo | 20 | fu/ hu, su |
| 6 | rowo/ riwi, ji, hili | 100 | 20*5, 10*10 |

[a]Please note that hypothetically related forms are separated by a slash (/), whereas unrelated ones are separated by a comma.

### 3.1.2.5 Igboid

This is a small group consisting of several languages. The forms which could be found in modern Igboid languages are listed in Table 3.33:

Table 3.33: Igboid numerals

| 1 | tù, ŋìnέ (Ekpeye)? | 7 | saà |
|---|---|---|---|
| 2 | bɔ́ | 8 | 5+3 |
| 3 | tɔ́ | 9 | totu /tolu |
| 4 | nɔ́ | 10 | ɗì/ ri/ li |
| 5 | sé | 20 | gwɔ̋ / ɣʰɔ̄, kpɔrɔ |
| 6 | ʃĩi | 100 | 20*5 |
| | | 1000 | puk(w)u |

Interestingly, the terms for 'one' attested in the Igboid languages (as found in Koelle 1963[1854]) are subject to significant variation. The following forms are noteworthy: '1' – Īsóāma *oo-te*, Íṣiēle *mfuu*, Ábādṣa *na*, Aro *mbɔ*, Mbọ́fīa *mpoŋ* (the transcription of the forms and languages follows Koelle). The rest of the numerals quoted by Koelle are essentially the same as the ones found in Table 3.34.

### 3.1.2.6 Jukunoid

Table 3.34: Jukunoid numerals

|  | 1. Bete (Juk.) | 2. Central | 3. Yukuben-Kuteb | Proto-Jukunoid |
|---|---|---|---|---|
| 1 | ʃíʃe | (d)zun/ (d)zuŋ | nzo, ji?, yʊn?, ŋgēmé?, tɔ́ŋ? | *d)zun? ʃíʃe? tɔ́ŋ? |
| 2 | há | pye(na) | pa(n)/ fa(n) | *pa(n) / fa(n) |
| 3 | tà | (t)sara | ta | *ta |
| 4 | ɲè | nye(na) | ɲi, nje/ nzì | *nye |
| 5 | tsòŋ | (t)swa(na) | t(s)oŋ | *tsoŋ |
| 6 | 5+1 | 5+1 | 5+1 | *5+1 |
| 7 | 5+2 | 5+2 | 5+2 | *5+2 |
| 8 | 5+3 | 4 redupl., 5+3 | 5+3 | *4 redupl., 5+3 |
| 9 | 5+4 | 5+4? | 5+4 | *5+4 |
| 10 | wo | dub (< Hausa?), dz(w)e | kur? kuwub, bji/ bzi, jwēr | *jwe, wo? kur? |
| 20 | ? | 'body' (á-dì) | kam / k(w)om | *'body' (di) |
| 100 | ? | 20*5 | 20*5, Hausa | *20*5 |
| 1000 | ? | < Hausa | Hausa | < Hausa |

Tentative reconstructions for the three major branches of this relatively small family are presented in the table above. The terms for 'one' and 'ten' vary significantly.

### 3.1.2.7 Kainji

The comparative analysis of the Kainji group is hindered by the fact that there is no linguistic description for the majority of its languages. However, there is a great range in numerical terms within those languages, for which reliable data is available. The following analysis is based on thirty pertinent sources, including the comparative list of forms compiled by **Dettweiler1993**. What follows is a step-by-step analysis of the available data that will hopefully yield some answers.

#### 3.1.2.7.1 'One'

The grouping principles for the forms included in this table are admittedly haphazard. On the one hand, the relationship between some of the forms arranged into the same column (e.g. *hĩn, tʃə̄:n* and *dɛn* or *dínkā* and *ʼlu-ruŋ*) is not immediately apparent. On the other hand, some of the forms placed in separate columns might be etymologically related (e.g. *dɨin* Giro and *dínkā* Iguta). In these circumstances it seems reasonable to go back to the reconstruction of the Kainji term for 'one' on the basis of the data provided by other Benue-Congo branches (see 3.1.4).

#### 3.1.2.7.2 'Two'

The above considerations regarding the term for 'one' are applicable to the term for 'two' as well. The inventory of forms found in Table 3.36 is neither helpful for the reconstruction of the Proto-Kainji term for 'two', nor suggestive of the morphemic analysis of the pertinent forms within each of the branches. As we hope to demonstrate below, additional information that may prove useful for the reconstruction of the term for 'two' can be obtained through the analysis of the term for 'seven'.

#### 3.1.2.7.3 'Three', 'Four' and 'Five'

Unlike the terms for 'one' and 'two', the numerals covering the sequence from 'three' to 'five' are quite homogeneous and thus can be reliably reconstructed (just as in the majority of other NC branches). The provisional forms suggested for 'three', 'four', and 'five' are *ʼtat, *ʼnas,* and *ʼtan* respectively. The latter form can also be reconstructed for Eastern Kainji on the basis of the Amo evidence. Thus *ʧibi* (*ʧi-bi*?) 'five' is an innovation of the Jera subgroup.

Table 3.35: Kainji stems for '1'

|  | Language | '1' | '1' | '1' | '1' |
|---|---|---|---|---|---|
| Eastern | | | | | |
| Jera | Iguta | | | dínkā | |
| Jera | Janji | | | diŋkɛ | ɪnde |
| Jera | Bunu | | ù-ŋ̀ŋíní | díŋkà | |
| Jera | Buji | | | díŋkà | |
| Amo | Amo | | | *lu-ruŋ | |
| Western | | | | | |
| Basa | Basa | hĩn | | | |
| Duka | C'lela | tʃĩ́ | | | |
| Duka | Hun-Saare(Duka) | cɔɔn | | | |
| Duka | Ut-Ma'in | tʃə̄:n | | | |
| Duka | Rijau | tʃoon | | | |
| Duka | Darangi | tʃooɾ | | | |
| Duka | Bunu | dɨɨ | | | |
| Duka | Iri | dən | | | |
| Duka | Dukku | dɛn | | | |
| Duka | Giro | dɨɨn | | | |
| Kambari | Tsishingini (Kambari) | | íyyán | | |
| Kambari | Agaushi (Tsikimba) | | | | '-tè |
| Kambari | Kambali (Koelle) | | ííɲa | | |
| Kamuku | Western Acipa (Cicipu) | | | | tô: |
| Kamuku | Kamuku (dial.) | | ĩ́já | | |
| Kamuku | Hungworo (Hungwere) | | ĩ́:jə̃́ | | |
| Kamuku | Pongu (Pangu) | hĩ́: | | | |
| Kamuku | Kamuku (Koelle) | híía͜ | | | |
| Kamuku | Fungwa | hĩ̌ | | | |
| Reshe | Reshe (Tsureshe) | tsúnnɛ̀ | | | |

Table 3.36: Kainji stems for '2'

|  |  | '2' | '2' | '2' | '2' |
|---|---|---|---|---|---|
| **Eastern** |  |  |  |  |  |
| Jera | Iguta |  |  | rὲːpú |  |
| Jera | Janji |  | tɪ-rɛ (~wa-~a-) | -rὲὲpó |  |
| Jera | Bunu |  |  |  |  |
| Jera | Buji |  |  | rὲpó |  |
| Amo | Amo |  |  |  | im-ba |
| **Western** |  |  |  |  |  |
| Basa | Basa | jὲbí (yééwi) |  |  |  |
| Duka | C'lela |  | ʔílɨ̀ |  |  |
| Duka | Hun-Saare(Duka) |  | yoor |  |  |
| Duka | Ut-Ma'in |  | jə̄ːr |  |  |
| Duka | Rijau |  | jooɾ |  |  |
| Duka | Darangi |  | jooɾ |  |  |
| Duka | Bunu |  | jɔɔɾ |  |  |
| Duka | Iri |  | jooɾ |  |  |
| Duka | Dukku |  | juuɾ |  |  |
| Duka | Giro |  | jooɾ |  |  |
| Kambari | Tsishingini (Kambari) |  | ì-ɾὲ |  |  |
| Kambari | Agaushi (Tsikimba) |  | -rὲ |  |  |
| Kambari | Kambali (Koelle) |  | íí-lɛ |  |  |
| Kamuku | Western Acipa (Cicipu) | jápù |  |  |  |
| Kamuku | Kamuku (dial.) | ⁿdə́ɰə̀ |  |  |  |
| Kamuku | Hungworo (Hungwere) |  | ʔⁱə̃̀-dʒə̀ |  |  |
| Kamuku | Pongu (Pangu) |  | ɾêːnù |  |  |
| Kamuku | Kamuku (Koelle) |  |  |  | wúúlee |
| Kamuku | Fungwa | jóːgó |  |  |  |
| Reshe | Reshe (Tsureshe) |  |  |  | rìsə̄ |

Table 3.37: Kainji stems for '3'-'5'

| | | '3' | '4' | '5' | '5' |
|---|---|---|---|---|---|
| Eastern | | | | | |
| Jera | Iguta | tààrū | nà:nzī | | ʃù:bì |
| Jera | Janji | | tɪ-naze | | ʧibi |
| Jera | Bunu | | nà:zé | | ʃí:bì |
| Jera | Buji | | nàzé | | ʃíbí |
| Amo | Amo | | nnas | n-ntaun | |
| Western | | | | | |
| Basa | Basa | tàtɔ | néʃì (nááʃii) | táná | |
| Duka | C'lela | tɨ́:ʧù | ná:sé | tǎ | |
| Duka | Hun-Saare(Duka) | tett | náss | táán | |
| Duka | Ut-Ma'in | tə̄t | ná:s | tán | |
| Duka | Rijau | tɪtʰ | nəss | taan | |
| Duka | Darangi | tɪtʰ | nas | taan | |
| Duka | Bunu | tɪtʰ | nas | tan | |
| Duka | Iri | tɪɪt | nass | taan | |
| Duka | Dukku | tɨɨt | nas | taan | |
| Duka | Giro | tɨtʰ | nass | taan | |
| Kambari | Tsishingini (Kambari) | tàʔàtsú | nə́ꜜʃín | tá:ꜜwún | |
| Kambari | Agaushi (Tsikimba) | | '-nə́ʃì | '-tə́ũ | |
| Kambari | Kambali (Koelle) | tááatsu | nóóʃin | tááu̱ | |
| Kamuku | Western Acipa (Cicipu) | tâ:tù | nósì | tə̃u̱ | |
| Kamuku | Kamuku (dial.) | tátɔ̀ | nə́ʃì | táɑ̀ | |
| Kamuku | Hungworo (Hungwere) | tâ̰tɔ̀ | ùnə́sĩ́ | sàtá | |
| Kamuku | Pongu (Pangu) | tâ:tù | nɔ̌:ʃĩ | tá | |
| Kamuku | Kamuku (Koelle) | tááto | náʃii | taa ~ tááa | |
| Kamuku | Fungwa | | nó:ʃi | tá | |
| Reshe | Reshe (Tsureshe) | tàtswā | nāʃě̃ | tɔ̃̌ | |

### 3.1.2.7.4 'Six' and 'Seven'

Table 3.38: Kainji stems and patterns for '6'-'7'

| | | '1' | '2' | '5' | '6' | '7' | '7' |
|---|---|---|---|---|---|---|---|
| | **Eastern** | | | | | | |
| 1 | Jera | Iguta | | | | twà:sì | | súnā:rí |
| 2 | Jera | Janji | | tɪ-rɛ | | tase | | sunare |
| 3 | Jera | Bunu | | | | tá:sè ~tà:sé | | súnà:ré |
| 4 | Jera | Buji | | | | tásé | | súnárí |
| 5 | Amo | Amo | | | n-ntaun | ku-toʧin | kuzor | |
| | **Western** | | | | | | |
| 6 | Basa | Basa | hĭn | | táná | ʧĭhin | ʧéndʒe | |
| 7 | Duka | C'lela | ʧĭ | *ʔí-lì | tǎ | ʧĭhĭ | tǎʔìlì | |
| 8 | Duka | Hun-Saare | cǫǫn | * yoo-r | táán | cînd | tá'yoor | |
| 9 | Duka | Ut-Ma'in | ʧə̄:n | *jə̄:-r | tán | ʃiʃin | tàʔèr | |
| 10 | Duka | Rijau | ʧoon | *joo-ɾ | taan | ʧiin | ta'jooɾ | |
| 11 | Duka | Darangi | ʧooɾ | *joo-ɾ | taan | ʧìn | taŋ'joɾ | |
| 12 | Duka | Bunu | dɨɨ | *jɔɔ-ɾ | tan | ʧiin | ta'juu | |
| 13 | Duka | Iri | dən | *joo-ɾ | taan | ʧinnd | ta'jooɾ | |
| 14 | Duka | Dukku | dɛn | *juu-ɾ | taan | ʧɪŋ | ta'jaar | |
| 15 | Duka | Giro | dɨin | *joo-ɾ | taan | ʧind | ta'jooɾ | |
| 16 | Kambari | Tsishingini | | ì-ɾɛ | tá:wún | tò:lí | ʧìndèɾé | |
| 17 | Kambari | Agaushi | -tè | -rè | -tǎū | -tə̀:lì | ʧìndèrè | |
| 18 | Kambari | Kambali | | íí-lɛ, *rɛ | tááu̯ | tɔ́ɔli | tsíndɛɛɾɛ | |
| 19 | Kamuku | West.Acipa | | *jà | tǎu̯ | tóríhĭ̃ | tíndàjà | |
| 20 | Kamuku | Cinda | | *ɰə̀ | táɒ̀ | tánɔ́hì | tándɔ́ɰə̀ | |
| 21 | Kamuku | Hungworo | | ʔʲɔ̃́-dʒə̀, *rʲə̀ | sàtá | ū-t̪únìhĭ̃ | ū-tándə́rʲə̀ | |
| 22 | Kamuku | Pongu | hĭ̃: | rê:nù, *rə̀ | tá | ʧíníhì | tə̃́ndɔ́rə̀ | |
| 23 | Kamuku | Kamuku | híía | *lee | taa ~ tááa | túnui | tandále | |
| 25 | Kamuku | Fungwa | hĭ | *lò | tá | ʧĭhĭ̃ | tíndàlò | |
| 25 | Reshe | Reshe | tsúnnè | | tɔ̃́ | tēnzɔ̄ | tànsã̀ | |

Some of the previously discussed terms for 'one', 'two' and 'five' are quoted in the table above alongside the terms for 'six' and 'seven'. Such grouping might facilitate a better understanding of compound numerals (if 'six' and 'seven' are indeed compounds) as well as the methodological and theoretical aspects behind their reconstruction. In addition, it might help to establish whether parts of compound numerals can be used to enhance the reconstruction of the primary numerical terms such as 'one', 'two', and 'five'.

The compound nature of the term for 'seven' is betrayed by its 'length': the forms quoted in the table normally have two to three syllables, whereas the primary numerals are as a rule mono- or (rarely) bisyllabic.

At the same time, in some of the cases the pattern '7=5+2' is immediately ap-

parent (cf. languages 7-11, 13-15).

At this point, however, we will deal with those languages that show only faint (or no) traces of the pattern in question ('7=5+2'). E.g. in Tsishingini (16) we have to assume the pattern '7=X+2', where 'X' is an unknown element, whereas in language 12 the pattern is '7=5+X' (the relationship between 'X' and the term for 'two' is questionable).

Let us assume that the Proto-Kainji terms for 'two' and 'five' are *CL-**re** (cf. e.g. Duka *jo-re > joor*) and *tan* respectively. In this case, the compound term for 'seven' would be *tan-(CL)-re* or *tan-X* (connector)-(CL)-re. The most typical diachronic scenarios for the emergence of the 'X'-patterns effective on the synchronic level are as follows:

1) Both basic elements of the compound 'seven' (i.e. reflexes of the terms for 'two' and 'five') are preserved in the language, as is the compound itself (sometimes slightly modified in accordance with the relevant phonotactic rules). Cf. e.g. the Darangi (11) evidence: *jo-re > joor* '2', *tan > taan* '5', *taan-jo-re > taŋ'joɾ* '7'. In this case, the reconstruction comes down to the simple statement that in the Darangi language '7=5+2'.

2) The compound 'seven' (even if slightly modified) is preserved in the language, while the term for 'two' is replaced with an innovation. Let us assume that in the Basa language (6) *jèbí* (Koelle: *yééwi*) '2' < *jo-bi* (innovation), *táná* '5' (the reflex of *tan*), *tʃéndʒe* < *tan-re* '7'. In this case, *tan-re > tan-dʒe > tendʒe* (regressive assimilation) > *tʃendʒe* (palatalization before the front vowel). Hypothetical as it may be, this example is phonetically plausible.

Any of these model processes may result in the loss of phonetic resemblance between a derived form and its source. This may lead to a situation where a derivation pattern is no longer recognizable by speakers. As a consequence, the term for 'seven' becomes opaque on the synchronic level and can no longer be analysed as '5+2'.

This means that the replacement of the original term for 'two' by an innovation does not affect the compound term for 'seven', i.e. that its second part is not automatically replaced. Moreover, in case there is sufficient evidence that the second of the aforementioned scenarios was applied, we may enhance the reconstruction of the primary term for 'two' on the basis of the compound term for 'seven'. E.g. the form *tʃéndʒe* suggests that the original Basa root for 'two' was *dʒe / re* and not *bi* as in the majority of the Kainji languages.

The available pertinent forms point toward the reconstruction of the Proto-Kainji form as *tan-da-re* ('5'-connector-'2'). The reconstructed forms for 'two' (marked with [*] in Table 3.38) suggest a Proto-Kainji form *re* '2' and the pattern

*'7=5+2'. The Eastern Kainji forms for 'seven' are probably innovations.

However, some of the forms attested for 'seven' may point toward the reconstruction of 'two' as *ba/bi* in Proto-Kainji. In this case our reference list should be expanded by adding dialects that were not included for reasons of space: it is not possible to quote every single NC source every time. E.g. Cawai (Eastern Kainji) *a-ba* '2', *a-tar-ba* '7', Ngwoi (Hungworo) *e-bia* '2', *sa-bia* '7' (the root *ba/bi* is also suggested by Eastern: Gure *pi-ba*, Gyem *ve*, Piti *ba*, Surubu *ka-va*).

The forms for 'six' are more problematic since they may go back to a primary root (or roots). They may be tentatively reconstructed as *ci(hi)n, *tas,* and *tel*. We will come back to these forms in order to enhance their reconstruction in case similar forms are detected in other BC branches.

### 3.1.2.7.5 'Eight'

### 3.1.2.7.6 'Nine' and 'Ten'

There are several forms and patterns for 'nine' whose reconstruction is equally plausible: '9=5+4', *tor(b)oj* (possibly < *'10-1'), *jiro.* Each of the forms/patterns is characteristic of a particular sub-group of languages. The term for 'ten' is reconstructed as *pwa*, with its reflexes attested in all Western Kainji branches. Three alternative forms (*turu, *kuri, *kup/ kpa*) are found in Eastern Kainji, where they are employed for counting and in quantity measures.

### 3.1.2.7.7 'Twenty' and 'Hundred'

The diversity of patterns for 'hundred' may indicate the absence of the term in Proto-Kainji. The term for 'twenty' likely followed the pattern '20=10*2'. At the same time, the form *ʃin/ ʃík* attested in three of the Western Kainji branches is noteworthy.

It should be noted that a full reconstruction of the Kainji numeral system is not presently achievable for a number of reasons: some of the forms have multiple alternative variants, many terms are not attested outside Kainji (or have an obscure morphological structure), the elements of the compound terms are not always identifiable (e.g. in the patterns '7=X+2' or '7=5+X'), etc.

The numerals attested within this group are so peculiar (at least for a non-specialist in the Kainji languages like myself) that one may wonder whether the Kainji group should indeed be treated as a branch of Benue-Congo. In any case, it seems reasonable to record all the forms reconstructable within the Kainji sub-groups. These forms and patterns are represented in the table below (Table 3.42):

Table 3.39: Kainji stems and patterns for '8'

|  |  | '8' | '8' | '8' |
|---|---|---|---|---|
| **Eastern** | | | | |
| Jera | Iguta | ùrū | | |
| Jera | Janji | uro | | |
| Jera | Bunu | ùrú | | |
| Jera | Buji | úrú | | |
| Amo | Amo | | | kuliv |
| **Western** | | | | |
| Basa | Basa | | tɔndatɔ (5+3) | |
| Duka | C'lela | jɨːɾù | | |
| Duka | Hun-Saare(Duka) | yéér | | |
| Duka | Ut-Ma'in | éːr | | |
| Duka | Rijau | eeɾ | | |
| Duka | Darangi | eɾ | | |
| Duka | Bunu | ɛɛɾ | | |
| Duka | Iri | ɪɪɾ | | |
| Duka | Dukku | ɛɛɾ | | |
| Duka | Giro | ɛɛɾ | | |
| Kambari | Tsishingini (Kambari) | | | kùnlə̀ |
| Kambari | Agaushi (Tsikimba) | | | kúnlə̀i |
| Kambari | Kambali (Koelle) | | | kúnlo |
| Kamuku | Western Acipa (Cicipu) | | | kùríl:ò |
| Kamuku | Kamuku (dial.) | | tántátɔ̀ (5+3) | |
| Kamuku | Hungworo (Hungwere) | | ū-tátàt̪ɔ̄(5+3) | |
| Kamuku | Pongu (Pangu) | | tə̃ndáːtù(5+3) | |
| Kamuku | Kamuku (Koelle) | | túndaat(5+3) | |
| Kamuku | Fungwa | | tíndátù(5+3) | |
| Reshe | Reshe (Tsureshe) | | dálànzɔ̀ | |

The Eastern Kainji and Duka forms (if related) suggest that the primary root *-ru* should be reconstructed for 'eight' in Proto-Kainji. At this point, let us reserve a preliminary form *u-ro/ ji-ru* for further comparison. In most of the Kamuku languages the pattern '8=5+3' is traceable (but note the Western Acipa form that is comparable to those attested in Kambari and possibly Amo (Eastern)). This points towards an alternative form of uncertain morphological structure (*kunle(v)/ kunlo* '8').

Table 3.40: Kainji stems and patterns for '9' and '10'

| | | '9' | '9' | '9' | '10' | '10' |
|---|---|---|---|---|---|---|
| **Eastern** | | | | | | |
| Jera | Iguta | | tɔ̀rbɔ̀ (10-1) | | | bū-tú:rú |
| Jera | Janji | | toroəi (10-1) | | | turo, kɪrəu |
| Jera | Bunu | | tò:rêj (10-1) | | | bì-tú:rú; rú-kúrí |
| Jera | Buji | | toroj (10-1) | | | bì-túrú; rì-kùrì |
| Amo | Amo | | ku-tivi | | | ku-lidir *li-kure |
| **Western** | | | | | | |
| Basa | Basa | tʃíndʒìʃì (5+4) | | | | úḿpwá |
| Duka | C'lela | | | dó:ɾè | ʔó:pá | |
| Duka | Hun-Saare(Duka) | | | jíɾɔ̠̀ | o̠pp | |
| Duka | Ut-Ma'in | | | dʒʷɔ̄:r | ɔ̄p | |
| Duka | Rijau | | | dʒiɾɔ | ɔpʰ | |
| Duka | Darangi | | | dʒiɾɔ | 'ɔpʰ | |
| Duka | Bunu | | | dʒiɾɔ | ɔpʰ | |
| Duka | Iri | | | dʒɪɾɔ | ɔpʰ | |
| Duka | Dukku | | | dʒɪɾɔ | ɔpʰ | |
| Duka | Giro | | | dʒɛdɔ | ɔp | |
| Kambari | Tsishingini (Kambari) | kùttʃí | | | kùppá | |
| Kambari | Agaushi (Tsikimba) | kùtʃì | | | kùpà | |
| Kambari | Kambali (Koelle) | kúciici | | | hókpa̠ | |
| Kamuku | Western Acipa (Cicipu) | kùtít:í (5+4) | | | ùkúp:à | |
| Kamuku | Kamuku (dial.) | tə́ndə́ʃì (5+4) | | | òpá | |
| Kamuku | Hungworo (Hungwere) | ūtə́nə̀sĩ̄ (5+4) | | | īkópʲè | |
| Kamuku | Pongu (Pangu) | tṹndúʃì (5+4) | | | úpwá | |
| Kamuku | Kamuku (Koelle) | tándaaʃii (5+4) | | | ópaa | |
| Kamuku | Fungwa | tíndíʃì (5+4) | | | úpá | |
| Reshe | Reshe (Tsureshe) | tānāʃě̃(5+4) | | | úpwà | |

Table 3.41: Kainji stems and patterns for '20' and '100'

| | | '20' | '20' | '20' | '100' |
|---|---|---|---|---|---|
| Eastern | | | | | |
| Jera | Iguta | | | 12+8 | 12*8+4 |
| Jera | Janji | | | | |
| Jera | Bunu | | | | rì:mú |
| Jera | Buji | | | 10*2 | *ri-nu |
| Amo | Amo | | | akut-2 | li-kalt |
| Western | | | | | |
| Basa | Basa | wéʃi (K:wóóʃi) | | | dupu íjèbi (50*2) |
| Duka | C'lela | dᵊkʷὲzὲ | | | kʷὲtʃtǎ / vzíŋgù |
| Duka | Hun-Saare(Duka) | ɛr-kwọọz | | | kwọọz-ɛt táán (20 * 4 ), o-zùngu |
| Duka | Ut-Ma'in | | ɔrʃĩk | | ɔʔʃĩkɔʔtán (20 * 5 ) |
| Duka | Rijau | | | | |
| Duka | Darangi | | | | |
| Duka | Bunu | | | | |
| Duka | Iri | | | | |
| Duka | Dukku | | | | |
| Duka | Giro | | | | |
| Kambari | Tsishingini (Kambari) | | ú:ʃín | | ? |
| Kambari | Agaushi (Tsikimba) | | | kà-màngà | |
| Kambari | Kambali (Koelle) | | úʃi̱ | | |
| Kamuku | Western Acipa (Cicipu) | | | 10*2 | 10*10, mándá |
| Kamuku | Kamuku (dial.) | | | 10*2 | ɗərí (< Hausa) or dè òpá |
| Kamuku | Hungworo (Hungwere) | | | 10*2 | íhɔ̄ŋgʷà, 10*10 |
| Kamuku | Pongu (Pangu) | wɔ́ʃí | | | bìjĭnʒ̃ |
| Kamuku | Kamuku (Koelle) | | | 10*2 | |
| Kamuku | Fungwa | | kùʤìjò | | ìkwà:ku, < Hausa |
| Reshe | Reshe (Tsureshe) | | | álə̀sə̀ | ránɛ̄kū |

Table 3.42: Kainji summarized data for BC reconstruction

| 1 | *tsin, hin, din, jan/ yan, *te ... | 7 | *5+2 |
|---|---|---|---|
| 2 | *re, *ba/ bi, -pu? | 8 | *ro/ ru, *5+3, *kunle(v)/ kunlo |
| 3 | *tat | 9 | *5+4, *10-1, *jiro |
| 4 | *nas | 10 | *pwa, *turu, *kuri, *kup/ kpa |
| 5 | *tan | 20 | *10*2, *ʃín/ ʃík |
| 6 | *ci(hi)n, *tas (< 3?), *tel | 100 | ? |

### 3.1.2.8 Platoid

#### 3.1.2.8.1 'One' (Table 3.43)

Table 3.43: Platoid stems for '1'

| | | | | | | |
|---|---|---|---|---|---|---|
| 1. | Alumu-Tesu | Tesu | | | | à-nyimbere |
| 2. | Ayu | Ayu | ɪ-dɪ | | | |
| 3. | Biromic | Birom | | gw-īnìŋ/(d)-īnìŋ | | |
| 3. | Biromic | Eten | dáy | | | |
| 4. | Cental | Izere | | z-iníŋ | | |
| 4. | Cental | Irigwe | | | | ˀzrú |
| 4. | Cental | Kaje (dial.) | | | | yiruŋ/yirəŋ |
| 4. | Cental | Tyap | | | a-nyuŋ | |
| 5. | Hyamic | Hyam | | ʒ-ìnì | | |
| 6. | Ninzic | Mada | | *nɛn | | gyə̄r |
| 6. | Ninzic | Ninzo | | *nì | | jír |
| 7. | Northern | Ikulu | | | | íɲjí |
| 8. | Southeastern | Fyam | | kʲ-éŋ, *in | | |
| 9. | Southern | Lijili | lǫ̀ | | | |
| 10. | Taroid | Tarok (dial.) | | | ù-zɨ̀ŋ, *ɗɨ́ŋ? | |
| 11. | Western | Yeskwa (dial.) | | | | è-nyí |
| 11. | Western | Rukuba (dial.) | | gy-ín | | |
| 11. | Western | Eggon (dial.) | | | | á-kiə́n |
| 11. | Western | Eggon (dial.) | ò-rí | | | |
| 11. | Western | Hasha | | nʸ-ìnāŋ | | |
| ? | Sambe | | ɲ-íɲínā | | | |

1. Tesu data are taken from **BlenchKato2012**.

The grouping of roots here is admittedly provisional, because their morphological structure is often obscure. In addition, phonetic changes that may have taken place are unknown. It is very difficult to propose any etymological interpretation for the forms represented in the table. Which of them could be attributed to the Proto-Platoid is unclear (*(y)in represents a possibility, in case noun class markers are indeed incorporated into the numerical terms).

#### 3.1.2.8.2 'Two', 'Three' and 'Four' (Table 3.44)

The roots for 'two' containing voiced and voiceless labials are attested in the Platoid languages (as well as in some other BC branches). They may be tentatively reconstructed as *pa/ fa/ ha and *ba/ wa.

Table 3.44: Platoid stems for '2', '3' and '4'

| | | | '2' | '2' | '3' | '4' |
|---|---|---|---|---|---|---|
| 1. | Alumu-Tesu | Tesu | | à-hùrwi | à-taatɔ | a-anɛ |
| 2. | Ayu | Ayu | ahwa/ afah | | a-taar | a-naŋaʃ |
| 3. | Biromic | Birom | | -bā | -tāt | -nā:s |
| 3. | Biromic | Eten | fà | | tàt/tʃàt | nà:s |
| 4. | Cenral | Izere | fà | | taar | nààs |
| 4. | Cenral | Irigwe | | ʾʍʲè | ʾtsʲè | ʾni |
| 4. | Cenral | Kaje (dial.) | ʼ-hwa | | ʼ-tat | -nai |
| 4. | Cenral | Tyap | a-feaŋ | | a-tat | a-naai |
| 5. | Hyamic | Hyam | feri, *fo | | taat | naaŋ |
| 6. | Ninzic | Mada | | y-wā, *gba | tar | nlyḛ̄ |
| 6. | Ninzic | Ninzo | há | *gba | tár | nā̰(s) |
| 7. | Northern | Ikulu | íń-pààlá | | íń-táá | íń-nāā |
| 8. | Southeastern | Fyam | por | | táár | naas |
| 9. | Southern | Lijili | | à-bḛ̄ | à-tʃɛ́ | à-nàrɔ̰́ |
| 10. | Taroid | Tarok (dial.) | ù-pàrɨ́m | | ù-ʃádɨ́ŋ | ù-nèdɨ́ŋ |
| 11. | Western | Yeskwa (dial.) | | èn-và | èn-tât | èn-nà |
| 11. | Western | Rukuba (dial.) | ʼ-hàk | | -tát | -nàs |
| 11. | Western | Eggon (dial.) | à-hàà | | à-tráá | ù-ɲí |
| 11. | Western | Eggon (dial.) | ɔ̀-hà | | ɔ̀-cá | ò-ɲì |
| 11. | Western | Hasha | à-pʷò | | ā-tāt | à-nìŋ |
| ? | Sambe | bèkà-fà | kà-tú | kà-tār/ béká-tār | kà-nὲ/ bèkà-nὲ | |

The roots for 'three' and 'four' are more stable. Some of their reflexes suggest that the Proto-Platoid forms must have been close to the NC forms: *tat* '3' and *nai / *nas* '4'.

### 3.1.2.8.3 'Five' and 'Six' (Table 3.45)

Table 3.45: Platoid stems and patterns for '5' and '6'

|  |  |  | '5' | '5' | '6' | '6' |
|---|---|---|---|---|---|---|
| 1. | Alumu-Tesu | Tesu | a-túŋgú |  | térékífí (<3?) |  |
| 2. | Ayu | Ayu | a-tugen |  | a-tɛɛr (3PL) |  |
| 3. | Biromic | Birom | -tūŋūn |  |  | -tī:mìn |
| 3. | Biromic | Eten |  | wí | tà:rà (<3) |  |
| 4. | Cenral | Izere | tùwùn |  | ìgà-rà:r (3PL) |  |
| 4. | Cenral | Irigwe | ˀtɕʷòô |  | rí-tsʲɛ́ (3PL) |  |
| 4. | Cenral | Kaje (dial.) |  | -pfwɔn | kə-tat (3PL) |  |
| 4. | Cenral | Tyap |  | a-fwuon | a-taa (3PL) |  |
| 5. | Hyamic | Hyam | twoo |  | twaa-ni (5+1) |  |
| 6. | Ninzic | Mada | tun |  | tān-nèn (5+1) |  |
| 6. | Ninzic | Ninzo | tʷí |  | tā-nì (5+1) |  |
| 7. | Northern | Ikulu | íń-cūū |  | íń-cúnú (5+1?) |  |
| 8. | Southeastern | Fyam | tóón |  | táár-in (5+1) |  |
| 9. | Southern | Lijili | à-sɔ́ |  | mìn-zí (3PL?) |  |
| 10. | Taroid | Tarok (dial.) | ù-túkún |  | ù-kpɔ́-d́ɪ́ŋ (X+1?) |  |
| 11. | Western | Yeskwa (dial.) | èn-tyúò |  | èn-cí (5+1) |  |
| 11. | Western | Rukuba (dial.) | -túŋ |  | tàiŋ |  |
| 11. | Western | Eggon (dial.) | ò-tnó | *fúúŋ | ù-fín (5+1?) |  |
| 11. | Western | Eggon (dial.) | ɔ̀-tnɔ̂ | *fɔ̆ŋ | ə̀-fĩ́(5+1?) |  |
| 11. | Western | Hasha | ā-tūkūn |  |  | à-kʷìp |
| ? | Sambe | kà-tûn |  |  | kù-hɔ̀/ dɔ̀gɔ̀-hɔ̀ |  |

The term for 'five' is reconstructed as *tu(ku)n*. It is likely that there was no primary term for 'six' in the Proto-Platoid group: in all pertinent languages (except for Eggon, Hasha and Sambe) the term in question either follows the pattern '5+1' or is built by adding a plural class to the term for 'three'.

### 3.1.2.8.4 'Seven' and 'eight' (Table 3.46)

Word-building patterns for the term for 'seven' are normally quite transparent: '7=5+2' is attested in the majority of the sub-groups, whereas '7=4+3' is more rare. The same can be applied to the term for 'eight', which either follows the pattern

Table 3.46: Platoid stems and patterns for '7' and '8'

| | | | '7' | '8' | '8' |
|---|---|---|---|---|---|
| 1. | Alumu-Tesu | Tesu | térékífí naɲí (6+X) | | tsyátsyá |
| 2. | Ayu | Ayu | a-taraŋaʃ (3+4) | a-na-ba-bog (4+X) | |
| 3. | Biromic | Birom | -tā:mà (5+2) | | -rwī:t |
| 3. | Biromic | Eten | nìtà (4+3) | nàràs (4+X) | |
| 4. | Cenral | Izere | kà-nàsàtáár (4+3) | | ì-kárá |
| 4. | Cenral | Irigwe | natsʲɛ́ (4+3) | | klaǹvà |
| 4. | Cenral | Kaje (dial.) | ti:ruŋ (cf. yiruŋ '1') | nai-mʊwak (4+X) | |
| 4. | Cenral | Tyap | a-natat (4+3) | a-ninai (4 redupl.) | |
| 5. | Hyamic | Hyam | twarfo (5+2)? | naaraŋ (4+X) | |
| 6. | Ninzic | Mada | tāmgbā (5+2) | tāndà (5+3) | |
| 6. | Ninzic | Ninzo | tāŋgbā (5+2) | tāndàr (5+3) | |
| 7. | Northern | Ikulu | tɔ́ɔ̀pāā (5+2) | níǹnāā (4 redupl.) | |
| 8. | Southeastern | Fyam | támor (5+2) | | tʃínít |
| 9. | Southern | Lijili | mú-tá | | rúnɔ̧́ |
| 10. | Taroid | Tarok (dial.) | ù-fàŋ-ʃát (X+3) | ù-nɘ̀nnè (4 redupl.) | |
| 11. | Western | Yeskwa (dial.) | tònvà (5+2) | tóndát (5+3) | |
| 11. | Western | Rukuba (dial.) | taŋbák (5+2) | ta:rat (5+3) | |
| 11. | Western | Eggon (dial.) | à-fóhà (5+2) | à-fóté (5+3) | |
| 11. | Western | Eggon (dial.) | ɔ̀-fɔ́hà (5+2) | ɔ̀-fɔ́tɛ́ (5+3) | |
| 11. | Western | Hasha | à-kʷìp nʸīnāŋ (cf. 6, 4) | nànìŋ (4 redupl.) | |
| ? | Sambe | kɔ̄rɔ̄nkɛ́rā / kúrkɘ́nrā | ī-tɔ́r | |

'8=5+3' or is built by partial reduplication of 'four' (4 redupl.). Sometimes the archaic primary terms for 'two' and 'five' are traceable in the forms for 'seven' and 'eight' (such forms are marked with an asterisk in the respective tables).

3.1.2.8.5 **'Nine' and 'Ten'** (Table 3.47)

It is likely that the term for 'nine' attested in Ikulu, Yeskwa and Sambe (*toro/cora*) is primary. The hypothetical inter-relationship of these roots may be of interest for the Proto-Platoid reconstruction, because these languages do not belong to the same sub-group. The forms of 'nine' in the majority of the languages show traces of 'five', 'four', 'ten' and 'one', which suggests that two alternative patterns ('9=5+4' or '9=10-1') could have been in use. Some rare patterns (e.g. '9=12-3' (Birom) and '9=8+X (Tesu)) are of interest for the linguistic typology.

According to **Bouquiaux1962** the term for 'twelve' (*kūrū*) is attested in Birom. In this language '21' (*kūrū ná syā:-tāt*) = '12+9' (*syā:-tāt*), while '80' (*bākūrū bātī: mìn ná rwī:t*) = '12*6' (*-tī: mìn*) + '8'(*-rwī:t*). The pattern '9=12-3' is not totally unexpected within this context. A similar system can be traced in the Mada language. As stated in our source (Abiel Barau Kato), "Like many languages in Platoid area, Mada has an old duodecimal numeral system up to 24."[10] The Mada terms for 'twelve' and 'twenty-one' are *tsɔ* and *tsɔtīyār* (*tīyār* '9') respectively. The same root for 'twelve' (*tsó* '12') is found in Ninzo for which our source notes that "In the traditional counting system, to count beyond twelve (12), that is from thirteen onwards, entails counting in sets of twelve."[11] Moreover, the same root is attested in Tesu (*tsɔ* '12'). According to Uche Aaron, a primary root *ɔ̀-cʷɔ́* '12' is discernible in Eggon (beside the composite term '12=10+2'). This root is also found in Rukuba (Che) in *u-sɔ́k* '12'. The duodecimal numeral system as attested in this language is of the utmost sophistication. According to Luc Bouquiaux: "There are two words for number '72', *kitu* and *atu*, 144 can be expressed as *atu ahak* and 200 is *atu ahak ni isɔk inas ni hak ni ta:rat* ( 72 * 2) + (12 * 4) + 8."[12] Other languages in this group normally use less exotic systems. In some of them, however, e.g. in Eten, "The highest number that can be counted in traditional way is 144,"[13] i.e. '12*12'. To sum up, it seems that a primary term for 'twelve' can be reconstructed on the Proto-Platoid level, hence the pattern for 'nine' should most probably be reconstructed as *'9=12-3'.

---

[10] https://mpi-lingweb.shh.mpg.de/numeral/Ninzo.htm

[11] https://mpi-lingweb.shh.mpg.de/numeral/Ninzo.htm

[12] https://mpi-lingweb.shh.mpg.de/numeral/Rukuba.htm

[13] https://mpi-lingweb.shh.mpg.de/numeral/Aten.htm

Table 3.47: Platoid stems and patterns for '9' and '10'

| | | | '9' | '9' | '10' | '10' | '10' |
|---|---|---|---|---|---|---|---|
| 1. | Alumu-Tesu | Tesu | tsyátsyá naɲí (8+X) | | | | gòròmàvɔ |
| 2. | Ayu | Ayu | a-tu-lu-bog (5+4?) | | | i-ʃog/ a-ja-la-bog | |
| 3. | Biromic | Birom | syā:-tāt (12-3) | | | | 12-2 |
| 3. | Biromic | Eten | dù:dʒàŋ (10-X) | | | | dù:bɔ̀ |
| 4. | Cenral | Izere | kàtúbɔ́k (5+X?) | | | kù-sɔ́k | |
| 4. | Cenral | Irigwe | | kruvájá | | ʃʷá | |
| 4. | Cenral | Kaje (dial.) | kumʊwi:ruŋ (10-1?) | | *ku? | swak | |
| 4. | Cenral | Tyap | akubunyuŋ (10-1?) | | *kub? | swak | |
| 5. | Hyamic | Hyam | mbwan kɔb (10-1) | | kɔ́b | | |
| 6. | Ninzic | Mada | tīyār (X-1?) | | | | gùr |
| 6. | Ninzic | Ninzo | tīr(s) (3-X?) | | | | wūr |
| 7. | Northern | Ikulu | | tɔ́ɔ̀llāā | nù-kɔ̄p | | |
| 8. | Southeastern | Fyam | téres (3-X?) | | | | dukút |
| 9. | Southern | Lijili | zà-tʃɛ̣́ (X-3?) | | | | zà-bẹ̀ |
| 10. | Taroid | Tarok (dial.) | ùfàŋzíŋtíŋ (X+4) | | ù-gbɔ́pei | | |
| 11. | Western | Yeskwa (dial.) | | tyúôrá | ó-kóp | | |
| 11. | Western | Rukuba (dial.) | ta:ras (3-X?) | | | | u-wùruk |
| 11. | Western | Eggon (dial.) | àfúúɲí ( 5+4) | | ó-kpo | | |
| 11. | Western | Eggon (dial.) | ɔ̀fɔ̂ɲí ( 5+4) | | ɔ̀-kbɔ́ | | |
| 11. | Western | Hasha | nànìŋ màrēŋ (4+X) | | | | ā-wūk |
| ? | Sambe | | tōrō/ kà-tɔ́rɔ́ | | | jɔ̀-wō | |

The system outlined above adds a new perspective to the forms with the meaning 'ten'. Presumably, there was a Proto-Platoid primary term for 'ten' that may be tentatively described as \**kop*. The alternative forms *sok/swak* may be etymologically related to the forms for 'twelve' cited above. If so, their change of meaning may have resulted from the adoption of a decimal system. The root *gur/wur* is distinguished as well.

The specific nature of the Platoid numeral system prevents us from providing separate forms for 'twenty' and 'hundred'. The pattern \*'20=12+8' traceable in a number of pertinent languages is reconstructed for Proto-Platoid. A compound nature is also assumed for 'hundred'.

The results pertaining to the advanced reconstructions of numerals in Proto-Platoid are summed up in the table below (Table 3.48):

Table 3.48: Proto-Platoid numeral system (*)

| 1 | (y)in, di(n), jir, nìŋ | 7 | 5+2, 4+3 |
|---|---|---|---|
| 2 | pa/ fa/ ha, ba/ wa. | 8 | 4 redupl., 5+3 |
| 3 | tat | 9 | 5+4, 10-1, 12-3, tu(ku)n |
| 4 | nai/ nas | 10 | kop, gur/ wur |
| 5 | tu(ku)n | 20 | 12+8 |
| 6 | 5+1, 3PL | 100 | ? |

### 3.1.2.9 Nupoid

Let us try to reconstruct the Proto-Nupoid numeral system.

Table 3.49: Nupoid numerals and Proto-Nupoid (*)

| Nupoid | Ebira | Gbari | Kakanda | Nupe | *Nupoid |
|---|---|---|---|---|---|
| 1 | ɔ̀ɔ̀-nyɪ̄ | gbᵐaː-rí,*wĩ | gú-ní | ni-ní | **ni/ nyi, wi? ri?** |
| 2 | ɛ̀ɛ̀-vā | ŋʷã̄-ba | gú-bà | gú-bà | **ba** |
| 3 | ɛ̀ɛ̀-tá | ŋʷã̄-tʃa | gú-tá | gú-tá | **ta** |
| 4 | ɛ̀ɛ̀-nà | ŋʷã̄-ɲi | gú-ni | gú-ni | **na/ ni** |
| 5 | ɛ̀ɛ̀-hí | ŋʷã̄-tⁿù | gú-tũ | gú-tsũ | **tun/ tnu/tsun, hi?** |
| 6 | hǐ-nɔ̌-nyɪ̄ (5+1) | tⁿú-wĩ (5+1) | gú-tua-ɲ̀ì (5+1) | gú-tswà-ɲĩ (5+1) | **5+1** |
| 7 | hǐ-m̀-bā (5+2) | tⁿâ-ba (5+2) | gú-tua-bà (5+2) | gú-twà-bà (5+2) | **5+2** |
| 8 | hǐ-ń-tá (5+3) | tⁿã̄-tʃa (5+3) | gú-tò-tá (5+3) | gú-to-tá (5+3) | **5+3** |
| 9 | hǐ-ǹ-nà (5+4) | tⁿâ-ɲi (5+4) | gú-tua-ni (5+4) | gú-twã̀-ni (5+4) | **5+4** |
| 10 | ɛ̀ɛ̀-wɔ́ | ŋʷã̄-wò | gú-wo | gú-wo | **wo** |
| 20 | ɔ̀ɔ̀-hū,*tʃɛ̃́ | wo-ʃì | e-ʃĩ | e-ʃi | **ʃi, hu?** |
| 100 | ɛ̄-tʃɛ̃́-hí (20x5) | 40*2+20 | ʃìt-ũ (20*5) | ʃit-sũ (20*5) | **20*5** |
| 1000 | 400*5??? | 100*10 | | kpá-tsũ (200*5) | ? |

The Nupoid group is relatively small and homogeneous and poses no problem for reconstruction.

### 3.1.3 Isolated BC languages

#### 3.1.3.1 Ikaan

The following description of the Ikaan numeral system (Table 3.50) is based on the analysis of data from a number of its dialects:

Table 3.50: Proto-Ikaan numeral system (*)

| | | | |
|---|---|---|---|
| 1 | ʃí | 7 | h-ránèʃì ('6+1') |
| 2 | wà | 8 | nà:ná$^j$ (4 redupl.) |
| 3 | tā:s / h-rāhr | 9 | h-ráòʃì (X-1) |
| 4 | nā$^j$ / nā/ náhį́ | 10 | ò-pú/ fú |
| 5 | tò:n/ h-rɤ̀:n/ sɔ̀n/ cɤ̀nṿ | 20 | ù-gbɔ́rɔ́ (< 'sack'), * à-gbá |
| 6 | h-ràdá/ sàdá/ sàrá | 100 | à-gbá à-h-ruǹ(20*5) |

#### 3.1.3.2 Akpes

Table 3.51: Akpes numerals

| | | | |
|---|---|---|---|
| 1 | í-gbōn, ē-kìnì | 7 | ī-tʃēnētʃ(ì) |
| 2 | ī-dīan(ì) | 8 | ā-nāānīŋ(ì) (4 redupl.) |
| 3 | ī-sās(ì) | 9 | ɔ̀-kpɔ̄lɔ̀ʃ(ì) |
| 4 | ī-nīŋ(ì) | 10 | ī-yōf(ì), *t-ēfī |
| 5 | ī-ʃōn(ì) | 20 | ɔ̄-gbɔ̄(lɔ̄) |
| 6 | ī-tʃānās(ì) | 100 | ī-gbɔ́ ʃōnì (20*5) |

The original BC forms for 'five' (*tan) and 'one' may have been preserved in the term for 'six'. These forms will be treated below as hypothetical.

### 3.1.3.3 Oko

Table 3.52: Oko numerals

| 1 | ɔ̀-ɔ́rɛ, ɔ̀-jɛ́rɛ | 7 | ú-fɔ́mbɔ̀rɛ̀ (5+2) |
|---|---|---|---|
| 2 | ɛ̀-bɔ̀rɛ̀ | 8 | ɔ̀nɔ́kɔ́nɔkɔ́nɔ(4 redupl.?) |
| 3 | ɛ̀-ta | 9 | ù-bɔ́ɔ̀rɛ̀(10-1) |
| 4 | ɛ̀-na | 10 | ɛ̀-fɔ |
| 5 | ù-pi | 20 | ɔ́-gbɔlɔ |
| 6 | ɔ̀-pɔ́nɔ̀ɔ́rɛ (5+1) | 100 | í-pì |

### 3.1.3.4 Lufu

Table 3.53: Lufu numerals

| 1 | ù-tí | 7 | 5+2 |
|---|---|---|---|
| 2 | (ba)-máhà | 8 | 5+3 |
| 3 | bá-tá | 9 | 5+4 |
| 4 | ba-ɲì | 10 | ú-wó |
| 5 | bá-tsó | 20 | e-ce |
| 6 | 5+1 | 100, 1000 | ? |

### 3.1.4 Proto-Benue-Congo

#### 3.1.4.1 'One'

The reconstruction of the term for '1' is objectively the most challenging (the term is especially difficult to reconstruct in languages with noun classes and complex systems of determinatives). This situation is even more complicated in the Benue-Congo languages, since more than one reconstruction of the term has been suggested. The existing hypotheses must be studied here, especially because the ones pertaining to the etymology of the term were proposed by Kay Williamson, the leading specialist in NC comparative studies. Moreover, Kay Williamson (1989a) used her reconstruction of the term for 'one' as an argument in favor of triconsonantal structure of Niger-Congo roots. This hypothesis has been actively developed by Roger Blench (**Blench2012a** etc.).

It should be noted that our evidence does not support Kay Williamson's reconstruction. Furthermore, her hypothesis regarding the triconsonantal nature of Niger-Congo roots is, in my opinion, untenable. The Bantoid data utilized by Williamson was discussed above. Now let us review the evidence she uses in support of her hypotheses. Originally she treated the root *#-kani* '1' as one of the basic BC roots ('old root', Williamson 1989a: 255). Later she changed her approach (on the basis of a wider NC context, namely on the data from the Ịjọ languages) suggesting a derivation of BC froms from a triconsonantal root *\*\*- 'kə'gəni* '1', for which she assumed a different set of reflexes (**Williamson1992**). The changes introduced by Williamson in this article are significant. She adds the reflexes of the reconstructed root in Akpes and Nupoid, includes its additional reflexes in Esimbi and Bekwarra (Bantoid), adjusts its reflexes in Cross and Platoid (e.g. by reinterpreting PUC *gá-ni/ \*-gwá-nì* previously analysed as an isolated form as a reflex of the root in question), and, finally, omits Kanji and Jukunoid reflexes.

In further interpretation of the BC numeral systems we will use a template chart representing the fourteen branches of BC (Table 3.54). It should be noted that Bantu (as the largest sub-branch of the BC family with the most detailed reconstruction) is treated separately. This means that the Bantoid field will only include non-Bantu forms. The chart below reproduces the data published by Kay Williamson (middle sections) as well as the relevant forms obtained as a result of our step-by-step reconstruction (the rightmost section):

1. Different colors are used in the charts to distinguish between the Eastern and the Western BC languages. A special marking is used for the Bantu languages due to their overall importance for the reconstruction. The abbreviations in the middle sections follow Williamson op. cit. with PLC-

Table 3.54: BC *kin/ cin '1' and alternative reconstructions

| Benue-Congo | | |
|---|---|---|
| Nupoid | Oko | Kainji |
| Defoid | Akpes | Platoid |
| Edoid | Ikaan | Cross |
| Igboid | Lufu | Jukunoid |
| Idomoid | Bantu | Bantoid |

Williamson 1989a: *#-kani* '1'

| | | |
|---|---|---|
| Yoruba ɔ̀-kɔ̃ | | Basa kə |
| | | Pyem kēŋ |
| | | Bete-Bendi ì-kə̄n, Bokyi kì̀n, PLC *-kèèn |
| | | Jukun kā |
| Eloyi kònzé | | Tiba a-kina |

Williamson1992: Proto-Atlantic-Congo **-'kə̱'gəni'1'

| | | |
|---|---|---|
| Gbagyi gmànyi | | |
| Yoruba ɔ̀-kɔ̃ | Ikeram ɛ-ki | PP2-J -gini, PP4 -ɣan |
| | | PUC gá-ni? , PLC -kèèn |
| Eloyi kònzé | | Tiba a-kina, Esimbi keni, Bendi: Bekwarra o-kin |

*kin-/cin- forms for '1' (step-by-step data)

| | | |
|---|---|---|
| | ē-kìnì, *si | tsin, hin |
| | ʃí | (y)in, kyeŋ, gyin |
| | | kin/ cin |
| | | ʃíʃe? |
| | | cin (Mambiloid) |

Proto-Lower Cross, PUC – Proto-Upper Cross, PP – Proto-Platoid.

It should be noted that the difference in the results achieved by means of our step-by-step reconstruction (see above) and those of Williamson is significant. According to our evidence, the postulation of the root \*\*- ʼkə̱ʼgə̱ni̱ '1' for Western Benue-Congo is unsustainable. The existence of this root in Bantoid is also questionable. In her earlier publication, Kay Williamson quoted its only Bantoid reflex (*a-kina* '1') supposedly attested in Northern Bantoid Tiba (Williamson 1989a: 255). However, the affiliation of Tiba with the Bantoid languages is debatable (a connection with the Adamawa languages is suggested in **Boyd1999**). In the article that followed, Williamson quoted another Bantoid form, this time the one attested in Southern Bantoid Esimbi (*keni* '1'). As noted above, this form was probably misinterpreted, becaused it includes the root *-ni/-nə̄*. At the same time, as I tried to demonstrate above, a number of related forms may be attested in the Mambiloid languages (Northern Bantoid): Twendi (Cambap) *tʃínī*, Mambila *tʃέn*. Thus, we are possibly dealing with Proto-Eastern Bantoid *\*cin/kin*. In order to decide whether this form is an innovation or a reflex of an inherent Niger-Congo root (as Kay Williamson says) we need to place it into a wider linguistic context. This issue will be addressed later. At this point we will deal with another root for 'one' postulated by Williamson. According to her, the root is a Benue-Congo innovation.

Since the root *nə̄ / ni* is distinguishable in Esimbi, it seems logical to treat it together with another set of terms for 'one' (*#-diiŋ*). This data (termed BC innovation by Williamson) compared to the results of our step-by-step reconstruction is quoted in the table below (Table 3.55):

Let us review the distribution of this root within the Benue-Congo branches.

Western Benue-Congo.

This root can be reliably reconstructed in Nupoid and Defoid, but not in Edoid. In Igboid it might be attested in Ikpeye: *ŋì-nέ (ŋ-ìnέ?)*. The root is possibly found in some of the Idomoid languages as well: Etulo *o-ɲiī*, Agatu *ó-yè*, Idoma *é-yè*, Alago *ó-je*, Eloyi (dial.) *ò-nzé*, *ńgwò-nzé*.

Eastern Benue-Congo.

Several Kainji forms deserve closer attention. The Gurmana form quoted by Williamson is unfamiliar to me. It may be related to the Bunu form, but the root itself is uncommon for Kainji and thus cannot be reconstructed. Moreover, the root is only marginally attested in the Platoid languages (single occurrences include Eskwa *è-nyí* '1' and possibly Ikulu *í-ń-jí* '1', and *kɔ̀p-ìrì-zīŋ* '11'). Another rare form is *di(n)* with an initial oral consonant (e.g. Ayu *ɪ-dɪ* '1', Eggon *ò-rí* '1'

Table 3.55: BC *ni '1' and alternative reconstructions

| Benue-Congo | | |
|---|---|---|
| Nupoid | Oko | Kainji |
| Defoid | Akpes | Platoid |
| Edoid | Ikaan | Cross |
| Igboid | Lufu | Jukunoid |
| Idomoid | Bantu | Bantoid |
| Williamson 1989a: BC innovations: ***#-diiŋ*** | | |
| Gwari ǹ-ɲī | Oko ɔ́ɔ́rɛ | Gurmana nɪ |
| PY *ɪ-nɛ̃́ | | PP2K *-niiŋ |
| | | OG ɛ̀-nɛ̃, CD #-niin |
| Ikwere ń-ním | | PJ *-yiŋ |
| PId *-nyí | | Lamja nūnɛ́, Ekoid #-jid, -jiŋ |
| ***ni*** forms for '1' (step-by-step data) | | |
| *ni/ nyi | | Bunu ù-ǹɲínì |
| *ɲɛ́ | | nìŋ, (y)in, di(n) |
| | | *ni(n) |
| ŋìnɛ́? | | *-jin? |
| nze/ je/ nye/ ye | | Esimbi -nə/ -ni |

and its palatalized variant *tʃíŋ* – cf. *ɔ̀-kbɔ́ à-tʃíŋ* '11', *ɔ̀-kβɔ́há là-tʃíŋ* '21'). These (etymologically unrelated?) forms, however, should not be reconstructed for Proto-Platoid, because the root *kin* (see above) is clearly distinguishable in the majority of the Platoid branches. At the same time, the Platoid data discredits the reconstruction of the root as *kin/cin*. Multiple arguments can be adduced in favor of the interpretation of the initial velar as a reflex of an archaic noun class prefix, which would yield a Proto-Platoid form *k-in*. This invites the possibility of an etymological connection between the Benue-Congo roots studied above, namely *-in and *-ni. The analysis of the Platoid compound numerals points toward the same conclusion. A number of noteworthy forms can be quoted in support of this, cf. Hyam *ʒìnì* '1' but *twaa-ni* '6' ('5+1', *twoo* '5'), Mada *tānn-ɛ̀n* '6' ('5+1', *tun* '5'), Ninzo *tānì* '6' ('5+1', *ʧʷí* '5'), Rukuba *tàiŋ* '6' ('5+1', *-túŋ* '5'). These Platoid forms bring to mind the case of the Jukonoid term for 'six'. Kay Williamson quotes a Proto-Jukunoid root *-yiŋ. The reasons behind this reconstruction are not immediately apparent, since in the majority of the languages other forms are reserved for this meaning. Her reconstruction may be based on the compound terms for

'six' that follow the pattern '5+1' (or rather '5+X', with X ≠ 1), cf. e.g. Jibu *sùn-jin* '6' (*swana* '5', *zyun* '1'), *cìn-jen/ ʃì-ʒen* (*tswana* '5', *dzun* '1'). As noted above, the root in question is not reconstructable for the Platoid languages. The reconstruction of *\*ni(n)* is assured only for the Eastern Benue-Congo branch (Cross), where it is systematically attested in at least three branches out of five, cf. Proto-Upper Cross (*\*ni*), Central-Cross (*nin*), and Ogoni (*nɛ*). Since *\*ni* can be safely reconstructed for Nupoid, Defoid and Cross, its further comparison to the pertinent roots attested in the languages that belong to other NC branches is required.

In conclusion, it should be noted that regardless of whether a conservative or a more speculative reconstruction (i.e. *\*kin* and *\*ni* vs. *\*k-in/ ni*) is preferred, the resulting root (or roots) is not tri- or disyllabic but rather monosyllabic.

In addition to this, several isolated roots for 'one' are attested in Benue-Congo. Undoubtedly, they represent local innovations. At first glance, this is applicable to the most common Bantoid roots for 'one', including the Bantu forms *mòì/mòdì mòtí*. This, however, may not be entirely correct for reasons that will be discussed in the next chapter. Another noteworthy root that may be tentatively described as *\*jir* is attested in both Oko and Platoid.

The table is subject to further interpretation. We will return to it later after the evidence from the other Niger-Congo branches has been collected. A few remarks are in order here:

1) Both Akpes terms for 'one' (*ē-kìnì, í-gbōn*) find close parallels in the Cross languages (*\*kin/cin, \*ni(n), \*gboŋ/gwan*). The Icheve form *à-mɔ́ɔ̀* is probably borrowed from one of the Bantu languages;

2) The Kainji term finds parallels in the Platoid languages (Ayu, Eten, Tarok, Eggon) and may be etymologically related to the Bantu and Nupoid terms (the morphological structure of the Proto-Bantu form is, however, unclear: *\*mòdì? \*m-òdì? \*mò-dì?*);

3) The Oko form is reminiscent of another Platoid form that is tentatively reconstructed as *\*jir*. The Akpes root *gbōn* '1' finds parallels in the Cross (*gboŋ*) and possibly Edoid languages (*gwo/ wo/ wu*).

### 3.1.4.2 'Two'

Table 3.56: BC stems for '2'

|       |                    | '2'           | '2'     | '2'           |
|-------|--------------------|---------------|---------|---------------|
| East  | Bantu              |               |         | bà-dí / bì-dí |
| East  | Bantoid (–Bantu)   | pa/ fe        | ba      |               |
| East  | Cross              | po/pa         | bae     |               |
| East  | Jukunoid           | pa(n) / fa(n) |         |               |
| East  | Kainji             | -pu?          | *ba/ bi | re            |
| East  | Platoid            | pa/ fa/ ha    | ba/ wa  |               |
| West  | Defoid             |               |         | jì            |
| West  | Edoid              |               | va/və   |               |
| West  | Idomoid            | pa            |         |               |
| West  | Igboid             |               | bɔ́      |               |
| West  | Nupoid             |               | ba      |               |
| West  | Akpes              |               |         | ī-dīan(ì)     |
| West  | Oko                |               | ɛ̀-bɔ̀rɛ̀  |               |
| West  | Ikaan              |               | wà      |               |

The root *pa (also found in the Idomoid languages) is reconstructable for Eastern Benue-Congo, but is not systematically attested in Bantu.

The Bantu form (as represented above) does not seem to be related to other Bantoid forms. At the same time it finds parallels in Defoid and possibly Akpes and Kainji. The most common BC form (*ba/ bai) may go back to *ba-i, with *ba- being a noun class prefix. In this case, the BC form may be reconstructed as *ba-di / ba-ji > bai > ba, which would make the Bantu form the most archaic within Benue-Congo.

These hypotheses will be discussed below, after the evidence from the other BC branches has been reviewed.

### 3.1.4.3 'Three', 'four', 'five'

Table 3.57: BC stems for '3', '4' and '5'

|  |  | '3' | '4' | '5' | '5' |
|---|---|---|---|---|---|
| East | Bantu | tat | nàì / (nàí) | táànò |  |
| East | Bantoid (−Bantu) | tat | nai | tan |  |
| East | Cross | ta(t)/ ca(t) | na(n) | tan | *gbo(k) |
| East | Jukunoid | ta | nye | tsoŋ |  |
| East | Kainji | tat | nas | tan |  |
| East | Platoid | tat | nai/ nas | tu(ku)n |  |
| West | Defoid | tā | lɛ(n), ne, je | tu(n) / lú(n) |  |
| West | Edoid | sa | ni | sien / su(w)on |  |
| West | Idomoid | ta/ la | nɛ̀, ndo, he | do/ lo, ho, ro/rwo |  |
| West | Igboid | tɔ́ | nɔ́ | sé |  |
| West | Nupoid | ta | na/ ni | tun/ tnu/tsun, hi? | hi? |
| West | Akpes | ī-sās(ì) | ī-nīŋ(ì) | ī-ʃōn(ì), *tan |  |
| West | Oko | ɛ̀-ta | ɛ̀-na |  | ù-pi |
| West | Ikaan | tā:s / h-rāhr | nāʲ / nā/ náhḭ́ | tò:n/ h-rɔ̀:n/ sò̰n/ cɔ̰̀nv̰ |  |

This is the most stable group of numerical terms within BC. It comprises the roots *tat '3', *nai '4', and *tan/ ton '5' that are very well-known among the specialists in NC studies. Issues pertaining to the phonetic realization of their reflexes will be treated in the next chapter.

**3.1.4.4 'Six'**

Table 3.58: BC stems and patterns for '6'

| | | | | | | |
|---|---|---|---|---|---|---|
| East | Bantu | 3 redupl. | | | | |
| East | Bantoid (−Bantu) | < 3 redupl.? | | | | |
| East | Cross | 3+3 | 5+1 | diʔ | | |
| East | Jukunoid | | 5+1 | | | |
| East | Kainji | < 3? | | | ci(hi)n | tel |
| East | Platoid | 3PL | 5+1 | | | |
| West | Defoid | | | | | fà |
| West | Edoid | 3PL, 3+3 | | | | |
| West | Idomoid | | | riwi/ rowo | ji | hili |
| West | Igboid | | | | ʃĭi | |
| West | Nupoid | | 5+1 | | | |
| West | Akpes | | 5+1? | | | |
| West | Oko | | 5+1 | | | |
| West | Ikaan | | | | | h-ɾàdá/ sàdá/ sàrá |

As the table shows, there was probably no primary Proto-Benue-Congo root for 'six'. Two alternative patterns are traceable, namely '3PL' ('3 redupl.', '3+3') and '5+1'. Other forms are marginal. The phonetic resemblance of the Kainji and Igboid forms is noteworthy.

### 3.1.4.5 'Seven'

Table 3.59: BC stems and patterns for '7'

| | | | | | |
|---|---|---|---|---|---|
| East | Bantu | càmbà (<**c/saN+2?) | | | |
| East | Bantoid (−Bantu) | samba (5+2?) | | | |
| East | Cross | 5+2 | | | |
| East | Jukunoid | 5+2 | | | |
| East | Kainji | 5+2 | | | |
| East | Platoid | 5+2 | | | 4+3 |
| West | Defoid | | byē | | |
| West | Edoid | | ghie? | | |
| West | Idomoid | 5+2 | | renyi | |
| West | Igboid | | | | saà |
| West | Nupoid | 5+2 | | | |
| West | Akpes | | | | ī-tʃēnētʃ(ì) |
| West | Oko | ú-fɔ́mbɔ̀rɛ̀ (5+2) | | | |
| West | Ikaan | | | h-ránèʃì ('6+1) | |

A primary root for 'seven' is also indistinguishable. The form \*camba/ samba may have lost any phonetic resemblance to its Benue-Congo prototype \*7=5+2 in Proto-Bantoid. The Defoid and Edoid forms are phonetically comparable (a shared innovation?).

### 3.1.4.6 'Eight'

Table 3.60: BC stems and patterns for '8'

| | | | | | |
|---|---|---|---|---|---|
| East | Bantu | nai-nai (< 4 redupl.) | | | |
| East | Bantoid (−Bantu) | na-nai (< 4 redupl.) | | | |
| East | Cross | 4+4 | | | |
| East | Jukunoid | 4 redupl. | 5+3 | | |
| East | Kainji | | 5+3 | ro/ ru | kunle(v)/ kunlo |
| East | Platoid | 4 redupl. | 5+3 | | |
| West | Defoid | | | jo/ ro | |
| West | Edoid | 4 redupl. | | | |
| West | Idomoid | | 5+3 | | |
| West | Igboid | | 5+3 | | |
| West | Nupoid | | 5+3 | | |
| West | Akpes | ā-nāānīŋ(ì) (4 redupl.) | | | |
| West | Oko | ɔ̀-nɔ́kɔ́-nɔkɔ́nɔ (4 redupl.?) | | | |
| West | Ikaan | nà:ná$^{j}$ (4 redupl.) | | | |

In this case, the pattern *\*nai* '4' >*\*na(i)-nai* '8' fits the reconstruction better than its alternative. The similarity between Kainji and Defoid is peculiar and may be due to innovations.

### 3.1.4.7 'Nine'

Table 3.61: BC stems and patterns for '9'

| | | | | | |
|---|---|---|---|---|---|
| East | Bantu | | | bùá | |
| East | Bantoid (−Bantu) | | | bukV | |
| East | Cross | 5+4 | 10-1 | | |
| East | Jukunoid | 5+4 | | | |
| East | Kainji | 5+4 | 10-1 | | jiro |
| East | Platoid | 5+4 | 10-1 | | 12-3, tu(ku)n |
| West | Defoid | | | sá(n) | dà |
| West | Edoid | | | cien/ sin | |
| West | Idomoid | 5+4 | | | |
| West | Igboid | | | | totu /tolu |
| West | Nupoid | 5+4 | | | |
| West | Akpes | | | | ɔ̀-kpɔ̄lɔ̀ʃ(ì) |
| West | Oko | | ù-bɔ́ɔ̀rɛ̀ (10-1) | | |
| West | Ikaan | | h-rá�ò̀ʃì (X-1) | | |

The rightmost column of the table includes many isolated forms (among them some primary ones). The term *buka*, which may appear as an important BC innovation, is reconstructed for Proto-Bantoid. In addition, the pattern '9=5+4' is distinguishable in Proto-Benue-Congo. Like for '8', Defoid and Edoid forms closely resemble each other.

### 3.1.4.8 'Ten'

Table 3.62: BC stems for '10'

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| East | Bantu | | kǒmì/ kámá | | | | |
| East | Bantoid (–Bantu) | fu | kum/ kam | | | | |
| East | Cross | fo? | | kpo/ kop | ʔo? | job | |
| East | Jukunoid | | | wo? | kur? | jwe | |
| East | Kainji | pwa | | kup/ kpa | kur? | | turu |
| East | Platoid | | | kop | gur/ wur | | |
| West | Defoid | | | | gwá | | |
| West | Edoid | | | kpe | gbe | | |
| West | Idomoid | (fu '20') | | | gwo/ wo | jwo | |
| West | Igboid | | | | | | dî/ ri/ li |
| West | Nupoid | (hu '20) | | | wo | | |
| West | Akpes | | | | | ī- yōf(ì), *t- ēfī | |
| West | Oko | ὲ-fɔ | | | | | |
| West | Ikaan | ò-pú/ fú | | | | | |

This is a heterogeneous group of forms. The root *\*pu/fu* attested in both Eastern and Western BC is the most likely candidate for BC reconstruction. At the same time, it is missing from Bantoid, for which the term *\*kum/kam* is reconstructable. The latter form must be a Bantoid innovation. However, assuming that the second consonant may have undergone nasalization in Proto-Bantoid, this form is comparabale to a number of other roots, suggesting that *\*kup/ kop* should be reconstructed for Eastern Benue-Congo. As the table shows, other roots should not be neglected either. They will be treated in combination with the evidence from other NC branches.

### 3.1.4.9 'Twenty'

Table 3.63: BC stems and patterns for '20'

| Region | Branch | | | | | |
|--------|--------|------|----------|----------------|-------------|------------|
| East | Bantu | 10*2 | | | | |
| East | Bantoid (–Bantu) | 10*2 | | | | |
| East | Cross | | *ti/ ci? | | | dip? |
| East | Jukunoid | | | | | 'body' (di) |
| East | Kainji | 10*2 | ʃín/ ʃík | | | |
| East | Platoid | | | | | 12+8 |
| West | Defoid | | | gwú(n), gbolo | | |
| West | Edoid | | | gie/ jie, gboro | | |
| West | Idomoid | | | | fu/ hu, su? | |
| West | Igboid | | | gwŏ̃ / ɣʰõ̃, kpɔrɔ | | |
| West | Nupoid | | ʃi | | hu? | |
| West | Akpes | | | ɔ-gbɔ̄(lɔ̄) | | |
| West | Oko | | | ɔ́-gbɔlɔ | | |
| West | Ikaan | | | ù-gbɔ́rɔ́ (<'sack'), *à-gbá | | |

It is highly unlikely that the Proto-BC term followed the pattern reconstructed for Proto-Bantoid (*'20=10*2'). In all likelihood there was no root for 'twenty' in Proto-BC at all. It should be noted that numerous branches of Western BC use the root (*g*)*bolo* (possibly related to the lexical root with the meaning 'sack') to make 'twenty'. A shorter root (**gba/ gwe*) is reconstructable in the same Western BC branches as well. Its source is likely lexical: it is well-known that the term for 'twenty' in the NC languages often goes back to lexemes with the meaning 'man', 'leader', and 'body' (cf. Jukonoid). The resemblance between the reconstructed Idomoid and Nupoid forms is noteworthy. However, these forms might be etymologically related to the term for 'ten'.

### 3.1.4.10 'Hundred' and 'thousand'

Table 3.64: BC stems and patterns for '100' and '1000'

|      |                    | '100'              | '100' | '100'              | '1000'              |
|------|--------------------|--------------------|-------|--------------------|---------------------|
| East | Bantu              |                    | kámá  | gànà, tʊa, jànda   | nùnù, pɔ̀mbì, kɔ́tɔ̀ |
| East | Bantoid (–Bantu)   | 20*5?              | kam?  | gbi? ki?           | ?                   |
| East | Cross              | 20*5               |       |                    |                     |
| East | Jukunoid           | 20*5               |       |                    | < Hausa             |
| East | Kainji             | ?                  |       |                    |                     |
| East | Platoid            | ?                  |       |                    |                     |
| West | Defoid             | 20*5               |       |                    |                     |
| West | Edoid              | 20*5               |       |                    | du, ria/li          |
| West | Idomoid            | 20*5, 10*10        |       |                    |                     |
| West | Igboid             | 20*5               |       |                    | puk(w)u             |
| West | Nupoid             | 20*5               |       |                    | ?                   |
| West | Akpes              | ī-gbɔ́ ʃōnì (20*5) |       |                    |                     |
| West | Oko                |                    |       | í-pì               |                     |
| West | Ikaan              | à-gbá à-h-ruǹ(20*5)|       |                    |                     |

If Proto-Benue-Congo did not have the term for 'twenty', it probably did not have the term for 'hundred' either, because the only pattern it could follow is *'100=20*5'. In this respect the Proto-Bantoid innovation (*kam) is noteworthy. It resembles another Proto-Bantoid innovation, namely the term for 'ten' (*kum/ kam), which is hardly a coincidence. The possibility that in the cases of 'ten' and 'hundred' we are dealing with alignment by analogy cannot be excluded. This could explain the irregular nasalization of the root for 'ten', cf. Proto-Bantoid *kup '10' → kum by analogy with *kam '100'. The term for 'thousand' was certainly nonexistent in BC.

Taking this into account, the segmental reconstruction of the Proto-BC numeral system may be suggested (Table 3.65):

This table gives an overview of the BC evidence that will be used for further comparison with other NC branches.

Table 3.65: Proto-Benue-Congo numeral system (*)

| 1 | ni, kin/ cin (<k-in?), gbon, (o-)di(n)?, (o-)ti? | 7 | 5+2 |
|---|---|---|---|
| 2 | ba-di / ba-ji, pa? ba(i)? | 8 | 4 redupl. |
| 3 | tat | 9 | 5+4 |
| 4 | nai | 10 | pu/ fu, kup/ kop, gwo / jwo |
| 5 | tan/ ton | 20 | absent? gwa/ gwe? < 'person'? |
| 6 | 3PL/ 3 redupl./ 3+3, 5+1 | 100 | absent? 20*5 |

## 3.2 Kwa

More than eighty Kwa sources were used for the reconstruction. They are representative of the major groups and sub-groups of this family, which consists of about seventy languages. A plausible internal classification of the Kwa languages does not exist. A step-by-step reconstruction of numerals may well be viewed as another important step in this direction. Our preliminary survey of the pertinent evidence is based on the traditional classification that distinguishes five major Kwa branches. We will start with the study of the numerical terms by branch. Then, individual reconstructions will be evaluated with regard to their potential for the general reconstruction of the Proto-Kwa numeral system.

### 3.2.1 Ga-Dangme

Table 3.66: Ga-Dangme numerals

|   | Dangme | Ga | Dangme | Ga |
|---|--------|-----|--------|-----|
| 1 | kákē | é-kòmé | 7 kpà-à-gō (6+1)? | kpà-wo (6+1?) |
| 2 | é-ɲɔ̀ | é-ɲɔ̀ | 8 kpà-a-ɲɔ̄ (6+2) | kpà-a-ɲɔ̃ (6+2?) |
| 3 | é-tɛ̄̃ | é-tɛ̃ | 9 nɛ̀ɛ̄́ | nɛ̀ɛhű |
| 4 | é-ywɛ̀/é-wìɛ̀ | é-ɟwɛ̀ | 10 ɲɔ̀ŋmá (PL: ɲìŋmí) | ɲɔ̀ŋmá |
| 5 | é-nṹɔ̃̄ | é-nùmɔ̃ | 20 ɲìŋmí éɲɔ̀̃ (10*2) | ɲɔ̀ŋmá -í éɲɔ̀ (10*2) |
| 6 | é-kpà | é-k͡pàa | 100 làfá | ò-há, pl. -ì |
|   |  |  | 1000 à-kpé | à-kpé, pl.-ì |

These two languages exhibit isolated forms of the term for 'one'. Both terms will be preserved for further comparison (note that the first syllable of the Dangme term probably represents a noun class prefix). The term for 'eight' is undoubtedly constructed as '6+2'. The term for 'six' is primary, hence the term for 'seven' must be formed of '6+1'. This would suggest the existence of an additional term for 'one' (*-gō/-wo). Two separate forms are attested for 'hundred'. Apart from that, the Dangme and Ga numeral systems are quite homogeneous.

The Adampe system is in many respects different, so there may be doubts as to whether it indeed belongs together with Dangme. The Adampe evidence will be treated later in this chapter.

### 3.2.2 Gbe

The reconstruction of the Proto-Gbe numeral system is straightforward, since alternative forms are few (Table 3.67). It is based on the available evidence from twelve of the Gbe dialects.

Table 3.67: Proto-Gbe numerals and patterns (*)

| | | | |
|---|---|---|---|
| 1 | è-ɖe/ɖe-kpo | 7 | 'hand'+2, 5+2 |
| 2 | è-ve/e-wè | 8 | e-ɲí, 'hand'+3 |
| 3 | è-tɔ̀ | 9 | 8+1, 5+4 |
| 4 | è-nè | 10 | e-wó, *bula |
| 5 | à-tɔ́ɔ̃ | 20 | 10*2, ko |
| 6 | à-dɛ́/zɛ̃́ | 40 | e-kà |
| 100 | 40*2+20 | 1000 | à-kpé, kotokũ |

The Gbe term for 'six' is primary. Its form, however, differs significantly from the (also primary) one attested in the languages of the Ga-Dangme group.

The term for 'eight' seems to be derived from 'four', whereas the term for 'nine' follows the pattern '8+1'.

The forms for 'twenty' follow the pattern 'X*2' in Aja (*bulaa-ve*), Waci-Gbe (*blá-ve*) and Ewe (*blá-vè*), which suggests an alternative form for 'ten' (**bula*).

The etymological relationship between the term for 'fifteen' and a lexical root with the meaning 'foot' attested in two of the dialects is an apparent innovation: Maxi-Gbe *à-fɔ̀-tɔ́* ('foot', '3') and Kotafon-Gbe *fɔ́-tɔ̀* ('foot','3'). This pattern is attested in a number of the NC languages (including Atlantic).

A primary term for 'forty' is distinguishable (hence '50=40+10', '60=40+20', '70=40+30', '80=40*2', '90=40*2+10').

### 3.2.3 Ka-Togo

Ka-Togo is a quite diverse group of the Left Bank languages. The reconstructions for each of its three branches are provided in the table below (Table 3.68). Its rightmost column lists forms and patterns that are the most likely candidates for the Proto-Ka-Togo reconstruction.

Table 3.68: Proto-Ka-Togo numeral system (**)

|  | *Avatime-Nyangbo | *Kebu-Animere | *Ikposo-Ahlo-Bowili | **Proto-Ka-Togo |
|---|---|---|---|---|
| 1 | o-le | ʈɛ́-ì, bɛ-ɹi | ɛ̀-dɪ/ɛ̀-dɪ-gbo | di |
| 2 | ɛ-bha | din/ji | ɛ̀-va /ɛ̀-fwa | bha, din |
| 3 | ɛ-ta | tha | ɛ̀-ta/ɛ̀-la | ta |
| 4 | ɛ-nɛ́ | nie | ɛ̀-na | na/nɛ |
| 5 | ɛ-tí, ɛ-cu | thu(ŋ) | ɛ̀-tɔ | tu(N) |
| 6 | golo/holo | kòrɛ̃̀ŋ | ɛ̀-gɔlu/ɛ̀-wɔlu | golo/ koro |
| 7 | 6+1 | 10-3 | 6+1, kɔ̀nɔ̀, ù-zòni | 6+1 |
| 8 | 10-2? a-nsɛ | 4*2 | ɛ̀-lɛ?,<4 | 4*2, nsɛ/lɛ? |
| 9 | 10-1? zi+3? | 5+4? | 8+1, 10-1? | 8+1? 10-1 |
| 10 | kɛ-fɔ | the | wa/wu, i-jo, *bula | fo/wo, te, bula |
| 20 | 10*2 | 10*2? | bula-2, lye-2, ŋué-2, tɛ́ɛ́yá? | 10*2 |
| 100 | a-lafa (< Ewe) | tùùrù, sala | gbɔwa | lafa? |
| 1000 | a-kpe (< Ewe?) | lààfā | a-kpe | a-kpe |

It needs to be stressed that the forms marked with /**/are only suggestive and should not be taken at face value. They are not reconstructions in the strict sense and only serve for comparative purposes, so the absence of a tonal marker in a reconstructed form should not be considered meaningful. It only shows that at this point the available evidence does not allow reconstructing a tone in the pertinent case.

### 3.2.4 Na-Togo

An overview of numerical terms as attested in the branches of Na-Togo and some isolated languages is provided below (Table 3.69). A tentative reconstruction of the Na-Togo numeral system can be found in the rightmost column.

Table 3.69: Proto-Na-Togo numeral system (**)

|  | Adele | Anii | *Lelemi | *Likpe-Santrokofi | Logba | **Proto-Na-Togo |
|---|---|---|---|---|---|---|
| 1 | è-kí | dĭŋ, *mi | ù-nwi/ɔ̀-wɛ̃ | nɔ̀ɛ́/nwîî (lɛ̀wɛ́) | i-kpɛ | **i-wɛ/kpɛ? , di(N)?** |
| 2 | è-nyɔ̀ɔ̀n | ī-ɲīȭ | í-ɲɔ́ | ɲɔ́/núɔ̀ | i-nyɔ | **i-nyɔ** |
| 3 | à-sì | ī-rīū | è-tɛ | tiɛ́ | i-ta | **i-ta** |
| 4 | è-nàà | ī-nāŋ | í-na | na | i-na | **i-na** |
| 5 | tòn | ī-nȭŋ | è-lɔ | nɔ́ | i-nú | **i-no(N)** |
| 6 | kòòròn | ī-kōlōŋ | è-ku | kua | i-gló | **golo/kolo, ku** |
| 7 | 6 + 1 | kūlūmī (6+1?) | 4+3? | 6+1? | 6+1 | **6+1** |
| 8 | nìyè | 4PL | 4PL? | 4PL? | 4PL | **4PL** |
| 9 | yè-1 | tʃīīnī | 10-1 | nase | X-1 | **10-1** |
| 10 | fò | tə̄b | vu/we | fo/wo? | u-ɖú | **fo, ɖu, təb** |
| 20 |  |  | 10*2 | 10*2 |  | **10*2, ɔ-ɖɔ(n), ā-kōō, dìkpìlìn** |
| 50 | 20*2+10 | 20-PL+10 | ti | 10*5 | 10*5 | **20*2+10** |
| 100 | 20*5 | 20*5, gā-sə̄wā | 50*2, lafa | kò-lòfá | u-ga | **20*5, lofa, u-ga** |
| 1000 | 200*5 | ū-fə̄lə̄, kōtōkū | pim, ka-kpi | kɔ̀-kpí | a-kpi | **a-kpi, pim?** |

The Lelemi term for 'fifty' (*lì-tì*) is peculiar because it is a likely source of 'hundred': *è-tì á-ɲɔ́* ('50*2').

### 3.2.5 Nyo

The Nyo group, which is comprised of dozens of languages, is the most representative within the family. For this reason (even though the Nyo numeral systems are closely related to each other) they will be studied separately (by sub-group) and then compared to each other.

#### 3.2.5.1 Agneby (Abbey, Abiji, Adioukru)

Alternative sources representative of these three languages are quoted below (Table 3.70). Significant variation of forms is sporadically attested.

Table 3.70: Proto-Agneby numeral system (*)

|  | Abbey1 | Abbey2 | Abiji1 | Abiji2 | Adioukru1 | Adioukru2 | *Proto-Agneby |
|---|---|---|---|---|---|---|---|
| 1 | ŋ̀kpɔ̄ | ŋ̀kpɔ̄ | ń ˈnɔ́ | ŋ́nɔ̀ | ɲâm | ɲâm | N-kpɔ, ɲ-âm, *a-ri |
| 2 | āɲɔ́ | āɲɔ̌ | áá ˈnɔ́ | áānʊ̄ | yóɲ | ɲóɲ | a-ɲʊ/nʊ̄ |
| 3 | ārí | ārí | ɛ̃́ɛ̃́ ˈtɪ̃́ | ɛ̃́ɛ̃̄tī | ɲâhn̩ | ɲâhǹ | a-ti(N)/ ri |
| 4 | ālɛ́ | àlɛ́ | ã́ã́ ˈlã̌ | ã́ãlā | yâr | jâr | a-nɪ́/la, jar |
| 5 | ōnɪ́ | ōní | ɛ̃́ɛ̃́ ˈnɛ́ | éēnē | yên | jên | o-ne,lɔ̀hʊ̀̃, jên |
| 6 | lɔ̀hʊ̀̃ | lɔ̀hɔ̌̃ | náhɔ̀à | náhɔ̌̃ã̀ | nɔ̂hǹ | nɔ̂hǹ | hu(n) |
| 7 | lɔ̀hʊ̀̃-árī | lɔ̀hm̃ã̄rí | nɔ̃́bɔ̀ | nɔ̃́ᵐbɔ̀ | lɔ́bɪ̀ | lɔ́bɪ̀ | 6+1, bu(n) |
| 8 | èpyè | èpʲè | nówò | nówò | níwǹ | níwǹ | è-pyè, wo(n) |
| 9 | ɲâkó | ɲāàkó | nɛ̃́ ˈbrɛ̃́ | nɛ̃́ᵐbrɛ̀ | líbárm̩ | líbárm̩ | bare(-n) |
| 10 | ènʊ̰̀ | ǹnɛ̀ | ńdíɔ̀ | ńdíɔ̀ | lêw | lêw | nɛ(n) (< 5PL?), diw/ liw |
| 20 | ēbrá-ɲʊ̀̃ | òbrāɲɔ̌̃ | àbrúáɪ̰́ | àbrǘáɪ̰́ | líkɪ̀ | líkɪ̀ | <'hand' *2?,li-kŋ |
| 100 | yā | jā | yǎ | jǎ | ékɪ̀-yén | ékŋ jên (20*5) | ja, 20*5 |
| 1000 | àkpī | àkpī |  | àkpǐ |  | fándí (Engl.?) | a-kpi |

The presence of the primary terms for 'seven', 'eight' and 'nine' is an important

characteristic of this sub-group.

### 3.2.5.2 Attié

Internal reconstruction of the Attié numeral system yielded the following results (Table 3.71):

Table 3.71: Attié numeral system (*)

| 1 | kə(n) | 7 | nson |
|---|-------|------|-------------------|
| 2 | mwə(n) | 8 | ma-4? 2 de 10? |
| 3 | ha(n) | 9 | ŋgwan |
| 4 | dʒí(n) < *kɥe? | 10 | kɛŋ |
| 5 | bə(n) | 20 | 'hand' (bwa?)*2? |
| 6 | mu(n) | 100 | ja |
| | | 1000 | a-kpi |

### 3.2.5.3 Awikam-Alladian

No numerical terms (except for 'one' and 'nine') are reconstructable on the sub-group level. This raises doubts as to whether these languages should indeed be grouped together. A representation of the pertinent forms is presented in the table below (Table 3.72) and may serve as a starting point for further discussion.

Table 3.72: Awikam-Alladian numerals

| | Awikam | Alladian | Awikam-Alladian | | Awikam | Alladian | Awikam-Alladian |
|---|--------|----------|-----------------|------|--------|----------|-----------------|
| 1 | ɛ́tɔ̰́ | ɛ̄tɔ̰̀ | **ɛ-to̰** | 7 | ɛ́byɔ́ | ɛ̄bwɛ̀ | **ɛ́-byɔ́, ɛ̄-bwɛ̀** |
| 2 | áɲɔ̰́ | ā̰yrɛ̀ | **á-ɲɔ̰, ā̰-yrɛ** | 8 | ɛ̀tyɛ́ | ēɥrì | **ɛ̀-tyɛ́, ē-ɥrì** |
| 3 | ázá̰ | ā̰ò̰ | **á-zá̰, ā̰-ò̰** | 9 | ɛ́mrɔ̰́ | ɛ̄mwrɔ̀ | **ɛ́-mrɔ̰́** |
| 4 | àná̰ | ā̰zò | **à-ná̰, ā̰-zò** | 10 | èjú | ɛ̄và | **è-jú, ɛ̄-và** |
| 5 | àɲú | ɛ̄nrì | **à-ɲú, ɛ̄-nrì** | 20 | èvé | ɛ̄ɥá, *ɛ̄kòɥì | **è-vé, ɛ̄-ɥá** |
| 6 | áwá | ɛ̄wrè | **á-wá, ɛ̄-wrè** | 100 | àkpá ˈ-ɲú | 20*5 | **20*5, àkpá ˈ-ɲú** |

**3.2.5.4 Potou-Tano**

3.2.5.4.1 Potou

The following forms are distinguishable in the Potou sub-group (Table 3.73):

Table 3.73: Potou numerals

|   | Ebrie | Mbato | *Potou |   | Ebrie | Mbato | *Potou |
|---|-------|-------|--------|---|-------|-------|--------|
| 1 | bɛ̰̀ /brɛ̀ | lóɓō | **bɛ̰̀ /brɛ̀, ló-ɓō; ce/se** | 7 | ákʰwácʰè | óɓīsɛ̰́ | **6+1** |
| 2 | mɔ̰̀ | óno̰ɔ́ | **no̰ɔ́** | 8 | áɓyá | ógɓī | **ɓyá/ gɓī** |
| 3 | ɓwàɗyá | néjē /nójē | **ɗyá/je** | 9 | áɓrɔ̀ | ótrū | **ɓrɔ̀, trū** |
| 4 | ɓwèɗí | nḛ́ní/nó̰ní | **ɗi/ni** | 10 | áwɔ́ | ówā | **wɔ** |
| 5 | mwàná̰ | nḛ́nā̰ | **nā̰** | 20 | ápʰɛ̰̀ | ópɛ̰̄ | **pɛ** |
| 6 | ákʰwá | ókoā | **kwa** | 100 | àyà | yǎ | **ya** |

3.2.5.4.2 Tano

The Tano branch consists of nearly thirty languages. It seems reasonable to treat them by sub-groups.

**3.2.5.4.2.1. Western Tano**

Table 3.74: Western Tano numerals

|      | Abure1    | Abure2   | Eotile | **Western Tano**      |
|------|-----------|----------|--------|-----------------------|
| 1    | okuè      | ókúè     | ìkʊ̀   | **o-kue**             |
| 2    | aɲù       | áɲṵ̂     | àɲɔ́   | **a-ɲu(n)**           |
| 3    | nŋà       | ŋ́ŋâ     | àhá̰   | **n-ha(n)**           |
| 4    | nnàn      | ńnâ̰     | ànɛ̰̀  | **n-na(n)**           |
| 5    | nnú       | ǹnú     | ànù    | **n-nu(n)**           |
| 6    | ncɪ̀ɛ̀   | ɲ́cɪ́ɛ̀  | àhɪ́ɛ̀ | **n-cíɛ̀/híɛ̀**       |
| 7    | ncʊ̀n     | ɲ́cʊ̰̂   | àfà    | **n-cʊ̀n, à-fà**      |
| 8    | mɔ̀kʊ̀ɛ́  | mɔ̀kúè   | ànɛ̀mrɔ̀ | **mɔ̀-kʊ̀ɛ́, à-nɛ̀mrɔ̀** |
| 9    | puáléhʊ̀n | pʊ̀àlɔ̀hɔ̰̂ | brúkú | **puáléhʊ̀n, brúkú**  |
| 10   | óblún     | òbùlṵ́   | èdí    | **ò-bùlú, è-dí**      |
| 20   | ɛ́fɪ́n    | ɛ́fɪ̰̀   | èfè    | **ɛ́-fɪ(n)**          |
| 100  | ὲvá okuè  | ὲyǎ kűè  | átá    | **ὲ-vá /ὲ-yǎ, átá**   |
| 1000 | akpí okuè |          |        | **a-kpi**             |

### 3.2.5.4.2.2. Central Tano    Akanic (Table 3.75):

Table 3.75: Akanic numerals

|      | Akan1 (Twi dial.)   | Akan2      | Abron1  | Abron2   | *Akanic      |
|------|---------------------|------------|---------|----------|--------------|
| 1    | baakó~              | baakő      | bakũ    | bìàkɔ́ʔ    | **ba-kó(n)** |
| 2    | ə̀bìé-ń              | mmie-nú    | mie-nu  | mḭ̀ènṵ́ʔ   | **mie-nú**   |
| 3    | ə̀bìè-sá~            | mmeɛ-nsǎ̃   | mie-nsá | mḭ̀ènzá̰ʔ  | **mie-nsá(n)** |
| 4    | à-náń               | (ɛ)náń     | nain    | ńná̰ḭ́     | **náín**     |
| 5    | ə̀-núḿ              | (e)núḿ     | num     | ǹnṵ́m     | **núḿ**      |
| 6    | ə̀-sìá~              | (e)nsĩ̂ã̂    | nsiã    | ǹzḭ̀á̰     | **sìá(n)**   |
| 7    | ə̀-sɔ́ń              | (ɛ)nsóń    | nsɔ     | ǹzɔ̰́ɔ̰́     | **só(n)**    |
| 8    | à-wòtɕɥé /tw/       | nwɔtwé     | ŋɔt͡ʃwie | wɔ̀cɥí    | **twé/cué**  |
| 9    | à-króń              | (ɛ)nkróń   | ŋkrɔŋ   | ǹgɔ́nɔ́   | **n-króń**   |
| 10   | dú                  | (e)dú      | du      | dúʔ      | **dú**       |
| 20   | ə̀dùònú             | aduonú     | edu enu | àdừ̰ònṵ̀   | **10*2**     |
| 100  | ɔ̀hà                | ɔha        | ɔha     | hà       | **ɔ-ha**     |
| 1000 | àpíḿ                | apéḿ       | apim    |          | **a-píḿ**    |

### 3.2.5.4.2.3. Bia    

The numeral systems in these languages (Agni, Baoule, Sefwi, Nzema, Ahanta, and Jwira-Pepesa) are virtually identical and can be described as follows (Table 3.76):

Table 3.76: Proto-Bia numeral system (*)

| 1 | ko(n)          | 7    | su(n)           |
|---|----------------|------|-----------------|
| 2 | nu, ɲɔ̀(n)      | 8    | cʋɛ/twɛ         |
| 3 | sa(n)          | 9    | ǹgɔ̀lɑ̃̀, nkróń  |
| 4 | na(n)          | 10   | bulu            |
| 5 | nu(n)/nu(m)    | 20   | 10*2            |
| 6 | sia(n)         | 100  | ya              |
|   |                | 1000 | akpi            |

**3.2.5.4.2.4. Guang** This sub-group has two branches, Southern and Northern Guang which consist of four and eleven languages, respectively). Despite, the Guang numeral systems do not differ significantly, hence quoting individual forms seems unreasonable. Our reconstructions for both branches, as well as the general Guang reconstruction, are given below (Table 3.77):

Table 3.77: Guang numerals

|      | *Northern Guang | *Southern Guang | **Guang |
|-----:|------------------|------------------|---------|
| 1    | kɔ́              | kɔ               | **kɔ** |
| 2    | ɲɔ́              | ɲɔ́              | **ɲɔ́** |
| 3    | sá              | sa(n)            | **sa(n)** |
| 4    | ná              | nɛ(n)/na         | **na(n)** |
| 5    | nú(n)           | nu/ni            | **nu(n)** |
| 6    | síyé            | siɛ(n)           | **siɛ(n)** |
| 7    | súnɔ́            | súnɔ̃            | **súnɔ(n)** |
| 8    | bùrùwá, kwé     | twi/cwi          | **bùrùwá, kwé/cwi** |
| 9    | kpɔnɔ, sàngɔ́ɔ́ʔ | kpunɔ            | **kpunɔ, sàngɔ́ɔ́ʔ** |
| 10   | dú              | du               | **du** |
| 20   | o-ko, 10*2      | 10*2             | **10*2, ko?** |
| 100  | lafa (< Akan?)  | ɔ̀lɔ̀fɛ́/lafa    | **lafa** |
| 1000 | kpíŋ, pim       | a-kpe            | **kpi(N), pim** |

**3.2.5.4.2.5. Krobu; Basilia-Adele; Ega**    To make our presentation complete, the evidence of these three isolated Tano languages is presented in the table below (Table 3.78):

Table 3.78: Numerals in Tano isolated languages

|      | Krobu          | Basilia-Adele        | Ega        |
|------|----------------|----------------------|------------|
| 1    | kɔ́             | kɔ̃̂, li/diŋ          | ì-lō-ɓó    |
| 2    | ń-ɲɔ́           | ɲṵ́à                  | ì-ɲɔ̀       |
| 3    | ń-sá̰           | sa                   | ì-tà       |
| 4    | ń-ná̰           | na                   | ì-lὲ       |
| 5    | ń-nṵ̀           | ton, nun             | ì-ŋwè      |
| 6    | ń-sỹɛ̰̄         | koron                | 5+1        |
| 7    | ń-sô̰           | 6+1?                 | 5+2        |
| 8    | mɔ̀-kwέ         | 4-4, cɰ́έ            | 5+3        |
| 9    | ɲ̀-grɔ̀ā̰        | -1, gwalan           | 5+4        |
| 10   | brú            | fo, teb, bulu        | ì-zù       |
| 20   | à-brūā̰ɛ́ (10*2?) | dikpilin, koo, bulV | ú-glū      |
| 100  | yǎ             | 20*5                 | 20*5       |
| 1000 |                | kpen?                |            |

### 3.2.6 Proto-Kwa

Intermediate reconstructions suggested above should be compared in order to reconstruct the forms of the Proto-Kwa numerals. It seems reasonable to group potentially related forms (or patterns) together. The rightmost column contains isolated forms attested in one particular group only.

#### 3.2.6.1 'One'

Table 3.79: Kwa stems for '1'

|  | 1 | 1 | 1 | 1 |
|---|---|---|---|---|
| *Ga-**Dangme** | ká-kē, *go/wo |  |  | é-kòmé |
| *Gbe | ḍe-kpo | è-ḍe |  |  |
| *Ka-Togo |  | di |  |  |
| *Na-Togo | i-wɛ/kpɛ? | di(N)? |  |  |
| *Nyo: |  |  |  |  |
| *Agneby | N-kpɔ | *a-ri |  | ɲ-âm |
| Attié | kə(n) |  |  |  |
| Awikam |  |  | ɛ́-tɔ́ |  |
| Alladian |  |  | ɛ̄-tò̰ |  |
| *Potou-Tano* |  |  |  |  |
| Potou | *ce/se |  |  | bɛ̰̀ /brɛ̀, ló-ɓō |
| *Tano* |  |  |  |  |
| Western | o-kue |  |  |  |
| *Central* |  |  |  |  |
| Akanic | ba-kó(n) |  |  |  |
| Bia | ko(n) |  |  |  |
| Guang | kɔ |  |  |  |
| Krobu | kɔ́ |  |  |  |
| Ega | ì-lō-gɓó | ì-lō-gɓó (< *li-kpo?) |  |  |

The Awikam-Alladian term for 'one' is definitely an innovation.

The root *di is attested in four branches out of five and thus is likely reconstructable at the Proto-Kwa level.

The forms given in the left column are more problematic. Each of them contains a velar consonant (the Potu form *ce may have resulted from the palatalization of a velar before a front vowel, *ce < *kue* – cf. Western Tano).

Regular phonetic correspondences between these languages have not been established and therefore cannot be used for purposes of reconstruction. In any

case, the following considerations might prove useful for the NC reconstruction. The inventory of forms attested in the eighty Kwa idioms may seem rather diverse. However, only two of them may be considered for the Proto-Kwa reconstruction, namely *di* and *k(p)o* (or the compound form *di-kpo* suggested by the Gbe (*ɖe-kpo*) and Ega (*li-gɓó?*) forms).

### 3.2.6.2 'Two'

Table 3.80: Kwa stems for '2'

|              | '2'        | '2'   | '2'     | '2'  |
|--------------|------------|-------|---------|------|
| *Ga-Dangme  | é-ɲɔ̀(n)    |       |         |      |
| *Gbe        |            |       | è-ve/e-wè |     |
| *Ka-Togo    |            | din   |         | bha  |
| *Na-Togo    | i-nyɔ      |       |         |      |
| *Nyo*        |            |       |         |      |
| *Agneby    | a-ɲʊ̠/nɔ̄   |       |         |      |
| Attié        |            |       | mwə(n)  |      |
| Awikam       | áɲɔ̰́       |       |         |      |
| Alladian     |            | ā̰yrɛ̀ |         |      |
| *Potou-Tano* |            |       |         |      |
| Potou        | no̠ɔ̰́      |       |         |      |
| *Tano*       |            |       |         |      |
| Western      | a-ɲu(n)    |       |         |      |
| *Central*    |            |       |         |      |
| Akanic       | mie-nú     |       |         |      |
| Bia          | nu, ɲɔ̀(n)  |       |         |      |
| Guang        | ɲɔ́        |       |         |      |
| Krobu        | ɲ́-ɲɔ̰́     |       |         |      |
| Ega          | ì-ɲɔ̀      |       |         |      |

The only form reconstructable at the Proto-Kwa level is evidently *ɲɔ*.

### 3.2.6.3 'Three' and 'Four'

Table 3.81: Kwa stems for '3' and '4'

|  | '3' | '4' | '4' |
|---|---|---|---|
| *Ga-Dangme | é-tɛ̃ |  | é-ɟwɛ̀ |
| *Gbe | è-tɔ̃̀ | è-nɛ̀ |  |
| *Ka-Togo | ta | na/nɛ |  |
| *Na-Togo | i-ta | i-na |  |
| *Nyo |  |  |  |
| *Agneby | a-ti(N)/ri | a-ní̃/la | jar |
| Attié | ha(n) |  | dʒí(n) <* kɥe? |
| Awikam | ázá̃ | àná̃ |  |
| Alladian | ā̰ò̰ |  | ā̰zɔ̀ |
| Potou-Tano |  |  |  |
| Potou | ɗyá/je | ɗ̃ĩ/ni |  |
| Tano |  |  |  |
| Western | n-ha(n) | n-na(n) |  |
| Central |  |  |  |
| Akanic | mie-nsá(n) | náín |  |
| Bia | sa(n) | na(n) |  |
| Guang | sa(n) | na(n) |  |
| Krobu | ń-sá̰ | ń-ná̰ |  |
| Ega | ì-tà̃ | ì-lɛ̀ |  |

Just as in the majority of the NC branches, the roots for 'three' and 'four' are the most persistent. Suggested Proto-Kwa reconstructions are *ta* and *na* respectively.

### 3.2.6.4 'Five'

Table 3.82: Kwa stems for '5'

|  | '5' | '5' | '5' |
|---|---|---|---|
| *Ga-Dangme |  | é-nùɔ̃ |  |
| *Gbe | à-t́ɔ́ɔ̃ |  |  |
| *Ka-Togo | tu(N) |  |  |
| *Na-Togo |  | i-no(N) |  |
| *Nyo |  |  |  |
| *Agneby |  | o-ne | lɔ̀hʋ̃̀, jên |
| Attié |  |  | bə(n) |
| Awikam |  | àɲú |  |
| Alladian |  |  | ɛ̄nrì |
| Potou-Tano |  |  |  |
| Potou |  | na̱ |  |
| Tano |  |  |  |
| Western |  | n-nu(n) |  |
| Central |  |  |  |
| Akanic |  | núḿ |  |
| Bia | nu(n)/nu(m) |  |  |
| Guang |  | nu(n) |  |
| Krobu |  | ń-nù |  |
| Ega |  | ì-ŋwɛ̀̃ |  |

The root *tan ('five') is only traceable in the Left Bank languages. Another root, commonly attested in other languages (*nun), is found in these languages as well. Both roots should be considered for the reconstruction (note that the former is comparable to the pertinent form reconstructed for Proto-Bantu).

### 3.2.6.5 'Six'

Table 3.83: Kwa stems for '6'

|  | '6' | '6' | '6' | '6' |
|---|---|---|---|---|
| *Ga-Dangme |  | é-kpà |  |  |
| *Gbe |  |  | à-dɛ́̃/zɛ́̃ |  |
| *Ka-Togo | golo/koro |  |  |  |
| *Na-Togo | golo/kolo | ku |  |  |
| *Nyo |  |  |  |  |
| *Agneby |  | hu(n) |  |  |
| Attié |  |  |  | mu(n) |
| Awikam |  |  |  | áwá |
| Alladian | ɛ̄-wrè |  |  |  |
| Potou-Tano |  |  |  |  |
| Potou |  | kwa |  |  |
| Tano |  |  |  |  |
| Western |  |  | n-cíɛ̀/híɛ̀ |  |
| Central |  |  |  |  |
| Akanic |  |  | sìá(n) |  |
| Bia |  |  | sia(n) |  |
| Guang |  |  | siɛ(n) |  |
| Krobu |  |  | ń-sỹ̰ɛ̰ |  |
| Ega |  |  |  | 5+1 |

The evidence presented in Table 3.83 is inconclusive. At this stage our task is to process the complex Kwa data so that it can be compared to the evidence of other NC languages. In this respect, three provisional Kwa forms are noteworthy: *golo/kolo, *kua,* and *ciɛ.* In any case, as the forms for 'seven' suggest, the Proto-Kwa term for 'six' was probably primary.

**3.2.6.6 'Seven'**

Table 3.84: Kwa stems and patterns for '7'

|  | '7' | '7' | '7' | '7' |
|---|---|---|---|---|
| *Ga-Dangme | 6+1 | | | |
| *Gbe | | | | 5+2, 'hand'+2 |
| *Ka-Togo | 6+1 | | | |
| *Na-Togo | 6+1 | | | |
| *Nyo* | | | | |
|   *Agneby | 6+1 | | bu(n) | |
|   Attié | | nson | | |
|   Awikam | | | ɛ́byɔ̰́ | |
|   Alladian | | | ɛ̄bwɛ̰̀ | |
|   *Potou-Tano* | | | | |
|     Potou | 6+1 | | | |
|     *Tano* | | | | |
|       Western | | n-cùn | | |
|       *Central* | | | | |
|         Akanic | | só(n) | | |
|         Bia | | su(n) | | |
|       Guang | | súnɔ(n) | | |
|       Krobu | | ń-sô | | |
|       Ega | | | | 5+2 |

The forms presented in the table above point toward the pattern '6+1' being used for the Proto-Kwa term for 'seven', whereas Proto-Nyo developed the primary term *sun*.

### 3.2.6.7 'Eight'

Table 3.85: Kwa stems and patterns for '8'.

|  | '8' | '8' | '8' | '8' | '8' |
|---|---|---|---|---|---|
| *Ga-Dangme |  |  |  |  | 6+2 |
| *Gbe |  | e-ɲí | 'hand'+3 |  |  |
| *Ka-Togo | 4*2 | nsɛ/lɛ? |  |  |  |
| *Na-Togo | 4PL |  |  |  |  |
| *Nyo |  |  |  |  |  |
| *Agneby |  |  |  | è-pyè | wo(n) |
| Attié | ma-4? |  |  |  | 10-2? |
| Awikam |  | ɛ̀tyɛ́ |  |  |  |
| Alladian |  | ēɥrì |  |  |  |
| Potou-Tano |  |  |  |  |  |
| Potou |  |  |  | ɓyá/gɓī |  |
| Tano |  |  |  |  |  |
| Western |  | mɔ̀-kʊ̀ɛ́ |  |  | à-nɛ̀mrɔ̰̀ |
| Central |  |  |  |  |  |
| Akanic |  | twé/cué |  |  |  |
| Bia |  | cʊɛ/twɛ |  |  |  |
| Guang |  | kwé/cwi |  |  |  |
| Krobu |  | mɔ̀-kwɛ́ |  |  |  |
| Ega |  |  | 5+3 |  |  |

Based on the evidence attested in the table above, the Proto-Kwa term for 'eight' may be reconstructed as either primary (*kwe/ kye*) or derivative, in which case it must have been based on 'four' (*'4PL').

**3.2.6.8 'Nine'**

Table 3.86: Kwa stems and patterns for '9'

| | '9' | '9' | '9' | '9' | '9' | '9' |
|---|---|---|---|---|---|---|
| *Ga-Dangme | | | | | | nɛ̌ɛ́(hɛ́) |
| *Gbe | 8+1 | | 5+4 | | | |
| *Ka-Togo | 8+1? | | 10-1 | | | |
| *Na-Togo | | | 10-1 | | | |
| *Nyo* | | | | | | |
| *Agneby | | bare(-n) | | | | |
| Attié | | | | | ŋgwan | |
| Awikam | | ɛ́mrɔ̰́ | | | | |
| Alladian | | ɛ̄mwrɔ̀ | | | | |
| *Potou-Tano* | | | | | | |
| Potou | | ɓrɔ̀ | | | | trṹ |
| *Tano* | | | | | | |
| Western | | brúkú | | | | puálɛ́hʊ̀n |
| *Central* | | | | | | |
| Akanic | | | | n-króń | | |
| Bia | | | | nkróń | ǹgɔ̰̀lɑ̰̀ | |
| Guang | | | | | | kpunɔ, sàngɔ́ɔ́ʔ |
| Krobu | | | | | ǹ-grɔ̀ɑ̄ | |
| Ega | | | 5+4 | | | |

This is the hardest form to interpret. A rare pattern '8+1' is attested in the Left Bank languages. In contrast to this, the Togo pattern is '10-1', while the Nyo term (*brɔ*/*mrɔ*) is 'primary'. The latter is probably connected to the term for 'ten', although this connection does not necessarily imply a derivation ('10-1') and could be explained by analogy. All three forms/patterns are considered for reconstruction.

### 3.2.6.9 'Ten'

Table 3.87: Kwa stems for '10'

| | '10' | '10' | '10' | '10' | '10' | '10' |
|---|---|---|---|---|---|---|
| *Ga-Dangme | | | | | | ɲɔŋmá |
| *Gbe | e-wó | *bula | | | | |
| *Ka-Togo | fo/wo | bula | | | te | |
| *Na-Togo | fo | | ɗu | | təb | |
| *Nyo* | | | | | | |
| *Agneby | | | | diw/liw | | nɛ(n)<5PL? |
| Attié | | | | | | kɛŋ |
| Awikam | | | èjú | | | |
| Alladian | ɛ̄-và | | | | | |
| *Potou-Tano* | | | | | | |
| Potou | wɔ | | | | | |
| *Tano* | | | | | | |
| Western | | ò-bùlṹ | | è-dí | | |
| *Central* | | | | | | |
| Akanic | | | dú | | | |
| Bia | | bulu | | | | |
| Guang | | | du | | | |
| Krobu | | brú | | | | |
| Ega | | | ì-zù | | | |

Isolated forms are attested in Ga-Dangme and Attié. The root *tə(b)* is traceable in the Ghana–Togo Mountain languages (Togo-remnant) and is not found elsewhere. Thus we are dealing with another isogloss suggesting that these languages belong to the same branch. The stem *\*du* supported by R. Blench could be proposed for Proto-Kwa. This stem is indeed attested in the majority of the groups that do not belong to the Left Bank languages (including Na-Togo).

The stem *\*bula* (Left Bank)/*\*bulu* (Tano) is distributed fairly evenly.

Finally, a Niger-Congo root reflected in Kwa as *\*fo/wo* can be reconstructed in a number of languages.

### 3.2.6.10 'Twenty'

Table 3.88: Kwa stems and patterns for '20'

|  | '20' | '20' | '20' | '20' | '20' | '20' |
|---|---|---|---|---|---|---|
| *Ga-Dangme | 10*2 | | | | | |
| *Gbe | 10*2 | ko | | | | |
| *Ka-Togo | 10*2 | | | | | |
| *Na-Togo | 10*2 | ā-kōō | dìkpìlìn | | | ɔ-ḍɔ(n) (<10?) |
| *Nyo* | | | | | | |
|   *Agneby | 'hand' (bra)*2? | | li-kŋ | | | |
|   Attié | 'hand' (bwa?)*2? | | | | | |
|   Awikam | | | | ɛ̀-vɛ́ | | |
|   Alladian | | *ɛ̄kòɥì | | ɛ̄-ɥá | | |
|   *Potou-Tano* | | | | | | |
|    Potou | | | | | pɛ̱ | |
|    *Tano* | | | | | | |
|     Western | | | | | ɛ́-fɪ(n) | |
|     *Central* | | | | | | |
|      Akanic | 10*2 | | | | | |
|      Bia | 10*2 | | | | | |
|     Guang | 10*2 | ko? | | | | |
|     Krobu | à-brūāɛ̱́ (10*2?) | | | | | |
|    Ega | | | | | | ú-glū |

The pattern '10*2' attested in the majority of the branches. The root *ko is also to be taken.

### 3.2.6.11 'Hundred' and 'thousand'

Table 3.89: Kwa stems and patterns for '100' and '1000'

|  | '100' | '100' | '100' | '100' | '1000' | '1000' |
|---|---|---|---|---|---|---|
| *Ga-Dangme | làfá | | ò-há | | à-kpé | |
| *Gbe | | | | 40*2+20 | à-kpé | |
| *Ka-Togo | lafa? | | | | a-kpe | |
| *Na-Togo | lofa | 20*5 | u-ga | | a-kpi | pim? |
| *Nyo | | | | | | |
| *Agneby | | 20*5 | ja | | a-kpi | |
| Attié | | | ja | | a-kpi | |
| Awikam | | | | àkpá ˈ-2 | | |
| Alladian | | 20*5 | | | | |
| Potou-Tano | | | | | | |
| Potou | | | ya | | | |
| Tano | | | | | | |
| Western | | | ὲ-vá /ὲ-yǎ | átá | a-kpi | |
| Central | | | | | | |
| Akanic | | | ɔ-ha | | | a-píḿ |
| Bia | | | ya | | a-kpi | |
| Guang | lafa | | | | kpi(N) | pim |
| Krobu | | | yǎ | | | |
| Ega | | 20*5 | | | | |

In addition to the pattern '20*5', the roots *lafa/lofa* and *ya/ja* (Nyo) are reconstructable for 'hundred'. The latter may be etymologically related to *ga/ha*.

The term for 'thousand' is commonly attested as *a-kpi*. Its less common byform is *pim*.

The following table lists provisional Proto-Kwa reconstructions based on the evidence discussed above:

Table 3.90: Proto-Kwa numeral system (*)

| | | | |
|---|---|---|---|
| 1 | di-kpo | 7 | 6+1 |
| 2 | ɲɔ, **di? | 8 | 4PL, kwe/kye |
| 3 | ta | 9 | 10-1? |
| 4 | na | 10 | fo/wo, bula, du |
| 5 | nu(n), ton | 20 | 10*2, ko |
| 6 | golo/kolo, kua, ciɛ | 100 | 20*5, lofa, ja/gya? |
| | | 1000 | kpi, pim |

The remaining roots and patterns are probably innovations that developed separately within a branch/language. They may help to adjust the internal classification of the Kwa languages.

## 3.3 Ijo

According to traditional classification, the Ijo family is comprised of the Ijaw languages and the Defaka language. Some scholars express doubts as to whether the latter indeed belongs to this family. According to Roger Blench, "The Ijo languages constitute a well-founded group, but the membership of Defaka (constituting Ijoid) remains problematic. Defaka has numerous external cognates and might be an isolate or independent branch of Niger-Congo which has come under Ijo influence" (Blench 2013).

Ijaw languages consist of the Eastern and the Western groups (the latter is sometimes called Central).

The following reconstruction is based on the evidence of all three Ijo branches (Table 3.91):

Table 3.91: Proto-Ijo numeral system

|  | Defaka | *East | *West | **Ijo |
|---|---|---|---|---|
| 1 (qualifying) | gbérí | gbérí | ? | ? |
| 1 (counting) | ? | ŋgèi | kɛnɪ | *n-kèni |
| 1 in 6 (5+1) | – | die/ie | die/zie | *die |
| 2 | mààmà | màmì | maamʊ | *mamV |
| 3 | táátó | tárú | tărʊ | *tató |
| 4 | nɛ̀ì | i-neĩ | nɛ́ín/nóín | *nɛ́ín |
| 5 | túúnɔ̀ | sónɔ́ | sɔ̃nɔ̃-rɔ̃ | *túnɔ́ |
| 6 | mààngò | 5+1 | 5+1 | *5+1 |
| 7 | 5+2 | 5+2 | 5+2 | *5+2 |
| 8 | 5+3 | 4+4 | 4+4 | *4+4 |
| 9 | 5+4 | 5+4 | 5+4? | *5+4 |
| 10 | wóì | ójí /àtìé | ójí | *(w)ójí |
| 15 | 10+5 | jìé | dié | *dié |
| 20 | sîì | sí | síí | *síí |

Both qualifying and counting terms for 'one' are attested in the Eastern Ijo languages (e.g. in Ibani). The Defaka form may be a borrowing. An unexplained allomorph for 'one' is attested as a part of the term for 'six' in Ijaw (?).

The root for 'two' (*mam*) is an Ijo innovation. It has no parallels outside this language family. Its phonetic similarity to several other forms is a mere coincidence, e.g. *ma-* in the Jaad (Atlantic) *maaɛ* does not belong to the root and can be

explained as a class prefix. The lexical meaning 'twin, pair' (as attested in Nembe (East) according to (Kaliai 1964)) may underlie the Ijo term. However, no reliable parallels for this term with the meaning 'twin, pair' are establishable in NC.

The root for 'three' is apparently of NC origin, with its most archaic form attested in Defaka.

The term for 'four' is undoubtedly a reflex of the NC root.

The term for 'five' probably goes back to the NC root *tan(o)*. As in the case of 'three', its most archaic form is found in Defaka.

The terms for 'six', 'seven', and 'nine' follow the common patterns ('5+1', '5+2', and '5+4' respectively).

The Ijaw term for 'eight' must have derived from 'four' by means of partial reduplication (*ni-nɛ́ín*). This pattern is reconstructable on the Proto-NC level and will be discussed at length in the next chapter.

A specific counting term for 'ten' is reconstructable in the Eastern Ijo languages (*àtìé*). The Defaka form is comparable to those found in the Ijaw languages.

A special form for 'fifteen' is reconstructable in Ijaw (*dié*), cf. e.g. the Nembe evidence: *dìé-èsí* '300' (='15*20'). This form may go back to Ijaw *díɛ̀* 'divide; separate into parts; split or break up into parts; share', 'distribute, donate', cf. Nembe *dìɛ̀*, Ibani (Koelle 1963[1854]) *dìɛ̀-, dìé.*

As in a number of other languages that belong to different families within NC, a special form is attested for the term for 'twenty' (*síí*). The term itself has several functions. It serves as a basis for a number of other terms for tens (also in Defaka), e.g. '40=20*2', … '100=20*5'. The Ijaw terms for 16-19 are based on it as well, e.g. '16=20−4', etc.

## 3.4 Kru

Our analysis of the Kru numerals is based on nearly forty sources representative of five major groups and eleven major subgroups of the family. Preliminary reconstructions of the pertinent numerical terms (by sub-group) are represented in commented tables below.

### 3.4.1 'One', 'Two' and 'Three'

Table 3.92: Kru stems for '1'-'3'

|  | '1' | '1' | '1' | '2' | '2' | '3' |
|---|---|---|---|---|---|---|
| Aizi |  | mu̱mɔ̱ | yre | i-ʃɪ |  | i-ta |
| *Eastern* |  |  |  |  |  |  |
| Bakwe/Wané | ɗô |  |  | sɔ̂ |  | ta |
| Bete/Godié |  | ɓlo/gbolo |  | sɔ |  | ta |
| Dida/Neyo |  | bolo |  | sɔ́ |  | ta |
| Kodia |  | gbɣlɣ/ɓɣlɣ |  | sɔː |  | taː |
| Kuwa | dee |  |  | sõr |  | tãã̀ |
| Seme | dyuɔ̃ |  | byɛ́ɛ̃ |  | nĩ | tyáār |
| *Western* |  |  |  |  |  |  |
| Bassa[14] | doo | (g)boo? |  | sɔ̃́ |  | tã |
| Grebo[15] | do(o) |  |  | sɔ̃̌ | hwɛ̃/hɔ̃ | taa(n) |
| Klao/Tajuasohn | do |  |  | son |  | tan |
| Wee[16] | due/too |  |  | sɔn |  | taan |

As in the majority of the NC languages the term for 'three' is the most persistent: the root *taa(n)* can be reliably reconstructed for Proto-Kru.

The same is applicable to the root for 'two' reconstructed as *so(n)* in Proto-Kru (isolated forms are attested in the Seme and Grebo sub-groups only). It should be noted that in general the Seme numeral system is peculiar in many respects. These peculiarities (e.g. Seme being the only language with a full set of primary terms covering the sequence from 'one' to 'ten') may be due to the isolated status of the language. In his recent article entitled "Le sèmè/siamou n'est pas kru" Vogler argues that Seme is not a Kru language (see Vogler 2015). On the basis of a comparison between Kru, Gur and Mande (Samogo) morphology and lexicon he concludes that Seme is either remotely related to the Mande languages or represents a separate branch of Niger-Congo. As we hope to demonstrate below, Seme shows systematic correspondences with neither Kru nor Mande (including the contact Mande languages – Samogo and Jowulu).

---

[14]Bassa, Dewoin, Gbii.

[15]Grebo, Krumen, Glio-Oubi.

[16]Wee is a Western Kru group which includes (among other languages) Sapo, Krahn, Nyabwa, Wobe.

'One'. It is likely that the root *\*do* should be reconstructed on the Proto-Kru level. However, there is enough evidence for reconstructing the alternative root *\*(g)bolo.*

### 3.4.2 'Four' and 'Five'

Table 3.93: Kru stems for '4' and '5'

|  | '4' | '4' | '4' | '5' | '5' | '5' |
|---|---|---|---|---|---|---|
| Aizi |  |  | yeɓi | yu-gbo |  |  |
| *Eastern* |  |  |  |  |  |  |
| Bakwe/Wané |  | hɪɛ̃⁴ | mɾɔ̄: | ɠbɔ̀ɔ̄, ŋʷũ |  |  |
| Bete/Godié |  |  | mʊ-wana | gbu/gbi |  |  |
| Dida/Neyo | na |  |  | gbɪ́ |  |  |
| Kodia | na |  |  | ⁿgbɤ |  |  |
| Kuwa | ɲìjèhɛ |  |  |  |  | wàyɔ̀ɔ |
| Seme |  |  | yur |  |  | kwɛ̃̀l |
| *Western* |  |  |  |  |  |  |
| Bassa | hɪ̀-nyɛ(n) |  |  |  | h-mm |  |
| Grebo |  | hɛn |  | gbə | mm | hun |
| Klao/Tajuasohn | nyìɛ̀ | hɛn |  |  | mù, hoom? (< m?) |  |
| Wee | nyìɛ |  |  |  | mm |  |

The forms for 'four' in the left column apparently are the reflexes of the NC root that is preserved in its archaic form *\*na* in Eastern Kru, whereas in Western Kru it changes into *nyìɛ̀*.

Two major forms are observable for 'five', namely *\*gbə/ gbo* and *\*mm* (Western).

### 3.4.3 'Six' to 'Nine'

Table 3.94: Kru stems and patterns for '6'-'9'

|  | '6' | '6' | '7' | '7' | '8' | '8' | '8' | '9' | '9' | '9' |
|---|---|---|---|---|---|---|---|---|---|---|
| Aizi |  | fɔ | fri+2 |  |  |  | patɛ |  |  | fi |
| *Eastern* |  |  |  |  |  |  |  |  |  |  |
| Bakwe/Wané | 5+1 |  | 5+2 |  | 5+3 |  |  | 5+4 |  |  |
|   Bete/Godié | 5+1 |  | 5+2 |  | 5+3 |  |  | 5+4 |  |  |
|   Dida/Neyo | 5+1 |  | 5+2 |  | 5+3 |  |  | 5+4 |  |  |
|   Kodia | 5+1 |  | 5+2 |  | 5+3 |  |  | 5+4 |  |  |
| Kuwa | 5+1 |  | 5+2 |  | 5+3 |  |  | 5+4 |  |  |
| Seme |  | kpãâ |  | kĩ̂ |  |  | kprɛ̃n̂ |  |  | kɛ̄l/kal |
| *Western* |  |  |  |  |  |  |  |  |  |  |
|   Bassa | 5+1 |  | 5+2 |  | 5+3 |  |  | 5+4 |  |  |
|   Grebo | 5+1 |  | 5+2 |  | 5+3 |  |  | 5+4 |  |  |
|   Klao/Tajuasohn | 5+1 |  | 5+2 |  |  | 4PL |  |  | 10-1 |  |
|   Wee | 5+1 |  | 5+2 |  | 5+3 |  |  | 5+4 |  |  |

It is immediately apparent that these numerals already followed the pattern '5+X' in Proto-Kru. As noted above, the Seme forms are innovations.

### 3.4.4 'Ten' and 'Twenty'

Table 3.95: Kru stems for '10' and '20'

|  | '10' | '10' | '20' | '20' | '20' |
|---|---|---|---|---|---|
| Aizi | bɔ |  | gu |  |  |
| *Eastern* |  |  |  |  |  |
| Bakwe/Wané | pʊ̀, bu? |  | grʊ̀, gˀlɔ |  |  |
| Bete/Godié |  | kʊ́gba | gwlʊ́ /gɔlɔ |  |  |
| Dida/Neyo |  | kʊ́gba | glʊ́/góló |  |  |
| Kodia |  | kʊgba | ǧalo |  |  |
| Kuwa |  | kowaa |  | 10*2 |  |
| Seme | fu |  |  |  | kār |
| *Western* |  |  |  |  |  |
| Bassa | ɓaɖa-bùè, |  |  | <10 |  |
|  | puuɛ, vu |  |  |  |  |
| Grebo | pu |  | gōrō/wlʊ̀ |  |  |
| Klao/Tajuasohn | pue/punn |  | wlòh-2 |  | quilar-2 |
| Wee | pue/bue |  | gwlʊ-2 |  | kwela 2 |

The root *kʊgba* is attested beside the common NC root for 'ten' (*pu/fu*) in Eastern and Kuwa. The root for 'twenty' is attested as *golo* in both Eastern and Western.

### 3.4.5 'Hundred' and 'Thousand'

Table 3.96: Kru stems and patterns for '100' and '1000'

|  | '100' | '100' | '1000' | '1000' | '1000' |
|---|---|---|---|---|---|
| Aizi |  | juyugbo |  |  |  |
| *Eastern* |  |  |  |  |  |
| Bakwe/Wané | 20*5 |  | 400*2+20*10 |  |  |
| Bete/Godié | 20*5 |  | 400*2+200 |  |  |
| Dida/Neyo | 20*5 |  | 400*2+200 |  |  |
| Kodia |  |  |  |  |  |
| Kuwa |  | kɔlɛh? |  | 100*10 |  |
| Seme | 20*5 |  |  |  | lit: 'goat one' |
| *Western* |  |  |  |  |  |
| Bassa | 20*5 |  |  |  | borrowed |
| Grebo | 20*5 |  |  |  | borrowed |
| Klao/Tajuasohn | 20*5 |  |  |  | borrowed |
| Wee | 20*5 |  |  |  | ? |

All Kru sub-groups are characterized by the lack of a primary term for 'hundred'.

The form for 'thousand' in Western Kru was borrowed from the Mande languages. A primary term for '400' (*$dwi$) that developed in the Eastern Kru served as the basis for a rare pattern for 'thousand' attested in these languages ('400*2+200').

The reconstruction of the Proto-Kru numeral system is as follows:

Table 3.97: Proto-Kru numeral system (*)

| | | | |
|---|---|---|---|
| 1 | do, (g)bolo | 7 | 5+2 |
| 2 | so(n) | 8 | 5+3 |
| 3 | taa(n) | 9 | 5+4 |
| 4 | na | 10 | pu, kʊgba? |
| 5 | gbə/gbo, mm | 20 | golo |
| 6 | 5+1 | 100 | 20*5 |
| | | 1000 | 400*2+200 |

## 3.5 Kordofanian

The evidence of about twenty Kordofanian languages does not permit reconstructing the Proto-Kordofanian numeral system (assuming that Proto-Kordofanian existed). Comprehensive data for each of the four major groups is represented below (Table 3.98). Forms and patterns traceable in at least two groups are in bold. The forms are grouped within the lines in a more or less ad hoc manner, e.g. there is no special reason to believe that Talodi *lu(k)/ li(k)* 'one' corresponds to the forms with initial **t-**/**ʈ-** attested in other groups.

The systematic presence of the final velar -**k** in some of the terms can also be found in the Atlantic languages (especially in North Atlantic).

The term for 'ten' appears in numerous forms in the Kordofanian languages, which is rare. At the same time, no root for 'ten' is represented in at least two languages simultaneously. Moreover, nearly every language in a group has its own term for 'ten'.

A primary term for 'eight' is distinguishable[17] in the Heiban and Rashad languages.

---

[17]I used data from the following Kordofanian languages and dialects: Aceron, Dagik, Heiban, Jomang, Katla, Koalib, Lafofa, Laro, Logol, Lumun, Moro, Nding, Orig, Rere, Shirumba, Tagoi, Talodi, Tegali, Tegem, Tima, Tira, Tocho, Utoro, Warnang.

Table 3.98: Kordofanian numerals 1–5

|   | *Heiban | *Katla | *Rashad | *Talodi | *Kordofanian |
|---|---------|--------|---------|---------|--------------|
| 1 | kwɛ-(ʈ)ʈɛ(k) | ʈí-ʈʌk | -tta | lu(k)/li(k) | **ʈe(k)/ lu(k)** |
| 1 | ŋɔ-(ʈ)ʈɔ | ʌ-ʈeen/ʈɪɪn |  |  | **ʈɔ(n)** |
| 1 | *-lel? |  |  | tleidi | **lel/ led?** |
| 2 |  | cik/heek | (k)ko(k) |  | **kok/kek/ cik** |
| 2 | -can /-ɽan, rɔm |  |  | we-ɽʌk/-tta | (can/ɽan, rak, rɔm) |
| 3 | tɔɽɔl/ʈeɽel | ʈʌʈ | tta | wa-ʈʈak | **tat/tɔ̀ɽ/ ʈak** |
| 3 | -ɽɪcɪn/-gɪtʃɪn | i-hwʌy |  |  | (ɽitin/ɽicin, hwʌy) |
| 4 | k(w)ɔ-ɽɔŋɔ/ma-ɽŋan/-rlon/-ɬɽʊ |  | ya-rem/wa-rʊm | -ɽandɔ | **-ɽɔŋ/-ɽandɔ/-ranto/-rʊm?** |
| 4 |  | ʌ-gʌlʌm/i-hʌlʌm |  | kekka | (-gʌ́lʌm, kekka) |
| 5 | tʊ-dìní/-ðɛnɛ | i-duliin |  |  | **dinin/ dulin?** |
| 5 | ŋer-/ɲer- |  | *ɲer- |  | **ŋer-/ ɲer-** ('hand', …) |
| 5 |  | ɟɔ-gbəlɪn | wʊ-ram, ma | 'hand'-'1', ki-liəgum |  |

Table 3.99: Kordofanian numerals >5

|      | *Heiban | *Katla | *Rashad | *Talodi | *Kordofanian |
|------|---------|--------|---------|---------|--------------|
| 6 | 5+1 | <5 | ɲere(-r/-l/-y) (< *5+1?) | 5+1 | **5+1** |
| 6 | 3+3? 3PL | | | | (3+3) |
| 7 | 5+2 | 5+2 | 5+2 | 5+2 | < A5+2 |
| 7 | 4+3 | 3PL+1 | | | (4+3, 3PL+1) |
| 8 | duuba(ŋ) | | dubba/tuppa | | **dubba** |
| 8 | 5+3, 4redupl.? | | | 5+3, 4 redupl. | **5+3**, 4 **redupl.** |
| 8 | bɔ | t̪ʌ́ŋgìl/t̪ŋɛrɛy | | | (bɔ, t̪əɲi-) |
| 9 | 10-1 | 10-1 | 10-1 | | **10-1** |
| 9 | 5+4 | ɟʌ́lbʌ̀t̪ín (<5?) | | 5+4 | **5+4** |
| 10 | di/ɗi/ri | *t̪ʊʌ, ɔ-rɔ | kʊ-man (5PL) | ma-tu(l) | ? |
| 10 | | rakpac, i-hedʌkun | fəŋən (fə-ŋən?) | tiəɽum, n̪ipɽa, gurruŋ) | ? |
| 20 | 10*2 | 10*2 | 10+10 | 10*2 | 10*2 |
| 20 | tuɽí ('grain'), 'big figure' | | | 'body', (a-rial, a-(na)ttu) | ('body', …? ) |
| 100 | 20*5, < Arabic | 10*10 | 10*10 | 10*10, 20*5 | 10*10, 20*5 |
| 1000 | rabic, 20*2*10 | absent | 10*10*10 | ɑ-ðɑɾ | ? |

## 3.6 Adamawa

Adamawa is the most divergent of the NC families. The variety of numeral systems attested in the Adamawa languages confirms this statement. This can be observed not only in cases of forms that belong to different groups, but often within groups and sub-groups as well, which makes the reconstruction of its numeral system quite problematic. In other words, it is not a rare case that small Adamawa branches consisting of only a pair of languages show incomparable forms. Some examples are in order here.

Let us compare the terms from 'one' to 'ten' in the Kim branch that is commonly attributed to the Mbum-(Day) group (Greenberg 14) (Table 3.100):

Table 3.100: Numerals in the Kim branch

|    | Besme | Kim |
|----|-------|-----|
| 1  | mōndā/mbírāŋ | ɗú |
| 2  | tʃírí | zí |
| 3  | hā̰sī (hā̰-sī?) | tā |
| 4  | ndày | ndà |
| 5  | ndìyārá | nūwḛ̄y |
| 6  | māngùl | mènèngāl |
| 7  | ɗīyārā | ɓēálā/ɓēálār |
| 8  | ndā-sì (4+3?) | tīmāl/wá-zì-zí(10-2) |
| 9  | nòmīnā | làmāɗō/wá-zì-ɗú (10-1) |
| 10 | wàl | wòl |

Only the terms for 'four', 'six', and 'ten' are comparable in these systems.

The Longuda language constitutes a separate branch of Waja-Jen (Greenberg 10). The table below gives an overview of the first ten numerical terms as attested in two dialects of Longuda (Table 3.101). The evidence for both dialects was collected by the same scholar (Ulrich Kleinewillinghöfer[18]). Morphological analysis of the forms is given according to Longurama of Koola (Longuda1) and Wala Lunguda (Longuda2).

Although we are dealing with two dialects of the same language, the roots for 'one', 'two', 'three', 'six', and 'ten' attested in them are different. At the same time, the terms covering the sequence from 'six' to 'nine' follow patterns com-

---

[18]https://mpi-lingweb.shh.mpg.de/numeral/Niger-Congo-Adamawa.htm

Table 3.101: Longuda numerals

|    | Longuda1                        | Longuda2                  |
|----|---------------------------------|---------------------------|
| 1  | laa-twè                         | naa-khal                  |
| 2  | nàà-kwɛ̃́                        | naaa-shir                 |
| 3  | nàà-tsə́r                        | naa-kwáí                  |
| 4  | nèé-nnyìr                       | naa-nyìr                  |
| 5  | nàà-nyɔ́                         | nàà-nyó                   |
| 6  | tsààtə̀n                         | na-khí-nà-kwáí (2*3)      |
| 7  | í-néé-nyìr i-nàà-tsə́r( 4+3)     | nyi-na-kwáí (4+3)         |
| 8  | nyíí-tìn (<4?)                   | nyí-thìn (<4?)            |
| 9  | é-nàà-nyɔ́ í-néé-nyìr( 5+4?)     | nyi-na-nnyó (4+5)         |
| 10 | koo                             | nɔ̂m                      |

monly attested elsewhere. Thus the differences between these dialects appear to be greater than those between the languages within Mande or Bantu families. This raises the question as to whether a Proto-Kim or Proto-Longuda reconstruction is indeed relevant.

Moreover, the reconstruction is additionally hindered by the fact that numerical terms in the majority of the Adamawa languages are subject to the alignment by analogy more frequently than in other NC languages. General considerations regarding this problem can be found in Chapter 2. This is of special significance for the Adamawa languages since it affects etymological interpretations. The evidence from a number of languages belonging to the Duru sub-group of Leko-Nimbari (Greenberg 4) may serve as a case study (Table 3.102).

Table 3.102: Duru numerals

| Peere      | Doyayo     | Gimme     | Gəunəm     | Vɔmnəm        | Momi      | Longto    |
|------------|------------|-----------|------------|---------------|-----------|-----------|
| 1 dɵ́ə     | gbúnú      | wɔɔna     | mani       | màn           | muzoz     | wɵ́ŋŋá     |
| 2 iro      | ééré́      | idtɨgè    | tɛk        | ètên          | ìttɵ́z     | sittó     |
| 3 tã́ãro    | taarɛ      | taagè     | taarək     | tāán          | tàáz      | tã́ãbó     |
| 4 naro     | násɔ       | náágè     | náárə́k     | nānnò         | náz       | nabbó     |
| 5 núuno    | nooné      | nɔɔnɨ̀gè   | nɔɔnɔ̀k     | gbà náárò     | gbanáá    | nɔ̃ɔ̃mó    |
| 6 nón-dɵ́ə | nɔ̀ɔn-gbúnú | nɔngè     | nɔɔ-waŋgə  | gbāā-sɵ̀ mâl  | bámbɵ́z    | sáámɛ     |

Matching final segments of the first few numerical terms in each of these languages are highlighted in red. I agree with Larry Hyman that "it might not be analogy, rather the use of a marker" (p.c.) but it should be noted that though these segments are different in each case (i.e. they do not match even within a pair of languages), they are present in each language under discussion.

In Mumuye-Yandang, which is another branch of Leko-Nimbari (Greenberg 5), an additional sub-morpheme (-t) is attested that is not present in Duru (Table 3.103):

Table 3.103: Analogical alignments in Mumuye-Yandang

|   | Mumuye | Bali | Yendang (dial.) |
|---|--------|------|-----------------|
| 2 | zi**ti** | i-ye | í-nī |
| 3 | ta:**ti** | taa**t** | tâ:**t** |
| 4 | dɛ̀:**tì** | naa**t** | nâ:**t** |

The following conclusions with regard to the Proto-Duru numeral system can be reached upon the basis of this evidence. First, the final segments (whatever their phonetic difference) should not be viewed as a hinderance to the comparison of numerical terms. This means that Momi *tàáz* 'three' can (and should be) compared to Longto *tã̃ãbó*. The question of whether their final segments should be analysed as morphemes or sub-morphemes is of secondary importance for our purposes. At the same time, the quality of the second consonant in Proto-Leko-Nimbari is obscure, so we have to reconstruct the form as *\*taaX*, where X is an unknown consonant.

As demonstrated above, numerical terms are exceptionally divergent within the family. In addition to this, systematic (diversified) alignment by analogy is often employed in the languages under study. Both factors make the reconstruction a challenging task, even though an attempt at reconstruction of the Adamawa numerals by a highly competent scholar is available (see Boyd 1989). His results, however, are of limited relevance for our comparative purposes, as the following example shows. According to Boyd, the Proto-Adamawa term for 'one' is to be reconstructed as *\*ku-di-n* (the root *\*di*) with *\*kwin* being its later development. His ideas on how this proto-form is reflected in particular branches of the Adamawa family are summarized in the table below (Table 3.104). Notations in the first column refer to Grinberg's grouping of the Adamawa languages.

Even if Boyd's reconstruction of the Proto-Adamawa form is correct, a diachronic interpretation that impies an etymological relationship between *bim-*

Table 3.104: *kwin- reflexes in Adamawa according to Boyd

|     | *Proto | Reflexes |
| --- | --- | --- |
| **G1** | kwin | kun |
| **G1** | kwin | kwaan |
| **G2** | kwin | gu-(a)s(a) |
| **G4** | kwin | gun, gbun, bin, wun-ga, guu |
| **G5** | kwi(t) | gbet, gorV |
| **G5** | kwin | in(d)i |
| **G6** | kwin-k | soŋ |
| **G7** | kwin | indi > fa-ndi |
| **G8** | kwin-kwin | bimbimi |
| **G8** | kwi(n) | gwi > ju |
| **G9** | kwin | tsuŋ/tsiŋ, cɔŋ |
| **G10** | kwi-t | > kwat > kal |
| **G13** | kwit | ɓuru, gulu |
| **G13** | kwit | > kwat > bara(k) |
| **G13** | kwin | ʈoŋ |
| **G14** | kwin | ɗu |
| **Day** | kwin-k | ngoŋ |
| **Day** | kwin | (k)wan > mɔn |

*bimi*, *cɔŋ*, *ɗu* and *gbet* does not fit the purpose of our integral comparative study of NC numerical terms because it can be used to justify nearly any etymological connection. In view of this, the Adamawa numerical terms will be treated in the same way as those from the preceding language families. First, the main forms of the numerical terms will be established, with no attempt at tracing them down to a provisional proto-form. Then the numeral systems of each of the Adamawa branches will be studied separately. Finally, an integral analysis of the available evidence pertaining to each of the terms will be offered. This approach will enable us to treat the Fali languages and even Laal together with the Adamawa languages, although their relationship to the latter is often questioned (in the case of Laal, doubts are raised as to whether it belongs to NC at all).

### 3.6.1  Fali-Yingilum (G11)

Table 3.105: Fali-Yingilum numerals103

| 1 | kpolo/bʌlo (< *lo?) | 7 | jɔ̠rɔs |
|---|---|---|---|
| 2 | cuk, gbara | 8 | 4 redupl. |
| 3 | taan (< taaX) | 9 | 10-1/ŋgʌs kàm(kàn) k͡pòlò 'rest hand one' |
| 4 | naan | 10 | ra |
| 5 | kɛ̃rɛw | 20 | 10*2 |
| 6 | yira/yilo | 100 | < Fula |
|  |  | 1000 | < Fula |

It should be noted that after a nasal, -*r*- in the Fali forms regularly corresponds to -*N*- in those of Yingilum, cf. '5' Fali *kɛ̠rɛw* ~ Yingilum *kɛ́ɲàu*, '7' *jɔ̠rɔs* ~ Yingilum *jɔ́nɔ̀s*. An alignment by analogy is probably attested in the terms for 'three' and 'four' (*taaX* > *taan* may have changed by analogy with *naan*).

### 3.6.2 Kam (Nyimwom, G8)

Table 3.106: Kam numerals

| | | | |
|---|---|---|---|
| 1 | b<u>ii</u> (Meek: bimbini) (< *b-ii?) | 7 | jùp yi-raak (6,2 - 'second six'?) |
| 2 | yi-raak (i-ra) | 8 | sâl |
| 3 | càr | 9 | níízaa |
| 4 | nár (< *naX) | 10 | bóò |
| 5 | ŋwún | 20 | kpáímí ,*ǹkpǫ́ |
| 6 | jù:p | 100 | 20*5 |
| | | 1000 | ? |

Within the NC context, a reversive alignment by analogy may be considered: *na*X '4' > *nar* by analogy with *car* '3'. As Boyd rightfully observes, in the case of 'one' it is often unclear whether the initial consonant is a part of the root, or a reflex of the noun class prefix.

The term for 'seven' simulates the pattern '7=6+2' (this phenomenon is not infrequent in NC). Sometimes (e.g. in some of the Mande languages) this impression is due to the fact that the term for 'six' originally derived from '5+'. Over time, an innovation replaced the original term for 'five', which was only preserved in the derived term for 'six'. Alternatively, the term for 'seven' could be explained as 'the other six' (or 'a big six' is some languages), as perhaps in Kam, assuming that *jù:p* does not go back to the term for 'five'.

### 3.6.3 Leko-Duru-Mumuye (G4, G2, G5)

#### 3.6.3.1 Duru (G4)

Table 3.107: Duru numerals

| 1 | dɔ́ə, gbúnú, wɔ́-ŋŋá/wɔɔna/dá(ŋ)gá/*nge, man(i)/*mal | 7 | 5+2, (gútambe, 6+'odd', dɔ́msàrà, 4+3) |
|---|---|---|---|
| 2 | du/ru/to, te/re | 8 | 4PL/4+4, 5+3,( < Hausa) |
| 3 | tãã tó/tãã ro | 9 | ' one finger is left ', níŋsínè, 5+4, 10-1 |
| 4 | nató/naró (< *naX) | 10 | bōʔ, kob/kop/fób |
| 5 | núno/nɔɔnì, gbà náárò/gbanáá, sáá | 20 | gbɛg/gbàhsɨ́ (='staff'), *wɔ́ɔ́g ('head'), zul/zur ('head'), (10*2, ráárò, jùgúyɔ), |
| 6 | gúú, 5+1 | 100 | tɛmere < Fula, 20*5 |
|  |  | 1000 | uzinere < Fula, (dukə) |

This table provides an overview of forms and patterns attested in eleven sources for this sub-group. This degree of variety is not normally attested within a single sub-group, which raises doubts as to whether these languages should be grouped together.

#### 3.6.3.2 Leko (G2)

Our study of this sub-group is based on the evidence of two languages. The summary table above is not descriptive of the language-specific mechanisms of the alignment by analogy. An overview of the numerical terms covering the sequence from 'two' to 'five' by language is provided below (Table 3.4. 107):

Apparently, the terms from 'three' to 'five' in these two languages are related to each other. At the same time, two groups of terms ('2-3' and '3-4') with an

---

[19]The group is often labeled Leko-Nimbari. Here we follow Raimund Kastenholz and Ulrich Kleinewillinghöfer, who note that 'The term 'Nimbari' should not to be used as a classificatory term, nor should the scarce and surely in large parts erroneous data be given central significance in any comparative approach to Adamawa languages' (Kastenholz & Kleinewillinghöfer 2012).

Table 3.108: Leko numerals

| 1 | nín̄a/níiá (<ŋa?) | 7 | 5+2 |
|---|---|---|---|
| 2 | nnú, ra?, *-i? | 8 | 5+3, < Hausa |
| 3 | toorà/toonú | 9 | 5+4,' one is left ' |
| 4 | naarà/nɛɛr-əb | 10 | kób/kóp |
| 5 | núúnà/núnn-ub | 20 | nɛd níi gbɛd, laa-1 |
| 6 | nɔ̂ŋgɔ̂s/núŋgɔ́ɔs | 100 | 20*5, < Fula |
| | | 1000 | 20*10?,< Fula |

Table 3.109: Analogical alignments in two Leko languages

| | Kolbila (Zurá) | Samba Leko |
|---|---|---|
| 2 | in**nú** | ii**rà** |
| 3 | too**nú** | too**rà** |
| 4 | nɛɛr**əb** | naa**rà** |
| 5 | núnn**ub** | núú**nà** |

alignment by the ultima are observable in Kolbila. This is applicable to a group of Samba Leko terms as well, namely '2-4' (possibly also '5'; the fact that the Samba Leko terms are adjusted by both the vowel quality and the tone is noteworthy). This means that the seemingly unrelated roots for 'two' may have derived from a common etymon (still unknown to us) by means of alignment by analogy. The source form of 'two' remains obscure. Assuming that it was similar to the one reconstructed for the Duru sub-group (e.g. *ru), it is likely that the same form is to be reconstructed for Leko as well: *ru > Kolbila *nu* by analogy with *toonu* '3' ; *ru > Samba Leko *rà* by analogy with *toorà* '3'. However, the evidence in favor of this reconstruction is inconclusive. Alternatively, the initial vowel of the term for 'two' (***ii-/in-**) may reflect the source root, while the final segment is potentially explained via an alignment by analogy with '3'.

### 3.6.3.3 Mumuye-Yandang (G5)

This sub-group is represented by three languages that show different forms of 'two'. The terms for 'three' and 'four' are adjusted by analogy. Studying them in a wider NC context reveals that the final consonant in 'four' was adjusted by anal-

Table 3.110: Numerals in Mumuye-Yandang

| 1 | ɓīntī/ɓini (*< nti/ni?) , gbétè | 7 | 5+2 |
|---|---|---|---|
| 2 | ziti, ye, nī | 8 | 5+3 |
| 3 | taat | 9 | 5+4 |
| 4 | naat | 10 | kop/kob |
| 5 | mǎ:ni, nɔng/ghìnān | 20 | mba-1, kar-1, mim-1 |
| 6 | 5+1 | 100 | 20*5 |
| | | 1000 | derived |

ogy with 'three'. The alignment itself must have occurred already at the Proto-Mumue-Yandang level, which explains our provisional reconstructions suggested for this proto-language in the table above.

No evidence pertaining to the Nimbari numerals is available to us. The forms of 'one' given by Boyd (Boyd 1989) are noteworthy (Nimbari *(n)yeme/ geme/ (ʒeme?)*).

### 3.6.4 Mbum-Day (G13, G14, G6, Day)

#### 3.6.4.1 Bua (G13)

This is very divergent branch that has been poorly documented. I'd like to thank Pascal Boyeldieu who has provided me with his personal data on Ɓa (Bua) and Lua (Niellim), as well as some other rare sources. The main forms and patterns are shown in Table 3.111.

Numerals in the Bua group can be presented as follows (Table 3.112)

#### 3.6.4.2 Kim (G14)

The first ten terms of Besme and Kim are given in the table above (Table 3.100). The term for 'twenty' in these languages follows the pattern '10*2', whereas the Kim term for 'hundred' is borrowed from Arabic. The Besme term for 'hundred' is borrowed from the French *sac* 'sack', whereas the term for 'thousand' is borrowed from Bagirmi.

Table 3.111: Bua numerals

| | Fanya | Niellim | Tunya | Bua | Zan Gula | Kulaal (Gula) | Bolgo | Koke |
|---|---|---|---|---|---|---|---|---|
| 1 | do/lo | ɓúdū/ɓúrū | sèlì | gúlu | sammā, saado | ʈóŋ | ba(k)ra, silla | barak |
| 2 | i-ru/ li-ru | ndīdí/ndīrí | à-rī | i-li/í-rīː | ɾisːi/lissi | ròk | lēti, retè | lēdi |
| 3 | taro | tērí | à-tā | í-tēr | toːɾi | tòòs | teri | tēri |
| 4 | nagi/ naro | niǎ̃ní/néni | à-nā | í-pāw/paõ | naːsɪ | nòrò̧ | har | hār |
| 5 | lugni | lùní | à-lōnī | í-lwār | tɛ(r) | lúŋ | tisso | tisó |
| 6 | kaba | tá:r | nānò | tãr | 5+1 | lú-én-ʈóŋ | tipsi | dípsil |
| 7 | 5+2 | longa | lúlú | lür | 5+2 | lú-é-ròk | 5+2 | tiglén |
| 8 | <4 | 3+4', < Bagirmi | kɔ̀ntā̧ | <* 4 PL? | 5+3 | | orhor (4 redupl.), 5+3 | 4 redupl. |
| 9 | 10-X | < Bagirmi | à-tī | lór-lor | 5+4 | sàkɔ́línnòrò̧ | diar, 6+3 | jār |
| 10 | teba | < Bagirmi, hulõa | kùtù | húlil/lor filo:le/filori poo | yíppà | | do(k) | dog |
| 20 | 10*2 | doksap | 10*2 | <10PL | ʊ-faːlɛ | | a-rep, a-hun, tehu | |
| 100 | | ro/ru | à-rū | a-ru | < Arabic | míà/míè | | ae léd |
| 1000 | | dubu | dūbú | dubu | < Arabic | hálìf | | ae har |

Table 3.112: Bua numerals (summarized)

| | | | |
|---|---|---|---|
| 1 | *do, *de?, bara(k), (ʈóŋ) | 7 | 5+2, 3+4, lúlú/lòŋgɔ̄/lur, (tiglen) |
| 2 | *di, *ri?, *ru?, (ròk), (rete) | 8 | 4 redupl., 5+3 |
| 3 | tar/tori/teri | 9 | ti, jar, 5+4, 10-X |
| 4 | na/nagi/niani, har | 10 | do(k), (kùtù), (filo:le), (yíppà), (teba) |
| 5 | luni/loni/*lu,tɛ(r), *kɔn?, (tiso) | 20 | 10*2, do-ksap, fa:lɛ, (a-rep), (a-hun) |
| 6 | 5+1, tá:r, (nānò), (kaba), tipsi | 100 | ro/ru |
| | | 1000 | < Bagirmi |

### 3.6.4.3 Mbum (G6)

Table 3.113: Mbum numerals

| 1 | mbew/mbiew, bɔ̄ɔ̄ŋ/búónó/bóm/vaŋno | 7 | 10-3, rɪŋ, (rënām, tàrnɑ́̃gà) |
|---|---|---|---|
| 2 | seɗe/sere, gwa/ɓɔ̀-gë, ɓà-tì | 8 | 10-2, nama/namma/nènmàʔä |
| 3 | say | 9 | 10-1, doraŋ |
| 4 | nìŋ, nai | 10 | boo, dʒama/dʒémà, (dùɔ, hù-wàlë ) |
| 5 | ndiɓi/ndēɓē/dūwēe/dápì | 20 | 10*2, '2 hands', 10+10 |
| 6 | ze(y)/ye(a), (tɔ́tɔ́klɔ́, bì-gírò) | 100 | sɔ́ɗ/sɔt, < Fula, < Arabic |
| | | 1000 | 'sac', bag', < Fula, < Bagirmi |

This sub-group is represented by a dozen languages. Unlike Leko-Duru-Mumue no alignment by analogy is attested. Some forms of 'two' are of unclear morphological structure.

### 3.6.4.4 Day

Table 3.114: Day numerals

| 1 | ngɔ̄ŋ́, *mon | 7 | 4+3 |
|---|---|---|---|
| 2 | dĭí | 8 | 4 redupl.? |
| 3 | tà | 9 | 'lacking one' |
| 4 | ndà, *bī-yām | 10 | mɔ̀ |
| 5 | sɛ̄rì | 20 | 10*2 |
| 6 | 5+1 | 100 | tù |
| | | 1000 | < Bagirmi |

This branch is comprised of an isolated language. Its attribution to Mbum-Day has been a subject of scholarly debate. The form *mon '1' is postulated on the basis of *sɛ̄rì mòn* 'six', whereas the reconstruction of *bīyām (*bī-yām?) '4' is based on *bīyām tà* 'seven'.

### 3.6.5  Waja-Jen (G9, G10, G1, G7)

#### 3.6.5.1  Jen (G9)

Table 3.115: Jen numerals

| 1 | kwín/*ʃín/tsɨng | 7 | 5+2 |
|---|---|---|---|
| 2 | ráb/*re, bwə-ng, bwa-yung | 8 | 4PL, 5+3 |
| 3 | gbunuŋ, bwa-tə | 9 | 5+4 |
| 4 | net, bwa-nyə | 10 | ʃóób, bwa-hywə |
| 5 | nóob/*na, bwa-hmə/*hwĩ | 20 | fa-1, ngwu-1 |
| 6 | 5+1 | 100 | 20*5 |
|   |  | 1000 | ʃik-1, 20-fe |

This branch is represented by two languages: Burak and Jenjo (Dza). The evidence from this group is among Boyd's best arguments for the reconstruction of *kwin (< *ku-di-n) 'one'. The primary term *li* (*bwa-li*) 'fifteen' is attested in Jenjo. Accordingly, the term for 'sixteen' follows the pattern '15+1' (*bwali ji tsɨng*). Interestingly, in Burak the term for 'hundred' is *li* (*li kwín*).

The form *hwĩ 'five' is traceable in Jenjo compound terms covering the sequence from 'six' to 'nine' (*hwĩ-tsɨng* 'six', *hwĩ-yung* 'seven', etc.) as is the corresponding Burak form *na 'five' (*naa-ʃín* 'six', *náá-re* 'seven', *ná-tát* 'eight'). The form *re 'two' is observable in *náá-re* 'seven', whereas *ʃín 'one' is traceable in *naa-ʃín* 'six'.

#### 3.6.5.2  Longuda (G10)

The evidence for the first ten numerals in two Longuda dialects can be found in the table above (Table <span style="color:red">3.101</span>). The term for 'twenty' in these languages follows the pattern '10*2'. The forms of 'hundred' are *pùlò(wé)/phulewe*.

#### 3.6.5.3  Waja (G1)

Some languages in this sub-group are characterized by a sub-morphological alignment of the terms for 'three' and 'four' well-attested in Adamawa: Dadiya *tal* '3' ~ *nal* '4', Bangunji (dial.) 1 *táát* '3' ~ *náát* '4', Bangunji (dial.) 2 *taar* '3' ~ *naar* '4', Tula (Kɨtule) *jí-t:à* '3' ~ *já:-nà* '4'. As a result, these terms are treated as minimal contrastive pairs in the paradigm. Within the NC context, forms with the final -*t* should be considered prototypical in the case of both terms. This means

Table 3.116: Waja numerals

| | | | |
|---|---|---|---|
| 1 | w-in/d-in/kw-an/g-ɛɛn/*k-un? | 7 | ni-bir/ni-ber/ni-bil/ni-bi(y) |
| 2 | yɔ́-rɔ́b/rɔɔp/yob/yo, (su) | 8 | na-rib/na-lib/na-rub (4*2) |
| 3 | taat, kunuŋ, (bwanbí) | 9 | 10-1, teer/teet/tɔɔrɔ |
| 4 | naat, (gwár) | 10 | kɔ́b/kub/kwab/kpop/kwu |
| 5 | nu(ŋ), (fwáːd) | 20 | 10*2, '2 hands' |
| 6 | nu-kun (<5+1?) | 100 | <10?, wɔn, (bwa-tigɛ) |
| | | 1000 | kʊʊl, nèe/kú-néŋ, 100*10, bi-kate, tedu |

that *naaX 'four' (final consonant unknown) may have evolved into *naat by analogy with 'three' in Proto-Waja. Later, an innovative form for 'three' developed in Awak and Waja: Awak *kunúŋ*, Waja *kunoŋ*. The Dijim-Bwilim *bwanbí* is apparently an innovation.

Interestingly, the froms for 'six' attested throuought the sub-group resemble the Awak and Waja forms for 'three'. At the same time, the forms for 'six' can be explained as '5+1' (assuming that they include an allomorph of *kun 'one').

### 3.6.5.4 Yungur (G7)

Table 3.117: Yungur numerals

| | | | |
|---|---|---|---|
| 1 | fini/fandi/pə́ndə́ŋ (< *ndi?), wunú | 7 | nbutu |
| 2 | raap, fətə/fiicì (< *tə/ci?) | 8 | 4 redupl. |
| 3 | táákə́n/(tɑɑrə́n) | 9 | 5+4 |
| 4 | kurun | 10 | bú(u), (kutun) |
| 5 | wonon/wonun | 20 | (10*2) |
| 6 | mindike | 100 | (-ru) |
| | | 1000 | (100*10) |

The terms for 'twenty', 'hundred' and 'thousand' are attested in only one source (Kaan (Libo)) out of the eight sources available for this branch, hence they are quoted in brackets. Morphological analysis of the terms for 'one' and 'two' is unclear: *ƒV may be a reflex of the original noun class prefix.

### 3.6.6 Laal

Finally, let us turn to the Laal numeral system. Laal's attribution to the Adamawa languages (as well as its attribution to NC) is debatable. Today it is assumed that it is an isolated case within Niger-Congo. Comparative study of its numerical terms may shed light on its genealogical relationship (Table 3.118).

Table 3.118: Numerals in Laal

| | | | |
|---|---|---:|---|
| 1 | ɓɪ̀díl (ɓɪ̀-díl?) | 7 | 5+2 |
| 2 | ʔīsī (ʔī-sī?) | 8 | 4 redupl. |
| 3 | māā | 9 | yàŋjáŋ |
| 4 | ɓīsān (ɓī-sān?) | 10 | tūū |
| 5 | sāb, *swa- | 20 | 10*2 |
| 6 | cìcààn | 100 | 10-'big' |
| | | 1000 | < Baguirmi < Hausa |

As in many other NC languages, the major problem with Laal numerals is the obscurity of their morphological structure. Pascal Boyeldieu established that traces of noun class suffixes are observable in Laal forms as their comparison to SG and PL forms show (see Boyeldieu 1982). At the same time, as I tried to demonstrate elsewhere (Pozdniakov 2010), some traces of noun class prefixes had been preserved in this language as well. At this point, it seems reasonable to set the alternative variants aside for further comparison.

What follows is an attempt to synthesize the Adamawa evidence.

### 3.6.7 Proto-Adamawa

#### 3.6.7.1 'One'

The main forms are given in Table 3.119.

In accordance with Boyd's hypotheses discussed above, the forms in the first two columns may be related in view of the reconstruction of the root *\*di* (possibly also *\*-in*), the noun class prefix *\*ku-* and the suffix *\*-n* (*\*ku-di-n* '1')

The last column lists forms that are attested in one of the branches only. The roots that can be tentatively reconstructed as *\*do*, *\*nga/ngɔ*; *\*(g)bunuand* and *\*mon* are noteworthy.

#### 3.6.7.2 'Two'

The main forms of this root are quoted in Table 3.120. The grouping of forms is admittedly not substantiated enough. The variety of forms within this family is striking, even when unrestricted phonetic grouping is applied.

#### 3.6.7.3 'Three'

Comparative evidence for this root points to its reconstruction as *\*taat* (with further alignment by analogy within each of the branches). As in the other NC families, the root is exceptionally stable, in contrast to the roots for 'one' and 'two' that demonstrate a wide variety of forms. A shared innovation in Jen and Waja (attested in Burak, Awak and Waja) is noteworthy.

Table 3.119: Adamawa stems for '1'

| | '1' | '1' | '1' | '1' '1' | '1' | '1' | '1' | '1' |
|---|---|---|---|---|---|---|---|---|
| Fali | | | | | | | | |
| Kam | b-ịị | | | | | | | |
| *Leko* | | | | | | | | |
| Duru | dɔə | | | -(ŋ)gá/ -naʔ / *nge níɲa/ niíá (<ɲaʔ) | gbúnú | man(i)/*mal | | |
| Leko | ɓi-nti/ɓi-ni (*< nti/ni?) | | | | | | | |
| Mumuye | | | | | | | | gbɛ̀tɛ̀ |
| *Mbum* | | | | | | | | |
| Bua | *dɛ | *do | | | | | bara(k) | tɔ́ŋ, *si? |
| Kim | | dù | | | | | mbírãŋ | |
| Mbum | | | | | | mõndã | | |
| Day | | | | ngɔ̀ŋ | bɔ̀ɔ̀ŋ/ búónó | *mon | | mbew/ mbiew |
| *Waja* | | | | | | | | |
| Jen | kw-in/ *ʃ-in/ ts-ing (< *in) | | | | | | | |
| Longuda | w-in/ d-in/ g-ɛɛn/ *k-un? | | | | | | | |
| Waja | fi-ni/ | | | | wunú | | | |
| Yungur | fa-ndi/ pá-ndɔ̀ŋ (< *ndi?) | | | | | | | |
| Laal | ɓìdíl (ɓì-díí?) | | | | | | | khal, twɛ̀ |

*-lo

Table 3.120: Adamawa stems for '2'

| | '2' | '2' | '2' | '2' | '2' | '2' | '2' | '2' | '2' |
|---|---|---|---|---|---|---|---|---|---|
| Fali-Yingilum | | | | | | | | | |
| Kam | yi-raak (i-ra) | | | | gbara | cuk | | | |
| *Leko-Duru-Mumuye* | | | | | | | | | |
| Duru | | du/ ru, to | | | | | | | |
| Leko | ra? | ii-/ in-? | te/ re | | | | | | |
| Mumuye | | ye | | ziti | | | nmú | nì | |
| *Mbum-Day* | | | | | | | | | |
| Bua | *ru, (rɔ̀k) | di/ ri | | (rete) tʃɪrí | | | | | |
| Kim | | | zí | | | | | | |
| Mbum | | díí | ɓà-tì | sede/sere | gwa/ɓɔ̀-gɛ̈ | | | | |
| Day | | | | | | | | | |
| *Waja-Jen* | | | | | | | | | |
| Jen | ráb/ *re, | | | | | | | | |
| Longuda | | | | | | | | | |
| Waja | yɔ́-rɔ́b/ rɔɔp/ yob/ yo | | | shir | kwɛ̈ | (su) | | | |
| Yungur | raap | | fɘta/fiicì (< *tɘ/ cì?) | | | | | | |
| Laal | | | ʔisì (ʔì-sì?) | | | | | | bwə-ng, bwa-yung |

Table 3.121: Adamawa stems for '3'

| | | | | |
|---|---|---|---|---|
| Fali-Yingilum | taan (< taaX) | | | |
| Kam | càr | | | |
| *Leko-Duru-Mumuye* | | | | |
| Duru | tã́ãtó/tã́ãro | | | |
| Leko | toorà/toonú | | | |
| Mumuye | taat | | | |
| *Mbum-Day* | | | | |
| Bua | tar/tori/teri | | | |
| Kim | tā | | | hā̰sī |
| Mbum | say | | | |
| Day | tà | | | |
| *Waja-Ɉen* | | | | |
| Jen | bwa-tə | ɡbunuŋ | | |
| Longuda | tsə́r | | kwáí | |
| Waja | taat (bwanbí) | kunuŋ | | |
| Yungur | táákə́n/(tɑɑrə́n) | | | |
| Laal | | | | māā |

### 3.6.7.4 'Four'

Table 3.122: Adamawa stems for '4'

| | | | | |
|---|---|---|---|---|
| Fali-Yingilum | naan | | | |
| Kam | nár | | | |
| | (< *naX) | | | |
| *Leko-Duru-Mumuye* | | | | |
| Duru | nató/naró | | | |
| | (< *naX) | | | |
| Leko | naarà/nɛɛr-əb | | | |
| Mumuye | naat | | | |
| *Mbum-Day* | | | | |
|   Bua | na/nagi/niani | | har | |
|   Kim | | | ndà(y) | |
|   Mbum | nai | nìŋ | | |
|   Day | | | ndà | *bī-yām |
| *Waja-Ɉen* | | | | |
|   Jen | net | bwa-nyə | | |
|   Longuda | nnyìr/nyìr | | | |
|   Waja | naat | | | gwár |
|   Yungur | | | | kurun |
| Laal | | | | ɓīsān |
| | | | | (ɓī-sān?) |

The main NC form *\*na*X is predominant here, its second consonant being subject to alignment by analogy. The same root is likely to be reconstructed at the Proto-Adamawa level as well.

### 3.6.7.5 'Five'

The main root (*nun*) may be the same as in the Gur languages and may be etymologically related to the term for 'hand'. It is likely that the isolated forms quoted in the rightmost column go back to similar terms as well. The Jen root *hmə* could be a borrowing from Chadian Arabic: *xamsa* '5'. The Mbum forms *ndēɓē/ dūwēe* may be influenced by Fula (*jowi* 'five').

Table 3.123: Adamawa stems for '5'

| Fali-Yingilum | | kɛ̃rɛw | | |
|---|---|---|---|---|
| Kam | ŋwún | | | |
| *Leko-Duru-Mumuye* | | | | |
| Duru | núno/ nɔɔnɨ̀, | | | gbà náárò/ gbanáá, sáá |
| Leko | núúnà/ núnn-ub | | | |
| Mumuye | nɔng/ ghìnān | | | mǎ:ni |
| *Mbum-Day* | | | | |
| Bua | | | | luni/ loni/ *lu,tɛ(r), *kɔn?, (tiso) |
| Kim | nūwḛy | | ndìyārá | |
| Mbum | | | ndiɓi/ dūwēe/ dápì | |
| Day | | | | sērì |
| *Waja-Ɉen* | | | | |
| Jen | nóob/ *na | -hmə/ *hwĩ | | |
| Longuda | nyɔ́ | | | |
| Waja | nu(ŋ) | | | fwá:d |
| Yungur | wo-non/ wo-nun | | | |
| Laal | | | | sāb, *swa- |

**3.6.7.6 'Six'**

Table 3.124: Adamawa stems and patterns for '6'

| Fali-Yingilum | | | yira/yilo |
|---|---|---|---|
| Kam | | jù:p | |
| *Leko-Duru-Mumuye* | | | |
| Duru | 5+1 | gúú | |
| Leko | | | nɔ̂ŋgɔ̂s/núŋgɔ́ɔs |
| Mumuye | 5+1 | | |
| *Mbum-Day* | | | |
| Bua | 5+1 | | tá:r, (nānò), (kaba), tipsi |
| Kim | | | māngùl/mènèngāl |
| Mbum | | | ze(y)/ye(a), tɔ́tɔ́klɔ́, bì-gírò |
| Day | 5+1 | | |
| *Waja-Ɉen* | | | |
| Jen | 5+1 | | |
| Longuda | | tsààtə̀n | 2*3? |
| Waja | nu-kun (<5+1?) | | |
| Yungur | | | mindike |
| Laal | | cìcààn | |

The most frequently attested pattern is '5+1'. At the same time, there is a great variety of isolated forms (see the last column). The similarity between the Laal and Longuda forms is noteworthy; both may go back to Chadian Arabic *sit:e* 'six'. The Kim (and also Yungur?) form could be a borrowing from Bagirmi (*mìká* '6').

**3.6.7.7 'Seven'**

As in the case of 'six', the predominant pattern ('5+2') for 'seven' is rather plain. It co-exists with a variety of isolated forms of uncertain etymology.

Table 3.125: Adamawa stems and patterns for '7'

| Fali-Yingilum | | | | | jɔrɔs |
|---|---|---|---|---|---|
| Kam | | | 'second six' | | |
| *Leko-Duru-Mumuye* | | | | | |
| Duru | 5+2 | 4+3 | 6+'odd' | | gútambe, dɜ́msàrà |
| Leko | 5+2 | | | | |
| Mumuye | 5+2 | | | | |
| *Mbum-Day* | | | | | |
| Bua | 5+2 | 3+4 | | | lúlú/lòŋgɔ̄/lur, (tiglen) |
| Kim | | | | ɓēálā/ ɓēálār | dǐyārā |
| Mbum | | | | | 10-3, rɪŋ, rënām, tàrnɑ̋gà |
| Day | | 4+3 | | | |
| *Waja-Ɉen* | | | | | |
| Jen | 5+2 | | | | |
| Longuda | | 4+3 | | | |
| Waja | | | | ni-bir/ -bil/-bi(y) | |
| Yungur | | | | | nbutu |
| Laal | 5+2 | | | | |

### 3.6.7.8 'Eight'

Table 3.126: Adamawa stems and patterns for '8'

| Fali-Yingilum | 4 redupl. | | | |
|---|---|---|---|---|
| Kam | | | | sâl |
| *Leko-Duru-Mumuye* | | | | |
|    Duru | 4PL/4+4 | 5+3 | | < Hausa |
|    Leko | | 5+3 | | < Hausa |
|    Mumuye | | 5+3 | | |
| *Mbum-Day* | | | | |
|    Bua | 4redupl. | 5+3 | | |
|    Kim | ndāsì (4PL?) | | wázìzí (10−2) | tīmāl |
|    Mbum | | | 10-2 | nam(m)a/nènmàʔä |
|    Day | 4 redupl.? | | | |
| *Waja-Ɉen* | | | | |
|    Jen | 4PL | 5+3 | | |
|    Longuda | | | | nyíthìn |
|    Waja | 4*2 | | | |
|    Yungur | 4 redupl. | | | |
| Laal | 4 redupl. | | | |

The pattern '8=4 redupl.' is to be reconstructed at the Proto-Adamawa level.

**3.6.7.9 'Nine'**

Table 3.127: Adamawa stems and patterns for '9'

| Fali-Yingilum | | 10-1/ŋgʌs kàm(kàn) kpòlò | |
| | | 'rest hand one' | |
| Kam | | | níízaa |
| *Leko-Duru-Mumuye* | | | |
| Duru | | 'one finger is left', nɨ̀ŋsɨ́nè, 5+4, 10-1 | |
| Leko | 5+4 | 'one is left' | |
| Mumuye-Yandang | 5+4 | | |
| *Mbum-Day* | | | |
| Bua | 5+4 | 10-X | ti, jar |
| Kim | | 10-1 | nòmīnā |
| Mbum | | 10-1 | doraŋ |
| Day | | 'lacking one' | |
| *Waja-ɟen* | | | |
| Jen | 5+4 | | |
| Longuda | 5+4 | | |
| Waja | | 10-1 | teer/teet |
| Yungur | 5+4 | | |
| Laal | | | yàŋjáŋ |

A primary term for 'nine' was apparently non-existent in Proto-Adamawa. A comparison between Bua *diar* and Kanuri *ləɣár* may be suggestive if a borrowing is considered. The same applies to the terms for 'nine' in Waja (*tɔɔrɔ*) and Hausa (*tara*).

**3.6.7.10 'Ten'**

Two alternative roots for 'ten' (Table 3.128) are distinguishable (\**boo* and \**kob* attested in four and two groups respectively). The root *d(u)o* is observable in two Mbum-Day sub-groups. Finally, the root *kutu(n)* is found in two languages, namely in Tunya (Bua) and Kaan (Yungur). Assuming that *ku-* is a class prefix, this root may prove to be related to *tūū* (Laal).

Table 3.128: Adamawa stems for '10'

| | | | | | | |
|---|---|---|---|---|---|---|
| Fali-Yingilum | | | | | | ra |
| Kam | bóò | | | | | |
| *Leko-Duru-Mumuye* | | | | | | |
| Duru | bōʔ, | kob/ kop/ fób | | | | |
| Leko | | kób/ kóp | | | | |
| Mumuye | | kop/ kob | | | | |
| *Mbum-Day* | | | | | | |
| Bua | | | do(k) | kùtù | | (filo:le), (yíppà), (teba) |
| Kim | | | | | wàl/ wòl/ wàr/ *wèy | |
| Mbum | boo | | dùɔ | | hù- wàlë | dʒama/ dʒémà |
| Day | mɔ̰̀ | | | | | |
| *Waja-Ɉen* | | | | | | |
| Jen | | ʃóób | | | | bwa- hywə |
| Longuda | | koo/ kù | | | | nɔ̂m |
| Waja | | kɔ́b/ kub/ kwab/ kpop/ kwu | | | | |
| Yungur | bú(u) | | | kutun | | |
| Laal | | | | tūū | | |

### 3.6.7.11 'Twenty'

The term for 'twenty' (Table 3.129) in the Duru languages either follows the pattern '20=10*2' or goes back to the lexical roots for 'head' and 'staff'. The Niellim term *do-ksap* was likely borrowed from Bagirmi *dùg sap* 'twenty'.

Table 3.129: Adamawa stems and patterns for '20'

| | | | | |
|---|---|---|---|---|
| Fali-Yingilum | 10*2 | | | |
| Kam | | | | *ǹkpɔ́, kpáímí |
| *Leko-Duru-Mumuye* | | | | |
| Duru | 10*2 | | gbɛg/ gbàhsɨ́ ('staff'), *wɔ́ɔg ('head'), zul/ zur ('head') | ráárò, jùgúyɔ |
| Leko | | | laa-1 | nɛd níi gbɛd |
| Mumuye | | | | mba-1, kar-1, mim-1 |
| *Mbum-Day* | | | | |
| Bua | 10*2 | | fa:lɛ | do-ksap, a-rep, a-hun |
| Kim | 10*2 | | | |
| Mbum | 10*2 | '2 hands', 10+10 | | |
| Day | 10*2 | | | |
| *Waja-Ɉen* | | | | |
| Jen | | | fa-1 | ngwu-1 |
| Longuda | 10*2 | | | |
| Waja | 10*2 | 2 hands' | | |
| Yungur | 10*2 | | | |
| Laal | 10*2 | | | |

### 3.6.7.12 'Hundred'

Table 3.130: Adamawa stems and patterns for '100'

| | | | | | |
|---|---|---|---|---|---|
| Fali-Yingilum | | | | | < Fula |
| Kam | 20*5 | | | | |
| Leko-Duru-Mumuye | | | | | |
| Duru | 20*5 | | | | < Fula |
| Leko | 20*5 | | | | < Fula |
| Mumuye | 20*5 | | | | |
| Mbum-Day | | | | | |
| Bua | | | | ro/ru | |
| Kim | | | | | < Arabic |
| Mbum | | | sɔɗ/sɔt | | < Fula, < Arabic |
| Day | | | tù | | |
| Waja-Jen | | | | | |
| Jen | 20*5 | | | | |
| Longuda | | | | | pùlò(wé)/phulewé |
| Waja | | <10? | | | wɔn, bwa-tigɛ |
| Yungur | | | | (-ru) | |
| Laal | 10-'big' | | | | |

The fact that this term was massively borrowed (most likely simultaneously) from Fula and Arabic suggests that it was lacking in Proto-Adamawa. It can be assumed that the root *ru* attested in Bua and Yungur is also a borrowing, this time from Bagirmi *àrú* 'hundred'.

### 3.6.7.13 'Thousand'

Table 3.131: Adamawa stems and patterns for '1000'

| | | |
|---|---|---|
| Fali-Yingilum | | < Fula |
| Kam | ? | |
| Leko-Duru-Mumuye | | |
|   Duru | | < Fula, < Hausa |
|   Leko | 20*10? | < Fula |
|   Mumuye | ? | |
| Mbum-Day | | |
|   Bua | | < Bagirmi |
|   Kim | | < Bagirmi |
|   Mbum | 'sack', bag' | < Fula, < Bagirmi |
|   Day | | < Bagirmi |
| Waja-Jen | | |
|   Jen | ʃik-1, 20-fe | |
|   Longuda | ? | |
|   Waja | kʊʊl, nèe/kú-néŋ, 100*10, bi-kate, tedu | |
|   Yungur | (100*10) | |
| Laal | | < Baguirmi, < Hausa |

The term for 'thousand' was massively borrowed from Fula, Bagirmi and Hausa, which points to its absence in the proto-language.

## 3.7 Ubangi

What follows is a preliminary analysis of the evidence of five separate language groups including Ubangi-Banda, Gbaya-Manza-Ngbaka, Ngbandi, Sere-Ngbaka-Mba (A. Ngbaka-Mba, B.Sere), and Zande.

### 3.7.1 Banda

Table 3.132: Numerals in Banda130

| 1 | bàlē (bà-lē?) | 7 | 5+2 |
|---|---|---|---|
| 2 | biʃi (bi-ʃi?) | 8 | 5+3, ngebeɗeɗe |
| 3 | vɔ-tɑ | 9 | 5+4, 8+1 |
| 4 | và-nā | 10 | mó-rófō, bu-fu, 'two hands ','all the fingers',*gba |
| 5 | mī-ndū | 20 | 'one person', 'the whole person', 'body-person-all' |
| 6 | 5+1, gɑzɑlɑ | 100 | ngàmbɔ̀/ngbàngbò,'five    persons' , < Sango , < Lingala? |
|   |   | 1000 | < French 'sack', < Lingala? |

The form *gba* 'ten' is traceable in the Mbanza (Mabandja) terms for tens.

### 3.7.2 Gbaya-Manza-Ngbaka

Table 3.133: Numerals in Gbaya-Manza-Ngbaka

| | | | |
|---|---|---|---|
| 1 | *kpɔ́k/kpóm ;ndáŋ | 7 | *5+2 |
| 2 | *bùà, *ʝíítò; bùwá (bù-wá?)/vàχ, -too | 8 | *5+3; 4PL |
| 3 | *tàr(à) | 9 | *5+4;kùsì |
| 4 | *nár(á) | 10 | *ɓú/ɓú-kɔ̰́ |
| 5 | *mɔ̀ɔ̀rɔ́/mɔ̀r-kɔ̰́ | 20 | *10*2 |
| 6 | *5+1, (gàzèlè) | 100 | *gɔ́m-màá ; < Lingala |
| | | 1000 | < French 'sack', < Lingala |

Ives Moñino's reconstructions (Moñino 1995) are quoted in the table under an asterisk. Selected noteworthy forms are also included.

In the diachronical perspective, the forms *ʝíítò and *bùà 'two' probably included noun class prefixes. They go back to *-too and *-wa respectively (cf.vàχ '2' in Gbaya Mbodomo).

In his discussion of *mɔ̀ɔ̀rɔ́ Moñino states that "La variante *mɔ̀ɔ̀rɔ́ semble être une contraction de *mɔ̀r-kɔ̰́, dans laquelle on peut reconnaître l'élément kɔ̰́ 'main' … " (Moñino 1995: 655). He also makes the folowing observation regarding the reconstruction of the term for 'ten': "*ɓú 'dix' est en relation avec *ɓú 'façonner, faire un cercle, joindre les mains'; la série partielle ɓú-kɔ̰́ est encore plus explicite, et décrit le geste qui accompagne l'énonciation du chiffre 10 chez tous les locuteurs" (Moñino 1995: 656)[20]. This is an important point, especially in view of the relatively frequent occurrence of *bu* in the NC languages and the possible etymological relationship between *ɓú and phonetically similar forms attested in other branches. However, such a relationship would be doubtful within Moñino's etymological hypothesis.

The following etymology is suggested for 'hundred' by Thomas Elvis Guenekean: "The word gɔ̰̀m means 'cut' or 'gathered' and n̄m̄à: means 'things'."[21] According to Moñino, the form literally means 'frapper-l'une l'autre (les mains)' (Moñino 1995: 657).

---

[20]However, in some Gbaya languages, these forms differ by tone: Gbaya (Roulon-Doko) ɓú '10' ~ ɓu 'to tap; to applaud, to roll'.

[21]https://mpi-lingweb.shh.mpg.de/numeral/Gbaya-Bossangoa.htm

### 3.7.3 Ngbandi

The Ngbandi and Yakoma evidence points toward the reconstruction outlined in the table below (Table 3.134):

Table 3.134: Numerals in Ngbandi

| | | | |
|---|---|---|---|
| 1 | kɔ(i) | 7 | mbara-mbara |
| 2 | sɛ | 8 | miambe/myɔ̀mbè |
| 3 | ta | 9 | gumbaya |
| 4 | siɔ/syɔ | 10 | sui, bàlé |
| 5 | kɔ̃/kṵ̄ | 20 | 10*2 |
| 6 | mana, mèrɛ̄ | 100 | ngbangbo |
| | | 1000 | < Lingala, Arabic |

### 3.7.4 Sere-Ngbaka-Mba

Since the languages within this group are extremely divergent, it seems reasonable to treat the evidence from its two major sub-groups separately.

Ngbaka-Mba (Table 3.135)

Table 3.135: Numerals in Ngbaka-Mba

| | | | |
|---|---|---|---|
| 1 | kpó-/kpáà-, ɓa-wɨ, ɓī-nì/bì-rì, ú-ma | 7 | 5+2, (mā-nāníkà, lè-rɵzi, zyálá, sáɓá), sílànā/sélènā/ʃíēnā (<4?) |
| 2 | bīʃ-ì/ɓī-sī, ɓi-né/bí-de, gbwɔ̀ | 8 | sénā (2*4?), gba-dzena/mā-dʒénà, (5+3, 10-2) |
| 3 | ba-ta/ba-la | 9 | 5+4, 10-1, (me-newá) |
| 4 | ba-na/ba-ɗa/ba-la | 10 | nzò kpā̰('head-hand')/àngbà, a-busa |
| 5 | bu-ruwe/bu-luve/θuwe, ʔeve/ve/vue | 20 | 10*2 |
| 6 | ʃí-tà/si-ta (2*3), mā-ɗíà/ká-zyá, 5+1 | 100 | < Sango, < Lingala, 20*5, (mya, kʉ́lɔ́, kpode, ngūndāngū) |
| | | 1000 | gyu, kutu, < Arabic, < French ('sack'), 100*10 |

Sere (Table 3.136)

Table 3.136: Numerals in Sere

| 1 | njẽe | 7 | 5+2 |
|---|------|----|-----|
| 2 | so | 8 | 5+3 |
| 3 | táʔò | 9 | 5+4 |
| 4 | nàʔò | 10 | ɓǐ-kürü , muʔɓì ('on hands') |
| 5 | vo | 20 | 'kill-person-one' |
| 6 | 5+1 | 100 | 'kill-persons-five', < Arabic |
| | | 1000 | 100*10 |

Sere-Ngbaka-Mba (Table 3.137)

Table 3.137: Sere-Ngbaka-Mba numeral system (*)

| 1 | kí-lī, sa | 7 | 5+2 |
|---|-----------|----|-----|
| 2 | ī-jō/ī-yō/úé | 8 | 5+3 |
| 3 | bíá-tá/ā-tā | 9 | 5+4 |
| 4 | lu, bīà-ngì ~ bīà-mà | 10 | ŋgbɔ̃/bà-wē |
| 5 | ì-sìbē/bī-sùè | 20 | 'people one' |
| 6 | 5+1 | 100 | ndɔ̄ŋgbʉ́, ngbàngbù< Sango |
| | | 1000 | sákì/sākè (< Sango < French) |

178

### 3.7.5 Proto-Ubangi

The evidence pertaining to each of the numerical terms is summarized below.

#### 3.7.5.1 'One'

Table 3.138: Ubangi stems for '1'

| | | | | | |
|---|---|---|---|---|---|
| Banda | bàlē (bà-lē?) | | | | |
| Gbaya-Manza-Ngbaka | | kpɔ́(k)/ (kpém) | ndáŋ | | |
| Ngbandi | kɔ(i) | | | | |
| *Sere-Ngbaka-Mba* | | | | | |
|   Ngbaka-Mba | ɓī-nì/ bì-rì | kpó-/ kpáà- | | | ɓa-wɨú-ma |
|   Sere | | | | njẽe | |
| Zande | kí-lī | | | | sa |

Two competing roots (*le/ne* and *k(p)o(k)*) are distinguishable here.

#### 3.7.5.2 'Two'

Table 3.139: Ubangi stems for '2'

| | | | | |
|---|---|---|---|---|
| Banda | biʃi (bi-ʃi?) | | | |
| Gbaya-Manza-Ngbaka | | bùwá (bù-wá?)/vàχ | -too | |
| Ngbandi | sɛ | | | |
| *Sere-Ngbaka-Mba* | | | | |
|   Ngbaka-Mba | bī-ʃì/ɓī-sī | gbwɔ̀ | | ɓi-né/bí-de |
|   Sere | | | so | |
| Zande | | | | ī-jō/ī-yō/úé |

The only root widely attested within this family is *si/ʃi*.

Table 3.140: Ubangi stems for '3' and '4'

|  | '3' | '4' | '4' |
|---|---|---|---|
| Banda | vɔ-tɑ | và-nā |  |
| Gbaya-Manza-Ngbaka | tààr | náár |  |
| Ngbandi | ta |  | siɔ/syɔ |
| *Sere-Ngbaka-Mba* |  |  |  |
|    Ngbaka-Mba | ba-ta/ba-la | ba-na/ba-ɗa/ba-la |  |
|    Sere | táʔò | nàʔò |  |
| Zande | bíá-tá/ā-tā |  | lu, bīà-ngì ~ bīà-mà |

### 3.7.5.3 'Three' and 'four'

The roots for 'three' and 'four' can be securely reconstructed as *taar* and *naar* respectively (with an alignment by analogy applied).

### 3.7.5.4 'Five'

Table 3.141: Ubangi stems for '5'

| Banda | mī-ndū |  |  |
|---|---|---|---|
| Gbaya-Manza-Ngbaka |  | mɔ̀r-(k)ɔ́ |  |
| Ngbandi |  | kɔ̃/kṵ̄ |  |
| *Sere-Ngbaka-Mba* |  |  |  |
|    Ngbaka-Mba | bu-ruwe/-luve/θuwe | ʔeve ~ ve/vue |  |
|    Sere |  | vo |  |
| Zande |  |  | ì-sìbē/bī-sùè |

The Proto-Ubangi form is unclear, since the term for 'five' is based on the lexical root meaning 'hand' (*kɔ*) in two groups out of five. The only root whose attestations are not limited to a single group is *du(w)/lu(w)*.

**3.7.5.5 'Six'**

Table 3.142: Ubangi stems and patterns for '6'

| | | | |
|---|---|---|---|
| Banda | 5+1 | ɡɑ-zɑlɑ | |
| Gbaya-Manza-Ngbaka | 5+1 | gà-zɛ̀lɛ̀ | |
| Ngbandi | | | ma-na, mɛ̀-rɛ̄ |
| *Sere-Ngbaka-Mba* | | | |
|    Ngbaka-Mba | 5+1 | mā-ɗíà/ká-zyá | ʃí-tà/si-ta (2*3) |
|    Sere | 5+1 | | |
| Zande | 5+1 | | |

In addition to forms that follow the common pattern '6=5+1', a number of other forms of uncertain etymology are attested in the first two groups (and possibly in Sere-Ngbaka-Mba as well, assuming that our morphological analysis of pertinent forms is correct).

**3.7.5.6 'Seven'**

Table 3.143: Ubangi stems and patterns for '7'

| | | |
|---|---|---|
| Banda | 5+2 | |
| Gbaya-Manza-Ngbaka | 5+2 | |
| Ngbandi | | mbara-mbara |
| *Sere-Ngbaka-Mba* | | |
|    Ngbaka-Mba | 5+2 | mā-nāníkà, lè-rɵzi, zyálá, sáɓá, sílànā/sélènā/ʃíēnā (<4?) |
|    Sere | 5+2 | |
| Zande | 5+2 | |

The variety of forms attested in Ngbaka-Mba is noteworthy.

### 3.7.5.7 'Eight'

Table 3.144: Ubangi stems and patterns for '8'

| Banda | 5+3 | | ngebeɗeɗe |
|---|---|---|---|
| Gbaya-Manza-Ngbaka | 5+3 | 4PL | |
| Ngbandi | | | miambe/myɔ̀mbè |
| *Sere-Ngbaka-Mba* | | | |
| Ngbaka-Mba | 5+3 | sɛ́nā (2*4?) | ɠɓa-dzena/mā-dʒɛ́nà, 10-2 |
| Sere | 5+3 | | |
| Zande | 5+3 | | |

### 3.7.5.8 'Nine'

Table 3.145: Ubangi stems and patterns for '9'

| Banda | 5+4 | | 8+1 |
|---|---|---|---|
| Gbaya-Manza-Ngbaka | 5+4 | | kùsì |
| Ngbandi | | | ɡumbaya |
| *Sere-Ngbaka-Mba* | | | |
| Ngbaka-Mba | 5+4 | | 10-1, (me-newá) |
| Sere | 5+4 | | |
| Zande | 5+4 | | |

Apparently, at the family level the common pattern '5+' should be assumed for the terms from 'six' to 'nine'. Isolated forms attested in groups and sub-groups are quoted here (as well as in the cases of other families) in order to collect exhaustive evidence for further etymological analysis. Moreover, a small chance that the Niger-Congo proto-form is traceable within only a single branch should not be ignored.

### 3.7.5.9  3.7.5.9. 'Ten'

The reconstruction of the term for 'ten' is so problematic that it raises doubts as to whether it was present in Proto-Ubangi at all. In view of the convincing internal etymology suggested by Ives Moñino, the root *\*bu* alternating with *\*pu*

Table 3.146: Ubangi stems for '10'

| Banda | bu-fu | *gba | mó-rófō, ' two hands', 'all the fingers' |
|---|---|---|---|
| Gbaya-Manza-Ngbaka | 'personne' ('joindre les mains') | | |
| Ngbandi | | | sui, bàlé |
| *Sere-Ngbaka-Mba* | | | |
|   Ngbaka-Mba | | nzò-kpā 'head'-'hand')/à-ngbà | a-busa |
|   Sere | | | ɓĩ-kürü, 'on hands' |
| Zande | | ŋgbɔ̄/bà-wē | |

and *fu* in some of the NC families is an unlikely candidate. The reconstruction of *gba/ kpa* is worth considering. However, the root may not be primary.

### 3.7.5.10 'Twenty'

Table 3.147: Ubangi stems and patterns for '20'

| Banda | 'one person', 'the whole person', 'body-person-all' | |
|---|---|---|
| Gbaya-Manza-Ngbaka | | 10*2 |
| Ngbandi | | 10*2 |
| *Sere-Ngbaka-Mba* | | |
|   Ngbaka-Mba | | 10*2 |
|   Sere | 'kill-person-one' | |
| Zande | 'people one' | |

Two reconstruction possibilities are available here, i.e. the pattern '20=10*2' commonly attested in NC, and a derivation from the lexical term meaning 'person'.

Table 3.148: Ubangi stems and patterns for '100'

| Banda | ngàmbɔ̀/ngbàngbò | 'five persons'< Sango, < Bangala (< Lingala?) |
|---|---|---|
| Gbaya-Manza-Ngbaka | | 'cut/gathered'-'things'? 'clap hands'?, < Lingala |
| Ngbandi | ngbangbo | |
| *Sere-Ngbaka-Mba* | | |
|   Ngbaka-Mba | | < Sango, < Lingala, 20*5, (mya, kúlɔ́, kpode, ngūndāngū) |
|   Sere | | 'kill-persons-five', < Arabic |
| Zande | ngbàngbù < Sango | 'ndɔŋɓú |

### 3.7.5.11 'Hundred'

Most of the forms are apparent borrowings which suggests that the term for 'hundred' was absent in Proto-Ubangi.

### 3.7.5.12 'Thousand'

Table 3.149: Ubangi stems and patterns for '1000'

| Banda | < French, < Lingala? | |
|---|---|---|
| Gbaya-Manza-Ngbaka | < French, < Lingala, tɔ́maỳ | |
| Ngbandi | < Lingala, < Arabic | |
| *Sere-Ngbaka-Mba* | | |
|   Ngbaka-Mba | < Lingala, < Arabic, < French, 100*10 | gyu |
|   Sere | 1000*10 | |
| Zande | < Sango < French | |

The absence of the term for 'thousand' in Proto-Ubangi is even more evident than the absence of the term for 'hundred.'

## 3.8 Dogon and Bangime

A step-by-step reconstruction of Dogon numerals does not seem reasonable because the family is relatively homogeneous. In addition, the formal differences between the numerical terms do not seem to correlate with the internal genealogical classification of the Dogon languages. The table below offers an overview of the pertinent data (Table 3.150) and is followed by a brief commentary.

Table 3.150: Dogon numerals

| | | | |
|---|---|---|---|
| 1 | túrú/tumɔ, ti(i) | 7 | suli/soli/soye |
| 2 | lɛ́(y)/lɔ́(y)/nɛ́(y)/nɔ́(y) | 8 | gá(a)rà, sagi, sele (< Mande?) |
| 3 | taan | 9 | túwɔ́ |
| 4 | nay(n), kɛɛso | 10 | pɛ́rú/pɛ́lú |
| 5 | núnɛ́ɛ́(n)/nǔː(yn)/nûm | 20 | 10*2 |
| 6 | kuro/kule | 100 | 80 (síìŋ/súŋ) +20, < Fula |
| | | 1000 | 800 (múɲú) +200 |

'One': Najamba-Kindige: *kúndé* '1', Mombo *yɛ̀ːtáːŋgù* '1'.

'Two': The forms with the nasal **n-** attested in several dialects are variants of the basic form with ***l-***. It should be noted that the final palatal element is systematically attested in other numerical terms, e.g. in Ben Tey (Table 3.151):

Table 3.151: Final palatal in '2'

| | | | |
|---|---|---|---|
| 2 | yěy | 6 | kúròy |
| 4 | nǐːy$^{n}$ | 7 | súy$^{n}$ɔ̀y$^{n}$ |
| 5 | nùmǔy$^{n}$ | 8 | gá:rày |

Regardless of whether this element is a morpheme or not, we are certainly dealing with a phonetic alignment by the final segment. Thus the final -**y** should not be reconstructed even in those forms that show its presence in the majority of languages.

'Three': This is a persistent form with only minor modifications applied to it (e.g. *taandu*, *taali*).

'Four': This is the only term for which the final palatal (probably nasalized) is potentially reconstructable. If so, systematic alignments by analogy attested in final segments of other numerals are probably based on the form of 'four'. The

root *kɛɛso/ kɛ́:jɔ́/ kɛ́:jɛ̀y/ cɛ́zɔ̀/ yè-cɛ́zɔ́* is probably an innovation (see, however, Jeff Heath who argues for its archaic nature[22]). The term may be etymologically connected to the term for 'eighty', cf. Najamba-Kindige *sîm*, *kɛ̀:sŭm*, Tommo So *kɛ̀ɛ̀súm* and a number of other related forms (Yorno So *dɔ̀gɔ̀-sŭm*'80', "Dogon hundred", Valentin Vydrin, p.c., Perge Tegu *dɔ̀gɔ̀-sŭŋ* '80', Yanda Dom *sìŋ* '80' etc.).

'Five': The etymological connection of this term with the lexical root meaning 'hand' *nùmà/ nùmó/ nùmɔ́/ nŏŷ* is immediately apparent.

'Six' and 'seven' are probably primary terms.

'Eight'. The root *sagi* attested in Najamba and Yanda Dom was probably borrowed from Mande. The forms *sila, seele* observable in a number of dialects may be related to it. The root *gá(a)rà* is commonly attested in the majority of languages of this group, sometimes with a partial reduplication (Donno So/Yorno So/Toro So *ga-gara/ga-gira*). Partial reduplication is a popular means of deriving 'eight' from 'four' commonly attested throughout NC. In view of the fact that the Dogon counting system is based on 8, this root should probably be compared to *gàrá,* meaning 'big, large, a large quantity, a lot, go beyond (limit), more, to a greater extent'. Tonal differences may be neglected in this case, especially since the derived forms tend to be formally marked, e.g. tonally.

'Hundred'. The basic 'large number' in Dogon is 'eighty' rather than 'hundred', so this meaning should probably be reconstructed for *siiŋ/suŋ.* In view of this, the fact that the term for 'hundred' was borrowed from Fula in nearly all Dogon languages is not a coincidence.

'Thousand'. Similarly, the root *muɲu* (var. *mùsú / mùdʒú*) '800' incorporated into the pattern '1000=800+200' is reconstructed in Dogon.

The Bangime numeral system should also be considered here, since most of the numerical terms attested in this isolated language are comparable to those found in Dogon (Table 3.152).

As in Dogon, the terms covering the sequence from 'six' to 'nine' are primary. An isolated root for 'forty' (also represented in some of the Dogon languages) is attested in Bangime. Interestingly, the root is the same as the one found in some of the Mande languages, cf. Bangime *dɛ̀vɛ́*, Dogulu Dom (Dogon) *dɛ̀ɛ́*, Mombo (Dogon) *dê:*, Marka Dafing *dɛbɛ*, Bozo *dɛ̀bɛ́/ lɛ́wɛ̀*, Bamana *dɛ̀bɛ́*.

The root for 'ten' does not correspond to the one attested in Dogon. The latter finds a direct parallel in Boko (East Mande *kuri* 'ten'.

---

[22]http://dogonlanguages.org/

Table 3.152: Bangime numerals

| 1 | tòré/tĭyɛ́ (in counting) | 7 | kĭjé |
|---|---|---|---|
| 2 | jíndò | 8 | sàágín (< Mande?) |
| 3 | táárù | 9 | tɛ́gò |
| 4 | nìjɛ́ | 10 | kúrɛ́ |
| 5 | nŭndí | 20 | tàáw̃á |
| 6 | kĕré | 100 | tɛ̀ɛ̀mɛ̀dɛ́rɛ́ (< Fula ) |
|  |  | 1000 | mŭʒú |

## 3.9 Gur

It should be noted that the Gur languages are extremely divergent in the majority of their numerical terms (including those that prove to be fairly persistent in other families). The approach we took for the evidence studied above (i.e. the establishing of the most common forms and their further comparison to the data from other branches) may not appear fruitful in the case of the Gur languages.

To deal with the problem, we are going to use the classification of the Gur languages found in *Ethnolog*, namely A. Bariba, B. Central, C. Kulango, D. Lobi, E. Senufo, F. Teen, G. Tiefo, H. Tusia, I. Viemo, J. Wara-Natioro[23] . The Gur family comprises nearly a hundred languages. In terms of the classification outlined above, their distribution is uneven. Seven groups (Bariba, Kulango, Lobi, Teen, Tiefo, Tusia, Viemo) have an isolated language as their only member. Similarly, Wara-Natioro is represented by only three idioms. This means that the majority of the Gur languages are split between the two remaining groups, i.e. Senufo and Central. The former is comprised of about fifteen languages and is relatively homogenous. Its affiliation to Gur is often considered doubtful. Compared to Central, which embraces the majority of the Gur languages (nearly seventy), this group is relatively small. Two major sub-groups are identifiable within Central, i.e. Northern (38 languages) with Oti-Volta (33 languages) as the dominant branch and Southern (31 languages) with its dominant branch of Grusi (23 languages). In other words, 71 of the Gur languages (out of a total of 91) belong to either Oti-Volta, Grusi or Senufo. In addition to that, there are more than ten branches represented by a single isolated language each. No evidence points to their possible affiliation with the major branches or to their inter-relationship.

---

[23]This classification is accepted here with slight modifications based on recent studies. For instance, Dyan and Lobi are treated as members of the same branch.

The same can probably be said about several isolated languages affiliated (often uncritically) with the Central group (the Bwamu, Kurumfe, Dogoso-Khe, Gan-Dogosé, and Kirma-Tyurama branches). This already complex picture gets even more sophisticated in view of the following:

1) Branches represented by one or two languages (e.g. Buli-Konni, Notre, Yom-Nawdm) are distinguishable even within the most reliably established bodies of genetically related languages of this family.

2) According to Ulrich Kleinewillinghöfer (p.c.), who is a renowned expert in both Gur and Adamawa comparative linguistics, a border between these two families is not clear at all. This means that some of the Gur branches may prove to be more closely related to Adamawa.

Our reconstruction of the Gur numeral system is based on nearly 120 sources that vary in regards to the evidence they offer (cf. our considerations above). By addressing one of the most problematic cases (i.e. the reconstruction of the Gur term for 'one') we hope to work out a general approach that will eventually allow further comparison of the Gur evidence to that of other NC families.

'One'

The table below lists several forms of the term for 'one' in smaller Gur branches (Table 3.153):

Table 3.153: Diversity of stems for '1' in Gur

| Gurma | Grusi-Eastern | Grusi-Western |
|---|---|---|
| Akaselem: m̀-bá | Bago-Kusuntu: ŋʊrʊkpákpá | Chakali: dígímáná |
| Bimoba: yènn | Chala: -re-, -dɔ́ndʊlɔŋ | Deg: beŋ-kpaŋ/kpee |
| Miyobe: n-ni (-sɛ) | Delo: daale | Phuie: déò/dùdúmí |
| Nateni: -cɔ̃́, dèn | Kabiye: kɔ́-yɔ́m | Sisaala: kʊ̀-bàlá/dìáŋ |
| Ngangam: mi-kpìɛkm | Lama: kó-ɖɔ́m | Winyé: n-do |

A brief study of these examples raises doubts as to whether the Gur numeral system is reconstructable at all (not to mention the Grusi-Northern system or those of the more isolated Gur branches).

Even if we consider one syllable roots of the CV(C)-type only, the impression will remain that every concievable root for 'one' is attested in the Gur languages. At the same time, none of these roots is traceable in at least half of the Gur groups. This situation is reflected in the matrix below (Table 3.154).

Table 3.154: Distribution of the CV(C)- forms for '1' in the Gur languages

|  | **I** | **A** | **U** |
|---|---|---|---|
| **P (p/f)** | – | – | – |
| **B (b/w/m)** | 3/5 | 1/4 | 1/1? |
| **T (t)** | 1/1 | 2/2 | – |
| **D (d/l/r/n)** | 3/16 | – | 3/13 |
| **C (c/s)** | – | – | 1/1 |
| **J (j/y/ny)** | 1/18 | 1/1 | 1/1 |
| **K (k/h/x)** | 2/5 | 1/2 | 2/4 |
| **G (g/ŋ)** | 1/5 | 1/1 | 1/1 |

The first figure refers to the number of groups where a form is attested (with a maximum of 10 groups), whereas the second one refers to the number of languages. Thus, **B-I** denotes a form comprising a voiced labial consonant (b, w or m) and a front vowel that is attested in five languages within three groups (Central, Lobi-Dyan and Senufo) (Table 3.155):

Table 3.155: BI- forms for '1' in Gur (3 groups, 5 languages)

| **béé** | Ditammari | B. Central | 1. Northern | C. Oti-Volta | ii. Eastern |
|---|---|---|---|---|---|
| **bìè-** | Lobi | D. Lobi-Dyan | | | |
| **bɛ̌g** | Dyan | D. Lobi-Dyan | | | |
| **nì-bín** | Cebaara | E. Senufo | | | |
| **nan-bin** | Shempire | E. Senufo | | | |

The remaining forms are quoted below as an illustration of their extreme divergency.

(1) **BA** (1/4) (Table 3.156):

Table 3.156: BA- forms for '1' in Gur (1 group, 4 languages)

| | | | | |
|---|---|---|---|---|
| ** M̀-bá** | Akaselem | B. Central | 1. Northern | C. Oti-Volta | Gurma |
| **bàa** | Konkomba | B. Central | 1. Northern | C. Oti-Volta | Gurma |
| **mi-ba** | Ngangam | B. Central | 1. Northern | C. Oti-Volta | Gurma |
| **ǹ.-bá /-bɔ́** | Ntcham | B. Central | 1. Northern | C. Oti-Volta | Gurma |

**BU** (1/1): only *pú-wò* (possibly *púw-ò*, **PU?**) in Wara (J.Wara-Natioro)
**TI** (1/1): only *tía* in Baatonum (A.Bariba)
**TA** (2/2) (Table 3.157).

Table 3.157: TA- forms for '1' in Gur

| | | |
|---|---|---|
| **ta, taà, tã̀ã̌** | Kulango (dial.) | C.Kulango |
| **tani** | Teen (dial.) | F.Teen |

**DI** (3/15) ( Table 3.158):

Table 3.158: DI- forms for '1' in Gur

| | | | | | |
|---|---|---|---|---|---|
| **dè** | Bwamu (Boore) | B. Central | 1. Northern | A. Bwamu | |
| **nni** | Miyobe | B. Central | 1. Northern | C. Oti-Volta | iii. Gurma |
| **dèn** | Nateni | B. Central | 1. Northern | C. Oti-Volta | iii. Gurma |
| **lé** | Khe Southern | B. Central | 2. Southern | A. Dogoso-Khe | |
| **í-lèŋ** | Khisa | B. Central | 2. Southern | C. Gan-Dogose | |
| **re-** | Chala | B. Central | 2. Southern | D. Grusi | i. Eastern |
| **dííŋ** | Paasaal | B. Central | 2. Southern | D. Grusi | iii. Western |
| **déò** | Phuie | B. Central | 2. Southern | D. Grusi | iii. Western |
| **dìáŋ** | Sisaala (dial.) | B. Central | 2. Southern | D. Grusi | iii. Western |
| **dìén** | Sisaala (dial.) | B. Central | 2. Southern | D. Grusi | iii. Western |
| **diigɛ** | Tampulma | B. Central | 2. Southern | D. Grusi | iii. Western |
| **déiŋ** | Kirma | B. Central | 2. Southern | E. Kirma-Tyurama | |
| **dẽ̀ẽn–** | Turka | B. Central | 2. Southern | E. Kirma-Tyurama | |
| **nɔ̀-ni** | Karaboro (dial.) | E. Senufo | | | |
| **dɛ̀** | Tiefo (dial.) | G. Tiefo | | | |

**DU** (3/13) (Table 3.159)

Table 3.159: DU- forms for '1' in Gur

| | | | | | |
|---|---|---|---|---|---|
| **dòǔ̃** | Bwamu | B. Central | 1. Northern | A. Bwamu | |
| **dòòn** | Bwamu | B. Central | 1. Northern | A. Bwamu | |
| **dò** | Láá Láá | B. Central | 1. Northern | A. Bwamu | |
| **rʊ** | Chala | B. Central | 2. Southern | D. Grusi | i. Eastern |
| **kà-lɞ̀** | Kasem (dial.)1 | B. Central | 2. Southern | D. Grusi | ii. Northern |
| **kà-lʊ** | Kasem (dial.)2 | B. Central | 2. Southern | D. Grusi | ii. Northern |
| **è-dù** | Lyele | B. Central | 2. Southern | D. Grusi | ii. Northern |
| **ù-dù** | Northern Nuni | B. Central | 2. Southern | D. Grusi | ii. Northern |
| **nɔ̀-dɞ̀** | Southern Nuni | B. Central | 2. Southern | D. Grusi | ii. Northern |
| **n-do** | Winyé | B. Central | 2. Southern | D. Grusi | iii. Western |
| **nú-nu** | Nafaanra | E. Senufo | | | |
| **dũde** | Viemo | I.Viemo | | | |

CU (1/2): only *mà-cɔ́̃* in Nateni (Central: 1. Northern: C.Oti-Volta: iii. Gurma
JI (1/19) (Table 3.160)

Table 3.160: CI- forms for '1' in Gur

| yéŋ/ wà-ɲī | Buli | B. Central | 1. Northern | C. Oti-Volta | i. Buli-Koma |
|---|---|---|---|---|---|
| yɛ̃n | Mbelime | B. Central | 1. Northern | C. Oti-Volta | ii. Eastern |
| yènn | Bimoba | B. Central | 1. Northern | C. Oti-Volta | iii. Gurma |
| yèn- | Gurma | B. Central | 1. Northern | C. Oti-Volta | iii. Gurma |
| jènǹ | Moba | B. Central | 1. Northern | C. Oti-Volta | iii. Gurma |
| bõ-ɣén | Birifor (dial.) | B. Central | 1. Northern | C. Oti-Volta | iv. Western |
| bo-yæn | Birifor (dial.) | B. Central | 1. Northern | C. Oti-Volta | iv. Western |
| bõ-yen | Dagaara (dial.) | B. Central | 1. Northern | C. Oti-Volta | iv. Western |
| yén- | Dagaara (dial.) | B. Central | 1. Northern | C. Oti-Volta | iv. Western |
| yén | Farefare | B. Central | 1. Northern | C. Oti-Volta | iv. Western |
| yé | Moore | B. Central | 1. Northern | C. Oti-Volta | iv. Western |
| bǒ-ŋjìŋ | Wali | B. Central | 1. Northern | C. Oti-Volta | iv. Western |
| yín | Dagbani (Dagomba) | B. Central | 1. Northern | C. Oti-Volta | iv. Western |
| yɪn- | Hanga | B. Central | 1. Northern | C. Oti-Volta | iv. Western |
| yín | Kamara | B. Central | 1. Northern | C. Oti-Volta | iv. Western |
| yén- | Kantosi | B. Central | 1. Northern | C. Oti-Volta | iv. Western |
| yín | Mampruli | B. Central | 1. Northern | C. Oti-Volta | iv. Western |
| nyɔ̆ŋ | Yom (Pila) | B. Central | 1. Northern | C. Oti-Volta | v. Yom-Nawdm |

JA (1/1) – only *à-yàʔ* in Safaliba (B. Central: 1. Northern: C.Oti-Volta: iv. Western)
JU (1/1) – only *yòn* in Waama (B. Central: 1. Northern: C.Oti-Volta: ii. Eastern)
KI (2/5) (Table 3.161)

Table 3.161: KI- forms for '1' in Gur

| ṃ-hén | Nawdm | B. Central | 1. Northern | C. Oti-Volta | v. Yom-Nawdm |
|--------|--------|------------|-------------|--------------|--------------|
| kpee | Deg | B. Central | 2. Southern | D. Grusi | iii. Western |
| kpéé | Vagla | B. Central | 2. Southern | D. Grusi | iii. Western |
| nì-kǐ | Sìcìté Senufo | E. Senufo | | | |
| nìŋ-kìn | Supyire Senufo | E. Senufo | | | |

**KA** (1/2) (Table 3.162)

Table 3.162: KA- forms for '1' in Gur

| beŋ-kpaŋ | Deg | | B. Central | 2. Southern | D. Grusi | iii. Western |
|----------|--------|---|------------|-------------|----------|--------------|
| kpáŋ | Vagla | | B. Central | 2. Southern | D. Grusi | iii. Western |

**KU** (2/3) (Table 3.163)

Table 3.163: KU- forms for '1' in Gur

| kpò | Khe (dial.) | B. Central | 2. Southern | A. Dogoso-Khe |
|------|-------------|------------|-------------|---------------|
| tì-kpóʔ | Dogose | B. Central | 2. Southern | C. Gan-Dogose |
| tʰì-kpo | Kaansá | B. Central | 2. Southern | C. Gan-Dogose |
| nú-kú | Toussian (dial.) | H. Tusia | | |

**GI** (1/5) (Table 3.164)

Table 3.164: GI- forms for '1' in Gur

| | | |
|---|---|---|
| **niŋ-gbe** | Palaka Senufo | E. Senufo |
| **nī-gbe** | Nyarafolo Senufo | E. Senufo |
| **ni-gì/ni-gĩ** | Mamara Senufo (Minyanka) | E. Senufo |
| **nin-gin** | Shempire Senufo | E. Senufo |
| **nu-gbe** | Tagwana Senufo | E. Senufo |

**GA** (1/1) – only *nuŋ-gba* in Djimini Senufo (E. Senufo).
**GU** (1/1) – only *gbú* in Northern Khe (B. Central: 2. Southern: A. Dogoso-Khe).

The only lacuna in this presentation is due to the lack of forms with voiceless labial consonants (this, however, may not prove true in the case of Wara-Natioro, as we hope to demonstrate below). It should be noted that the general distribution pattern is that a single form is attested in one branch out of ten, three forms are found in both two and three branches, and none of the forms is recorded in four or more branches. This makes an attempt at tracing them down to a source form (with its further comparison to the evidence of the other families) unreasonable. In view of the genetic classification of the Gur languages and the considerations presented above, the optimum solution to the problem probably lies within separate reconstructions of numerals in the following sixteen Gur branches that belong to ten major language groups of this family, assuming that each of them may shed some new light on the reconstruction of the Niger-Congo numeral system:

1. Bariba
2. Central: 1. Northern: A. Bwamu
2. Central: 1. Northern: B. Kurumfe
2. Central: 1. Northern: C. Oti-Volta
2. Central: 2. Southern: A. Dogoso-Khe
2. Central: 2. Southern: C. Gan-Dogose
2. Central: 2. Southern: D. Grusi
2. Central: 2. Southern: E. Kirma-Tyurama
3. Kulango
4. Lobi-Dyan

5. Senufo
6. Teen
7. Tiefo
8. Tusia
9. Viemo
10. Wara-Natioro.

Numerical terms as attested in each of these branches will be examined below.

### 3.9.1 Bariba

Table 3.165: Bariba numerals

| 1 | tiā | 7 | 5+2 |
|---|---|---|---|
| 2 | ru | 8 | 5+3 |
| 3 | i-ta | 9 | 5+4 |
| 4 | ǹ-nɛ | 10 | wɔ-kuru |
| 5 | nɔ̀ɔbù | 20 | yɛndu |
| 6 | 5+1 | 100 | 20*5 |
| | | 1000 | fɔ̀rɔ̀tɔ? |

### 3.9.2  Central Gur

#### 3.9.2.1  Northern Central Gur

##### 3.9.2.1.1  Bwamu

Table 3.166: Bwamu numerals

| 1 | do | 7 | 5+2 |
|---|---|---|---|
| 2 | ɲū | 8 | 5+3 |
| 3 | tĩ | 9 | dìiní/dènú |
| 4 | náa | 10 | pílú/píru/ˀɓúrúù |
| 5 | hò-nú | 20 | ɓóní/ɓénle/kēwēníì |
| 6 | 5+1 | 100 | kʰĭminù (< Mande keme ) |
|   |   | 1000 | 100*10, muaseé |

##### 3.9.2.1.2  Kurumfe

Table 3.167: Kurumfe numerals

| 1 | dom | 7 | pɛ̃ɛ̃ |
|---|---|---|---|
| 2 | hĩĩ | 8 | tɔɔ |
| 3 | tãã | 9 | fa |
| 4 | nãã | 10 | fɪ |
| 5 | nɔm | 20 | sofe (<10? ) |
| 6 | hʊrʊ | 100 | bɛrʊ |
|   |   | 1000 | tʊsrɪ < from Moore |

3.9.2.1.3  Oti-Volta

Table 3.168: Buli-Koma numerals

| | | | |
|---|---|---|---|
| 1 | yéŋ (adj.), ní (count) | 7 | yòpɔ̄āī, pɔ́ĩ̀ |
| 2 | yè, li | 8 | nāāniŋ/à-níì (<* 4redupl., 4PL?) |
| 3 | tà | 9 | nèūk/ŋ̀wɛ́ |
| 4 | nààsì/nísà | 10 | pī/bâŋ |
| 5 | nù | 20 | 10*2 |
| 6 | yùèbì/óbìŋ | 100 | kòòk, kobɪga/bóɾà |
| | | 1000 | < Engl. |

**i. Buli-Koma (Table 3.168)**

Table 3.169: Eastern Oti-Volta numerals

| | | | |
|---|---|---|---|
| 1 | cɔ̄rɔ̄, béé, dèǹnì (counting), yɛ̃nde/yòn, *de | 7 | pèléī/bérén, yīēkà/nyiekɛ, doodɛ̄ (6+1) |
| 2 | dyā, dɛ́ɛ́, diání/dɛɛni, yēdē/yɛ́ndí | 8 | nēī/nɛ̀í/ni/ninyɛ̄̃ |
| 3 | tâati/tâadi/tāārī | 9 | wáī/wɛi/wɛ̄ |
| 4 | naa(si) | 10 | pwígɔ̄/pííkà/piíkɛ/piitɛ , *pi |
| 5 | num(mu)/nun | 20 | 10*2 |
| 6 | kūà/kuɔ, dūo, hɛ̃̀dwàm, kpàrùn | 100 | kòɣɔ̄/kookɛ/kɔ́úkpà/kɔ̀ɔ̀tà |
| | | 1000 | túsírè |

**ii. Eastern (Table 3.169)**   Please note the extreme divergency of languages within this branch: the variety of forms presented in the table above are attested in only four languages, i.e. Biali, Ditammari, Mbelime and Waama.

Table 3.170: Gurma numerals

| 1 | bá, yènn(do), den (isol.: ni, cɔ̃) | 7 | lòlé/lèlé (isol.: sɛ́ɛi, yehì) |
|---|---|---|---|
| 2 | le/dɛ́/tɛ́ | 8 | ni(n) |
| 3 | tà | 9 | wɛʔ/wɛ́ɛ/wɔ̀i/wáī |
| 4 | nà(hì) | 10 | píík/pʷíʔ/fi/pita |
| 5 | mù/nùm̀/nu(pũ)/ŋùn | 20 | 10*2 (isol.: kòó, mùŋ̀kú < mande?) |
| 6 | loòb/luu, kɔ̀dì/kouulǘ | 100 | kúb (isol.: pílɛ, kɔ̀ta) |
|   |   | 1000 | < kùtùkúʻsackʼ, borrowing |

**iii. Gurma (Table 3.170)**

Table 3.171: Western Oti-Volta numerals

| 1 | yen/yin, dam?, (dàkɔ̃́ʔ) | 7 | yopoi (< yo-poi?) |
|---|---|---|---|
| 2 | yi(ʔ) | 8 | nii(n) |
| 3 | ta | 9 | way/wey |
| 4 | naasi/naar/nãan | 10 | pia/pie |
| 5 | nú | 20 | 10*2 |
| 6 | yobu | 100 | kob/kɔɔ |
|   |   | 1000 | tur/tudi (borrowed?) |

**iv. Western (Table 3.171)**

Table 3.172: Yom-Nawdm numerals

| 1 | hén, nyɞ̌ŋ- /nyɞ̌rɣə- | 7 | lèbléʔ (<6?), 5+2 |
|---|---|---|---|
| 2 | li/ɾéʔ/*rɣa? | 8 | nì:ndí; 10-2 |
| 3 | ta/tâʔ | 9 | wɛ́ʔ, 10-1 |
| 4 | naa/nɛ̀ɛ̀sɞ̀ | 10 | ʔɾíʔ, fɛɣa |
| 5 | nu | 20 | 2PL |
| 6 | m̀ɾò:ndí (X+1?), lèèwɞ̀r | 100 | lémú, wʊr- |

**v. Yom-Nawdm (Table 3.172)**

**Proto-Oti-Volta**    The evidence of five Oti-Volta branches (isolated forms excluded) is summarized in Table 3.173.

Table 3.173: Numerals in Proto-Oti-Volta

|  | i. Buli-Koma | ii. Eastern | iii. Gurma | iv. Western | v. Yom-Nawdm | *Proto-Oti-Volta |
|---|---|---|---|---|---|---|
| 1 | yéŋ, ní | dènnì, yẽnde/ yòn, *de | yènn(do), den, ni | yen/ yin, dam | hén, nyǎŋ | den/ yen, ni, de? |
| 2 | yɛ̀, li | dɛ́ɛ́(ni), yᵉēdē | le/ dɛ́ | yi(ʔ) | li/ ɾéʔ/ *rɣa? | li/ yi |
| 3 | tà | tâati | tà | ta | ta | ta(t) |
| 4 | nààsì | naa(si) | nà(hì) | naasi | naa/ nɛ̀ɛ̀sə̀ | naa(si) |
| 5 | nù | nun | nùm̀/ nu/ ŋùn | nú | nu | nu |
| 6 | yùèbì/ óbìŋ | dūo | loòb/ luu | yobu | lèèw-ər | lob/ yob |
| 7 | yòpɔ̄āī, pɔ́ì | doodē̃ (6+1) | lòlé/ lèlé | yopoi | lèblé? | *lob-le (6+1)? poi(n)? |
| 8 | nāāniŋ/ à-níì | nɛ̃́í/ ni/ ninyɛ̃̄ | ni(n) | nii(n) | nì:ndí | ni |
| 9 | nèūk/ ŋ̀wɛ́ | wáī/ wɛi/ wɛ̄ | wɛ̀ʔ/ wɛ́ɛ́/ wáī | way/ wey | wɛ́ʔ | wey/ weʔ |
| 10 | pī | pwígə̄/ pííkà/ *pi | píík/ pʷíʔ/ fi | pia/ pie | fɛɣa | pi(k) |
| 20 | 10*2 | 10*2 | 10*2 | 10*2 | 2PL | 10*2 |
| 100 | kòòk, kobɪga | kòɣə̄/ kookɛ/ kɔ́úkpà | kúb | kob/ kɔɔ | lɛ́mú, wʊr- | kob, kook |

The reconstruction of the Oti-Volta numeral system is surprisingly unproblematic. In addition to the expectedly persistent reflexes of 'three' and 'four', homogeneous forms for 'two', 'five', and 'ten' are noteworthy. The term for 'eight' seems to be based on 'four' (either via the partial reduplication or according to the '4PL' pattern). In addition to that, Oti-Volta is characterized by the presence of the primary (homogeneous) forms of 'six', 'eight', and 'nine'. The forms of 'seven' are probably derived and follow the pattern '6+1'. It appears that the derivative form *lob-le > lole is already reconstructable at the Proto-Oti-Volta level.

### 3.9.2.2 Southern Central Gur

3.9.2.2.1 Dogoso-Khe

Table 3.174: Dogoso-Khe numerals

| 1 | kpò, lé | 7 | 5+2 |
|---|---------|---|-----|
| 2 | jɔ(n) | 8 | 5+3 |
| 3 | thɔ | 9 | 5+4 |
| 4 | dáa | 10 | kpélé |
| 5 | nɔ(n) | 20 | cúkúrì/gʊ̀ʊsì |
| 6 | 5+1 | 100 | 20*5 |
|  |  | 1000 | kpɛ́ |

The forms pertaining to these languages that are not present in the main databases are quoted according to Kerstin Winkellmann in (Miehe; Reineke; Winkelmann 2007b: 181-210). Although the numerals attested within the two languages of this group are quite persistent, Kerstin Winkellmann stresses their grammatical difference: " … while Dɔgɔ-sʊ uses noun suffixes, sʊ-Khe is a prefixing language " (Winkellmann 2007d: 209).

3.9.2.2.2 Gan-Dogose

Table 3.175: Gan-Dogose numerals

| 1 | kpo/po, (lèŋ) | 7 | 5+2 |
|---|---------------|---|-----|
| 2 | yɔ́/ɲɔ̠/ʤʒ̀ŋ | 8 | 5+3 |
| 3 | sáa̠/tʰɔ́ʔ | 9 | 5+4, 10-1 |
| 4 | ɲee/ì-yì̠ị̠, (á-dàa) | 10 | (kpoogo, gbùnè, kpélé, sí-nõ̃y - 5PL) |
| 5 | mwã/wàa, nɔ̃̀n | 20 | gbeere, (tʃúkúrì) |
| 6 | 5+1 | 100 | 20*5 |
|  |  | 1000 | kpíɛ 'a goat' |

Three of the languages belonging to this branch show too many forms, suggesting that we are dealing with a heterogeneous branch. In view of its numerical terms, it is not immediately apparent why this branch has been singled out.

3.9.2.2.3  Grusi

Table 3.176: Eastern Grusi numerals (*)

| 1 | ḍə́m/lə̀m/yə́m, re/ɔ́ḍe | 7 | lʊbɛ, 6+1, 4+3, 10-3 |
|---|---|---|---|
| 2 | la/lɛ̀ | 8 | 4redupl., 4PL, 10-2, toozo, (k͡pèèrè) |
| 3 | tòòsó/tooro | 9 | 10-1, isolated forms |
| 4 | násá/naara | 10 | fu, (nʊ́á - 5PL, sàlá) |
| 5 | nʊ̌/nʊ̌ŋ, kpásì/gbǎ́nzì | 20 | ko/kuo/koowu, (sao, nɛɛ́lɛ̀, 10*2) |
| 6 | loḍò/looro/lɛ̀èjò, (3PL) | 100 | 20*5, < Ewe, ('guinea fowl') |
|   |   | 1000 | kòtòkó, kpoŋ |

**i. *Eastern Grusi (Table 3.176)**

Table 3.177: Northern Grusi numerals (*)

| 1 | du/lu, (téngí) | 7 | pɛ̀, (4+3, 5+2) |
|---|---|---|---|
| 2 | le/là/(ɲìí) | 8 | nānā (4 redupl.), (lyɛlɛ, bàndá) |
| 3 | tɔ̀/twà/cɔ́ɔ̀ | 9 | nògʊ, nìbu, (10-X) |
| 4 | na/nīān/nàas | 10 | fúgá, (fo) |
| 5 | nu | 20 | 10*2, (sāpōā, 10+10, swéní) |
| 6 | dɔ̌, (5+pi) | 100 | bi, (zɔ̌m) |
|   |   | 1000 | mòrò |

**ii. *Northern Grusi (Table 3.177)**

Table 3.178: Western Grusi numerals (*)

| 1 | kpáŋ/kpee, bala, do/deo/dííŋ/digi | 7 | lʊp,pɛ́ɛ́/piɛ , 5+2 |
|---|---|---|---|
| 2 | lɛ/nɛ/lìɛ | 8 | córí/kyórí, 5+3, (pɔɔ) |
| 3 | toro | 9 | nɛ́mɛ́/nìbí, 10-1, 5+4 |
| 4 | naa/naasi/naare | 10 | fi |
| 5 | nue/nwɔ́/nɔ̀ŋ | 20 | mɛ́rɛ́, mʊgɔ́ (< Mande?), |
| | | | (máágí, toko, ma-cu?) |
| 6 | lɔ̀rɔ̀/*lug/dɔ̀, 5+1, (go) | 100 | kɔ̀wá/kɔ̀ɔ́, zɔ́lɔ́, lafa |
| | | 1000 | gboŋ/bʊ́í |

**iii. *Western Grusi (Table 3.178)**    The most probable *Proto-Grusi reconstructions based on the roots attested in at least two Grusi branches are summarized in the table below (Table 3.179):

Table 3.179: Proto-Grusi numeral system (*)

| 1 | do/du/lu, de/re | 7 | pɛ/lʊ-pɛ/lʊ-bɛ, 5+2 |
|---|---|---|---|
| 2 | lɛ/le/ne/ɲi | 8 | 4 redupl. |
| 3 | toro/toso/tɔ | 9 | 10-1, nibi/nibu (ni-bi/bu?) |
| 4 | naare/naasi/na | 10 | fu/fi |
| 5 | nu/nʊ | 20 | 10*2? |
| 6 | dʊ/lo-ḍo/lo-ro, 5+1 | 100 | 20*5? bi? kɔwa/kɔɔ? |
| | | 1000 | kpoŋ/gboŋ |

### 3.9.2.2.4 Kirma-Tyurama

Table 3.180: Kirma-Tyurama numerals

| 1 | déiŋ/dẽ̄ẽná | 7 | 5+2 |
|---|---|---|---|
| 2 | hã́ĩ/hãl | 8 | 5+3 |
| 3 | síɛi/siɛl | 9 | 5+4, 10-1 |
| 4 | na(a) | 10 | nṹ́sɔ̃/cĩ́ŋcíelùò |
| 5 | dɨ | 20 | kómòrré/ɡuř̃ |
| 6 | 5+1 | 100 | ɡundi, 20*5 |
|   |   | 1000 | 200*5, 800+200 |

## 3.9.3 Kulango

Table 3.181: Kulango numeral system

| 1 | ta(a) < *ta̰à̰ | 7 | 5+2 |
|---|---|---|---|
| 2 | bila( < Mande), nyʊ̀ʊ̀ | 8 | 5+3 |
| 3 | sã̀ã̀be (< Mande) | 9 | 5+4 |
| 4 | na | 10 | nuunu (< *5redupl.), *ji/yi |
| 5 | tɔ | 20 | yipì-/dʒipi- |
| 6 | 5+1 | 100 | kɛmɛ̀ (< Mande) |
|   |   | 1000 | wulo (< Mande) |

The source form of the term for 'one' with a nasalized vowel is reconstructed on the basis of the evidence presented by Stefan Elders (2007: 323). As we have seen, the Gur term for 'five' is reconstructed as *nu on the basis of the evidence provided by the groups discussed above. It should be noted that this form goes back to the lexical root meaning 'hand' (Kulango *nu-gò*). The term for 'ten' in Kulango is a reduplicated *nu, whereas a different root is attested for 'five'. It is also noteworthy that the terms for 'two', 'three', 'hundred' and 'thousand' are borrowed from Mande.

### 3.9.4 Lobi-Dyan

According to Anthony Naden's classification (Naden 1989), these languages belong to different groups of the Gur languages, so their evidence will be presented separately.

"More recent classifications (Labouret and Manessy) regarded Lobi (Lobiri) and Ja̰a̰ne as closely related" (Miehe & Tham 2007: 212) (Table 3.182).

Table 3.182: Lobi-Dyan numerals

|      | Lobi        | Dyan                    | *Lobi-Dyan              |
|------|-------------|-------------------------|-------------------------|
| 1    | bìɛ̀l, *do   | bɛ̃g/ɓɛ̰̀(ŋ)kù/bɪɛle, *dṵ̀ | bɪɛ̀l, *dò              |
| 2    | nyɔ̀/nɔ̀     | nyɔ̰̀                    | nyɔ̀(n)                 |
| 3    | tʰɛ̆r        | thɛ̰̀s(i)                | thɛ̰̀s(i)/tʰɛ̆r          |
| 4    | nɑ̃́         | nàà̰                    | nɑ̃́                     |
| 5    | mɔì/*mà     | dìèmà, *mɔlɔ̀            | mɔì/*mà/*mɔlɔ̀, dìèmà,   |
| 6    | 5+1         | 5+1                     | 5+1                     |
| 7    | 5+2         | 5+2                     | 5+2                     |
| 8    | 5+3         | 5+3                     | 5+3                     |
| 9    | 10-1        | 10-1                    | 10-1                    |
| 10   | nyɔ̀ɔ́r      | ni-kpo                  | ni-kpo, nyɔ̀ɔ́r          |
| 20   | kpèle       | ceeru                   | kpèle, ceeru            |
| 100  | tàmâ        | tàmúgú                  | tàmâ                    |
| 1000 | gbɔ̀lanɪ    | 100*10                  | gbɔ̀lanɪ, 100*10        |

### 3.9.5 Senufo

Table 3.183: Senufo numerals

| 1 | nɔ̀n-, ni-ŋgbe/nuŋgba, nìkǐ/ningin | 7 | 5+2, 6+1 |
|---|---|---|---|
| 2 | sin/soin/sun/syen | 8 | 5+3, 6+2 |
| 3 | tǎ̃ã/taàr | 9 | 5+4, 10-1, 6+3 |
| 4 | tésyàr/sīcɛ̄rɛ̄/tityere | 10 | kɛ |
| 5 | bwa/bwɔ, guru/kuru (<'fist'), guno, (nɔ) | 20 | gbèɲ/ɠɓēy, fulo, toko/togo, nafa, isolated forms |
| 6 | kwaɲ̀/kwāy, gbaara, gɔlɔŋ , 5+1, (nõli) | 100 | 20*5, lafa (< Kwa) |
| | | 1000 | 200*5, (gben-, bɔlɔ, pwoo, sakere) |

Many of the forms are quoted in brackets, i.e. they are isolated forms attested within the Senufo group comprising about fifteen idioms. As in a number of other Gur branches, the last syllable/segment of a numerical term often represents a coordinating noun class suffix. Below is an excerpt from the table showing the inflection of numerals by class in Tenyer (Syer variety), as published by Klaudia Dombrowsky-Hahn in (Miehe; Reineke; Winkelmann 2007a:420) (Table 3.184):

Table 3.184: Tenyer numerals (a fragment)

| Class SG | u | li | ke | te dim. |
|---|---|---|---|---|
| 'one' | nun | nuni | nuŋ | nunge |
| Class PL | pi | ki | yi | te dim. |
| 'two' | syob ~ syou | syã | syii | syimbi |
| 'three' | trab | tar | tar | tarbi |
| 'four' | tikyireb | tihyɛr | tihyɛr | tihyɛrbi |

This presentation illustrates how problematic defining the numerical roots can be.

### 3.9.6 Teen

Table 3.185: Teen numerals

| 1 | tani | 7 | 5+2 |
|---|------|------|-------|
| 2 | nyor | 8 | 5+3 |
| 3 | sanr | 9 | 10-1 |
| 4 | nan | 10 | pɔrwɔ |
| 5 | tɔ | 20 | toko |
| 6 | 5+1 | 100 | 20*5 |
| | | 1000 | danyɛ |

### 3.9.7 Tiefo

Table 3.186: Tiefo numerals

| 1 | dɛ̰̀ | 7 | 5+2 |
|---|------|------|-------|
| 2 | jɔ̃ | 8 | 5+3 |
| 3 | sá̰ | 9 | 5+4 |
| 4 | ʔuʔɔ̰́/ŋɔɔ | 10 | támú, kɛ̃ |
| 5 | kà̰ | 20 | kpã |
| 6 | 5+1 | 100 | 20*5 |
| | | 1000 | waga (< Mande) |

### 3.9.8 Tusia

Table 3.187: Tusia numerals

| 1 | nɔ́nkì, *nɔ̃̌ŋ | 7 | 5+2 |
|---|---|---|---|
| 2 | nínɔ́, *nĩ̌ŋ | 8 | 5+3 |
| 3 | tɔ̃́nɔ́ | 9 | 5+4 |
| 4 | ńyɑ̃́h/jɑ̃̌ | 10 | gbãm/*gbɔ̃/bwɔ̀ |
| 5 | k(w)lɔ́ | 20 | túkúrí, *tiki |
| 6 | 5+1 | 100 | 20*5, kwɛ̌ |
|  |  | 1000 | < píy 'goat', nɑ̃́ˤ'cow' |

### 3.9.9 Viemo

Table 3.188: Viemo numerals

| 1 | dũde, *dun- | 7 | 5+2? |
|---|---|---|---|
| 2 | niinĩ | 8 | 4*2, 5+3 |
| 3 | sãsĩ | 9 | 10-1 |
| 4 | jumĩ | 10 | kwɔmũ |
| 5 | kuɛge, *kɔ̠ | 20 | fɛrɛyɔ |
| 6 | 5+1 | 100 | tãmõ |
|  |  | 1000 | vie-? |

### 3.9.10 Wara-Natioro

It should be noted that the most important evidence pertaining to this group is relatively recent. In his publication of the comparative lexical list Tasséré Sawadogo noted that Faniagara is radically different from both Wara and Natioro (Sawadogo 2002). Its similarity index with the Natioro and Wara dialects is 12 and 30 percent respectively (the SIL list? idem., p. 15). Thus he had every reason to postulate the existence of an isolated language (Palɛn) in the Wara-Natioro group.

Since the data collected by Tasséré Sawadogo is absent from the major databases that are now incorporated into the RefLex database by Guillaume Segerer, it

seems reasonable to present it below for each Wara-Natioro-Paleni idiom in order to suggest the reconstruction of numerical terms within each of the three sub-groups and within the group as a whole (Table 3.189):

Table 3.189: Wara-Natioro-Paleni numerals

|  |  | '1' | '2' | '3' | '4' | '5' |
|---|---|---|---|---|---|---|
| Natioro | Dinaoro | ká:bà | ɲíndí | táe | ŋnáe | sùsú |
| Natioro | Timba | ká:bà | ɲíndí | tá | nã́ | sùsú |
| Natioro | Kawara | kābà | ɲìdí | tá | ná | sùsú |
| *Natioro |  | ká:bà (ka-ba?)[24] | ɲíndí | tá(é) | ná (é) | sùsú |
| Wara? | Sourani | pɔ́ | bǒ | tã́ | nàsá | sùsú |
| Wara | Negeni | kàpó | bǒ | tí: | nã́:sṹ | sùsú |
| Wara | Niansogoni | pɔ́:wò | bǒ | tí: | ná:só | sùsú |
| *Wara |  | pɔ́ | bǒ, *nĩ́ntó | tã́(i) | naaso | sùsú, |
| Palɛn | Faniagara | káfā | bá | tã́:ré | ná:ré | sùsú |
| *Palɛn | Faniagara | ká-fā | bá, *nínté | tã́:ré | ná:ré | sùsú, *si/sɔ |
| **Wara-Natioro-Paleni** | | **ba/fa, pɔ** | **nínté, bǒ** | **ta(r)i** | **na(r)i** | **sùsú, sV** |
|  |  | '6' | '7' | '8' | '9' | '10' |
| Natioro | Dinaoro | ŋzàbɔ́ | té:ndé | nã́ŋgànã́ŋgánì | kâwó | pwɔ̀: |
| Natioro | Timba | ŋ̀zà:bɔ́ | dé:ndí | náŋgánáŋgánì | kāwɔ̀mṹ | pwɔ́: |
| Natioro | Kawara | nsàbɔ́ | tèndí | nàŋgānàŋgádí | kàwũ̀mò | pɔ́ |
| *Natioro |  | nsàbɔ́ (sa-1?) | téndí | 4+4 | kawo | p(w)ɔ́ |
| Wara? | Sourani | sùrpɔ́ | sūrùdó | sìntã́ | sìn:á | kã̀n:sú |
| Wara | Negeni | sírípò | sĩ́nĩ̄ntó | sĩntí | sīn:á:sṹ | kã̀:sã́ |
| Wara | Niansogoni | sírìpò | sùrùntó | sī:ntí: | sín:ã́:sṹ | kã̀:sã́ |
| *Wara |  | si-1 | si-2 | si-3 | si-4 | kã̀:sã́ |
| Palɛn | Faniagara | sĩ́nífā | sĩ́nínté | sɔ̄tá:ré | sɔn:á:ré | fɔ́ |
| *Palɛn | Faniagara | si-1 | si-2 | sɔ̄-3 | sɔ̄-4 | fɔ́ |
| **Wara-Natioro-Paleni** | | **5+1** | **5+2, téndí?** | **5+3, 4+4** | **5+4, kawo?** | **p(w)ɔ/fɔ, kã̀:sã́?** |

---

[24]Regarding the Natioro forms for 'one' André Prost remarks: '*puwolo* (après un substantif: *kaaba)*' (Prost 1968: 78). Thus, the opposition between the Wara and Natioro forms of 'one'

According to other sources, the forms *wɑ̃́/ nwõ, sɔ* are attested in Wara-Natioro for 'twenty'. The patterns '20*5' and '400*2+200' are attested for 'hundred' and 'thousand' respectively.

### 3.9.11 Proto-Gur

#### 3.9.11.1 'One'

The main forms of 'one' reconstructable in sixteen branches of Gur are as follows (Table 3.190):

Table 3.190: Stems for '1' in Gur

| | | | | |
|---|---|---|---|---|
| A. Bariba | | | tiā | |
| B. Central: | do | | | |
|   1. Northern | | | | |
|     A. Bwamu | | | | |
|     B. Kurumfe | dom | | | |
|     C. *Proto-Oti-Volta | | den/ yen, de? | | ni |
|   Southern | | le | kpò | |
|     A. Dogoso-Khe | | | | |
|     C. Gan-Dogose | | lèŋ | kpo/ po | |
|     D. *Proto-Grusi | do/ du/ lu | de/ re | | |
|     E. Kirma-Tyurama | | déiŋ/ dẽẽná | | |
| C. Kulango | | | ta(a) < *taà̰ | |
| D. Lobi-Dyan | *dò | | | |
| E. Senufo | | | ni- ŋgbe/ nu- ŋgba | nìkì̀/ ningin |
| F. Teen | | | tani | |
| G. Tiefo | | dɛ̃̀ | | |
| H. Tusia | | | | nɔ́nkì |
| I. Viemo | dũde, *dun- | | | |
| J. Wara-Natioro-Paleni | | pɔ | | |

---

reflected in the table may be purely functional (for Wara Prost quotes the *puwo* and *kapo* forms).

An attempt to reconstruct a Proto-Gur form is probably not reasonable at this point, since all the forms quoted above are important for comparative purposes.

### 3.9.11.2 'Two'

Table 3.191: Stems for '2' in Gur

|  | '2' | '2' | '2' | '2' | '2' |
|---|---|---|---|---|---|
| A. Bariba | ru | | | | |
| B. Central: | | | | | |
| 1. Northern | | | | | |
| A. Bwamu | ɲū | | | | |
| B. Kurumfe | | | | hĩĩ | |
| C. *Proto-Oti-Volta | | li/yi | | | |
| Southern | | | | | |
| A. Dogoso-Khe | jɔ(n) | | | | |
| C. Gan-Dogose | yɔ́/ ɲɔ̱/ dʒɔ̰̀ŋ | | | | |
| D. *Proto-Grusi | | lɛ/ le | ne/ ɲi | | |
| E. Kirma-Tyurama | | | | hã́ĩ/ hãl | |
| C. Kulango | nyʊ̀ʊ̀ | | | | bila (< Mande) |
| D. Lobi-Dyan | nyɔ̀(n) | | | | |
| E. Senufo | | | | | sin/ soin/ sun/ syen |
| F. Teen | nyor | | | | |
| G. Tiefo | jɔ̃ | | | | |
| H. Tusia | | | nínɔ́, *nĩ̃ŋ | | |
| I. Viemo | | | niinĩ | | |
| J. Wara-Natioro-Paleni | | | nĩ́nté | | bŏ |

Apparent isolates and obvious borrowings are presented in the rightmost column.

### 3.9.11.3 'Three' and 'Four'

Table 3.192: Stems for '3' and '4' in Gur

|  | 3 | 3 | 4 | 4 |
|---|---|---|---|---|
| A. Bariba | i-ta |  | ǹ-nɛ |  |
| B. Central: |  |  |  |  |
| 1. Northern |  |  |  |  |
| A. Bwamu | tĭ |  | náa |  |
| B. Kurumfe | tãã |  | nãã |  |
| C. *Proto-Oti-Volta | ta(t) |  | naa(si) |  |
| Southern |  |  |  |  |
| A. Dogoso-Khe | thɔ |  | dáa |  |
| C. Gan-Dogose | sá̰a̰/ tʰɔʔ |  | ɲee/ ì-yḭ̀ḭ̀, (á-dàa) |  |
| D. *Proto-Grusi | toro/ toso/ tɔ |  | naare/ naasi/ na |  |
| E. Kirma-Tyurama | síɛi/ siɛl |  | na(a) |  |
| C. Kulango |  | sããbe (< Mande) | na |  |
| D. Lobi-Dyan | thɛ̃̀s(i)/ tʰɛr |  | nã́ |  |
| E. Senufo | tã̀ã̀/ taàr |  |  | tésyàr/ sīcɛ̄rɛ̄/ tityere |
| F. Teen | sanr |  | nan |  |
| G. Tiefo | sã́ |  |  | ʔuʔɔ̃́/ ŋɔɔ |
| H. Tusia | tɔ́nɔ́ |  | ńyã́h/ jã̃ |  |
| I. Viemo | sãsĭ |  |  | jumĭ |
| J. Wara-Natioro-Paleni | ta(r)i |  | na(r)i |  |

The reflexes of the most persistent NC roots are observable in the majority of the branches.

### 3.9.11.4 'Five'

Table 3.193: Stems for '5' in Gur

| | '5' | '5' | '5' | '5' | '5' |
|---|---|---|---|---|---|
| A. Bariba | nɔɔbù | | | | |
| B. Central: | | | | | |
| 1. Northern | | | | | |
| A. Bwamu | hò-nú | | | | |
| B. Kurumfe | nɔm | | | | |
| C. *Proto-Oti-Volta | nu | | | | |
| Southern | | | | | |
| A. Dogoso-Khe | nɔ(n) | | | | |
| C. Gan-Dogose | nə̀n | mwã/ wə̀a | | | |
| D. *Proto-Grusi | nu/ nʊ | | | | |
| E. Kirma-Tyurama | | | | di | |
| C. Kulango | | | tɔ | | |
| D. Lobi-Dyan | | mɔ̀ì/ *mà/ *mɔ̀lɔ̀ | | dìèmà | |
| E. Senufo | guno, (nɔ) | bwa/ bwɔ | | | |
| F. Teen | | | tɔ | | |
| G. Tiefo | | | | | kə̀ |
| H. Tusia | | | | | k(w)lɔ́ |
| I. Viemo | | | | | kuɛge, *kɔ̱ |
| J. Wara-Natioro-Paleni | | | sùsú, sV | | |

The etymological relationship of *nu '5' and 'hand', is attested in Central Gur and possibly in Bariba and Senufo. Isolated bases may go back to this meaning as well. At the same time, the base preserved in Kulango, Teen and possibly Wara-Natioro-Paleni is comparable to *tan found in BC and some other families.

Table 3.194: Stems and patterns for '6' and '7' in Gur

| | '6' | '6' | '7' | '7' | '7' |
|---|---|---|---|---|---|
| A. Bariba | 5+1 | | 5+2 | | |
| B. Central: | | | | | |
| 1. Northern | | | | | |
| A. Bwamu | 5+1 | | 5+2 | | |
| B. Kurumfe | | hʊrʊ | | pɛ̃ɛ̃ | |
| C. *Proto-Oti-Volta | | lob/ yob | | poi(n)? | *lob-le (6+1)? |
| Southern | | | | | |
| A. Dogoso-Khe | 5+1 | | 5+2 | | |
| C. Gan-Dogose | 5+1 | | 5+2 | | |
| D. *Proto-Grusi | 5+1 | dʊ/ | 5+2 | pɛ/ | |
| | | lo-ɖo/ | | lʊ-pɛ/ | |
| | | lo-ro | | lʊ-bɛ | |
| E. Kirma-Tyurama | 5+1 | | 5+2 | | |
| C. Kulango | 5+1 | | 5+2 | | |
| D. Lobi-Dyan | 5+1 | | 5+2 | | |
| E. Senufo | 5+1, | kwaɲ̀/ | 5+2 | | 6+1 |
| | | kwāy, | | | |
| | | gbaara, | | | |
| | | nõli | | | |
| F. Teen | 5+1 | | 5+2 | | |
| G. Tiefo | 5+1 | | 5+2 | | |
| H. Tusia | 5+1 | | 5+2 | | |
| I. Viemo | 5+1 | | 5+2? | | |
| J. Wara-Natioro-Paleni | 5+1 | | 5+2 | | téndí? |

### 3.9.11.5 'Six' and 'Seven'

The patterns *'6=5+1' and *'7=5+2' can be safely reconstructed at the Proto-Gur level. The exeptionally wide range of forms for 'six' attested in Senufo is noteworthy.

### 3.9.11.6 'Eight' and 'Nine'

Table 3.195: Stems and patterns for '8' and '9' in Gur

| | '8' | '8' | '8' | '9' | '9' | '9' |
|---|---|---|---|---|---|---|
| A. Bariba | 5+3 | | | 5+4 | | |
| B. Central: | | | | | | |
|   1. Northern | | | | | | |
|     A. Bwamu | 5+3 | | | | | dɨ̀iní/ dènú |
|     B. Kurumfe | | | cɔɔ | | | fa |
|     C. *Proto-Oti-Volta | | | ni | | | wey/weʔ |
|   Southern | | | | | | |
|     A. Dogoso-Khe | 5+3 | | | 5+4 | | |
|     C. Gan-Dogose | 5+3 | | | 5+4 | 10-1 | |
|     D. *Proto-Grusi | | 4 redupl. | | | 10-1 | nibi/nibu (ni-bi/bu?) |
|     E. Kirma-Tyurama | 5+3 | | | 5+4 | 10-1 | |
| C. Kulango | 5+3 | | | 5+4 | | |
| D. Lobi-Dyan | 5+3 | | | | 10-1 | |
| E. Senufo | 5+3 | | 6+2 | 5+4 | 10-1 | 6+3 |
| F. Teen | 5+3 | | | | 10-1 | |
| G. Tiefo | 5+3 | | | 5+4 | | |
| H. Tusia | 5+3 | | | 5+4 | | |
| I. Viemo | 5+3 | 4*2 | | | 10-1 | |
| J. Wara-Natioro-Paleni | 5+3 | 4+4 | | 5+4 | | kawo? |

In addition to the common patterns '8=5+3' and '9=5+4', alternative ones are attested for 'eight' and 'nine' ('8=4 redupl.' and '9=10-1' respectively).

### 3.9.11.7  'Ten'

Table 3.196: Stems for '10' in Gur

| | | | | | |
|---|---|---|---|---|---|
| A. Bariba | | wɔ-kuru | | | |
| B. Central: | | | | | |
| 1. Northern | | | | | |
| A. Bwamu | pílú/ píru/ ˀɓúrúù | | | | |
| B. Kurumfe | fɪ | | | | |
| C. *Proto-Oti-Volta | pi(k) | | | | |
| Southern | | | | | |
| A. Dogoso-Khe | kpélé | | | | |
| C. Gan-Dogose | | kpoogo | nõy - 5PL | | gbùnè, kpélé, sí- |
| D. *Proto-Grusi | fu/ fi | | | | |
| E. Kirma-Tyurama | | | nűɔ́sɔ̀ | | cíŋcíelùó |
| C. Kulango | | | | | nuunu (< *5 redupl.), *ji/ yi |
| D. Lobi-Dyan | | ni-kpo | nyɔ̀ɔ́r | | |
| E. Senufo | | | | kɛ | |
| F. Teen | pɔrwɔ | | | | |
| G. Tiefo | | | | kɛ̃ | támú |
| H. Tusia | | | | | gbãm/ *gbɔ̃/ bwɔ̀ |
| I. Viemo | | kwɔmũ | | | |
| J. Wara-Natioro-Paleni | p(w)ɔ/ fɔ | | | | kằːsá? |

This term exhibits a variety of isolated (and possibly non-primary) forms. The main form has a voiceless labial as its initial consonant.

### 3.9.11.8 'Twenty'

Table 3.197: Stems and patterns for '20' in Gur

|  | '20' | '20' | '20' | '20' | '20' |
|---|---|---|---|---|---|
| A. Bariba |  |  |  |  | yɛndu |
| B. Central: |  |  |  |  |  |
|   1. Northern |  |  |  |  |  |
|     A. Bwamu |  | ɓóní/ ɓénle/ kēwēníì |  |  |  |
|     B. Kurumfe | sofe (<10?) |  |  |  |  |
|     C. *Proto-Oti-Volta | 10*2 |  |  |  |  |
|   Southern |  |  |  |  |  |
|     A. Dogoso-Khe |  | gʊ̀ʊsì | cúkúrì |  |  |
|     C. Gan-Dogose |  | gbeere | tʃúkúrì |  |  |
|     D. *Proto-Grusi | 10*2? |  |  |  |  |
|     E. Kirma-Tyurama |  | guř̃ |  |  | kómòrré |
| C. Kulango |  |  |  |  | yipì-/ dʒipi- |
| D. Lobi-Dyan |  | kpèle | ceeru |  |  |
| E. Senufo |  | gbèɲ/ gbēy, |  | toko/ togo toko | fulo, nafa |
| F. Teen |  |  |  | toko |  |
| G. Tiefo |  |  |  |  | kpã |
| H. Tusia |  |  | túkúrí |  | *tiki |
| I. Viemo |  |  |  |  | fɛrɛyɔ |
| J. Wara-Natioro-Paleni |  |  |  |  | wǎ/ nwõ, sɔ |

In view of the great variety of forms and patterns attested for this term, the existence of the term for 'twenty' in Proto-Gur is uncertain.

### 3.9.11.9 'Hundred'

Table 3.198: Stems and patterns for '100' in Gur

| | | | | |
|---|---|---|---|---|
| A. Bariba | 20*5 | | | |
| B. Central: | | | | |
| 1. Northern | | | | |
| A. Bwamu | | | | kʰĭminù (< Mande keme) |
| B. Kurumfe | | | bɛrʊ | |
| C. *Proto-Oti-Volta | | kob, kook | | |
| Southern | | | | |
| A. Dogoso-Khe | 20*5 | | | |
| C. Gan-Dogose | 20*5 | | | |
| D. *Proto-Grusi | 20*5? | kɔwa/ kɔɔ? | bi? | |
| E. Kirma-Tyurama | 20*5 | | gundi | |
| C. Kulango | | | | kɛmɛ̀ (< Mande) |
| D. Lobi-Dyan | | tàmâ | | |
| E. Senufo | 20*5 | | | lafa (< Kwa) |
| F. Teen | 20*5 | | | |
| G. Tiefo | 20*5 | | | |
| H. Tusia | 20*5 | kwĕ | | |
| I. Viemo | | tãmõ | | |
| J. Wara-Natioro-Paleni | 20*5 | | | |

**3.9.11.10** 'Thousand'

Table 3.199: Stems and patterns for '1000' in Gur

| | | | |
|---|---|---|---|
| A. Bariba | | | fɔ̀rɔ̀tɔ? |
| B. Central: | | | |
| 1. Northern | | | |
| A. Bwamu | | 100*10 | muaseé |
| B. Kurumfe | | | tʊsrɪ (< Moore) |
| C. *Proto-Oti-Volta | | | |
| Southern | | | |
| A. Dogoso-Khe | kpɛ́ | | |
| C. Gan-Dogose | kpíɛ 'a goat' | | |
| D. *Proto-Grusi | | | kpoŋ/ɠboŋ |
| E. Kirma-Tyurama | | 200*5, 800+200 | |
| C. Kulango | | | wulo (< Mande) |
| D. Lobi-Dyan | | 100*10 | gbʊ̀lanɪ |
| E. Senufo | | 200*5 | gben-, bɔlɔ, pwoo, sakere |
| F. Teen | | | danyɛ |
| G. Tiefo | | | waga (< Mande) |
| H. Tusia | < píy 'goat', ná̵ˤ 'cow' | | |
| I. Viemo | vie-? | | |
| J. Wara-Natioro-Paleni | | 400*2+20 | |

No evidence supports the reconstruction of the term for 'thousand' in this family.

## 3.10  Mande

The intermediate step-by-step reconstructions available for the Mande languages in Vydrin's Mande Etymological Dictionary (ms) and in Vydrin 2007[25] has made treatment of the data easier.

The genetic classification of Mande, outlined in the latter work, will serve as the basis for our analysis. This classification differs from the one suggested by R.Kastenholz and is accessible via *Ethnolog*. According to V.Vydrin, 'Its major innovations, in comparison with that of Kastenholz, are the following:

– the Susu–Jalonke group is put together with the Southwestern group, rather than with Kastenholz's "Central Mande" (in fact, it is a return to the proposal of André Prost (1958));

– Soninke–Bozo, Samogho and Bobo are no longer considered as branches of the same genetic unit (Kastenholz's "Northwestern Mande"), but rather as independent groups inside Western Mande;

– the Mokole group is put together with Vai–Kono, rather than with Manding;

– in the Southern Mande group, Mwan is separated from Wan and put together with the Guro–Yaure subgroup;

– San (Samo) is put together with Bisa, rather than with Busa-Boko.' (Vydrin 2016: 110).

Let us note an important fact: the numeral system of Jowulu differs considerably in certain points both from other Samogho languages and from Mande languages in general. It is interesting to outline that in R. Kastenholz's classification (based on the method of shared innovations, rather than on lexicostatistics) Jowulu is given a special status, more precisely, the first split in his Northwestern Mande branch (Bozo-Soninke + Bobo + Samogo + Jowulu).

Our further analysis will be based on the evidence from twelve branches of Mande represented in the diagram below.

### 3.10.1  'One'

Vydrin's preliminary reconstructions, as well as isolated forms resulting from the analysis of the numerical terms, are marked with an asterisk [*].

The isoglosses for 'one' suggest the existence of two alternative roots (*dǫ́* and *kelen*) attested in both major Mande groups. The latter root is distinguishable under the assumption that the forms with a voiced velar attested in the Eastern

---

[25]I would like to thank V. Vydrin for his suggestions and comments on the preliminary draft of this chapter.

Table 3.200: Scheme 3.1. Mande languages

| 1. Manding | | 7. Bozo-Soninke | | |
|---|---|---|---|---|
| 2. Jogo-Jeri | | 8. Bobo | | |
| 3. Mokole | 5. Susu | 9. Samogo | | 11. Eastern |
| 4. Vai-Kono | 6. SWM | 10. Jowulu | | 12. Southern |

Table 3.201: Mande stems for '1'

| | | | | |
|---|---|---|---|---|
| Manding | *dɔ́ | *kélen | | |
| Jogo-Jeri | *do | *kɛlɛ (?) | | díé(n)/ dúlì |
| Mokole | *dɔ́ndɔ̀ | *kéle̱ | | |
| Vai-Kono | *dɔ́ndɔ | *N-kélen | | |
| Susu | | *kédén | nde/ ǹdá | |
| SWM | | *gìláaŋ | *tà | |
| Bozo-Soninke | | kuɔn/ kɛnɛ/ ke̱/ ko̱ | | sa̱na̱ | bane, fie |
| Bobo | | | tàlá/ tèlé | |
| Dzuun (Samogo) | | *ké | | *so/ sɔʔi/ swɛ̃̄ |
| Jowulu | | | tẽ̄ẽna/ tenŋ | |
| SE-Eastern | *do | gɔ̀rɔ́/ gôon? | | |
| SE-Southern | *dɔ̣́ | | | |

branch of the South-Eastern group (Matya Samo *gɔ̀rɔ́*, Southern Samo (Maka) *gôon*) are related to the **k**-forms found in Western Mande.

The next two roots, if related, may be suggestive with regard to the classification of Western Mande (otherwise, they probably represent similar unrelated forms). It should be noted that the root *ǹdá* (Susu *nde* 'one, certain', *ndende* 'anybody, whoever; nobody', Jalonke *ǹdá* 'certain') attested, according to Vydrin, in Susu-Jalonke may be related to *\*dọ*. The determiner *\*dọ́*, which can be reconstructed at the Proto-Mande level, goes back to the root *\*do*.

The rightmost column of the table embraces the isolated forms.

### 3.10.2 'Two'

Table 3.202: Mande stems for '2'

| | |
|---|---|
| **Manding** | *fìlá |
| **Jogo-Jeri** | *fàlá |
| **Mokole** | *fìla |
| **Vai-Kono** | *fèLá |
| Susu | *fìdíń |
| SWM | *fèelé |
| Bozo-Soninke | pè:ndé, fíllò |
| Bobo | pálà |
| Dzuun (Samogo) | fĩ:(kí) |
| Jowulu | fúúli |
| SE-Eastern | *pela |
| SE-Southern | *pìì-lāŋ |

A common root for 'two' that may be tentatively recorded as *pila / fila* is attested in all Mande branches. Its precise phonetic reconstruction is beyond the scope of our investigation. The reader can refer to the works of specialists in the historical phonetics of Mande. A reference designation that will enable us to compare this root to the evidence of the other NC families is sufficient for our reconstruction purposes.

### 3.10.3 'Three'

Table 3.203: Mande stems for '3'

| | | |
|---|---|---|
| Manding | sàbá | |
| Jogo-Jeri | sègbá/sigbù | |
| Mokole | sàwa/saba | |
| Vai-Kono | sàkpá/sagba/sáwa | |
| Susu | sàxán/sàqáŋ/sawa | |
| SWM | sàwá/sāaɓā | |
| Bozo-Soninke | síkkò, sike | |
| Bobo | sàà (?) | |
| Dzuun (Samogo) | ʒiʔi/ʒìːgī ́/ʃwɛ̀/ɣei | |
| Jowulu | bʒei < *jɔnŋ/i? | |
| SE-Eastern | sɔɔ/cɔ́w? | ʔààkɔ̃ |
| SE-Southern | | *yààká |

The common root *sakpa/ sagba/ sawa is represented in all Western branches. The relationship between some of the forms attested in the Eastern group (Southern Samo (Maka) sɔ̄ɔ̄, Matya Samo tjɔwɔ) remains uncertain. The Jowulu form is especially peculiar. It should be noted that the forms of some numerical terms differ significantly depending on the source. Our study is based on four Jowulu sources that provide the following evidence[26] (Table 3.204):

The terms for 'seven', 'eight' and 'nine' follow the pattern '3,2,1+'to lose'' respectively (cf. their inaccurate interpretation in Hochstetler, see §3.10.9), hence the reconstruction of the term for 'three' with the initial palatal (*jɔ̀n). The forms quoted in Jowulu for 'three', 'four', and 'ten' are uncommon. If we were dealing with a language with a noun class system, we would have to conclude that a noun class marker (cl19?) with two allomorphs (**p**- and **b**- before voiced and voiceless respectively) is traceable in the pertinent forms. However, we are dealing with a language that undoubtedly belongs to Mande, so no class-related morphemes can be involved. This leaves the presence of the initial labial in the term for 'three' unexplained. A borrowing from Gur or Kru cannot be assumed since these languages lack the comparable forms. The only plausible solution is the alignment of 'three' and 'four' by analogy with 'ten' where it must have been originally

---

[26]Hochstetler https://mpi-lingweb.shh.mpg.de/numeral/Jowulu.htm, Djilla et al. (2004); Carlson (1993); Prost (1958).

Table 3.204: Jowulu numerals

| Source | '1' | '2' | '3' | '4' | '5' |
|---|---|---|---|---|---|
| Hochstetler (1996) | tẽẽna | fuuli | bʒei, *dʒɔ̃ | pʃɪrɛˈ | tãã |
| Djilla et al. (2004) | tenŋ | fúúli | byàŋ, *jɔ̀n | pyiiraŋ | táánŋ |
| Carlson (1993) | tɛ̀ɛ̀nì | fuˈuˈlī | byāɪ̄ *jɔ̄ɔ̠̄ | piˈiˈrēī | ta̠ˈa̠ˈ |
| Prost (1958) | têna | fole | dyue, *dyô | piœe | tâ |

| Source | '6' | '7' | '8' | '9' | '10' |
|---|---|---|---|---|---|
| Hochstetler (1996) | tãmãnɪ | dʒɔ̃m-pʊn | ful-pʊn | tẽm-pʊn | bʒĩ |
| Djilla et al. (2004) | táán-mání | jòn-pɔnni | fuuli-pɔnni | ten-pɔnni | byìnŋ |
| Carlson (1993) | ta̠ˈa̠ˈ-mānī | jɔ̄ɔ̠̄-poˈnì | fuˈl-poˈnì | tɛ̀ɛ̀-poˈnì | byì̠ |
| Prost (1958) | ton-te | dyômpônô | filepônô | têpônô | bî |

present.

A special term for 'three' appears in South-Eastern. In Eastern it can be reconstructed as *ʔààkɔ̃ or possibly **ʔàà-(kɔ̃), cf. Bisa kakʊ́, Boko ʔààɔ̃ (in Koelle 1963[1854] ááɣo̠), Bokobaru (Zogbẽ) ʔààgɔ̃, Busa ʔààkɔ̃, Maya Samo kàakú, Kyanga ˀāàː, and Shanga ʔà. The latter reconstruction is supported by the fact that the terms for 'three' and 'four' share the ultima, cf. the data are presented in the following table (Table 3.205):

Table 3.205: Final morphemes in the Boko-Busa numerals

| | Boko | Boko (Koelle 1963[1854]) | Bokobaru | Busa |
|---|---|---|---|---|
| '3' | ʔàà-ɔ̃ | áá-ɣo̠ | ʔàà-gɔ̃ | ʔàà-kɔ̃ |
| '4' | síí-ɔ̃ | síí-ɣo̠ | síí-gɔ̃ | ʃíí-kɔ̃ |

It should be noted that in these languages, the syllable in question is also present in the terms for 'eight' that are built according to the pattern '5+3' (cf. e.g. Bobo Karu sɔ́r-ààgɔ̃). Here we may be dealing with alignment by analogy, possibly with an additional final morpheme of uncertain meaning. It should be stressed that the ultima in 'three' and 'four' is never the same in the Eastern subgroup of the South-Eastern languages, whereas the medial velar is only attested in 'three' but not in 'four'. Assuming that the forms of the two Eastern branches are related, the term for 'three' can be reconstructed as *ʔààkɔ̃/yààká, whereas

the term for 'four' may be interpreted as resulting from the alignment by analogy with the forms of 'three' attested in the Eastern branch of South-Eastern Mande. The evidence in favor of its etymological connection with *\*sakpa* is inconclusive.

### 3.10.4 'Four'

Table 3.206: Mande stems for '4'

| | | |
|---|---|---|
| Manding | *náani | |
| Jogo-Jeri | náani | |
| Mokole | náani | |
| Vai-Kono | nááñì | |
| Susu | náání | |
| SWM | *nááñì | |
| Bozo-Soninke | na:na/nàtǻ/nà:rá/naxat- | |
| Bobo | náà/nìã̀ | |
| Dzuun (Samogo) | nã̀ãi/naai/nà:lẽ́ | |
| Jowulu | | pʃɪrɛˈ <ʃɪrɛˈ? |
| SE-Eastern | | sɪ̀/sííkɔ̃ |
| SE-Southern | | *yìì-sɪ̰̀ỳá: zɪ̀ɛ̃́/yîî-sīɛ̈ |

An easily recognizable NC form (*\*náání/ nã̀ãi*) can be reconstructed in Western Mande, whereas in South-Eastern Mande it is replaced with an innovation (*\*sɪ̰̀yɑ̰̀*). This innovation may also be attested in Jowulu.

### 3.10.5 'Five'

Table 3.207: Mande stems for '5'

| | | | | |
|---|---|---|---|---|
| Manding | dúuru/loolu | *wo (cf. '7') | | |
| Jogo-Jeri | sóólò/sóolo | | | |
| Mokole | lɔ́ɔlu | *wo (cf. '7') | | |
| Vai-Kono | dúʔu/sóó(ʔ)ú | | | |
| Susu | suuli/sùlù | *fò (cf. '7') | | |
| SWM | dɔ́ɔlú/lɔ́ɔlu | *wɔ/ngɔ̀ | | |
| Bozo-Soninke | | kɔ́lɔ́hɔ̀/kárágò | | |
| Bobo | | kō/kóò | | |
| Dzuun (Samogo) | | | | nṳ̀ |
| Jowulu | | | tãã | |
| SE-Eastern | *sodu: sɔ́ɔro/sɔ́ɔ̀ | | | |
| SE-Southern | sɔ́ɔɗú/sólú | | | |

There is a correspondence between **d-/ l-/ s-** within Western Mande, hence the Eastern forms with the initial **s-** should not necessarily be treated separately. A discussion of the exact phonetic reconstruction is better left to specialists in the field. For our purposes, it is sufficient to record that the Proto-Mande root for 'five' is reconstructed as *dúuru/ sɔ́ɔru.*

At the same time, the root(s) *wo, *ko are traceable in the compound numerical terms attested in Western Mande. They may be etymologically related to the lexical root meaning 'hand' (Vydrin, p.c.; cf. Proto-South-Mande *kɔ̀ 'hand'). The latter may be a NC root, cf. e.g. the term for 'hand' in Proto-Gbaya (*kɔ́*), Dida (Kru) (*kɔ̄*) and in other languages.

The Jowulu and Samogo forms are peculiar. As we hope to demonstrate in the next chapter, two alternative roots for 'five' can be reconstructed for NC, namely *tan/ ton* and *nu(n)*. Both roots are directly attested in these marginal groups. Is this enough to reconstruct the terms for 'five' traceable in NC for the Mande languages? We will return to this question in the last chapter of the book.

### 3.10.6 'Six'

Table 3.208: Mande stems and patterns for '6'

| | | |
|---|---|---|
| Manding | wɔ́rɔ (5+1) | |
| Jogo-Jeri | mɔ̀ɔ̀dó (5+1?)/mì:lù | |
| Mokole | wɔ́ɔrɛ/wɔyɔ (5+1) | |
| Vai-Kono | wɔ́ɔlɔ/wɔɔrɔ (5+1) | |
| Susu | sénní (5+1?) | |
| SWM | *5+1 | |
| Bozo-Soninke | goro? (5+1?) | tű̀mù/tũmi |
| Bobo | 5+1 | |
| Dzuun (Samogo) | | t(s)ù̃mɛ̃̀´/tsìì |
| Jowulu | 5+1 | |
| SE-Eastern | 5+1 | |
| SE-Southern | 5+1, wáŋ́? | |

The reconstruction of the Mande term for 'six' is problematic. The root *t(s)um* is worth considering, since it is attested in both Bozo-Soninke and Samogo (the root found in Susu is probably isolated). Its reconstruction at the Proto-Mande level is, however, unlikely. The common pattern '6=5+1' is attested in both major branches. The root *wɔrɔ* is non-primary and eventually goes back to the aforementioned pattern (or to the pattern '6'='hand'+1' to be precise). This hypothesis is supported by the forms of 'seven' as well.

### 3.10.7 'Seven'

Table 3.209: Mande stems and patterns for '7'

| | | |
|---|---|---|
| Manding | x+2 | |
| Jogo-Jeri | ma+2 | |
| Mokole | x+2 | |
| Vai-Kono | 5+2 | |
| Susu | 5+2 | |
| SWM | 5+2 | |
| Bozo-Soninke | ɲérù/jeeni | |
| Bobo | 5+2 | |
| Dzuun (Samogo) | ɲɛ̃̀:nɛ̃́ (<5?)/ɲɛ̀ɛ̀ | |
| Jowulu | | 3+'to lose' |
| SE-Eastern | 5+2 | |
| SE-Southern | 5+2 | |

A few remarks are in order before we turn to the discussion of the term for 'seven'. In the majority of the Mande branches, the term represents a compound. Its second element goes back to the term for 'two', cf. e.g. Jula *wólonfìlà* '7', *fìlà* '2'.

The relationship between the terms for 'six' and 'seven' is based on alignment by analogy. This bond sometimes results in unification of the terms, so that sources may explain 'seven' as '6+1' (despite the fact that 'two', not 'one', is manifestly present in 'seven'). This interpretation has become recurrent for the Mokole languages. According to Phillip Logan[27], the Kuranko evidence is as follows: *wɔrɔnfila* ('**6+1**') (?! –*K.P.*), *wɔrɔ* '6', *fila* '2', *kelen* '1'. The same idea is applied to Lele (cf. Marc Gebhard[28] : *wɔrɔŋ kela* ('**6+1**')[29] , *wɔɔrɔ* '6', *fela* '2', *kelɛŋ* '1') and Kakabe (cf. Daria Mishchenko[30]: *wɔ́rɔwila* ('**6+1**'), *wɔ́ɔrɔ* '6', *fila* '2', *kélen* '1'). Other scholars are more reserved, stating that 'Kono has a decimal system with special construction for 7'[31]. It is, however, quite evident that the forms in

---

[27]https://mpi-lingweb.shh.mpg.de/numeral/Kuranko.htm

[28]https://mpi-lingweb.shh.mpg.de/numeral/Lele-Mande.htm

[29]According to (Vydrine 2009), the Lele term for 'seven' is *wɔ́rɔncɛla* (or *wɔyɛnkela* in the Southern dialect).https://mpi-lingweb.shh.mpg.de/numeral/Jowulu.htmMende *núú ɠɓɔyɔ́ngo* '20' ('person finished'). https://mpi-lingweb.shh.mpg.de/numeral/Mende.htm

[30]https://mpi-lingweb.shh.mpg.de/numeral/kakabe.htm

[31]Raimund Kastenholz,https://mpi-lingweb.shh.mpg.de/numeral/kono.htm

---

question follow the pattern '5+2' (or at least 'X+2' with X being an unidentified component).

It is not a mere coincidence that the interpretation outlined above is recurrent in the Mokole languages, where the forms of 'six' and 'seven' have become partially unified. In a number of languages from other groups that have etymologically related terms for 'six' and 'seven', these terms differ in their second consonant, cf. Bamana (Manding): *wólonwula* '7', *wɔ́ɔrɔ* '6'.

In both groups of South-Eastern Mande the patterns '5+1' and '5+2' for 'six' and 'seven' respectively are still clearly recognizable (Table 3.210).

Table 3.210: Stems for '6' and '7' in South-Eastern Mande

|                    | '5'   | '1' | '6'     | '2'    | '7'      |
|--------------------|-------|-----|---------|--------|----------|
| SE: Eastern: Busa  | sɔ́o   | do  | sóo-do  | pia    | soo-pia  |
| SE: Southern: Beng | sɔ́-ŋ  | do  | sɔ́-do  | pla-ŋ  | sɔ́-pla  |

Taking all of this into consideration, the most likely evolution scenario for 'six' and 'seven' is as follows:

- At the most archaic Proto-Mande level the terms for 'six', 'seven' (and also 'eight' as we hope to demonstrate below) followed the pattern 'X+1,2,3' respectively. The X-element in this pattern possibly represented an archaic root with the meaning 'hand' (?) *ko(*N-ko > *go/wo?).

- Proto-Mande developed the root *dúuru/ sɔ́ɔru '5'.

- This new root served as the basis for the South-Eastern Mande terms for 'six', 'seven' and 'eight'.

- In Western Mande this process is only attested in single languages, e.g. in Vai (*sóóʔú* '5', *sɔ̂ŋ lɔ̀ndɔ́* '6' (*lɔ̀ndɔ́* '1'), *sɔ̂ŋ fɛ̀ʔá* '7' (*fɛ̀ʔá* '2')) and Looma (*dooluo* '5', *dɔzita* '6', *dɔfela* '7', *dɔ́sáwà* '8').

- The majority of the Western Mande languages retained the inherent forms for 'six' and 'seven', but their derivational motivation became unapparent (at least in the case of the first component, cf. Bandi *ndɔ̀ɔ́lú(ŋ)* '5', but *ngɔ̀hítáŋ* '6' (*hítàŋ* '1') and *ŋgɔ̀félàŋ* '7' (*feelé* '2') in contrast to Looma).

- This factor conditioned the partial unification of the terms for 'six' and 'seven' (by analogy) in some of the Western Mande languages (Mokole in particular).

### 3.10.8 'Eight'

Table 3.211: Mande stems and patterns for '8'

| | | | |
|---|---|---|---|
| Manding | séegi/ séki/ séyi | | |
| Jogo-Jeri | | ma+3 | |
| Mokole | sɛ́ɛn/ saɛn/ seyi | | |
| Vai-Kono | séi/ séin | 5+3 | |
| Susu | | 5+3 | |
| SWM | | wá-yákpá/ wɔ-yaagba/ ngɔ̀sákbá(n) (5+3) | |
| Bozo-Soninke | segi-/ seegu | | |
| Bobo | sɛ́kì/ tʃɛ̀kí | | |
| Dzuun (Samogo) | | | kàà, 4pl |
| Jowulu | | | 2+'to lose' |
| SE-Eastern | | *5+3 | síɲe, kíwísí (<4) |
| SE-Southern | | sɛ́ɛ̀-gā/ sálààkā/ sɔ̀làá/ sé-yā̰ (5+3?) | |

The pattern '8=4\*2'/'4PL' commonly found in the majority of the families dis-cussed above is barely attested in Mande. Meanwhile, the phonetic similarity between *naai* '4' ~ *ŋaai(n)* '8' (attested in the majority of the Samogo dialects) is hardly an accident.

The etymology of *kàà* (not found outside Seenku) is unknown.

The pattern '5+3' is inconclusive, because it often developss independently in various languages. The interpretation of the main Mande root (tentatively de-scribed as *seki/ segi*) is uncertain. On the one hand, its current forms suggest that this root can be reconstructed not only for Proto-Western Mande, but for Proto-Mande as well (cf. South-Eastern forms, in particular *sɛ́ɛ̀gā* '8'). On the other hand, such reconstruction is hindered by at least two issues.

Firstly, the second velar in the South-Eastern Mande forms does not belong to the root. It is part of a reduced segment that goes back to the term for 'three' (cf. Tura *yɛ̀ká* '3'), whereas the first segment goes back to the term for 'five' (cf. Tura *sɔ́lǔ, sɔ́ɔ́lǔ, sʊ́lʊ̌*). The comparative analysis of the forms of 'eight' attested in

the South-Eastern Mande languages (not quoted here in detail) strongly suggests that the South-Eastern Mande pattern for 'eight' is '5+3'.

Secondly, this reconstruction is problematic from a typological point of view. As has been demonstrated above, our evidence prevents us from reconstructing primary roots for 'six' and 'seven'. In terms of typology, a primary root for 'eight' would look highly unusual in this context. Such a root could be expected in those few numeral systems where 'eight' is a basic numeral (just like 'twelve' is a basic numeral in some of the Benue-Congo numeral systems described above, hence '100=12*8+4'). However, 'eight' has never been a basic unit of counting in Mande systems. The existence of a primary term for 'forty' (assuming that 'forty' is '8*5') in some of the Mande languages could be interpreted as a hint at a special status of 'eight'. However, this is not supported by any real evidence.

This raises a question about the etymology of the Western Mande term for 'eight' (*seki/ segi*). Its resemblance to the term for 'three' (especially in Bozo and Soninke, cf. Jenaama Bozo *síkɛ̀ũ* '3' ~ *sɛ̀k:í* '8') may be suggestive here. Is there enough evidence to reject the hypothesis that 'eight' in the Proto-Western Mande was built according to the pattern '8=plus 3' (this would assume a counting reference to 'five')?

Despite the doubts expressed above, these forms are worth comparing to other forms of 'eight' attested in other NC families.

### 3.10.9 'Nine'

Table 3.212: Mande stems and patterns for '9'

| Manding | | kɔ̀nɔntɔ (10-1?) | |
|---|---|---|---|
| Jogo-Jeri | ma+4 | | |
| Mokole | | kɔ̀nɔndɔn (10-1?) | |
| Vai-Kono | 5+4 | kɔ̀nɔ́ntɔn | |
| Susu | 5+4 | | |
| SWM | 5+4 | 10-1 | |
| Bozo-Soninke | | | kàp:í/káfì/kabi |
| Bobo | | kɔ̀rɔ̀nɔ̀ŋ | |
| Dzuun (Samogo) | | | kjè:rɔ̃́/kle:lo/kùòmɛ̀ |
| Jowulu | | 1+'lose' | |
| SE-Eastern | 5+4 | 10-1 | |
| SE-Southern | 5+4 | | |

Two competitive patterns are distinguishable here ('9=5+4' and '9=10-1'). In some of the branches (e.g. SWM, Vai-Kono) they are attested side-by-side.

At the same time, these patterns cannot be postulated for some of the languages without additional support. The pattern '9=10-1' seems to be apparent in South-Eastern Mande and some of the SWM languages only, cf. Boko '9': *kὲokwi* (lit: 'tear away 1 (from) 10'), *kwi* '10' ; in Busa '9': *kɛ́ndo/kĩ́ndokwi* (lit: 'tear away 1 (from) 10'), *kwi* '10', *do* '1'; in Bandi (SWM) *taá-vu* '9', *ìtá(ŋ)* '1', *púu* '10'. According to Robert Carlson (Carlson 1993: 30), the terms from 'seven' to 'nine' in Jowulu follow the pattern '1-3' + 'lose' (*fɔ́nì*), i.e. *jɔ̃ɔ̃-pɔ́nì* '7', *fúl-pɔ́nì* '8', and *tὲὲ-pɔ́nì* '9' (note that these terms are misinterpreted as 3+4, 2*4, 5+4[32] by Lee Hochstetler).

The root *kònonto/kɔ̀nɔ́ndɔ(n)* attested in Manding and Mokole is unclear and deserves discussion by specialists. On the contrary, the forms interpreted as the combination of '5+4' in the table below seem to be quite transparent (Table 3.213):

Table 3.213: '9 = 5+4' in Mande

| Language | '9' | '5' | '4' |
|---|---|---|---|
| Kyanga | sòòʃí | sɔ́ɔ́rū | ʃíí |
| Tura | sɔ́ɨ̀sɛ̄ | sólú | jɨ̀sɛ̄ |
| Susu | sólómánááni | súlí | nááni |
| Vai | sɔ̂ŋ náánì | sóó(ʔ)ú | náánì |
| Bobo Madare | kórónǯ | kóò | náà |

This section, however, is not unproblematic. The Jogo-Jeri non-primary terms for '6-9' are formed by two components. The second (i.e. the terms for 'one', 'two', 'three' and 'four' respectively) is easily recognizable, whereas the etymology of the first (**ma-**) is unclear.

### 3.10.10 'Ten'

This term is especially interesting in light of the fact that the distribution of the isoglosses of 'ten' served as the basis for Maurice Delafosse's early classification of the Mande languages including the *Mande-tan* and *Mande-fu* groups. These two roots are indeed the main Mande roots with this meaning. However, their distribution does not correspond to the two major branches of Mande as they are distinguished today. The root \**tan* is indeed found in all groups of the Western

---

[32]https://mpi-lingweb.shh.mpg.de/numeral/Jowulu.htm

Table 3.214: Mande stems for '10'

| Manding | *tán | *bî | |
|---|---|---|---|
| Jogo-Jeri | táà(n), ta | | |
| Mokole | tán | *bí | |
| Vai-Kono | tâŋ | | |
| Susu | *tòngó | fùú | |
| SWM | | *puu | |
| Bozo-Soninke | tan/téeŋ/cɛmi | | |
| Bobo | | fɔ̀̃ | m̩ḿ |
| Dzuun (Samogo) | t(s)eũ/ceũ | | |
| Jowulu | | | bʒĩĩ /byìnŋ |
| SE-Eastern | | *fu/*vu (<* pu) | kwi/kuri, wókòì |
| SE-Southern | | *bù | gɔ́ɔ̂(dō),kɔ̀ŋ |
| | | | sɔ́jɔlú, |

branch except for Bobo and SWM. At the same time, the attestations of the root *pu/fu are not limited to South-Eastern and extend to a number of the Western branches such as Bobo, SWM, Susu (and possibly Manding-Mokole, assuming that its reflex denotes tens in compound numerals). Isolated forms attested in South-Eastern and in peripheral Western languages are noteworthy.

The reconstruction of *pu/fu for Proto-Mande and the interpretation of *tan as the Proto-Western Mande innovation seem well-founded.

The etymology of *tan is obscure. Its similarity to the locally attested root *tan (cf. Soninke tàán 'foot, leg'; 'wheel'; 'time' (when counting), Bozo Tieyaxo tɔn 'foot, leg'; 'time' (when counting), Bozo Hainyaxo tǎ, Bozo Tiemacewe tawa, Bozo Sorogama taba) is likely a coincidence. Lexical roots with the meaning 'foot' are attested in NC numeral systems, usually as a basis for the non-compound terms for 'fifteen'. The logic behind this development is simple: 'ten' is 'two hands', 'twenty' means 'man', i.e. 'two hands and two feet', hence 'fifteen' is 'foot'. This seems to be the case for Boko and Busa, where a non-compound term for 'fifteen' (gè̃o/ gè̃ro) is attested (hence '16=15+1' in these languages). This root is etymologically related to 'foot, leg' in Duungoma (Samogo) gẽ, Dan gễ , Mano gà (it should be noted that within Mande a non-compound root for 'fifteen' is also attested in Ligbi, cf. tígán / tiga '15', tígá-ló '16').

In addition, a similarity to the term for 'one' as attested in some of the languages must be a coincidence.

A hypothesis assuming a semantic shift *NC *tan* '5' > Proto-Western-Mande *tan* '10' in parallel with the development of the Mande innovation *dúuru/ sɔ́ɔ́ru* 'five' seems to be a better explanation.

It bears reminding that the Bokobaru root *kuri* 'ten' (cf. also Boko *kúúli* recorded by Koelle) has a direct parallel in the isolated Bangime language (*kúrɛ́*).

### 3.10.11 'Twenty'

Table 3.215: Mande stems and patterns for '20'

| Manding | <'human'? | | |
|---|---|---|---|
| Jogo-Jeri | | | ɟālāmà/kèlèmó |
| Mokole | <'human'? | | |
| Vai-Kono | <'human' | 10*2 | |
| Susu | <'human' | | |
| SWM | <'human'? | 10*2 | |
| Bozo-Soninke | | 10*2 | |
| Bobo | | | kpòró, cɔ́rɔ̀ |
| Dzuun (Samogo) | <'human' | | fwɛ́ |
| Jowulu | | | kɔ̃ne/kɔnninŋ |
| SE-Eastern | | 10*2 | kèè-/ka |
| SE-Southern | <'human'[33] | 10*2 | yɔ |

There is every reason to believe that the term for 'twenty' was based on the lexical root(s) meaning 'human person' at the Proto-Mande level. The etymology of some of the isolated forms presented in the table should be sought with this in mind.

---

[33]Mende *núú ɠbɔyɔ́ngo* '20' ('person finished'). https://mpi-lingweb.shh.mpg.de/numeral/Mende.htm

### 3.10.12 'Hundred'

Table 3.216: Mande stems and patterns for '100'

| | | |
|---|---|---|
| Manding | *kɛ̀mɛ́ | |
| Jogo-Jeri | čĕmé/tʃímí | 20*5 |
| Mokole | kɛ̀mɛ | |
| Vai-Kono | kɛmɛ | |
| Susu | kɛ̀mɛ́ | |
| SWM | kɛmɛ(ŋ) | Kpelle: <'head' (ŋwúǹ) |
| Bozo-Soninke | kame/keme | 'islam'-60 |
| Bobo | | ɟɔ̄(lì)/zɔ̀(lɔ́) |
| Dzuun (Samogo) | | 20*5, 80+20 |
| Jowulu | | 'rope'*5 |
| SE-Eastern | | *20*5 |
| SE-Southern | *kɛ̀mɛ́? | kɛ̀ŋ/kǎ̀ɛ̌, la/lú |

The root *kɛmɛ*, widely attested throughout Western Africa, is noteworthy. Its original semantics deserve a separate study: it is well known that in some languages this root can be used for 'sixty' or 'eighty' and not for 'hundred' (the archaic Bamana counting system: *màninkɛ̀mɛ* '60', *bámanankɛ̀mɛ / kɛ̀mɛ* '80', *kɛ̀mɛ ní mùgan* '100' (80+20)) (Vydrin & Perekhvalskaya 2015: 360).

check keng/kai

### 3.10.13 'Thousand'

Table 3.217: Mande stems and patterns for '1000'

| Manding | wúlú/ wúli | wáa/ wá/ wà/ wága | bà |
|---|---|---|---|
| Jogo-Jeri | búlí, wúlú (< manding) | | |
| Mokole | | wàa/ wá/ waga | |
| Vai-Kono | wúl | | |
| Susu | wúlù/ wúlì | | |
| SWM | wùlù | wála/ wáá | |
| Bozo-Soninke | gulu | waxa | ('islam')-muso,wúdzǜnè |
| Bobo | | | |
| Dzuun (Samogo) | | gbàˀà, baa | bi 'goat', 800+200, <juula |
| Jowulu | | waˈaˈ | 800+200 |
| SE-Eastern | | wàà '200' | 200*5,vûû, ˈdúú, pàdí, pə, boro |
| SE-Southern | wúlù/ wlǚ/ gblǚɪ (?) | *wágá: wáá | kpi , kɛn |

The roots for 'thousand' attested in the Mande languages were borrowed from by the Western African languages. The original meaning of the Mande root *wáa/ wága* may be 'a basket of cola nuts' (Perekhvalskaja, Vydrin & Perekhvalskaya 2015: 361), cf. Bamana *wágá* 'panier à colas', Bobo *wágá* 'panier qui sert à transporter les colas ou wòlōwágá.'

\*\*\*

The table below gives an overview of Mande forms and patterns that will be used for further comparison to the evidence of other families (Table 3.210).

Table 3.218: Numerals in Proto-Mande

| 1 | do, kelen | 7 | wɔ-X-fila ('hand'+2?) |
|---|---|---|---|
| 2 | pila/fila | 8 | seki/segi (<'plus'-3?) |
| 3 | sakpa/sagba/sawa, ʔààkɔ̄/yààká? | 9 | kònonto/kɔ̀nɔndɔ(n) (10-1, 5+4) |
| 4 | náání/nɑ̃ɑ̃i | 10 | pu/fu, tan (< *'5'?) |
| 5 | dúuru/sɔ́ɔ́ru, wo? ko? **tan? (> '10'?), nṹ? | 20 | <'human' |
| 6 | wɔrɔ (wɔ-rɔ? 'hand'+1?), t(s)um? | 100 | kɛmɛ, 20*5 |
| | | 1000 | wulu, wa(g)a |

## 3.11 Mel

A narrow definition of the Mel family is preferred here (in accordance with the classification of the Atlantic languages suggested in (Pozdniakov & Segerer 2017). This family comprises two compact language groups, namely Northern (Temne, Landuma, and all Baga languages except for Baga Fore and Baga Mboteni, namely Baga Koba, Baga Maduri, Baga Sitemu and others) and Southern (Kisi, Sherbro, Mani, and Krim). Sua, Limba and Gola are not included within the Mel family and are viewed as isolated NC languages. The numeral systems of the two Mel groups comprised of the distant languages are treated separately below.

### 3.11.1 Southern Mel

Table 3.219: South Mel numerals

|  | Kisi | Sherbro | Bullom | Mani (Bullom So) | Krim |
|---|---|---|---|---|---|
| 1 | pìlɛ̀ɛ́/pilɔ, *pum? | bul | (nim)-bul | nìm-búl | yì-mo̲ |
| 2 | díŋ/C-íŋ/C-ɔ́ŋ,danyõ | tıŋ | (nin)-tsiŋ/tiŋ | nìn-cɔ́ŋ | yì-ɣın/yèèn, dím |
| 3 | ŋg-àá/y-àá | ræ | (niin)-ra | nìn-rá | yì-ɣa/gàà |
| 4 | hìɔ́ɔ́lú | hyo̲l | (nii)-hiɔɔl | nìŋ-nyɔ́l/-nyɔ́l | yì-hĭo̲n |
| 5 | ŋùèénú | mɛn | (nii)-man | nìmán    < niN-wán? | yì-wɛn/n-wén |
| 6 | 5+1 | 5+1 | 5+1 | 5+1 | 5+1 |
| 7 | 5+2 | 5+2 | 5+2 | 5+2 | 5+2 |
| 8 | 5+3 | 5+3 | 5+3 | 5+3 | 5+3 |
| 9 | 5+4 | 5+4 | 5+4 | 5+4 | 5+4 |
| 10 | tɔ́ | wāŋ | waan | wàm | wāŋ/wàn |
| 20 | bídìí(ŋ)/bélé | 'finished it is man' | u-tɔɔŋ | ù-tɔ̀ŋ | <'person' |
| 100 | < Mande | < English |  | pé, < Susu |  |
| 1000 | < Mande | < English |  | < Susu |  |

Noun class markers are usually positioned as suffixes in Kisi. At the same time, the first numerical terms in this language have noun class prefixes, which makes the forms look inconsistent, cf. *mùúŋ/ mìɔ́ɔ́ŋ / ŋìɔ́ɔ́ŋ / dìíŋ, tìɔ́ɔ́ŋ/là-tìɔ́ɔ́ŋ* 'two'.

The terms for 'hundred' and 'thousand' were probably absent in Proto-South-Mel. The similarity between Kisi *tɔ́* 'ten' and Bullom-Mani *tɔ̀ŋ* 'twenty' is noteworthy. 'Twenty' may follow the pattern '20=10PL'. If so, the original *tɔ̀ŋ* 'ten' should be viewed as an early borrowing from Western Mande (*tan* '10'). In this case, *wan* '10' is an innovation (probably based on *wan/wen* 'five') that developed in South Mel after Kisi had separated. The numeral system of modern Kisi exhibits no significant changes from the forms described by Koelle. It includes the form *ŋam-puum* '6' (Tucker Childs: *ŋ̌ŋpúm*) that may have retained an archaic

allomorph of 'one' (*pum*). The forms that will be used for further comparison are summed up in the table below (Table 3.220).

Table 3.220: Proto-South Mel numeral system (*)

| | | | |
|---|---|---:|---|
| 1 | pìlɛ̀/pilɔ (< *lɛ/lɔ?), bul, mɔ | 7 | 5+2 |
| 2 | tsiŋ/tiŋ | 8 | 5+3 |
| 3 | ra | 9 | 5+4 |
| 4 | hiɔl | 10 | 5PL? , < *West Mande? |
| 5 | wan/wen | 20 | 'person', 10PL? |
| 6 | 5+1 | 100, 1000 | absent |

### 3.11.2 Northern Mel

A higher degree of homogeneity observable in these languages allows an instant reconstruction of their numeral system at the Proto-Nothern Mel (Table 3.221)

Table 3.221: Proto-Northern Mel numeral system (*)

| | | | |
|---|---|---:|---|
| 1 | -in | 7 | 5+2 |
| 2 | -rəŋ | 8 | 5+3 |
| 3 | -sas | 9 | 5+4 |
| 4 | -ŋkɨlɛ/-nlɛ | 10 | tɔfʌt (< tɔ-f-ɔt?)/pu , wɨtʃɔ? |
| 5 | kə-ʈamaʈ (< * kə- ʈa 'hand'?) | 20 | 10*2, kə-gba (< *bay/bey 'chief'?) |
| 6 | 5+1 | 100, 1000 | absent |

### 3.11.3 Proto-Mel

The table below gives an overview of South Mel and North Mel forms (Table 3.222).

Table 3.222: Proto-Mel numeral system (*)

| | | | |
|---|---|---|---|
| 1 | -in,< *lɛ/lɔ? | 7 | 5+2 |
| 2 | díŋ/tsiŋ/tiŋ, -rəŋ | 8 | 5+3 |
| 3 | *tat (> sas, ra) | 9 | 5+4 |
| 4 | hiɔl, -ŋkɨlɛ/<-nlɛ? | 10 | *pu/fu, 5PL? |
| 5 | wan/wen, <'hand' | 20 | 'person', 10PL? |
| 6 | 5+1 | 100, 1000 | absent |

## 3.12 Atlantic

Our step-by-step reconstruction of numeral systems in the Atlantic languages will be based on their classification suggested in Pozdniakov & Segerer 2017 (forthcoming) that distinguishes two main groups within the Atlantic family, namely Northern and Bak.

### 3.12.1 Northern

The numeral systems of Northern Atlantic are treated below by sub-group.

#### 3.12.1.1 Cangin

Table 3.223: Proto-Cangin numerals (*)

| | | | |
|---|---|---|---|
| 1 | no | 7 | 5+2 |
| 2 | nak | 8 | 5+3 |
| 3 | haj/ʔéeyə | 9 | 5+4 |
| 4 | nik-iɬ < *nak-iɬ? | 10 | sabbo (< Fula), daːŋkah |
| 5 | jat (<'hand'), ʔiːp | 20 | 10*2 |
| 6 | 5+1 | 100, 1000 | < Wolof? Fula? |

Some of the reconstructions presented above are not immediately apparent and are in need of additional commentary. A detailed discussion of each of them

would be impossible here, so we will take the reconstruction suggested for 'four' (*nik-il̡*) as a sample.

At first glance, the forms of 'four' attested in the Cangin languages have nothing in common. Two of the five Cangin languages have *kinil* 'four' (Ndut-Palor), whereas in the remaining three (Laala, Noon, and Safin) *nikis* is used in this function. The easiest solution to the problem would be to postulate two alternative forms for this group. However, as the evidence of comparative-historical phonetics suggests, the final **-l** in Ndut-Palor regularly corresponds to the final **-s** in Laala-Ndut-Safin (Table 3.224).

Table 3.224: l ~ s regular correspondence in Cangin

| *-l̡ | 'eye' | 'black' | 'road' | 'four' |
|---|---|---|---|---|
| **Ndut** | ʔil | suul | wal | kinil |
| **Palor** | ʔil | suul | waal | kinil, enil |
| **Laala** | kɔs | *susus | was | nikis |
| **Noon** | kwas | *sujus | waz | nigis |
| **Safin** | xas | *suzus | was | nikis |

This fact alone urges closer examination of the forms quoted above. Further analysis shows that a fossilized noun class prefix **kV-** is present in some of the Palor numerals, cf. **ka**-*nak* 'deux', **ke**-*jek* 'trois', **ki**-*nil* 'quatre', **ki**p 'cinq. At the same time, the suffix **-Vs** is observable in the Noon numerals, cf. *jet-**us*** 'five'. This evidence combined suggests the following development of the forms for 'four' (Table 3.225):

Table 3.225: Development of *\*nik-Vl̡* '4' in Cangin

| Proto-Cangin | **\*nik-Vl̡** | | |
|---|---|---|---|
| Laala/Noon /Safin | *nik-**Vs** | | **nikis** |
| Ndut/Palor | *\***ki**-nik- Vl̡ | **ki**-nik-il | **kinil** |

### 3.12.1.2 Nyun-Buy

Numerical terms are highly divergent within this sub-group, so it seems reasonable to treat them by branch (Table 3.226).

Table 3.226: Nyun-Buy numerals

|  | **Nyun** | **Buy (Kobiana, Kasanga)** |
|---|---|---|
| 1 | duk | tee(na), -anɔʔ |
| 2 | nak | naŋ |
| 3 | lal | taar |
| 4 | ren(d)-ek | sannaŋ |
| 5 | ci-lax (<'hand'), -məkila | ju-roog (<'hand'?) |
| 6 | 5+1 | 5+1 |
| 7 | 5+2 | 5+2 |
| 8 | 5+3 | 4+4 |
| 9 | 5+4 | 5+4 |
| 10 | ha-lax (<'hands') | 5PL, ntaajã |
| 20 | <'king' | < Mande, 10*2 |
| 100 | < Mande | < Mande, < French |
| 1000 | < Mande | ŋ-kontu < Portuguese[34] |

The pattern '5'='hand' ~ '10'='hands' is immediately apparent in Nyun. In the case of Buy, it can be accepted only under the assumption that the derived term for 'five' became phonetically distant from its source form, cf. Kasanga *ji-rek*, Kobiana *ji-hak* 'hand' (these forms must be related to Nyun *ci-lax* 'hand'). In any case, the Kasanga term *ŋaa-roog* follows the pattern '5PL' that uses the same plural noun class as the one attested in *ŋa-rek* 'hands'.

The forms for 'ten' attested in Joola Ejamat (Atlantic Bak) *si-ntaaja* is important for the diachronic interpretation of the Kobiana form *ntaajã*. The evidence suggests that the latter was probably directly borrowed from Joola[35]  (as was -*anɔʔ* 'one').

### 3.12.1.3  Jaad-Biafada

The forms of 'one' (*ɲi/ nɛ*) are distinguishable in the compound numerals, cf. Jaad *ŋka-inɛ* '6' ('5+1'), Biafada *mpaaji nyi* '7' ('6+1'), etc. The term for 'five' goes back to the lexical root meaning 'hand' (Biafada *gə-bəda*, Jaad *ko-bəda*).

---

[34]Guillaume Segerer (p.c.).

[35]According to Guillaume Segerer (p.c.) it is possible that the Ejamat and Kobiana forms both come from Manjak.

Table 3.227: Jaad-Biafada numerals

| | | | |
|---|---|---|---|
| 1 | nnəmma, *ɲi/nɛ/-inɛ, -kkã | 7 | 5+2, 6+1 (< Manjak) |
| 2 | ke, ma-ae | 8 | 5+3, wose/wase |
| 3 | jo/tʃaw | 9 | 5+4, leberebo |
| 4 | n(n)e/nnihi | 10 | (p)po |
| 5 | bəda (<'hand') | 20 | 10*2 |
| 6 | 5+1, paaji (< Manjak), ŋka-? | 100, 1000 | < Fula |

#### 3.12.1.4 Tenda

The reconstruction of the Proto-Tenda numerals (K.Pozdniakov, ms) is based on a comparative analysis of five Tenda languages: Basari, Tanda, Bedik, Bapen, Konyagi.

Table 3.228: Tenda numerals (*)

| | | | |
|---|---|---|---|
| 1 | ɓɑt, ndi/riye/diye/iye, mbɔ | 7 | 5+2 |
| 2 | ki | 8 | 5+3 |
| 3 | taʈ | 9 | 5+4 |
| 4 | nə̀x | 10 | poxw |
| 5 | mbəɗ (<'hand'), cɔ/njɔ | 20 | 10*2 |
| 6 | 5+1 | 100, 1000 | < Fula, < Mande |

The etymology of the Konyagi term for 'five (*mbəɗ*) is based on the Jaad-Biafada evidence (these languages belong to the same sub-group as Tenda).

#### 3.12.1.5 Fula-Sereer

The numerical terms are highly divergent within this sub-group, so it seems reasonable to treat them by language (Table 3.229).

The fact that the Seerer terms covering the sequence from 'two' to 'five' have the same final segment is noteworthy. This could potentially be interpreted as a special morpheme or as a sub-morpheme that resulted from alignment by analogy. This discussion will be resumed below. Here it can only be stated that the

---

[36]Reviewing my first version of the book, Guillaume Segerer has advanced a new interesting etymology for Fula: *jow-i* '5' = *jun-ngo* <*jow-ngo* 'hand'. His hypothesis is quite possible.

Table 3.229: Fula-Sereer numerals

|  | Fula | Sereer |
|---|---|---|
| 1 | goʔo | leŋ |
| 2 | ɗiɗi | ɗik |
| 3 | tati | tadik |
| 4 | na(y)i | nahik |
| 5 | jo(w)i[36] | ɓe-tVk |
| 6 | 5+1 | 5+1 |
| 7 | 5+2 | 5+2 |
| 8 | 5+3 | 4+4 |
| 9 | 5+4 | 5+4 |
| 10 | sapp-o | xarɓ- |
| 20 | noogas/noogay | 10*2 |
| 100 | teeme- | < Fula |
| 1000 | < Mande, < Hausa | < Wolof? |

morphological analysis of the Sereer term for 'five' (*ɓe-tVk*) suggested in the table below is not immediately apparent and is thus debatable. Within this approach the element **ɓe-** is interpreted as a noun class prefix despite the fact that such a class is lacking in Sereer. Complex issues pertaining to the reconstruction of the term for 'five' will not be treated here. We shall only note that the plural animate class is reconstructable as **ɓe-** (class 2) in Proto-Fula-Sereer.

### 3.12.1.6 Wolof

Table 3.230: Wolof numerals

| 1 | CL-enn | 7 | 5+2 |
|---|---|---|---|
| 2 | ñaar (< *CL-(X)aar) | 8 | 5+3 |
| 3 | ñ-ett (< *CL-(X)ett) | 9 | 5+4 |
| 4 | ñ-ent (< *CL-(X)en(i)t) | 10 | fukk |
| 5 | jurom | 20 | < 'person', 10*2 |
| 6 | 5+1 | 100, 1000 | < Fula, < Mande |

The Wolof term for 'one' exhibits the agreement in noun class, cf. *k-enn nit*

'one person', *g-enn garab* 'one tree', *f-enn* 'somewhere', *l-enn* 'something', etc.
The same can be applied to the terms covering the sequence from 'two' to 'four'
as demonstrated in Pozdniakov 2015: 82. Nothing is known about the original
radical of the root (assuming there was one) since it was replaced by a noun
class consonant.

Speaking of 'twenty', it should be said that the form *nit(t)* (apparently related to
the lexical root *nit* 'person') is widely used alongside the common Wolof pattern
'10*2'.

### 3.12.1.7 Nalu-Baga Fore-Baga Mboteni

This sub-group is the most problematic within Northern Atlantic. Admittedly,
the evidence pertaining to their classification as Northern is inconclusive. More-
over, the sub-group itself is highly heterogeneous, which affects its numeral sys-
tems as well. The pertinent data for each of these languages is provided below
(Table 3.231):

Table 3.231: Numerals in Nalu, Baga Fore and Baga Mboteni

|  | **Nalu** | **Baga Fore** | **Baga Mboteni** |
|---|---|---|---|
| 1 | deːndɪk | ki-ben | mbɔ́ |
| 2 | bi-lɛ | ci-di | sà-lɛ́ |
| 3 | p-aat | ci-tɛt | n-dɛ́r |
| 4 | bii-naaŋ | ci-nɛŋ | í-nà |
| 5 | teedoŋ (< tɛ́ 'hand'?) | su-sɑ(n) | ì-rì$\beta$ɛ̆, *ba(x)? |
| 6 | 5+1 | 5+1 | 5+1 |
| 7 | 5+2 | 5+2 | 5+2 |
| 8 | 5+3 | 5+3 | 5+3 |
| 9 | 5+4 | 5+4 | 5+4 |
| 10 | 5*2, *a-lafaŋ? | ɛ-tɛ-lɛ (<'hands'+2) | tèn (< '*hand'? ) |
| 20 | 10*2 | 10*2 | 10*2 |
| 100 | m-laak | bɔ-1 | < Mande |
| 1000 | m-ɲaak (100pl?) < Susu | tɛngbeŋ-1 | ? |

### 3.12.1.8 Proto-Atlantic North

The prospects for the reconstruction of the Proto-North Atlantic numerals are discussed below.

3.12.1.8.1 'One' (Table 3.232)

Table 3.232: Numerals for '1' in Northern Atlantic

| Cangin | | no | | |
|---|---|---|---|---|
| Nyun | | | | duk |
| Buy | | nɔʔ | | tee(na) |
| Jaad-Biafada | *ɲi/nɛ | | | nnəmma,pakkã |
| Tenda | di(ye) | | mbɔ | bat |
| Fula-Sereer | leŋ | | | goʔo |
| Wolof | -enn | | | |
| Nalu | deendik | | mbɔ́ | ki-ben |

Isolated forms are quoted in the rightmost column. Direct parallels to some other forms are attested in Cangin – Buy (*nɔʔ*) and Konyagi – Baga Mboteni (*mbɔ*). The most common root is *di(n)/ li(n)/ ye(n)/ ne(n)* (assuming that these forms are related).

3.12.1.8.2 'Two', 'Three' and 'Four' (Table 3.233)

The forms of 'two' in Tenda-Jaad-Biafada can be explained as a shared innovation, since these two branches belong to the same sub-group. The forms quoted in the two leftmost columns could be related, but the pertinent evidence is inconclusive. The roots *nak* and *di(k)* are reserved for further comparison.

As in the majority of other NC branches, the terms for 'three' and 'four' (tentatively recorded as *taʈ* '3' and *nak* '4') are fairly consistent in North Atlantic. Thus it appears that the terms for 'two' and 'four' are the same (or phonetically similar) across the languages of this branch. Cangin is the only language that does not comply with the additional distribution, because in the case of Cangin both terms are reconstructed as *nak*. Interestingly, the form of 'four' bears a suffix, hence it could potentially be explained as a derivative of 'two'. At the same time, the root *nak* 'four' is reminiscent of one of the most persistent NC roots with this meaning.

Table 3.233: Numerals for '2'-'4' in Northern Atlantic

|  | '2' | '2' | '2' | '2' | '3' | '3' | '4' | '4' |
|---|---|---|---|---|---|---|---|---|
| Cangin | nak |  |  |  |  | haj | nik-iɭ < nak-iɭ? |  |
| Nyun | nak |  |  |  | lal |  |  | ren(d)-ek |
| Buy | naŋ |  |  |  | taar |  |  | sannaŋ |
| Jaad-**Biafada** |  |  | ke |  |  | jo/caw | n(n)e(hi) |  |
| Tenda |  |  | ki |  | taʈ |  | nǽx |  |
| Fula-**Sereer** |  | ɗik |  |  | tati(k) |  | na(y)i(k) |  |
| Wolof |  |  |  | X-aar | X-ett |  | X-en(i)t |  |
| Nalu |  | di/lɛ |  |  | tɛt/tat |  | naaŋ/ nɛŋ/ na |  |

In Jaad-Biafada we find the root *\*jow/caw* '3'. This is undoubtedly an innovation in the group which is represented by a remarkable isogloss. This is therefore an argument in favour of interpreting this group as part of the northern branch of the Atlantic family: Biafada *-njo / bíí-co/ bií-yo* '3', Jaad *ma-cao/ ma-caw/ má-cɔu* '3'. It is possible that we are dealing with an ancient borrowing of Proto-Jaad-Biafada from Mande (from *saba* 'three').

In theory, it is possible that forms attested in the Cangin languages (*ka-hay / \* ʔe-jɛʔ*), also originated from the Mande form (likely weakened to *\*habi / hawi*).

In this case, we find either reflexes of the Proto-NC form *\*tath* or borrowings (taking into account very ancient forms) – from the Mande languages in numerous Northern Atlantic languages.

3.12.1.8.3 'Four'

The root *\*na(h)i-k* can be securely reconstructed for Proto-Northern Atlantic. As has been demonstrated above, the initial **ñ-** of the Wolof term is a reflex of a noun class prefix that replaced the initial radical of the root. The final -t in the Wolof term probably resulted from the alignment by analogy with the term for 'three' that ends in -t, cf. *\*ñ-eenk* ? → *ñ-eent* '4' by analogy with *ñ-ett* '3'.

3.12.1.8.4 'Five' (Table 3.234) and the terms from 'six' to 'nine'

Table 3.234: Numerals for '5' in Northern Atlantic

| Cangin | jat (<'hand') | | | ʔiːp | |
|---|---|---|---|---|---|
| Nyun | ci-lax (<'hand') | | | | -məkila |
| Buy | | | ju-roog | | |
| Jaad-Biafada | bəda ('hand') | | | | |
| Tenda | mbəɗ (<'hand'?) | co/njo | | | |
| Fula-Sereer | | jo(w)i | * ɓe-tVk | | |
| Wolof | | jurom | | | |
| Nalu | teedoŋ/*tee ('hand'?) | | | ribə(l) | su-sa(n), *ba(x)? |

The North Atlantic languages are characterized by the term for 'five' being systematically derived from the lexical root meaning 'hand'. Interestingly, this development seems to post-date the replacement of the original root for 'hand' by an innovation in the majority of the branches. At least four independent formations of this kind are attested within eight branches (cf. the evidence quoted in the leftmost column of the table). Both Tenda and Jaad-Biafada terms for 'five' are of common ancestry: they seem to have developed from the root *ɓəda at the Proto-Jaad-Biafada level, since both languages belong to the same sub-group. This probably indicates that the pattern based on the term for 'hand' was used in the languages that belong to the Northern group at the proto-level (possibly as an alternative to the inherent NC root for 'five'). In view of this, the formal alterations of 'five' are easily explained as those automatically caused by the replacement of the inherent term for 'hand' by an innovation. As we hope to demonstrate in the next chapter, the derivational pattern 'hand' > 'five' is surprisingly rare in the NC languages. It is barely attested, for example, in Benue-Congo, thus being characteristic of the North Atlantic languages (and the Atlantic languages on the whole, see below).

In view of this, the reflexes of the inherent NC root for 'five' could have been preserved in only a minority of North Atlantic branches. The roots *jo/ co, *tVk/ rog and *rib/ ʔiːp unrelated to the term for 'hand' deserve special attention within this context.

The pattern '5+' ('hand'+) can be securely reconstructed for the terms covering the sequence from 'six' to 'nine'. The uncommon pattern '7=6+1' attested in

Biafada was borrowed from one of the Manjak languages (Atlantic Bak), as was the derived term for 'six' (*mpaaji*).

3.12.1.8.5 'Ten' and 'Twenty' (Table 3.235)

Table 3.235: Numerals and patterns for '10' and '20' in Northern Atlantic

| | '10' | '10' | '10' | '20' | '20' | '20' |
|---|---|---|---|---|---|---|
| Cangin | | | < Fula, da:ŋkah | 10*2 | | |
| Nyun | | <'hands' | | | <'king' | |
| Buy | | 5PL | ntaajã | 10*2 | | < Mande |
| Jaad-Biafada | (p)po | | | 10*2 | | |
| Tenda | pəxw | | | 10*2 | | lapɛm |
| Fula-Sereer | | | sapp-o, xarɓ- | 10*2 | | noogas/ noogay |
| Wolof | fukk | | | 10*2 | 'person' | |
| Nalu | | 5*2 | *a-lafaŋ? | 10*2? | | |

With the evidence of the three branches, the reconstruction of the term for 'ten' (tentatively recorded as *\*pok*) seems secure. Its attestations are admittedly limited, apparently due to its replacement with derived terms based on 'five' ('hand'). This reconstruction is also supported by the presence of the final velar: as we have seen, it is reconstructable in a number of other numerical terms at the proto-level.

The pattern for 'twenty' is reconstructable as '20=10*2'. Particular derivates based on the typologically widely attested patterns ('20' <'person', 20 <'king') seem to have formed independently.

3.12.1.8.6 'Hundred' and 'thousand'

The evidence points to the absence of these terms in Proto-North Atlantic. Attested forms are borrowings from 'influential' languages such as Fula, Wolof, Manding, Hausa (in the case of Niger Fulfulde). Interestingly, the terms in question are already borrowings in some of these source-languages.

Proto-North Atlantic numeral system (Table 3.236)

Table 3.236: Proto-North Atlantic numeral system (*)

| | | | |
|---|---|---|---|
| 1 | di(n)/li(n)/ye(n)/ne(n), mbɔ | 7 | 5+2 |
| 2 | di(k), nak | 8 | 5+3 |
| 3 | taʈ | 9 | 5+4 |
| 4 | nak | 10 | pok |
| 5 | <'hand', jo, tVk/rog, rib/ʔiːp | 20 | 10*2 |
| 6 | 5+1 | 100, 1000 | absent |

### 3.12.2 Bak

#### 3.12.2.1 Joola languages

Over a hundred sources covering the numeral systems of fifteen major Joola dialects have been made available to us courtesy of Guillaume Segerer. His collection of evidence may be labeled a 'dialect atlas' of numerical terms. These terms often exhibit significant variations not only in their phonetics but in the inventory of lexical roots as well[37] . The name Joola pertains to a group of at least seven related languages (including Bayot). A study of their numeral systems may help set a clearer distinction between these languages. Moreover, it might shed some light on their (hitherto unclear) internal classification.

Numerical terms as attested in ten major Joola languages are discussed below.

3.12.2.1.1 'One' (Table 3.237)

Table 3.237: Joola numerals for '1'

| Bliss Banjal | Kasa Mlomp | Fogny Karon | Keeraak Ejamat | Bayot Kwaatay |
|---|---|---|---|---|
| -anɔʔ | -anor | -anor | -anor | |
| -anor | -anor | -anor | -anor | |
| | (akon) | əkon | | (akon) |
| | (ta) | | | don |
| | | | yinka, (sia) | fɛnɛŋ |

---

[37]I wish to express my gratitude to G.Segerer for his assistance with regard to the dialectal attribution of sources.

The main form is reconstructed as *-anor*, with the initial vowel forming a part of the root. The only languages where this root is not present are Bayot (*don* '1') and Kwaatay (*fɛnɛŋ* '1'). The root *əkon* with a vocalic opening (sporadically attested in Kasa and Bayot) is found in Fogny alongside *-anor*.

3.12.2.1.2 'Two', 'three' and 'four' (Table 3.238)

Table 3.238: Joola numerals for '2'-'4'

| Bliss Banjal | Kasa Mlomp | Fogny Karon | Keeraak Ejamat | Bayot Kwaatay |
|---|---|---|---|---|
| '2' | | | | |
| si-lubəʔ | si-ɬubɣʔ | (liba) | sɪˈsubə | ʔi-ɾigəʔ/ tɪgga |
| si-rubə | sɪ-subəl | su-supək/ ɕi-ɕipəkʰ | si-luːbɜʔ | sɪˈsubə |
| '2' | | | | |
| | si-gäbä, (ku-mɛntɛn) | si-gäbäʔ | | |
| si-gabaʔ | | | si-gäbä | |
| '3' | | | | |
| si-həəji | si-hɣːɟiʔ | si-feegiir/ si-feːɟiʔ | sɪ-həːjɪ | i-fiigiʔ/ i-fəəʒi |
| gu-fɪːgɪr/ si-fɣɟiɾ | sɪ-həːjɪl | si-həːciːl | si-həəji, (fu-fooateen) | ki-hɣːɟiʔ |
| '4' | | | | |
| si-bäkir | si-bäːkiɾ/ si-bäkiʔ | si-bäkiɾ/ si-baːciːr | si-bacir | sɪ-bɐɣɪr |
| si-baagir | sɪ-bacɪl | ɕɪ-päkil/ si-baːciːl | si-bäkir | ki-bäkir |

Two alternative roots for 'two' are attested in Joola, namely *si-ɬubəʔ* and a relatively wide-spread *si-gabaʔ*.

The term for 'three' goes back to *\*si-feegir*, with its reflexes being attested in all dialects.

The term for 'four' is securely reconstructed as *\*si-bääkiɽ*.

### 3.12.2.1.3 'Five' and 'ten' (Table 3.239)

Table 3.239: Joola numerals for '5' and '10'

| Bliss Banjal | Kasa Mlomp | Fogny Karon | Keeraak Ejamat | Bayot Kwaatay |
|---|---|---|---|---|
| **'5'** | | | | |
| hu-tɔk | hu-tɔkʰ | fu-tɔk/ u-sɔk | hu-tɔk | o-to/ ɔ-ɫɔ/ ɔ-rɔ |
| fu-tɔk | | ɪ-ɕäkʰ/ i-sak | fu-tɔk/ hu-ʂok | hu-tɔk |
| **'5'** | | | | |
| | (naa-suan) ŋaa-suwaŋ | | | |
| **'5'** | | | | |
| | | *fu-tam | | |
| *tən | | | | |
| **'10'** | | | | |
| ku-ŋɛn <'hands' gu-ɲɛn <'hands' | ku-ŋɛn <'hands' | ku-ŋɛn <'hands' | ku-ŋɛn <'hands' ku-ŋɛn <'hands' | |
| **'10'** | | | | |
| | | | | gu-tie(pɔkɔ) 'hands' |
| | sɛ-bɛɛs 'hands' | ŋaa-suwan | | su-moŋu/ su-ŋɔmu 'hands' |

The Banjal form *\*tən* (reconstructed on the basis of the compound numerical terms) and the (related?) Fogny form *fu-tam* attested in a source dating to the seventeenth century (d'Avezac 1845) are of special interest.

The Mlomp form of 'five' (sporadically attested in Kasa as well) is identical to the Karon form for 'ten' (*ŋaa-suwan* in both cases). The etymology of these forms is unclear. At the same time, the majority of the forms for 'ten' (but not for 'five' as in the majority of the North Atlantic languages) go back to the lexical root meaning 'hands'. To illustrate this point, the lexical stems for 'hand' in the Joola languages are quoted in the table (Table 3.240):

As can be deduced from the presentation above, at least four lexical roots for 'hand' that serve as a basis for the terms for 'ten' are distinguishable in Joola. Interestingly, the source roots and the numerical terms that depend on them are not necessarily the same within a language. The main root is *\*ku-ŋɛn/ ku-ɲɛn* '10' <'hands'. At the same time, *bɛɛs* 'hand' yields *sɛ-bɛɛs* 'ten' in Mlomp. This derivative is not attested in in Kasa and Karon where *bees* 'hand' alternates with *ŋɛn/ ɲɛn* 'hand'. The base *\*ka-ʈe* 'hand' attested in Bayot and Kasa yields *gu-tie-* in Bayot. Finally, *ɛ-mɔŋu* 'hand' > *su-moŋu* 'ten' in Kwaatay (also *ɛ-ŋɔmu* 'hand' > *su-ŋɔmu* 'ten' with a metathesis).

As noted above, the root *ɛ-ntaaja* attested in Keeraak and Ejamat was possibly incorporated into Kobiana (North Atlantic). This root, admittedly very rare in the Joola cluster, is the only primary one for 'ten' and as such it deserves special attention (especially in view of its later replacement with the derivatives based on 'hand').

3.12.2.1.4 'Twenty', 'hundred', and 'thousand'

Two apparent derivational patterns are used for the term for 'twenty' in the Joola languages:

<'king':   Bliss *a-yuɣ*, Banjal *ə-vi/ə-vvi*, Kasa *a-yi/ ɔ-ji*, Karon *əwi*, Bayot *ə-y*;
<'person':   Kasa *an / bu-k-an*, Fogny *ka-banan* 'person finished'.

In Kwaatay the term for 'twenty' is based on 'mouth' (*bu-tum-an*).

The terms for 'hundred' and 'thousand' are borrowings from Mande or 'influential' Atlantic languages (often either Fula or Wolof) in the majority of the dialects, cf. *keme/teme* '100', *wuli, juni* '1000'.

In conclusion it should be added that the Joola terms covering the sequence from 'six' to 'nine' follow the common pattern '5+'.

Table 3.240: Joola stems for 'hand'

| Bliss Banjal | Kasa Mlomp | Fogny Karon | Keeraak Ejamat | Bayot Kwaatay |
|---|---|---|---|---|
| **'hand'** | | | | |
| ka-ŋɛn(ak) | ka-ŋɛn | ka-ɲen(ak)/ ka-ŋɛn | ka-ŋɛn | |
| ga-ɲɛn/ ka-ɲɛn(ak) | | ka-ɲɛn | ka-ŋɛn(ak) | ka-ŋyɛn(ak) |
| ' **'hand'** | | | | |
| | e-bɛɛs | | | |
| ɛ-pɛs | ɛ-bɛɛs | ɛ-pɛs/ ɛ-bɛs | | |
| **'hand'** | | | | |
| | | | | ɛ-mɔŋu/ ɛ-ŋɔmu |
| **'hand'** | | | | |
| | ka-seʔ | | | ka-te/ ga-te/ ʈe/ kə-se |
| **'hand'** | | | | |
| bu-lɛhɛj | | ɛ-lɛcɛs | | |
| 'hand' | | 'upper arm' | | |
| bi-lɛfɛj | | | bu-lɛfec | |
| | | | 'inner hand' | |
| **'hand'** | | | | |
| ka- ʂɛɲum(əku) | | | kə-ləɲum 'hand' | |

### 3.12.2.2 Manjak languages

This branch is represented by three closely related languages (Manjak, Mankanya, Pepel). Numerical terms attested in them are presented in the table below (Table 3.241):

Table 3.241: Manjak numerals

| 1 | lɔɔl(e)/lɔŋ | 7 | 6+1, jand/jaanʔ/ cand (Pepel) |
|---|---|---|---|
| 2 | -təb/-təw, -puguṭ/pugus (Pepel) | 8 | 4PL, koas/ʊʌs |
| 3 | wa-(y)anṭ/wa-jenṭ/ jens | 9 | 10-1, (8+1) |
| 4 | baakər/wakər | 10 | 5PL ('hands'), (n)taaja/taaya, taim (Pepel) |
| 5 | ɲɛɛn ('hand') | 20 | 10*2 |
| 6 | paagi/paaji | 100 | < French |
| | | 1000 | kʊnt |

As can be gleaned from the table, the Manjak stems for numerals are very different from those attested in Joola. At the same time, morphological and lexical evidence strongly suggests that these two branches are genetically the closest and belong to the same Bak sub-group.

This implies that the numeral system of one of these branches must have undergone systematic innovations. We will reserve our conclusions until the evidence from the other Bak sub-groups, i.e. Balant and Bijogo, is reviewed.

### 3.12.2.3 Balant

Despite the fact that Balant is usually treated as one language, we will present the evidence of Balant Ganja and Balant Kentohe separately (Table 3.242), because the difference between these two idioms is of key importance to our study.

The opening sequence of the Ganja terms is quoted according to Creissels & Biaye 2015. They form the most reliable part of the presentation. A few remarks pertaining to the differences in these Balant dialects are in order. First of all, the Balant Kentohe terms for 'one', 'two', 'three' and 'six' exhibit a final homorganic nasal of uncertain origin. The forms attested by Koelle in the 19th century sources suggest that we are dealing with a morpheme -**n** not assimilated to a preceeding consonant by point of articulation. Secondly, Koelle's evidence speaks in favor of 'six' being a base for a larger group of numerical terms. According to him, not only 'eight' and 'nine' but also 'ten' followed the pattern '6+'.

Table 3.242: Balant numerals

|  | Balant Ganja | Balant Kentohe |
|---|---|---|
| 1 | hɔ́dà/wɔ́dā/-ɔdaʔ, bɔ́ɔ́díbɔ́/wɔdibɔ (counting) | -ɔɔdn/ho:dn/fóóda |
| 2 | sìbí/-sebe | -sɪbm/-sebm/g-ʃííbn (Koelle 1963[1854]) |
| 3 | hàbí/yààbiĩ̀ | -habm/káábn (Koelle 1963[1854]) |
| 4 | tàllá/tàhàlā | -tasla/tahla/táʃiila (Koelle 1963[1854]) |
| 5 | jɪ̀ɪ́f/jéèf | cɪf/'-cef/kiif ~ ciif (Koelle 1963[1854]) |
| 6 | fááj/faac | mfaacɲ/faad (Koelle 1963[1854]), 5+1 |
| 7 | 6+1 | 6+1, 5+2 |
| 8 | táhtállà/tāntàhlā (4 redupl.), 6+2 | 5+3, 6+2 (Koelle 1963[1854]) |
| 9 | jíntàllá/jīntàhlā (5+4) | 5+4, 6+3 (Koelle 1963[1854]) |
| 10 | jímmín/jīnmīnn (<5?) | cɪfmɪɪn/f-cef meen (<5?), 6+4 (Koelle 1963[1854]) |
| 20 | 10*2 | <'person' |
| 100 | gɛ̀mɛ́/kɛmɛ (borrowed) | <'5 persons' |
| 1000 | wílí (borrowed), kont | f-ko:nti |

### 3.12.2.4 Bijogo

Let us examine an analysis of the Bijogo numeral system found in (Segerer 2002). According to him, the term for 'one' is *nɔɔd* ('cette forme est retenue pour l'énumération abstraite', ibid. 171). His interpretation of *-d* as the only true reflex of the etymon (with other segments ensuring the grammatical agreement) is immediately convincing, cf. the following examples quoted by him (ibid. 171):

(2)　a.　*o-to ɔ-nɔɔd* 'a person'

　　　b.　*e-booʈi ɛ-nɛɛd* 'a dog'

　　　c.　*u-gbe u-nɛɛd* 'a road'

　　　d.　*ka-jɔkɔ n-ka-d* 'a house'

　　　e.　*ŋɔ-katɔ ŋ-ŋɔ-d* 'a fish'.

Table 3.243: Bijogo numerals

|  | Bijogo Kagbaga (Bubaque) | Bijogo (other dialects) |
|---|---|---|
| 1 | n-ɔɔd (*-d) | |
| 2 | n-somb (Segerer, p.c.), n-sombɛnʈ | sòòbɛ́/súngb/cuuwɛ, ndank (Kamona) |
| 3 | ɲ-ɲɔ-ɔkɔ (<'fingers') | |
| 4 | ya-agɛnɛk | |
| 5 | n-de-ɔkɔ (dɛ 'to finish', -ɔkɔ 'hand') | nu-duβ-ɔkɔ (Kamona) |
| 6 | 5+1 | |
| 7 | 5+2 | |
| 8 | 5+3 | |
| 9 | 5+4 | |
| 10 | n-ruakɔ (ru 'to rise', -ɔkɔ 'hand') | |
| 20 | o-joko ('person'), -ansak-o-to ('to finish'+'somebody') | ŋɔjɛt oto (Kamona), Koelle 1963[1854]: ríaakɔ́ɔ́to/ŋórembaʃóóto |
| 100 | 20*5 | |
| 1000 | kuntu | |

Segerer justly observes that 'La forme générale de l'élément ayant pour valeur 'un (autre)' est donc **(V)-n-pC-d**, où **pC** est le préfixe de classe du nom déterminé' (ibid. 171).

He also quotes the form *dideeki* 'seul' (var. *deeki* 'tout seul'). A variant of this form probably appears as *èɖìgɛ́/ néédigɛ/ módiigɛ* 'one' in Wilson and Koelle.

As demonstrated by Segerer, the term for 'three' (*ɲ-ɲɔɔkɔ*) is a Bijogo innovation of a cultural origin, cf. SG *ɲɔ-ɔkɔ* - PL of *nɔ-ɔkɔ* 'finger' (dim. *<kɔ-ɔkɔ* 'hand'): 'Un roi bijogo ne se déplace jamais sans l'attribut symbolique de sa fonction, consitué par une sculpture de bois et de corne ... Cet objet, nommé u-ran kɔ-ɔkɔ, represente une main à trois doigts' (ibid. 172). It should be noted that this root is attested in all Bijogo dialects and is already accounted for by Koelle (*-ɲɔ́ɔ́gɔ*).

As established by Segerer, the same root is attested as *ɔkɔ* in the terms for 'five' and 'ten'.

### 3.12.2.5 Proto-Bak

Now we will compare the Bak numerals.

#### 3.12.2.5.1 'One' (Table 3.244)

Table 3.244: Bak numerals for '1'

| Joola | **don** | -anor, əkon, fɛnɛŋ, yinka, (sia), (ta) |
|---|---|---|
| Manjak | lɔɔl(e)/**lɔŋ** | |
| Balant | | -ɔdaʔ |
| Bijogo | **\*d** | -ed̪ìgɛ |

A comparison of the terms quoted in the leftmost column yields the form that can be tentatively recorded as *\*don*. The rightmost column gives an overview of roots attested in only one out of four branches.

#### 3.12.2.5.2 'Two' (Table 3.245)

Table 3.245: Bak numerals for '2'

| Joola | si-ɬubəʔ | si-gabaʔ |
|---|---|---|
| Manjak | | -təb/-təw, puguʈ/pugus |
| Balant | sɪbɪ/-sebe | |
| Bijogo | sòòbɛ́/súngb/cuuwɛ | |

The leftmost column presents the root attested in three sub-groups. It is traceable to *\*ɬubəʔ*.

#### 3.12.2.5.3 'Three' and 'four' (Table 3.246)

For the first time in our step-by-step analysis of numeral systems in the numerous NC families we observe the existence of a separate root for 'three' in each of the branches of a language group.

The term for 'four' exhibits an isolated Joola-Manjak innovation as well as isolated innovations in Balant and Bijogo.

Table 3.246: Bak numerals for '3' and '4'

|          | '3'                            | '4'         | '4'            |
|----------|-------------------------------|-------------|----------------|
| Joola    | si-feegir                     | si-bääkiɽ   |                |
| Manjak   | wa-(y)anɟ/wa-jenɟ/jens        | baakər/wakər |               |
| Balant   | habi/yabi                     |             | tasala/tahala  |
| Bijogo   | ɲ-ɲɔ-ɔkɔ (<'fingers')         |             | ya-agɛnɛk      |

3.12.2.5.4 'Five' (Table 3.247)

Table 3.247: Bak numerals for '5'

|          |                                           |                                          |
|----------|-------------------------------------------|------------------------------------------|
| Joola    |                                           | fu-tɔk, tən?, ŋaa-suwaŋ? (cf. '10')       |
| Manjak   | ɲɛɛn ('hand') (cf. Joola '10')            |                                          |
| Balant   |                                           | jìíf/jéèf                                 |
| Bijogo   | n-de-ɔkɔ (dɛ 'to finish', -ɔkɔ 'hand')    |                                          |

The pattern 'hand' > '5' is traceable within two branches. However, the roots involved are different in each case. Numerous isolated forms are grouped together in the rightmost column.

The terms from 'six' to 'nine' (Table 3.248)

Table 3.248: Bak numerals and patterns for '6'-'9'

|          | '6'   | '6'          | '7'                    | '8'           | '9'           |
|----------|-------|--------------|------------------------|---------------|---------------|
| Joola    | 5+1   |              | 5+2                    | 5+3           | 5+4           |
| Manjak   |       | paagi/ paaji | 6+1, jand/ jaan?/ cand | 4PL, koas/ ʊʌs | 10-1, (8+1)   |
| Balant   |       | fááj/ faac   | 6+1                    | 4 redupl., 6+2 | 6+3, 5+4     |
| Bijogo   | 5+1   |              | 5+2                    | 5+3           | 5+4           |

The form *paag/paaj* 'six' is a common Manjak-Balant isogloss[38] . It is not

---

[38]Guillaume Segerer is right to note (p.c.) that the Manjak-Balant form *paag-* '6' may be ralated to Joola *-feegir/-hɔɔji* '3'

surprising that the primary term for 'six' attested in these languages served as the basis for the '7=6+1' pattern. This pattern received further development in Balant where it was employed for terms up to 'ten' (i.e. '10=6+4') according to the 19th century sources. At the same time, the archaic pattern '8=4PL'/'8=4 redupl.' is attested in these languages alongside the pattern '8=6+2'.

3.12.2.5.5 'Ten' (Table 3.249)

Table 3.249: Bak numerals for '10'

| Joola | ɛ-ntaaja[39] | ku-ŋɛn/ ɲɛn 'hands' | 'hands' (bɛɛs, moŋu/ ŋɔmu, tie) | ŋaa-suwan |
|---|---|---|---|---|
| Manjak Balant | (n)taaja/ taaya | | 5PL ('hands') | taim jímmín, 6+4 |
| Bijogo | | | n-ruakɔ (ru 'to rise', -ɔkɔ 'hand') | |

In addition to the common pattern '10 = 'hands'', both branches share a common root (*ntaaja*) that could be interpreted as a shared Proto-Joola-Manjak innovation.

3.12.2.5.6 'Twenty', 'hundred' and 'thousand'

The term for 'twenty' is based on the lexical root meaning 'person' in all of the branches (except for Manjak, where it was replaced with the pattern '20=10*2'). The same development is observable in Balant Ganja as well.

The terms for 'hundred' and 'thousand' are most likely borrowings. However, the origin of *kont*/*kunt* 'thousand' attested in three of the Bak branches deserves special discussion (in North Atlantic this root (*ŋ-kontu*) is found in both of the Buy languages).

Overview of the Bak numerical terms (Table 3.250)

---

[39]The stem is attested only in Joola Feloup, so, it seems to be borrowed from Manjak.

Table 3.250: Bak numerals

| | | | |
|---|---|---|---|
| 1 | don/lɔŋ, -anor, əkon | 7 | 6+1, 5+2, jand/jaanʔ/ cand (Pepel) |
| 2 | ɫubəʔ, -təb/-təw, -puguṭ/pugus | 8 | 4PL/4 redupl., ʊʌs |
| 3 | feegir, yanṭ/jenṭ, habi/yabi | 9 | 5+4, 10-1, 6+3 |
| 4 | baakər/wakər, tasala/tahala | 10 | 5PL ('hands'), (n)taaj, taim, -suwan |
| 5 | 'hand', tɔk, tən? | 20 | 'person', 10*2 |
| 6 | paag/paaj, 5+1 | 100 | borrowed |
| | | 1000 | kʊnt (borrowed?) |

### 3.12.3 North Atlantic and Bak Atlantic numerals in the comparative perspective

It should be stressed that the Atlantic family is among the most divergent within Niger-Congo. Some of the numerical terms in both of the Atlantic groups exhibit a variety of forms potentially explained as Proto-NC reflexes. Moreover, the comparative evidence presented in Tables 3.225 (Proto-North-Atlantic) and 3.239 (Proto-Bak-Atlantic) points to the near total absence of common roots present in both groups. The only exception to this is the root *tɔk/ tVk* 'five'.

In view of this, the only available solution would be the study of the Atlantic evidence within a wider NC context (i.e. in contrast to the reconstructions available for other NC families). A comparison of the intermediate reconstructions within the macro-family will be offered in the next chapter.

## 3.13 Isolated languages vs. Atlantic and Mel

According to the traditional classification outlined in Sapir 1971, Limba, Sua and Gola belong to the Atlantic languages. However, as we tried to demonstrate in Pozdniakov & Segerer 2017 (forthcoming) this hypothesis is as ill-grounded today as it was half a century ago.

An overview of the pertinent data for each language is presented in the tables below.

### 3.13.1 Sua

Table 3.251: Sua numerals

| 1 | sɔn | 7 | 5+2 |
|---|-----|---|-----|
| 2 | cen | 8 | 5+3 |
| 3 | b-rar | 9 | 5+4 |
| 4 | b-nan | 10 | tɛŋi |
| 5 | sɔŋgun | 20 | 10*2 |
| 6 | 5+1 | 100 | kɛmɛ |
| | | 1000 | uŋ-kɔntu |

### 3.13.2 Gola

Table 3.252: Gola numerals

| 1 | guùŋ | 7 | 5+2 |
|---|------|---|-----|
| 2 | tì-yèe/tī-el/cel | 8 | 5+3 |
| 3 | taai/tāāl | 9 | 5+4 |
| 4 | tii-nàŋ | 10 | zììyà |
| 5 | nɔ̀ɔ̀nɔ̀ŋ | 20 | kpɛ̀(w)ùŋ |
| 6 | 5+1 | 100 | 20*5 |
| | | 1000 | < English |

### 3.13.3 Limba

Table 3.253: Limba numerals

| 1 | ha-nthe | 7 | 5+2 |
|---|---|---|---|
| 2 | ka-le/kaa-ye | 8 | 5+3 |
| 3 | ka-tati | 9 | 5+4 |
| 4 | ka-naŋ | 10 | kɔhi |
| 5 | ka-sɔhi | 20 | 10*2 |
| 6 | 5+1 | 100 | kɛmɛ, wuli (borrowed) |
| | | 1000 | wulu (< Mande) |

The chapter 3 includes 250 tables presenting the evidence by group, branch or sometimes a dialect of a certain language. Among them are summary tables that provide an overview of the numerical terms in twelve major families of Niger-Congo and in a number of isolated languages. Our attempt at reconstructing the Proto-Niger-Congo numeral system on the basis of this comprehensive evidence will be presented in the chapter 4.

# 4 Reconstruction of Numerals in Niger-Congo

## 4.1 'One'

The five stems present in Table 4.1 are the most likely candidates for the reconstruction of 'one' in NC (Table 4.1):

1. Commentary. The chart is used to demonstrate the distribution of roots across language families. It groups twelve families into five major branches, including Western NC (Atlantic, Mel), Northwestern NC (Dogon, Gur, Mande), Northern NC (Ubangi, Adamawa), Southern NC (Kru, Kwa, Ijo, BC), and Eastern NC (Kordofanian).

   It should be stressed that this grouping has no implication for the genealogical classification of the NC languages and merely serves as convenient means of display for the isoglosses that will hopefully help to adjust the existing classification.

The chart demonstrates a variety of possible reconstructions. At the same time, some positive knowledge can be gleaned from it. First of all, it should be stressed that a step-by-step analysis of the forms for 'one' attested in the families and branches of NC strongly suggests that no other candidates, except for those displayed in the chart above, can be reconstructed. It should also be noted that the reconstruction of a tri- or even disyllabic root on the basis of the available evidence seems highly improbable, since all potentially reconstructible roots are monosyllabic. Moreover, the inventory of these roots is limited and merits special discussion. Such a discussion is essential, since many of the quasi-reconstructions presented above are not immediately apparent. The problems pertaining to the reconstruction of these roots were to some extent treated in the previous chapter. What follows is a brief survey of the basic facts.

The root *di*. This well-known root has received much scholarly attention as the major candidate for the reconstruction of 'one'. It is manifestly absent only in Kru, Mande and Dogon. In addition to the families listed above, this root is

Table 4.1: Niger-Congo stems for '1'

| | | Dogon | | Kordofan |
|---|---|---|---|---|
| Atlantic | Mande | Gur | Ubangi | Adamawa |
| Mel | Kru | Kwa | Ijo | BC |
| **\*di** | | | | lel/ led? |
| di(n)/li(n)/ne(n | | de/ le/ re | le/ ne | di |
| -lɛ? | | di- | die ? | (o-)di(n), ni/nye |
| **\*in** | | | | |
| -in? | | | | in? |
| -in | | | | hin/kin/cin/-in |
| **\*do** | | do | | |
| don/ lɔŋ | do | | | do? |
| | do | | | |
| **\*ti** | | ti(i) | | ṭe(k)/ lu(k) |
| | | | | |
| | | | | (o-)ti |
| **\*gbo, \*kpo** | | | | |
| mbɔ | | (k)po | k(p)o(k) | *(g)bunu, (mon) |
| bul, mɔ | (g)bolo | | gbérí?/n-kὲni? | gbon, m-o? |

also attested in the Laal language isolate (ɓɨ̀díl (ɓɨ̀-díl?) '1'). At the same time, it is absent in the Sua, Gola and Limba isolates. It bears reminding that the reconstruction of this root in Benue-Congo and Bantu is only possible under the assumption that PB *mòdì < *m-ò-dì '1' (with **m-** being a Proto-Bantu CL1, and **-o-** being an archaic noun class marker (possibly < **\*ko-/ \*ʔo-**, i.e. NC class CL1 incorporated into the stem).

The root *in. Although this root is not attested outside Western NC, BC and possibly Adamawa, it is worth mentioning, especially in view of its possible etymological relationship with \**di* (see above).

The same is applicable to \**do* (best attested in Northern NC, Atlantic and Kru).

The reconstruction of \**ti* '1' is the least certain among the roots discussed above. The form *ha-nthe* '1' attested in the Limba language isolate is noteworthy.

The last root is a tentative representation of the forms with the initial labiovelar (or labial in the case of Western NC) that are not necessarily etymologically related. The root *guùŋ* '1' attested in the Gola isolate may belong here as well.

In addition to the five roots treated above, apparent innovations may be attested in particular families (or even in groups within them). Among these are Kordofanian *ʈɔn* (cf. Sua *sɔn*), Gur *túrú/ tumɔ*, Mande West *kelen*, and Atlantic Bak -*anor, ǝkon*.

## 4.2 'Two'

### 4.2.1 'Two'

A systematic comparison of the terms for 'two' attested in the NC families yields somewhat unexpected results. The only candidate for the reconstruction of the NC term is the root that can be tentatively recorded as *di. At the same time, nearly every family has its own root (or, more often, roots) for 'two' that finds no parallel outside the branch/family in question. The distribution of *di, as well as an overview of isolated roots, is presented in the chart below (Table 4.2):

Table 4.2: Niger-Congo stems for '2'

| | | Dogon | | Kordofan |
|---|---|---|---|---|
| Atlantic | Mande | Gur | Ubangi | Adamawa |
| Mel | Kru | Kwa | Ijo | BC |
| **\*di/ni** | | lɛ́(y)/ lɔ́(y)/ nɛ́(y)/ nɔ́(y) | | |
| di(k), nak | | nyi/ ne(n) | | du/ru, te/ re/ si |
| díŋ/tsiŋ/tiŋ/rəŋ | | | | ba-di / ba-ji |
| **isolated roots** | | | si/ ʃi, | kok/kek/cik, (can/ʈan, rak, rɔm) |
| ɫubəʔ, -təb/-təv -puguʈ/pugus | pila/ fila | nyu/ ju, hin/ han | (wa/gbwo, to/ so) | ra(k)/ ra(p), gba/ gwa |
| | so(n) | ɲɔ | mamV | pa ? ba(i) ? |

1. Commentary. The isolated forms are as follows: Laal *ʔīsī (ʔī-sī?)* (this root is comparable to that attested in Ubangi), Sua *cen*, Gola *tì-yèe/ tī-el/ cel* (the Gola and Sua terms may be related), Limba *ka-le/ kaa-ye* (this root may go back to NC *di*).

The unprecedented variety of forms exhibited by the term for 'two' is especially surprising because this notion has been viewed as one of the most persistent

in language history (it is the only numeral on the Swadesh list). As we will see below, this term is the least stable in the Niger-Congo languages. However, the NC root *\*di* is well-attested across the families.

### 4.2.2 'Two' = 'one' PL?

As can be gleaned from the evidence presented above, the only root for 'two' reconstructible in NC (*\*-di*) is suspiciously similar to the most likely reconstruction for 'one' (*\*-di*). This similarity was first observed by Raymond Boyd, one of the most renowned experts in the reconstruction of Adamawa. Before we turn to the discussion of the most promising (in terms of the NC reconstruction) forms, an overview of Raymond Boyd's hypothesis regarding Adamawa and some of the BC languages is in order. Here is what Boyd writes about the reconstruction of 'one': "A rather complicated hypothesis would, in fact, cover most of the Cross River/Platoid data: Let us assume a single root, \*DI (sometimes ~\*DU) and two affixes, (V)K(V) and (V)N(V), which can appear, separately or together, as either prefixes or suffixes, or both. <…> Some support for this hypothesis is provided by the frequently observed inversion of the coronal and velar features: in most cases, where we find a term with initial velar, we find a final coronal nasal; and where we find an initial coronal, we find a final velar nasal. This can be explained by assuming the prefixation of \*KV-N- in the former case, and suffixation of \*-N-K(V) in the latter." (Boyd 1989: 151-152). Boyd's proposal is to reconstruct the Proto-Adamawa terms for 'one' and 'two' as *\*n-di* and *\*bà-dí* (with class 2 prefix) respectively (Boyd 1989: 156). According to him, "It was suggested above that the Cross River/Platoid root for 'one' was \*DI. We may now hypothesize that the root for 'two' in the proto-language for these groups was the plural \*BA.DI, and that, when Proto-Bantu developed its more complicated class system, this term, whose prefix may have been invariable, was reinterpreted as mono-morphemic" (Boyd 1989: 157).

It should be stressed that Boyd's hypothesis explains the Proto-Bantu forms that underwent the following transformation over the course of time: *\*m* (CL1)-*o*(<*CL1)-*di* > *\*mɔ̀-òdì* / *mòì* '1' / *ba*(CL2)-*di* > *badi* '2' (the dialectal Proto-Bantu form *jòdè* (zones BH) (< *\*jò(*CL5 ?)-*di*?)). It bears reminding that our evidence favors the reconstruction of *(o-)di(n)* '1' / *ba-di* / *ba-ji* '2' at the BC level.

One of the major problems with this reconstruction is that synchronically the roots for 'one' and 'two' are the same in only a minority of the modern NC languages. This rare phenomenon is attested in the Ngabaka branch of Ubangi (Table 4.3):

As stated above, examples of this kind are exceptionally rare. A possible ex-

Table 4.3: The same stem in '1' and '2' (*di)

|  | 'one' | 'two' |
|---|---|---|
| Bayanga | bo-dé | bi-dé |
| Bomasa | bo-dé | bi-dé |
| Baka | kpó-de | bí-de |
| Gundi | po-dé | bi-dé |
| Ngombe | kpóo-de- | bí-de- |

planation for the overwhelming absence of the identical roots for 'one' and 'two' is that one of the classes is subject to the nasalization process (entailing further phonetic changes within the root), while the other is not. It bears reminding that, according to Boyd, a number of expanded forms such as *n-di (with further development to *-ni / -in 'one') is reconstructible along with *-di.

In view of this, the Oti-Volta numbers, thoroughly discussed in the previous chapter, are especially interesting. The pertinent Oti-Volta forms are as follows (Table 4.4):

Table 4.4: Potential reflexes of *di '1' = *di '2' in Gur

|  | i. Buli-Koma | ii. Eastern | iii. Gurma | iv. Western | v. Yom-Nawdm | **\*Proto-Oti-Volta** |
|---|---|---|---|---|---|---|
| 1 | yéŋ, ní | dèǹnì, yẽnde/yènn(do), yòn, *de | /yènn(do), den, ni | yen/ yin, dam | hén, nyǎŋ | **den/ yen, ni, de?** |
| 2 | yè, li | dɛ́ɛ́(ni), yēdē | le/ dɛ́ | yi(ʔ) | li/ ɾéʔ/ *ɾɣa? | **li/yi** |

The terms for 'one' and 'two' are similar within each of the branches, the differences between them being due to the presence of the nasal component in the term for 'one'.

## 4.3 'Three'

Table 4.5: Niger-Congo stems for '3'

| | | Dogon | | Kordofan |
|---|---|---|---|---|
| Atlantic | Mande | Gur | Ubangi | Adamawa |
| Mel | Kru | Kwa | Ijo | BC |
| | | taan | | tat/ tə̀ɽ/ ʈak |
| North: 'taʈ | | tat/ ta(n) | taar | taat |
| sas/ ra | taa(n) | ta | tato | tat |
| | | | | (ɽitin/ ɽicin, hwʌy) |
| Bak: feegir, yanʈ/ jenʈ, habi/yabi | sakpa/sagba/sa ʔààkɔ̃/yààká? | | | kunuŋ/ gbunuŋ |
| | | | | |

As is well known, the term for 'three' is exceptionally persistent, with the same root attested in all of the major NC branches (except for Mande). The same root is also present in the Western NC isolates, cf. Sua *b-rar*, Gola *taai/tāāl*, Limba *ka-tati*. At the same time, some languages exhibit what are apparently innovative forms (see the downmost segment of the chart). An isolated root is also attested in Laal (*māā* '3').

Although the relationship between the reflexes of the main root (\**tath*) is unquestionable, their phonetics pose a problem. The issue is that each family exhibits a great variety of reflexes, while some of them cannot be explained as going back to either the initial \**t*- or the final \*-*t* of the main root. In other words, reliable correspondences (with \***t** preserved) are traceable in the majority of families, but not in the case of 'three'. This forces us to assume that \*t may be irregularly reflected as **s**, **r**, **h** in particular families.

The table below (Table 4.6) provides an overview of the pertinent Bantu reflexes of \**tátʊ̀* (ABEFGHJKLMNPRS) / \**cátʊ̀* / \**cácʊ̀* (CD) 'three' (these recon-

structions follow BLR3):

Table 4.6: Reflexes of *tátʋ̀* '3' in Bantu

| zone | Language | Form | zone | Language | Form |
|------|----------|------|------|----------|------|
| A | Nyo'o | tá | *PB | *PB (dial.) | cátʋ̀ |
| A | Lundu | aru | D | Lega | sáro |
| A | Bonkeng | alu | E | Pokomo | hahu |
| A | Fang | lal | E | Embu | thatu |
| A | Ewondo | lá | E | Kahe | radu |
| A | Kpa | ráá | F | Sukuma | datu |
| A | Lombi | laso | G | Pemba | tatu |
| A | Bubi | cha | G | Tikuu | chachu |
| B | Yansi | taar | J | Konzo | satu |
| B | Mbere | tadi | J | Luganda | ssatu |
| B | Sira | reru | J | Nyankole | shatu |
| B | Kande | lato | K | Nyengo | ato |
| B | Galwa | ntʃaro | K | Mbwela | hatu |
| C | Bua | salu | L | Kete | sàcw |
| C | So | saso | S | Lozi | talu |
| C | Sakata | sâa | S | Venda | raru |
| C | Koyo | tsáro | S | Swazi | tsâtfu |

The Bantu forms should be discussed in order to determine which processes in Bantu (and in Niger-Congo in general) give rise to such a diversity of phonetic variants.

The root includes two consonants. Putting aside the problem of the vowel in the second syllable, we label the two consonants C- and -C respectively. Each of them may be dropped, yielding the Bantu forms **ta** and **at** (Scheme 4.1).
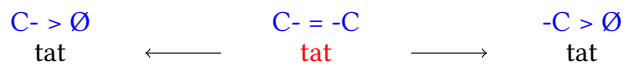
C- > Ø          C- = -C          -C > Ø
tat     ⟵     tat     ⟶     tat

Figure 4.1: Scheme 4.1

Each of them can be transformed, for example, with a spirantisation **\*t > s**, **or \*t > r, \*t > l**, can become voiced **\*t > d** and only after that can the second consonant be dropped. (Schemes 4.2-4.3).
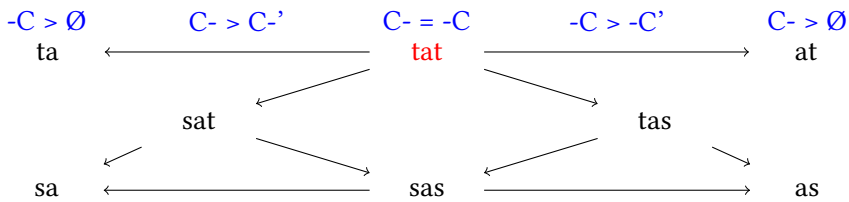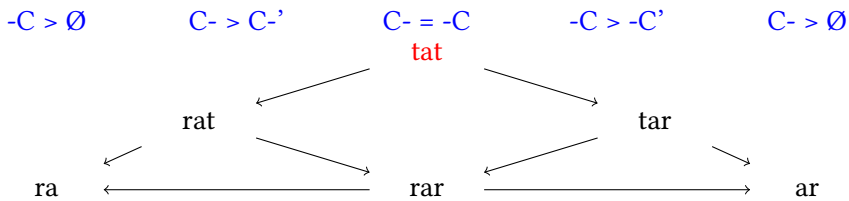
Figure 4.2: Scheme 4.2



Figure 4.3: Scheme 4.3

As a result, we have numerous forms, while the variation can be reduced to a very limited number of processes:

- Voicing (*t > d)

- Lenition – partial (spirantization: *t > s, *t > r) or full ( > Ø).

The following table (Table 4.7) provides a structured overview of the derived Bantu forms (with no arrows):

However, the resource for changes in Bantu is not limited to the above. The derivational schemes mentioned above are constructed not only on the basis of *tat*, but also from newly derived forms. For example, *\*tat > sat*, and others (Scheme 4).

This is where the following forms (Table 4.8), many of which are attested in Bantu, originate (forms without square brackets):

1. We often do not know how one or another derived form appeared. For example, the form *las* in the first line of the table could have originated from *\*tas* (as a result of the change in the first consonant – the variation in the line) or from *\*lat* (the change of the second consonant – column). Many of the forms which are predicted theoretically are not attested in Bantu; these are shown in square brackets.

Table 4.7: Phonetic variations of *tat-*

| -C | C- | C- -C | -C | C- |
|----|----|----|----|----|
|    |    | tat |    |    |
| ta |     |     |     | at |
|    | sat |     | tas |    |
| sa |     | sas |     | as |
|    | cat |     | tac |    |
| ca |     | cac |     | ac |
|    | rat |     | tar |    |
| ra |     | rar |     | ar |
|    | lat |     | tal |    |
| la |     | lal |     | al |
|    | hat |     | tah |    |
| ha |     | hah |     | ah |
|    | dat |     | tad |    |
| da |     | dad |     | ad |
|    | zat |     | taz |    |
| za |     | zaz |     | az |
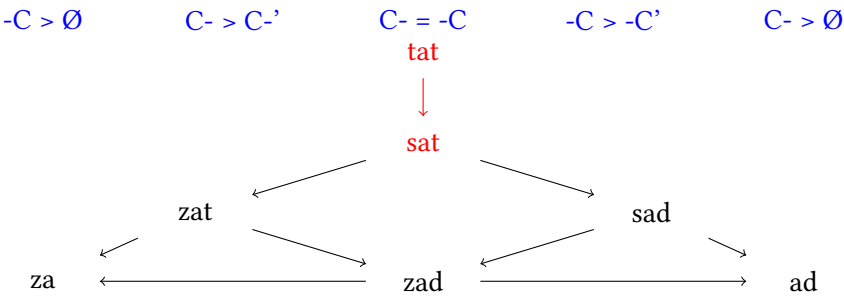


Figure 4.4: Scheme 4.4

Table 4.8: Reflexes of *tat-* attested in Bantu

|     | sat   | cat   | rat   | lat   | dat   | zat   |
|-----|-------|-------|-------|-------|-------|-------|
| tas | sas   | [cas] | [ras] | las   | [das] | [zas] |
| tac | sac   | cac   | [rac] | [lac] | [dac] | zac   |
| tar | sar   | car   | rar   | [lar] | dar   | [zar] |
| tal | sal   | [cal] | [ral] | lal   | [dal] | [zal] |
| tah | [sah] | [cah] | rah   | [lah] | [dah] | [zah] |
| tad | sad   | [cad] | rad   | [lad] | dad   | [zad] |
| taz | [saz] | [caz] | [raz] | [laz] | [daz] | zaz   |

The most amazing observation here is not the high degree of variation (which itself needs to be considered), but the fact that we find precisely the same variations in different branches of NC. As a result, in different branches of NC—that is—in languages with distant genetic relations, we find numerous identical forms, while in every branch taken separately we find an "antimagnetic" landscape of forms, which in closely related languages tend to be maximally differentiated.

Examples from seven branches of NC are given below and divided into two structurally identical tables (Table 4.9-4.10):

We see, for example, that roots **TAL** and **TAR** are observed in all seven branches.

To get a comprehensive idea of the presence of the forms in each branch we are attracting attention to the following chart, where the presence of the forms (at least in one language) is marked by a cross (the data is arranged in descending order in the summarising column as well as in the summary line) (Table 4.11):

The following chart represents the number of groups (within the 14 branches of Niger-Congo) presenting the respective combinations of the first (the line) and the second (the column) consonants (the data is presented in descending order) (Table 4.12):

As we can see, the most frequent consonants in the initial position are **t**- and **s**-, while the second consonant is one of the following three: -**Ø**, -**t**, or -**r**.

If we reconstruct *tat-* on the NC level, in line with the majority of linguists, we will have to contend with quite a mysterious picture. In the majority of younger proto-languages we will also have to reconstruct *tat-*, because, as it has already been shown, it descends into more or less the same variation of forms. It means that during thousands of years, from Proto-NC to the formation of proto-languages in separate branches, the form remained phonetically unchanged. Then, suddenly the root *tat* independently started to explode, giv-
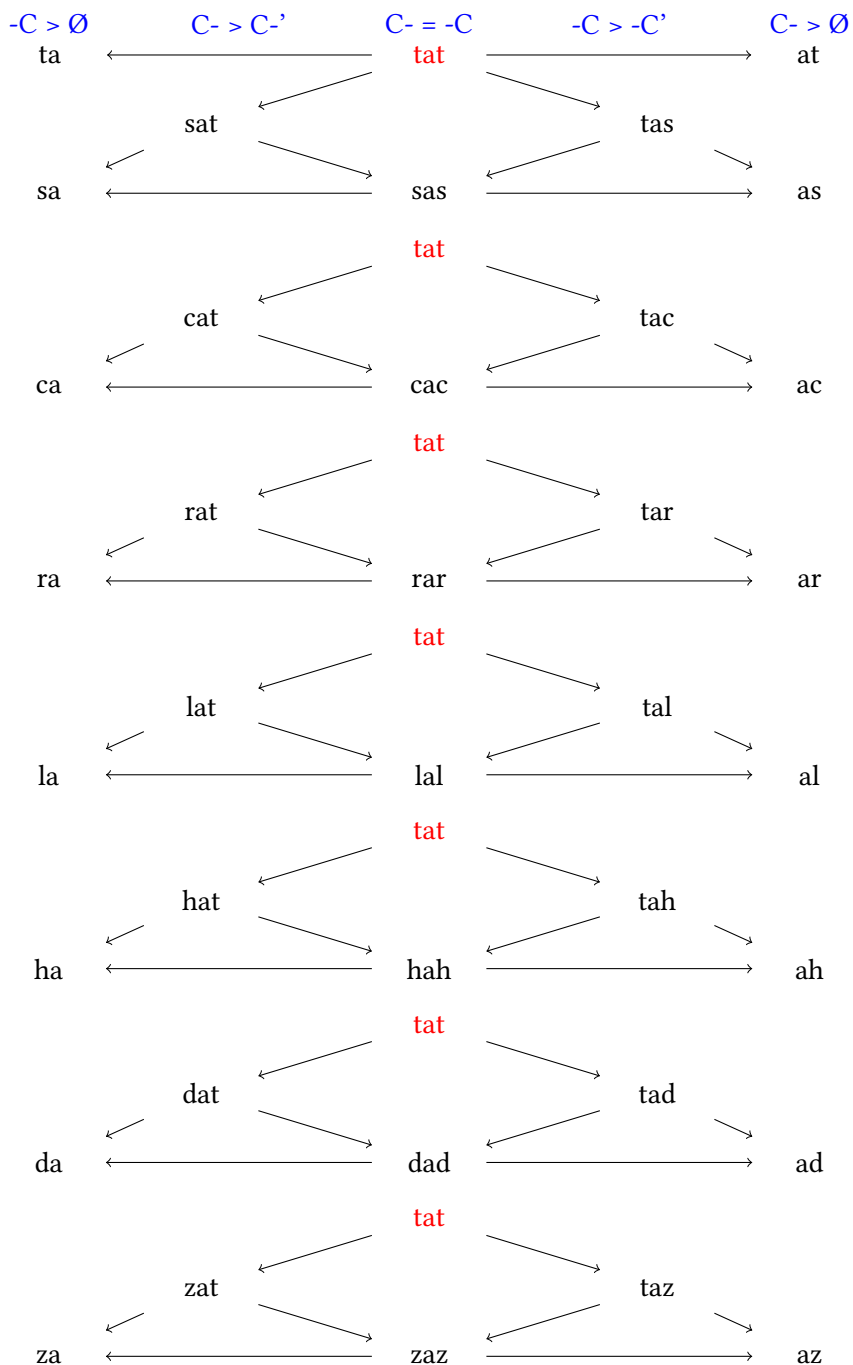
Figure 4.5: *Scheme 4.5*

Table 4.9: Reflexes of *tat-* in Niger-Congo (1)

|  | Bantu | | Adamawa | | Atlantic-Mel | |
|---|---|---|---|---|---|---|
| **TAT** | Rundi | **tatu** | Yendang | **tat** | Fula | **tat-** |
| **TAR** | Yansi | **taar** | Bangunji | **taar** | Buy | **taar** |
| **TAL** | Lozi | **-talu** | Dadiya | **tal** | Gola | **tāˈl** |
| **TAD** | Mbere | **-tadi** |  |  | Sereer | **tad-ak** |
| **TAS** |  |  | Kulaal | **tòòs** | Bapen | **ɓʌ-tas** |
| **TAZ** |  |  | Mom Jango | **tàáz** | Tanda | **-taaz** |
| **TA** | Nyo'o | **tá** | Tunya | **ta** |  |  |
| **SAT** | Bushong | **-satu** | Kumba | **sa:t** |  |  |
| **SAR** | Nzadi | **i-sár** |  |  |  |  |
| **SAS** | So | **-saso** |  |  | Temne | <span style="color:red">**pɛ̀-sās**</span> |
| **SA** | Sakata | **i\|sâa** | Mangbai | **bi-ssá-** |  |  |
| **AT** | Nyengo | **-ato** |  |  | Nalu | **-at** |
| **AR** | Lundu | **-aru** |  |  | Kasanga | **-ar** |
| **LAL** | Fang | **lal** |  |  | Nyun | **ha-lal** |
| **RAR** | Venda | **-raru** |  |  | Sua | **-rar** |
| **RA** | Kpa | **-ráá** |  |  | Sherbro | **ra** |
| **CAR** | Orungu | **tʃaro** | Kam | **tshar** |  |  |
| **CA** | Bubi | **-cha** | Galke | **cha-?a-** |  |  |
| **HAT** | Nkoya | **-hatu** |  |  | Manjak | **go-hant** |
| **DER** |  |  |  |  | Baga | **der** |
|  |  |  |  |  | Mboteni |  |

ing rise to much phonetic variation in its reflexes.

I think that a hypothesis stating that the root already contained close but not identical consonants in NC is far more typologically justified. The first consonant in that case was *t-, while the second one was represented by a specific phoneme for which no traces remain, for example, *-th ?, *-ṭ ?,*-tṣ?,*-c? As we tried to show in (**PozdniakovSegerer2007**), the phonotactics of many languages (not exclusively in Africa) demonstrates the same tendency: in CVC structures languages tend to avoid consonants constituting a minimal pair, for example, *fVp, bVp, sVz, lVr, rVl, sVʃ, etc.* In diachronic perspective, the existence of such combinations often leads to numerous irregular changes, in the course of which the consonants either become identical, for example, *lVr > lVl,* or, on the contrary, acquire a higher level of contrast, escaping the zone of "dangerous proximity",

Table 4.10: Reflexes of *tat-* in Niger-Congo (2)

|      | Bantoid |        | BC          |         | Dogon      |         | Gur       |              |
|------|---------|--------|-------------|---------|------------|---------|-----------|--------------|
| TAT  | Bankala | **tát**  | Birom       | **be-tat**  | kolum so    | **tã́ã́ti**   | Ditammari | **-tã́ã́tī**      |
| TAR  | Mambila | **tar**  | Jiru        | **i-tar**   | bangeri-me  | **ke-taro** | Senari    | **tãre**        |
| TAL  | Kom     | **tál**  | Olulumo     | **è-tál**   | toro tegu   | **taali**   | Nateni    | **tã́lī, tã́di**  |
| TAD  | Ngwe    | **tád**  | Upper-Cross | **\*-ttáD** | tommo so    | **tadu**    | Nateni    | **tã́di, tã́lī**  |
| TAS  |         |        | ikaan       | **tás**     |            |         |           |              |
| TAZ  |         |        |             |         |            |         |           |              |
| TA   | Abon    | **-ta**  | Ibibio      | **ì-tá**    |            |         | Dagbani   | **-ta**         |
| SAT  |         |        | Morwa       | **sat**     |            |         |           |              |
| SAR  | Mbe     | **bé-sár** | Kugbo     | **ì-sàr**   |            |         | Lorhon    | **sã:r**        |
| SAS  |         |        |             |         |            |         | Viemo     | **saasi**       |
| SA   | Ekoi    | **é-sá** | Oloma       | **e-sa**    |            |         | Kulango   | **sã**          |
| AT   |         |        | Kohumono    | **a-àtá**   |            |         | Hanga     | **ata**         |
| AR   |         |        |             |         |            |         |           |              |
| LAL  |         |        |             |         |            |         |           |              |
| RAR  |         |        | Abua        | **ì-rààr**  |            |         |           |              |
| RA   | Nkem    | **í-rá** | Ukue        | **è-rhá**   |            |         |           |              |
| CAR  |         |        | Ufia        | **kù-tshàr**|            |         |           |              |
| CA   |         |        | Bandawa     | **ni-ca**   |            |         |           |              |

for example, *sVsh > sVh, *bVp > bVf*. In other words, similar sounds being adjacent to one another are a constant zone of tension which provokes all possible irregular changes.

It is very likely that such a situation characterises the NC root for 'three'. In this case, the considerable phonetic variability of the root in all the stages of its development from Proto-NC to contemporary languages can be typologically – phonotactically – explained.

Table 4.11: Distribution of different reflexes of *tat- in the Niger-Congo families

| | Bantu | Benue-Congo | Atl | Adam. | Bantoid | Gur | Mel | Kwa | Ubangi | Dogon | Kordof. | Kru | Ijo | Mande |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TA | x | x | | x | x | x | | x | x | | x | x | | 9 |
| TAR | x | x | x | x | x | x | | | x | x | | | x | 9 |
| TAT | x | x | x | x | x | x | x | | | x | | | | 8 |
| TAL | x | x | | x | x | x | x | | x | x | | | | 8 |
| TAD | x | x | x | | x | x | | | | x | x | | | 7 |
| SA | x | x | | x | x | x | | x | | | | | x | 7 |
| AT | x | x | x | | | x | | x | | | x | | | 6 |
| RA | x | x | | x | | | x | | x | | | | | 5 |
| SAR | x | x | | x | x | | | | | | | | | 4 |
| SAS | x | | x | | x | | x | | | | | | | 4 |
| LA | x | x | | | | | | | x | | | x | | 4 |
| TAS | | x | x | x | | | | | | | | | | 3 |
| SAT | x | x | | x | | | | | | | | | | 3 |
| AR | x | | x | | | | | x | | | | | | 3 |
| HAT | x | | x | | | | | | | | x | | | 3 |
| RAR | x | x | x | | | | | | | | | | | 3 |
| CAT | x | x | | | x | | | | | | | | | 3 |
| CAR | x | x | | x | | | | | | | | | | 3 |
| TAZ | | | x | x | | | | | | | | | | 2 |
| HA | | | x | | | | | x | | | | | | 2 |
| LAL | x | | x | | | | | | | | | | | 2 |
| DAT | x | x | | | | | | | | | | | | 2 |
| CA | x | | | x | | | | | | | | | | 2 |
| SAL | x | | | | | | | | | | | | | 1 |
| AL | x | | | | | | | | | | | | | 1 |
| AS | | | | | | | x | | | | | | | 1 |
| HAH | x | | | | | | | | | | | | | 1 |
| THAT | x | | | | | | | | | | | | | 1 |
| TSAR | x | | | | | | | | | | | | | 1 |
| RAH | | | | | | | x | | | | | | | 1 |
| DAR | | | x | | | | | | | | | | | 1 |
| TAH | | x | | | | | | | | | | | | 1 |
| TAC | | x | | | | | | | | | | | | 1 |
| DAD | x | | | | | | | | | | | | | 1 |
| DAZ | | | | | | x | | | | | | | | 1 |
| RAT | | | | | x | | | | | | | | | 1 |
| RAD | x | | | | | | | | | | | | | 1 |
| LAT | x | | | | | | | | | | | | | 1 |
| LAS | x | | | | | | | | | | | | | 1 |
| SAD | | x | | | | | | | | | | | | 1 |
| SAC | x | | | | | | | | | | | | | 1 |
| CAC | x | | | | | | | | | | | | | 1 |
| ZA | | | | | | | | x | | | | | | 1 |
| ZAC | | | x | | | | | | | | | | | 1 |
| | 31 | 19 | 14 | 10 | 10 | 10 | 6 | 6 | 5 | 4 | 4 | 2 | 1 | 1 | 123 |

Table 4.12: Number of different phonetic structures for '3' in 14 NC branches

|       | Ø  | t  | r  | l  | d  | s | c | h | z |     |
|-------|----|----|----|----|----|---|---|---|---|-----|
| t     | 10 | 8  | 9  | 8  | 7  | 3 | 1 | 1 | 2 | 49  |
| s     | 7  | 3  | 4  | 1  | 1  | 4 | 1 |   |   | 21  |
| c, ts | 3  | 3  | 5  |    |    |   | 1 |   |   | 12  |
| Ø     |    | 6  | 3  | 1  |    | 1 |   |   |   | 11  |
| r     | 5  | 1  | 3  |    | 1  |   |   | 1 |   | 11  |
| l     | 4  | 1  |    | 2  |    | 1 |   |   |   | 8   |
| h     | 2  | 3  |    |    |    |   |   | 1 |   | 6   |
| d     |    | 2  | 1  |    | 1  |   |   |   | 1 | 5   |
| z     | 1  |    |    |    |    |   | 1 |   |   | 2   |
|       | 32 | 27 | 25 | 12 | 10 | 9 | 4 | 3 | 3 | 125 |

## 4.4 'Four'

Table 4.13: Niger-Congo stems for '4'

| Atlantic | Mande | Dogon / Gur | Ubangi | Kordofan / Adamawa |
|---|---|---|---|---|
| Mel | Kru | Kwa | Ijo | BC |
| | | nay(n) | | |
| Nord: nak | náání / nã̃ãi | naan | naar | naX, ɲɛ̄n/ nìŋ, nda |
| | na | na | nɛ́ín | nai |
| | | kɛɛso | | -ɽɔŋ /-ɽandɔ/ -rʊm?,(-gʌ́làm) |
| Bak: baakər/wa tasala | | | ( syɔ), lu) | |
| Nord: '-ŋkɨlɛ/-nlɛ, Sud: hiɔl | | | | |

Just like the term for 'three', the term for 'four' is exceptionally persistent in NC. It is represented by the same root in all the families (except for Mel and Kordofanian), as well as in the Western NC isolates, cf. Sua *b-nan*, Gola *tii-nàŋ*, Limba *ka-naŋ*. At the same time, a number of innovations are attested in some of the families (see the downmost segment of the chart) and in the Laal isolate, cf. *ɓī̄-sā̱n* (*ɓī̄-sā̱n*?) '4'.

This root is not present in Nilo-Saharan (including Songhai), nor in Afroasiatic or Khoisan. In light of this, the root can be viewed as one of the best isoglosses indicating the genetic relationship of languages within NC. Used together with the isogloss for 'three', it becomes a powerful means of classification, i.e. if the term for 'three' has (or goes back to) **t-** as the initial consonant in a given language, whereas the term for 'four' starts with **n-**, this language must belong to the Niger-Congo family. Hundreds of the NC languages match this description, while, as far as I am aware, none of the languages from other families meets these

requirements.

There will probably be no objection from the specialists in the field to the statement that the main root for 'four' begins with **\*na**-, e.g. this form is reconstructed for Proto-Potou-Akanic-Bantu by John Stewart. At the same time, many languages show that the root initially included two vowels, **\*i** being the second of the two. The major issue, however, is establishing whether the root included another consonant (i.e. whether \**nai* or \**naCi* should be preferred) and if so, what it was. Stewart suggests *\*na~ŋi~* '4' as the Proto-Potou-Tano-Congo form (**Stewart1983**), but his reconstruction is not applicable to NC.

At the same time, the reconstruction of the proto-form for 'four' is not an easy task. The problem is that a given form does not define the languages it is attested in as members of the same group. Nearly every group has an inventory of phonetically similar forms (just like in case of 'three'). The Bantu languages may provide a good illustration for this phenomenon.

The most frequently attested Bantu forms include *na, nai, nayi, ne, nei* and *ni* (six in total). They are found in 276 of 355 Bantu sources that include a form for 'four' available in our database. Their zonal distribution is as follows (Table 4.14):

As can be gleaned from the table, the six forms discussed above are commonly attested in our sources stemming from zones as diverse as C, F, J, M, and S. For instance, pertinent forms are attested in 26 out of 27 sources available in our database for the J zone (the last source, namely the Luganda language, has *nya* 'four' that probably goes back to the same root).

The problem, however, is that this (or a nearly identical) set of forms is attested within the other NC families as well, cf. e.g. the Kwa evidence (Table 4.15):

The Adamawa evidence is as follows (Table 4.16):

My suggestion is that the variety of similar forms attested in the majority of the NC branches may be due to the complex inter-relationship between the terms for 'four' and 'eight' in NC. We will return to this hypothesis later, in the section dealing with 'eight'.

Table 4.14: Distribution of the main n- forms for '4' in Bantu zones

| zone | na | nai | nayi | ne | nei | ni | SUM | SOURCES |
|---|---|---|---|---|---|---|---|---|
| A | 13 | 3 | 2 | 6 | 1 | 7 | 32 | 52 |
| B | 31 | 8 | 10 | 7 | 1 | 1 | 58 | 65 |
| C | 2 | 2 | | 2 | 18 | 1 | 25 | 28 |
| D | 1 | 1 | | 4 | | | 6 | 14 |
| E | 4 | | | 4 | | 1 | 9 | 19 |
| F | | | | 9 | | 3 | 12 | 13 |
| G | 2 | | | 18 | | 1 | 21 | 26 |
| H | 7 | | | | | | 7 | 11 |
| J | 10 | | | 15 | | 1 | 26 | 27 |
| K | 6 | | | 7 | | 1 | 14 | 15 |
| L | 6 | 1 | 2 | | | | 9 | 12 |
| M | 3 | 1 | | 11 | | 5 | 20 | 20 |
| N | 2 | 3 | 2 | 2 | | | 9 | 12 |
| P | 2 | 2 | | | | | 4 | 11 |
| R | | | | 3 | | | 3 | 7 |
| S | 7 | | | 14 | | | 21 | 23 |
| SUM | 96 | 21 | 16 | 102 | 20 | 21 | 276 | 355 |

Table 4.15: Main n- forms for '4' in Kwa

| | |
|---|---|
| Agni (Anyin) | n-na |
| Abron | n-nai |
| Baule | nu-ne |
| Eotile (Beti) | a-ni |

Table 4.16: Main n- forms for '4' in Adamawa

| | |
|---|---|
| Tupuri | na |
| Mundang | nai |
| Gula | nay |
| Waja | ni |

## 4.5 'Five'

The term for 'five' is typically based on the lexical term for 'hand' in Mel and Atlantic. At the same time, the term for 'ten' is often derived from 'five' or, like 'five', directly from 'hand' in the plural. Multiple examples illustrating this phenomenon will be provided below. At this point I will limit myself to merely stating that the attestation of this pattern throughout the NC branches is inconsistent. Thus, it is virtually unattested in Bantu (as well as in BC on the whole). According to **NursePhilippson1975**, the Usseri dialect of Rombo (Bantu E) is a unique exception in this respect, cf. *ku-oko* 'hand' (Proto-Bantu *\*bókò*) yielding *ku-oko* ('5') and *ku-oko ka-vili* ('10', '5\*2'). At the same time, the reflexes of the Proto-Bantu roots for 'five' (*tanu*) and 'ten' (*i-kumi*) are attested in this language along with the irregular forms discussed above. These two patterns are barely attested in Kwa, Gur, Kru, or Ijo. On the contrary, they are common not only in Atlantic and Mel but also in Ubangi (Gbaya in particular), in some of the Adamawa languages, in a number of Kordofanian branches and possibly in Mande. In view of this distribution, the existence of these patterns in NC seems unlikely. Apparently, the terms for 'hand' should be considered when trying to establish the NC etymology for 'five' and 'ten'.

Our discussion will start with the unrelated roots for 'hand' and 'five' attested within the same branch. Then we will turn to the evidence of those groups where both terms go back to the root for 'hand'. This approach will allow the accumulation of data that will enable us to suggest a likely diachronic explanation for the phenomenon.

We will start with the Bantu evidence. The Bantu languages (like the majority of the NC groups in general) are characterized by the presence of multiple roots for 'hand' and 'arm'. The most persistent of these according to BLR3 are the following roots (Table 4.17):

I would like to stress that these roots are virtually unattested in Bantu with the meaning 'five' or 'ten'. According to BLR3, the only primary root for 'five' commonly attested in Bantu is *\*táànò.* At the same time, the root *\*dòngò,* which probably goes back to *\*dòngò* 'line, row' (zones: ABCDEGHJKLMNRS) deserves our attention as well.

The initial consonant in *\*táànò* is the same as in *\*tátò* 'three', which is probably a coincidence. However, this fact can still be used for establishing the genetic relationship of the NC forms for 'five'. The possibility that the languages (or language groups) are related to the reconstructed Bantu forms is stronger if the terms for 'three' and 'five' attested in them have the same initial consonant. The

Table 4.17: Distribution of the stems for 'hand', 'arm' in Bantu zones

| PB | meaning | regions (5) | zones (16) |
|---|---|---|---|
| bókò | arm; hand; front paw | 5: NW SW Ce NE SE | 14: A B C D E G H J K L M N R S |
| gànjà | palm of hand; main | 5: NW SW Ce NE SE | 14: A B C D F G H J K L M N P S |
| pí | palm of the hand; slap | 5: NW SW Ce NE SE | 14: A B D E F G H J K L M N R S |
| kónò | forearm; arm; hand; leg; hoof | 4: SW Ce NE SE | 10: E F G J K L M N P S |
| nàmà jádà | limb: arm; leg; thigh nail (> finger > 'hand) | 4: NW SW Ce NE | 8: A B C E H L M R > 'hand' A D E F G J L N P S |

following Bantu evidence (Table 4.18) is illustrative of this admittedly unconventional approach (further BC evidence will be quoted later in this chapter).

Table 4.18: Identical initial consonants in '3' and '5' in Bantu

| | Language | '3' - *tátò | '5' - *táànò |
|---|---|---|---|
| Bantu-J | Rwanda | tatu | tanu |
| Bantu-B | Punu | reru | ranu |
| Bantu-E | Gusii | sato | sano |
| Bantu-G | Swahili | tatu | tano |
| Bantu-R | Herero | odatu | odano |
| Bantu-A | Bubi | ca | cio |
| Bantu-A | Tunen | lal | lan |

This rule is irreversible, i.e. the diversity of the initial consonants is not indicative of either form not being a Proto-Bantu reflex (Table 4.19):

The fact that the same consonants are reflected differently may have several explanations, e.g. that the noun class prefixes (especially the nasal marker of class 9) may have impacted the process. A number of other phonotactic factors may also be involved (some of which are treated in detail in the section dealing with 'three').

Table 4.19: Different initial consonants in '3' and '5' in Bantu

|         | Language | '3' - *tátò̱ | '5' - *táànò̱ |
|---------|----------|-------------|---------------|
| Bantu-F | Bungu    | tatu        | (zi)sano      |
| Bantu-G | Pogoro   | tatu        | mhanu         |
| Bantu-S | Sesotho  | taro        | hlano         |
| Bantu-G | Komoro   | traru       | canu          |
| Bantu-D | Holoholo | satu        | tano          |
| Bantu-J | Haya     | -satu       | i-tanu        |
| Bantu-K | Mbwela   | -hatu       | -tanu         |
| Bantu-E | Kahe     | si-radu     | si-tanu       |
| Bantu-A | Kpa      | -ra         | -tan          |
| Bantu-G | Tikuu    | -cacu       | -tano         |
| Bantu-K | Mwenyi   | -atu        | mu-tanu       |
| Bantu-A | Balong   | be-lal      | be-tan        |
| Bantu-B | Kele     | -lali       | -tani         |
| Bantu-L | Mbwera   | k-atu       | -tanu         |
| Bantu-E | Digo     | -hahu       | cano          |
| Bantu-E | Taita    | i-dadu      | i-sanu        |
| Bantu-N | Manda    | ji-datu     | mu-hanu       |
| Bantu-S | Ronga    | -rjarju     | tlhanu        |

The pairs of BC terms with the same initial consonant attested outside Bantu will be our primary concern in further discussion. Some of them are quoted in the table below (Table 4.20): As can be gleaned from the table, the root *tanV / *taVn is systematically attested in nearly every BC branch, hence its reconstruction at the Proto-BC level seems certain. Moreover, it is widely attested in many other NC branches as well. The following forms of 'three' and 'five' (with the same initial consonant) are comparable to *BC root (Table 4.21):

The Table 4.21 shows peculiar forms attested in one of the Southern Mel languages (Bom) that are virtually identical to the BC reconstructions. Thus, we have every reason to reconstruct the term for 'five' as *tan (unrelated to 'hand') at the NC level. The distribution of this root is illustrated in the following chart (Table 4.22):

The attestations of this root in Southern NC (namely in BC, Kwa and Ijo) are

---

[1] **Elugbe1987**.

Table 4.20: Identical initial consonants in '3' and '5' in Benue-Congo

| BC | Language | '3' - *taT | '5' - *tan |
|---|---|---|---|
| Bantoid | Tiv | -tar | -tan |
| Bantoid | Mambila | tar | tin |
| Bamileke | Bamun | i-tet | i-ten |
| Chamba | Chamba | tera | tuna |
| Daka | Dirrim | tara | tona |
| Daka | Gandole | tara | tuna |
| Bamileke | Kom | tal | tain |
| Beboid | Dumbo | te | ten |
| Grassfieldss | Mmen | ta | taiŋ |
| Jarawan | Jarawa | tat | towun |
| Nkambe | Mbe' | tei | tan |
| Idomoid | Gade | i-ta | i-to |
| Jukun | Proto-Jukunoid | *tat (i-) | *ton (i-) |
| Ikaan | Ikaan | tas | ton |
| Lower-Cross | Anaang | i-ta | i-tien |
| Upper-Cross | Olulumo | e-tal | e-tan |
| Kainji | Amo | n-tat | n-taun |
| Platoid | Horom | tat | ton |
| Ekoid | Nkem | i-ra | i-ron |
| Jarawan | Mboa | sai | sian |
| Edoid | Proto-Edoid | *i-caGi[1] | *i-ciNeni |
| Edoid | Ukue | e-rha | i-rhini |
| Edoid | Okpamheri | esa | iseni |
| Idomoid | Eloyi | e-la | e-lo |
| Jukun | Wapan | cara | cwana |
| Jukun | Jukun Jibu | sara | sona |
| Upper-Cross | Korop | bu-nan | bu-neg |
| Upper-Cross | Kiong | o-nan | o-nen |
| Platoid | Irigwe | ciæ | co |
| Platoid | Morwa | sat | suon |

Table 4.21: Identical initial consonants in '3' and '5' in Niger-Congo

| Family | Language | '3' | '5' |
|---|---|---|---|
| Kwa | Ewe | eto | ato |
| Kwa | Fon-Gbe | a-to | a-to, *ta |
| Kwa | Fon | a-tɔn | a-tɔɔ́n |
| Kwa | Tuwuli | ɛ-lalɛ | e-lo |
| Kwa | Kebu | ta | to |
| Kwa | Igo (Ahlon) | ita | uto |
| Adamawa-Bua | Gula | tar | tiŋ |
| Adamawa-Bua | Bolgo | teri | tiso |
| Adamawa-Bua | Koke | teri | tiso |
| Adamawa-Mbum | Mambai | bi-saa | bi-sape'e |
| Ijo | Defaka | tato | tuno |
| Mel | Bom | tat | tan |

Table 4.22: *tan '5' in Niger-Congo

| Atlantic | Mande | Dogon / Gur | Ubangi | Kordofan / Adamawa |
|---|---|---|---|---|
| Mel | Kru | Kwa | Ijo | BC |
| | | | | dinin/ dulin? |
| tɔk, tən? | **tan? (> '10'?) | tɔ | | sa? |
| kə-ʈamaʈ (<*kə-ʈa 'hand'?), tan? | | ton | túnɔ́ | tan/ ton |

more systematic. In Western NC the root is reliably attested as well, despite the fact that the Northern Mel form *kə-ʈamaʈ* allows a two-fold interpretation (i.e. as a derivative of either *ʈam-* or *\*kə-ʈa* 'hand').

The Bom form is a direct reflex of *tan* 'five'. It bears reminding that the final velar in the Northern-Atlantic forms is regular. In the Gur languages, the pertinent form is attested in particular branches only. As attested in Western Mande, the form implies a semantic innovation, i.e. *\*'5' > '10'*. The relationship of the Kordofanian forms is not immediately apparent.

The distribution of the alternative reconstructible root *\*nu/ nun* is described in the chart below (Table 4.23):

Table 4.23: *\*nun* '5' in Niger-Congo

| | | núnέέ(n)/ nŭ:(yn) / nûm | | |
|---|---|---|---|---|
| | | nu(n) | | nu(n) |
| | mm | nu(n) | | |

A comparison to Kru implies the labialization of dentals in the vicinity of a back vowel. As the Dogon and Gur evidence suggests, the root is possibly derived from the term for 'hand'. In Dogon the forms of 'five' and 'hand' differ in all languages / sources. Interestingly, the term that means 'five' in one Dogon language may be used with the meaning 'hand' in another (and vice versa, see **HochstetlerEtAl2004**, cf. the following evidence (Table 4.24):

Table 4.24: 'Hand' and '5' in Dogon

| Group | Language | 'hand' | '5' |
|---|---|---|---|
| Central | Tommo So | numɔ | nʔnɔ |
| Central | Donno So | numɔ | nɔʔ |
| Northern | Dogulu Dom | numɔ | nnɔ |
| South-East | Jamsay | numɔ | nui |
| Central | Toro So | nonnɔn | numonron |
| Central | Kolum So | nuwɛn | numu |

In light of this, the fact that, according to some sources, similar distribution of

the same root is attested in a number of Gur languages is intriguing, cf. e.g. the following data (Table 4.25):

Table 4.25: 'Hand' and potential reflexes of *nun* '5' in Gur

| Group | Source | Language | **'hand'** | 5 |
|---|---|---|---|---|
| Bariba | Koelle 1963[1854] | Baatonum | nóma | nǫ́wu |
| Bwamu | BloemartsdeRasilly2012 | | núumánnu | |
| Grusi | Koelle 1963[1854] | Tem | | nǫ́nūa |
| Grusi | CLNK1999 | Kabiye | | naanʋwa |
| Grusi | Koelle 1963[1854] | Kiamba | noon / noozi | noonuua |
| Grusi | Koelle 1963[1854] | Sisaala Tumulung | | ńnō̱m |
| Oti-Volta | Koelle 1963[1854] | Mosi | nurǫ | nu |
| Oti-Volta | Koelle 1963[1854] | Gurma | unu / inui | mu ~ mmu |

This raises the question, are we dealing with direct Dogon-Gur contact or with the reflexes of an additional NC root for 'hand'? The following roots may be considered potential correspondences: Proto-Bantu *nàmà* 'limb: arm; leg; thigh' (Regions 4: NW SW Ce NE ; Zones 6: ABEHMR) or *nǒe̩* 'finger, toe' (Regions 5: NW SW Ce NE SE ; Zones 9: ADJKLMPRS), (cf. Bantu, zones MN – Nyiha-Malila-Lambya (**NursePhilippson1975**) *i-nyove*, cf. (Koelle 1963[1854]) Aku (Defoid) *ɲɔwɔ* 'hand'. The Bak (Atlantic) root *ñen* 'hand', 'five' discussed above may belong here as well. The Gola root *nɔ̀ɔ̀nɔ̀ŋ* should also be mentioned here. The meaning 'hand' is not attested for this root in Kwa and Adamawa.

The following Atlantic roots attest to the semantic development of 'five' (and consequently 'ten') < 'hand' (Table 4.26).

This data is especially interesting in view of the BC evidence discussed above. As we have seen, the phenomenon of 'five' and 'ten' being based on the term for 'hand' is attested in both Atlantic groups (Bak and Northern). Moreover, this pattern is observable in a wide variety of roots with the meaning 'hand' attested in the languages under study (e.g. five roots with this meaning are attested in eight languages represented in the table above; the derivation pattern is the same

Table 4.26: 'Hand' > '5' in Atlantic

| Group | Language | 'hand' | '5' | '10' |
|---|---|---|---|---|
| Atlantic-Bak | Balant Kentohe | f-cef / k- | cef | f-cef meen ('whole hands') |
| Atlantic-Bak | Bijogo Kagbaaga | kɔ-ɔkɔ / ŋa-akɔ | nde-ɔkɔ | n-rua-kɔ |
| Atlantic-Bak | Bijogo Kamona | kɔ-kɔ / ŋa-kɔ | ŋu-duβ-kɔ | ŋɔ́-rúŋa-kɔ |
| Atlantic-Bak | Mankanya | ka-nyɛn | ka-nyɛɛn | e-nyɛn |
| Atlantic-Bak | Manjak | ka-ñen | ka-ñen | ka ñen |
| Atlantic-Bak | Pepel | ɲenɛ | ɲenɛ | dise-ɲenɛ |
| Atlantic-North | Nyun Djibonker | si-lax | ci-lax | haa-lax |
| Atlantic-North | Nyun Gujaxer | ci-lax / xa- | ci-lax | xa-lax |
| Atlantic-North | Biafada | gə-bəda / ma-bb- | gə-bəda | |
| Atlantic-North | Jaad | ko-bəda | ko-bəda | |

in each case). In view of this, it is not surprising that the reconstructed NC root is not traceable in Atlantic.

The same pattern is also attested in the Northern Mel languages (that are in contact with Bak) for 'five' (but not for 'ten'), cf. (Table 4.27).

Table 4.27: 'Hand' > '5' in Northern Mel

| Group | Source | Language | 'hand' | '5' |
|---|---|---|---|---|
| Temne-Baga-Landuma | **Wilson2007** | Baga Koba | kə-tsa/ɛ- | kə-tsa-mat |
| Temne-Baga-Landuma | **Ganong1998** | Baga Sitemu | kɛ-ca | kə-ca-mət |
| Temne-Baga-Landuma | **Wilson2007** | Landuma | kə-ca/cə- | kə-caa-mət |
| Temne-Baga-Landuma | **Wilson2007** | Temne | kə-ta/mə- | ta-math |

However, we may be dealing with the secondary alignment of the terms for 'hand' and 'five'. The pattern CV-stem-VC (with CV- and -VC being a noun class prefix and suffix respectively) is characteristic of this language group, e.g. the

Temne form may go back to *ta-m-ath* with the lexical root *-mV-* as its base. This pattern could also explain the similarity between the Temne terms for 'five' and 'ten': in this language *tɔfɔt* '10' probably goes back to *tɔ-f-ɔt* and hence to the NC root *fu* '10'.

Some of the Atlantic languages (e.g. various Joola and probably Proto-Joola as well) developed a separate root for 'five', while the term for 'ten' still remained a derivative of 'hand'. As expected, this root corresponds to Southern NC *tan/ ton* '5' discussed above (Proto-Atlantic: *tok* 'five': Kasanga-Kobiana *ju-roog*, Sereer *ɓe-tak / ɓe-tuk / ɓe-tik* (cf. also Limba *bi-sọhi*; Sua *sungun*), cf. Table 4.28.

Table 4.28: 'Hand' > '10' in Joola (Atlantic: Bak)

| Language | 'hand' | '5' | '10' |
|---|---|---|---|
| Joola_Banjal | ga-ɲen / gu-ɲen | fu-tox | gu-ɲen |
| Joola_Fogny | ka-ɲɛn / u-ɲɛn | fu-tɔk | u-ɲɛn |
| Joola_Gusilay | ga-ɲɛn / u-ɲɛn | fu-tɔk | u-ɲɛn |
| Joola_Kasa | ka-ŋɛn | hu-tɔk | ku-ŋɛn |
| Joola_Kasa_Esuulaalu | ka-ŋɛn | hu-tɔk | ku-ŋɛn |
| Keeraak | ka-ŋɛn-ak / ʊ-ŋɛn-aw | hʊ-tɔk | kʊ-ŋɛn |
| Joola_Kwaatay | ɛ-ŋɔmu | hu-tɔk | si-ŋɔmu |
| Joola_Kwaatay | ɛ-mɔŋo | hu-tɔk | su-muŋo |
| Joola_Mlomp | ɛ-bɛ:s | ŋa:-suwaŋ | sɛ-bɛ:s |

The etymological link between the terms for 'five' and 'ten' and their source ('hand') is not always explicit, e.g. different roots for 'hand' are attested in some of the sources for Mankanya-Manjak (Atlantic) and Temne (Mel), along with the derived form for 'five'. Such innovations are quoted in bold in the table below (Table 4.29).

Some of the forms of the term for 'five' go back to the root *ko* in a number of the Ubangi languages (and possibly in some of the Mande languages as well, see Chapter 3 for details). Here we may be dealing with a NC root, cf. e.g. 'hand': Proto-Gbaya *kɔ́*, Proto-South Mande *kɔ̃*, Proto-Eastern Mande *gɔn* (?), Dida (Kru) *kɔ̃*, etc.

The following Kordofanian terms that attest to the development of 'hand' > '5' are also noteworthy: Dagik (Kordofanian) *si-s-ɜlːʊ* '5' (lit: 'one hand'): "The *si* in 5 comes from the word 'hand'. So 5 is 'one hand'"[2], Acheron *zɜguŋ zulluk* (lit:

---

[2]John Vanderelst, https://mpi-lingweb.shh.mpg.de/numeral/Dagik.htm

Table 4.29: 'hand' > '5'/'10' in some Atlantic and Mel languages

| Branch | Language | 'hand' | '5' |
|---|---|---|---|
| Atl.-Centre-Manjak | Mankanya | ka-nyɛn | ka-nyɛɛn |
| Atl.-Centre-Manjak | Manjak | ka-ñen | ka-ñen |
| Atl.-Centre-Manjak | Manjak | **kádṣāg** | kányan |
| Atl.-Centre-Manjak | Mankanya | **úlōl** | kányēn |
| Atl.-Centre-Manjak | Manjak Bassarel | **pëndänd** | kañan |
| Atl.-Centre-Manjak | Manjak Tame | **wűepalōl,** | ké̲nyān |
| | | **pl. n·gípalōl** | |
| Temne-Baga-Landuma | Temne | kə-ta / mə- | ta-maṭ |
| Temne-Baga-Landuma | Temne | **a-loṅk (i), ma-** | ṭamaṭ |
| Temne-Baga-Landuma | Temne | **à.loŋk** | -tàmath |

'one hand' ): "The number 'five' is literally 'one hand': *zəguŋ* = 'hand', *z-ulluk* = 'one'"[3].

To summarize, the primary root for 'five' (\**tan*) probably existed in Proto-NC. Over time it was independently replaced with the derivatives of 'hand' in some branches and various languages. In turn, the original term for 'hand' was replaced with innovations (with the term for 'five' in particular) in a number of languages, cf. Atlantic *rib/ ʔi:p*, Mel *wan/wen*, Mande *dúuru/ sɔ́ɔru*, Kru *gbə / gbo*, Gur *mwan/ bwa*, Ubangi *du(w)/ lu(w)*, Kordofanian *ŋer-/ ɲer-*. As a rule, these innovations (not quoted here exhaustively) are only attested in particular branches of the families under study.

---

[3]Russell Norton, https://mpi-lingweb.shh.mpg.de/numeral/Acheron.htm

## 4.6 'Six'

The explicit pattern '6=5+1' is present in the vast majority of the families. Primary terms for 'six' are attested in some of the NC families (or, more precisely, in their particular branches). However, they cannot be reconstructed at the NC level (see Chapter 3 for their detailed treatment). Selected forms of this kind include Atlantic *paag/paaj* ('7=6+1'), Kwa *golo / kolo, kua, ciɛ* ('7=6+1'), Adamawa *jup, gu*, Ubangi *zala/ zya*, Dogon *kuro/ kule*, Gur *do(b)*, Mande *t(s)um*? (the examples are quoted by family without further detail). The pattern '6=3 redupl.' is rarely attested. It is found in BC (possibly as a Proto-BC innovation attested in Bantoid, Cross, Edoid, Kainji?, and Platoid) and Kordofanian only.

## 4.7 'Seven'

The main pattern is '7=5+2' (or '7=X+2' if the term for 'five' is replaced with an innovation). Primary roots are rare, being attested in BC (Defoid *\*byē* (cf. Edoid *ghie?*), Idomoid *renyi* (cf., however, Ikaan *h-ránèʃi* ('6+1')), Adamawa (*bir/ bil, rɪŋ, nbutu*), Ubangi (*sílànā, lè-rɵzi*), Dogon (*suli/ soli/ soye*), Gur (*pɛ(n)*)) and Atlantic Bak (*jand/ jaan?/ cand* (Pepel)).

The rare patterns of '7=6+1' and '7=4+3' are limited to Atlantic Bak, Kwa, BC Platoid, and Kordofanian.

## 4.8  'Eight' ('Four' and 'eight')

In the majority of the NC families the term for 'eight' is historically based on the term for 'four' (with the exception of Mel, Kru, Dogon, Mande and Western NC isolates).

The pattern '8=4+4' is normally implemented via the reduplication of the root for '4'. In some cases an 'entire' reduplication (affecting the conjunction and the noun class marker) is employed (Table 4.30).

The reduplication can also be 'partial' (as a rule the reduction of the first syllable is involved), cf. Table 4.31.

This pattern can also be used when the original root for 'four' is replaced by another one, cf. the Balant (Bak) evidence: *tahla* '4' ~ *ta-ta(h)la* '8'. The same is observable in Yungur (and possibly in Burak (Adamawa)), cf. *net* '4' ~ *nat-at* '8' (Boyd 1989).

Sometimes 'eight' is derived from 'four' not via the reduplication, but by means of a simple replacement of CL.SG with CL.PL (or by adding the Pl. marker), cf. Table 4.32.

In Dii (Adamawa-Duru) a step-by-step replacement of classes is used as a derivation mechanism, i.e. '2' > '4' > '8': *i-dú* '2' > *nda-ddɨ́* '4' > *ka-ʔa-nda-ddɨ́* '8'.

A rare pattern is '8=4*2', with the direct involvement of the term for 'two', cf. Viemo (Gur) *jumĩ* '4', *niinĩ* '2', *jumĩ-jɔ niinĩ* '8'.

When considering the reconstruction of 'four', it should be noted that if the term for 'four' (on which a reduplicated term for 'eight' is based) has any vowel other than [a] (typically [e] or [i]), the reduplicated form either preserves the vowel present in 'four' or has [a] in the first syllable. This mechanism is confirmed at least in the case of Bantu (Table 4.33).

The latter fact leads to at least two conclusions: 1) the reduplication mechanism was used to derive 'eight' from 'four' at the Proto-Bantu level; 2) [a] that which is preserved in 'eight' should be reconstructed in the first syllable of 'four', where it was lost.

Moreover, there is a considerable body of Bantu examples of a Proto-Bantu root being preserved in the reduplicated term for 'eight', but lost in the term for 'four' (Table 4.34).

One of the factors that could explain the emergence of the second nasal in the term for 'four' is the alignment of 'four' and 'eight' by analogy, followed either by the replacement of the term for 'eight' with a composite term ('5+3' or '10-2', see Table 4.35) or with an innovation (Table 4.36):

The evidence presented above strongly suggests that the pattern '8=4 redupl.'

Table 4.30: '8' < '4+4' (entire reduplication)

| Branch | Languages | '4' | '8' |
| --- | --- | --- | --- |
| Bantoid-Ekoid | Ekoi | ni | e-ni-ga-ni |
| Bantoid-Ekoid | Kwa | ni | a-ni-ka-ni |
| Bantoid-Ekoid | Ndoe | ne | be-ne be-ne |
| Bantoid-Ekoid | Nkem | ni | a-ni-gi-ni |
| Bantu-Central-E | Chaga | na | nana |
| Bantu-Central-E | Embu | nya | i-nyanya |
| Bantu-Central-E | Kamba | nya | nya-nya |
| Bantu-Central-E | Kikuyu | nya | i-nyanya |
| Bantu-Central-G | Sango | na | m-nana |
| BC-Edoid | Okpamheri | ni | e-ni-e-ni |
| BC-Edoid | Urhobo | ne | e-nene |
| Bantoid-Grass | viya | na | ge-nana |
| Bantoid-Jarawan | Mbula-Bwazza | i-ne | i-ne i-ne |
| Bantu-Central-D | Enya | na | ce-nana |
| Bantu-NW-B | kande | na | ge-nana |
| Bantu-NW-B | Lumbu | na | di-nana |
| Bantu-NW-B | Punu | na | i-nana |
| Bantu-NW-B | Sira | na | gi-nana |
| Bantu-Central-J | haya | na | omu-nana |
| Bantu-Central-J | Nyankole | na | om-nana |
| Bantu-Central-J | Nyoro | na | om-nana |
| Bantu-Central-J | Gwere | na | mu-nana |
| Bantu-Central-J | Nkore-Kiga | na | mu-nana |
| Bantu-Central-J | Soga | na | mu-nana |
| BC-Cross | Alege | ne | e-nene |
| BC-Cross | Bokyi | ɲe | ɲe-ri-ɲe |
| BC-Cross | Kukele | na | i-na-mi-na |
| BC-Bantoid | Esimbi | mō-ɲī | mō-ɲì-ō-ɲī |
| BC-Jukunoid | Mbembe | nyɛ | ɛ́-nyɛnyɛ |
| Bc-Ikaan | Ikaan | nāʲ / nā | nàːnáʲ / nàːná |
| Adamawa-Fali | Fali | náːn | nàn nán |
| Adamawa-Duru | Gǝunǝm | náárə́k | náárə́k àp náárə́k |
| Gur-Southern | Lamba | nasa | nasi-nasa |
| Gur-Southern | Lyele | na | nana |
| Laal | Laal | ɓīsān | ɓīsān.ɓīsān |

4 Reconstruction of Numerals in Niger-Congo

Table 4.31: '8' < '4+4' (partial reduplication)

| Branch | Language | '4' | '8' |
| --- | --- | --- | --- |
| Bantoid-Jarawan | Kulung | i-nin | i-ni-nin |
| Bantu-NW-B | Enenga | nai | e-na-nai |
| Bantu-NW-B | Myene | nayi | e-na-nayi |
| Bantu-NW-B | Orungu | nayi / i-nayi | e-na-nayi / na-nayi |
| BC-Eastern-Platoid | Boyawa | nas | na-nas |
| BC-Eastern-Platoid | Kwanka | nas | na-nas |
| BC-Eastern-Platoid | Idong | enar | na-nar |
| BC-Eastern-Platoid | Kadara | er-nar | ir-na-nar |
| Ijo | Nembe | i-nei | ni-nei |
| Atl-Centre | Balant | tahla- | ta-tahla- |
| Adamawa | Yungur | kurun | kun-kurun |

Table 4.32: '8' = 4PL

| Branch | Language | '4' | '8' |
| --- | --- | --- | --- |
| Kwa-Nyo | Lelemi | í-nɛ́ | máá-nɛ́ |
| Kordofanian Heiban | Warnang | ŋèlàmlàŋ | ŋelamlaaŋ-ɔ |
| BC Platoid | Ikulu | íń-nāā | nír̀-nāā |
| Adamawa Leko-Nimbari | Yendang | nâ:t | ɓɔ̄-lá-nā:t |
| Adamawa Mbum-Day | Niellim | ɲɛ̄ní | twāː-ɲɛ̄ní |
| Adamawa Waja-Jen | Waja | nʊ | wu-nii |
| Ubangi Sere-Ngbaka-Mba | Gbanzili | ɓɔ̄-nā | sá-nā |
| Gur Grusi | Delo | a-naara | gya-naara |
| Gur Grusi | Tampulma | a-naasi | ŋmɛ-naasa |

Table 4.33: ne/ *ni* '4' ~ *nane/ nani* '8' ( Bantu)

| Zone | Language | '4' | '8' |
|------|----------|-----|-----|
| Proto | PB | ne | nane |
| NW-B | Vove (Pove) | nai | nanai |
| NW-B | Sira | ne | gi-nane |
| NW-B | Punu | ne | yi-nane |
| NW-B | Lumbu | ne | nane |
| NW-C | Kela | nei | i-nane |
| NW-C | Kusu | nem | e-nanem |
| NW-C | Ombo | nei | i-nanei |
| Central-E | Pokomo | ne | nane |
| Central-E | Zanaki | i-nye | i-nyanye |
| Central-F | Bende | i-ne | mu-nane |
| Central-F | Kimbu | ji-ne | mu-nane |
| Central-F | Mbugwe (Irangi) | ne | i-nane |
| Central-F | Nyamwezi | ne | m-nane |
| Central-F | Sukuma | ne | nane |
| Central-F | Sumbwa | i-ne | m-nane |
| Central-G | Bondei | ne | nane |
| Central-G | CAsu (dial.) | ne | nane |
| Central-G | Kami | ne | nane |
| Central-G | Komoro | ne | nane |
| Central-G | Kutu | ne | nane |
| Central-G | Ngulu | ka-ne | m-nane |
| Central-G | Pangwa | i-ne | nane |
| Central-G | Shambala | ne | m-nane |
| Central-G | Swahili | ne | nane |
| Central-G | Tikuu | ne | nane |
| Central-G? E? | Tubeta (Taveta) | i-ne | nane |
| Central-G | Zigula | ne | m-nane |
| Central-J | Hunde | i-ne | mu-nane |
| Central-J | Konzo | ne | omu-nane |
| Central-J | Luhya | ne | mu-nane |
| Central-J | Masaba | ci-ne | si-nane |
| Central-J | Nande | ne | omu-nane |
| Central-J | Vinza | ka-ne | mu-nane |
| Central-M | Mambwe | vi-ni | ci-nani |
| Central-M | Pimbwe | i-ne | nane |
| Central-M | Rungu | vi-ni | ci-nani |

Table 4.34: '8' < '4' ~ '4' is lost (Bantu)

| Zone | Language | '4' | '8' |
|------|----------|-----|-----|
| Central-G | Mbugu | hahi | nane |
| Central-G | Bena | tayi | fi-mu-nana |
| Central-G | Hehe | tayi | i-mu-nana |
| Central-G | Ndamba | mceci | nani |
| Central-G | Pogoro | msesi | nani |
| Central-H | Kikongo | kuya | e-nana |
| Central-H | Yaka | ya | nana |
| Central-H | Yombe | ya | di-nana |
| Central-N | Manda | cece | nani |
| Central-N | Matengo | sesi | nani |
| Central-N | Mpoto | sesi | nani |
| Central-P | Matuumbi | sese | nani |
| Central-P | Ngindo | cece | nani |

Table 4.35: '8=4+4' > '8=5+3'

| Group | Language | '4' | '8 ' ('5+3') |
|-------|----------|-----|--------------|
| Atlantic | Baga Fore | si-neŋ / ci-neŋ | sak-tet |
| Atlantic | Baga Mboteni | i-neŋ | ib-ader |
| Atlantic | Wolof | ɲenet | jurom-ɲeta |
| Gur | Birifor (dial.) | anan | anu-ni-ata |
| Gur | Teen | nan | to sanr |
| Mande | Vai | nani | sog sakpa |
| Adamawa | Karang | niŋ | tòŋ ndɔ́k sé'de ('10-2') |

Table 4.36: '8=4+4' > '8' innovated

| Family | Languages | '4' | '8' |
|---|---|---|---|
| Bantu-A | Bafo | benin | wam |
| Bantu-A | Bankon | bi-nan | mwam |
| Bantu-A | Fang | ɲiɲ | mwom |
| Bantu-A | Ndambomo | li-naɲi | li-mwabi |
| Bantu-B | Kota | naɲi | mwabi |
| Bc-Platoid | Mabo | nen | hur |
| Dogon | Tene Kan | nani | sila |
| Dogon | Tene Kan | nani | sira |
| Kwa | Abron | nain | ŋocie |
| Kwa | Akan (Akuapem Twi) | anan | awotcye /tw/ |
| Kwa | Baule (Baoulé) | nan | nmocue |
| Kwa | Foodo | naŋ | dukwe / dukoi |
| Kwa | Mbato | ne-ni | o-gbi |
| Mande | Mandinka | náani | segi |
| Mande | Looma | náanĩ | dosawa |

was already in use at the Proto-NC level.

It should be noted that in those languages where this reduplication mechanism (or the pattern '8=4PL') is observable most clearly, another pattern is often used along with '8=4+4', namely '6=3+3' (or '6=3PL') (Table 4.37):

As expected, numerous languages that belong to different families exhibit a variety of patterns that are reused along with the one discussed above (including the general pattern '8=5+3' as well as '8=10-2' and even '8=6+2'). It seems, however, that such a wide distribution of this pattern ('8=4 redupl.') within the NC languages is genetic rather than typological.

Primary roots for 'eight' are also attested. However, their attestations are usually limited to one or two families or to particular branches within a family, cf. e.g. '8' in Defoid (BC) *\*jo/ ro* (cf. in Kainji *ro/ ru*), Kwa *kwe/ kye*, Kordofanian *bɔ, ʈəŋi-*, Mande *seki/ segi*, Dogon *sele/ sagi* (< Mande ?), *gá(a)rà*, Atlantic Bak *\*ʊʌs-*. These forms (as well as some additional ones) are interpreted as local innovations.

Table 4.37: '8' < '4', '6' < '3'

| Branch | Language | '3' | '6' | '4' | '8' |
|---|---|---|---|---|---|
| Bantoid-Ekoid | Ekoi | e-sa | e-sa-g-asa | e-ni | e-ni-ga-ni |
| Bantoid-Ekoid | Kwa | e-sa | a-sa-ka-su | i-ni | a-ni-ka-ni |
| Bantoid-Ekoid | Ndoe | be-ra | be-ra-ba-ra | be-ne | be-ne be-ne |
| Bantoid-Ekoid | Nkem | i-ra | i-ra-ra | i-ni | a-ni-gi-ni |
| Bantu-E | Embu | i-tatu | i-ta-tatu | i-nya | i-nya-nya |
| Bantu-E | Kamba | i-tatu | ta-tatu | i-nya | nya-nya |
| Bantu-E | Kikuyu | i-tatu | i-ta-tatu | i-nya | i-nya-nya |
| Bantu-F | Nyamwezi | datu | ta-dato | ne | m-na-ne |
| Bantu-F | Sukuma | datu | ta-datu | ne | na-ne |
| Bantu-G | Gogo | datu | m-ta-datu | ni | mu-na-ne |
| Bantu-<span style="color:red">G?E?</span> | Tubeta (Taveta) | tatu | ta-datu | i-ne | na-ne |
| Bantu-G | Zigula | ka-tatu | ta-datu | ne | m-na-ne |
| BC-Edo | Okpamheri | e-sa | e-sa-sa | e-ni | e-ni-e-ni |
| BC-Cross-River | Bokyi | bé-ciaat | ɲá-ciaat | bé-ɲįį | ɲí-rįį-ɲį |
| BC-Cross-River | Alege | é-cɛ | é-ce-e-ce | é-ne | ee-nɛ́-ne |

## 4.9 'Nine'

The main pattern for 'nine' ('9=5+4') is self-explanatory. This is the only pattern that can be reconstructed for Proto-Niger-Congo.

The alternative pattern '9=10-1' is much less common, whereas the pattern '9=6+3' (attested in Atlantic Bak) is exceptionally rare. The Platoid pattern '9=12-3' seems to be unique, cf. Birom, '15=12+3', '9=minus 3', '10=minus 2'. Primary roots are attested in those languages (branches) that have a full set of primary terms covering the sequence from 'one' to 'ten' (which is a rare case), e.g. Bantoid *bukV* (if indeed primary), Akpes *ɔ̀-kpɔ̄lɔ̂ʃ(ì)*, Defoid *\*sá(n), dà* (cf. Edoid *cien/sin*), Igboid *totu/tolu*, Ubangi *kùsì, me-newá*, Laal *yàŋjáŋ*, Dogon *túwɔ́*, Mande *kònonto/kɔ̀nɔndɔ(n)* (historically perhaps '10-1').

## 4.10 'Ten'

The root *pu/ fu is the most likely candidate for the NC reconstruction. The distribution of its reflexes is shown in the chart below (Table 4.38).

Table 4.38: *pu/ fu '10' in Niger-Congo

| Atlantic | Mande | Dogon | Ubangi | Kordofan |
|---|---|---|---|---|
| | | Gur | | Adamawa |
| Mel | Kru | Kwa | Ijo | BC |
| | | | | |
| pok | pu/ fu | fu/ po | ɓú / fu? | boo/ fu? |
| pu/ tɔ-f-ɔt ? | pu | fo / wo | | pu/ fu |

The roots listed in this chart are obviously related. The root is lacking in Kordofanian, where a variety of terms for ten are attested, e.g. *tu(l), rakpac, fəŋən, tiəɽum, 5PL*. This probably indicates that in Proto-Kordofanian the root for 'ten' was not present. The Dogon form *pérú/ pélú has the same initial consonant, but our evidence is inconclusive as to whether it is related to the roots above. Finally, the Ijo form *(w)ójí* allows a twofold interpretation. If it is taken as *(w)ó-jí* based on *ji, it is comparable to *zììyà* '10' attested in the Gola isolate. Alternatively, it can be analysed as a complex root *(w)o '10' plus *ji* (< *'1'). If so, it may be related to the roots quoted above (or at least to one of its allomorphs (?) attested in Kwa).

The presence of forms with the voiced **b-** in Adamawa-Ubangi requires an explanation. The evidence suggesting a connection between the **b-** and **f-** forms attested in these languages is insufficient. In view of this, it can only be noted that a similar phenomenon is observable within the Mande family: the form *bù is reconstructed in the Southern group of the South-Eastern Mande branch, whereas in Western Mande (as well as in the Eastern group of South-Eastern Mande) the reconstructed form is *pu/fu.

It should be noted that the Adamawa root with the initial voiceless labial is only marginally attested (e.g. in Munga (*fuə*) and Pere (*fób*)).

Raymond Boyd tentatively suggests that *fob* is to relatedhe tomain Adamawa root *kop*: «The Kutin group has *fóp* which may be related to *kóp»* (Boyd 1989:

162). However, an alternative explanation exists. A brief study of the Adamawa number systems shows that numerical terms attested within this family (unlike those found in other NC families) often end in **-p** or **-b**. The Tula system, one of the first quoted by Boyd in his excellent article, may serve as an example (Table 4.39).

Table 4.39: Labial suffix in Tula numerals

| 1 | -iǹ | 6 | nukuǹ |
|---|---|---|---|
| 2 | **rɔp** | 7 | nibiǹ |
| 3 | táa | 8 | **naˊá-rəp** |
| 4 | naa | 9 | **túrúkup** |
| 5 | nu | 10 | **kwɔp** |

The final **-p** in 'eight' is easily explainable (possibly due to '8=4*2). However, at least in the case of 'two' and 'ten', the final **-p** is attested in non-compound terms. In his discussion of the final **-p** in the Adamawa terms, Boyd suggested that we may be dealing with the suffix **\*-(a)p** (or **\*-(a)b**, with the devoicing characteristic of a reduced consonant inventory in the final position). < ...> The same suffix also appears in group 1 in *\*naar-ap* 'eight', derived from *\*naar* 'four'. < ...> Compare this situation with 'Bantoid' Vute: *'būrúp* 'two', *nà:sùp* 'four''* (Boyd 1989: 156). Furthermore, he challenges Kay Williamson's opinion on whether this morpheme was an original suffix or a suffix that developed out of a noun class prefix. The most important result of this discussion is that the suffix **\*-p/-b** found in numerical terms allows us to trace the Adamawa forms directly to NC *\*pu/po* without the intermediate *\*kop/kob*. As for the isolated Adamawa forms of *bo* 'ten', Boyd suggests a Chadic origin for them, although alternatively they may be related to the similar Ubangi root and reflect the NC root *\*pu / fu.*

The main Adamawa root *\*kop/kob* '10' should be discussed in a wider NC context as well. In view of the secondary nature of the final **-p/-b** in Adamawa (see above), this root is comparable to the NC roots *ko* 'ten ; hand'.

Direct BC parallels for this root (with the final labial) should be discussed first. We refer here to the hypothetical relationship of a number of forms discussed in Chapter 3, including Delta-Lower-Cross *-kɔp/ du-op/ du-ob* (**Dimmendaal1978** *\*lùgòp*) (cf. Bendi *kpu* '10', nearby *fo/ hwo*), Yukuben-Kuteb (Jukunoid) *kuwub*, Kainji *\*kop / ʔup / kpa* (together with *\*pwa/ pa*), and Platoid *\*kop*. This evidence suggests that more attention should be paid to the reconstruction of the allomorph *\*kop* in both Proto-BC and Proto-Adamawa. This root should probably be

compared to the Kru root *kʊgba* '10', unless it is a non-compound root that goes back to *ko* (see below).

In view of Boyd and Williamson's interpretation of the final labial as a suffix, the forms quoted above should probably be treated together with the root *ko* '10', which is sporadically attested in multiple families. As noted above, it most probably goes back to the lexical root *ko* 'hand', that represents one of the alternative Proto-NC reconstructions of this term. Its distribution with this meaning is as follows:

First of all, it is reconstructed by Moniño for Proto-Gbaya as *kɔ́* 'hand'. This root is also attested in Mande (at least in the Southern group of the South-Eastern Mande branch, cf. Vydrin's evidence: Proto-South-Eastern Mande *kɔ̃* 'hand, arm'). In Kru, this root is attested not only in the Eastern group (Dida *kɔ̄* 'hand'), but in the Western group as well (Glio-Oubi *hõ*, Krumen *hɔ̃’*). Finally, it is (admittedly only marginally) attested in Bantoid (as an alternative to the wide-spread root *kɔ́mì* '10'): according to Larry Hyman (in **Paulin1995**) this root is distinguishable in Kom (*ə̄-kæ̀*) and Narrow Bantu, e.g. in zones B (Mpur *kɔ*, Yansi *kwɔ*) and E (Mashami *oko*, Meru *uko*, **NursePhilippson1975**). The Limba root *koh-* '10' probably belongs here as well.

It is difficult to say whether this evidence is sufficient for the Proto-NC reconstruction. However, when choosing between the two possibilities for the reconstruction of the term for 'ten' (i.e. from **\*pu/ fu** and **\*ko**) the first one should be preferred.

Among other roots relevant to our discussion, the following two roots (whose attestations are not limited to one family) are of interest: Gur *gba/kpa* '10' (cf. the BC root *gwo/ jwo*) and Kwa *du* '10' (possibly related to the Adamawa root *d(u)o*; cf. also Kordofanian *ru* and Gur *nu/ nyu*?). The latter root may be compared to Bantu *\*dòngò* '10'. It is attested in seven zones (i.e. EGJMPR according to BLR3, but a number of attestations from D.62 are available, hence it is found in all five regions). BLR tentatively suggests a Bantu etymology for this root ('*spécialisation de "ligne" dòng?*'). However, it has parallels in other BC branches, namely in Cross River (**Connell1991**) and probably Idomoid (Table 4.40).

The use of numerous other roots for 'ten' is limited to one family, i.e. they are apparent innovations, such as in Bantoid *kum/kam* '10' (Bantu *kɔ́mì/ kámá*). The latter form (that sometimes coincides with the term for 'hundred') has an internal Bantu etymology: its tentative relationship to the lexical root meaning 'touch' is assumed in BLR 3 (BLR3: 'see also *kɔ́m* 'touch' - zones DHJLM'). However, the nasalization of the final segment in the Bantoid proto-form cannot be excluded. If this process indeed took place, this form becomes comparable to *\*ku(b)* as well

Table 4.40: Parallels for Bantu *dòngò* '10' in Cross River and Idomoid

| Branch | Language | Form |
|---|---|---|
| Cross River | Ebughu | lùgò |
| Cross River | Efai | dùgù |
| Cross River | Ekit | dùgò |
| Cross River | Enwang | lùgù |
| Cross River | Etebi | dùgù |
| Cross River | Ilue | lògù |
| Cross River | Okobo | lùgù |
| Cross River | Oro | lùwù |
| Cross River | Uda | lùgù |
| Idomoid | Eloyi | dǭn· & ndǫ́n· (Koelle 1963[1854]) |

as others discussed above.

Other isolated forms for 'ten' include Atlantic *(n)taaj, taim, -suwan*, Mel *wɨ-tʃɔʔ*, Western Mande *tan* (< *'5'?), Gur *kɛ(n)*, Kwa *bula* (cf. Ubangi *bale*), Ubangi *busa, sui*, Kordofanian *tu(l), di, rakpac, fəŋən, tiəɽum*, Adamawa *kutu(n)* (< *kutu(n)*, cf. Laal *tūū*, Kordofanian *ʈʌʌ*, Sua *tɛŋi* etc.

## 4.11 Large numbers ('twenty', 'hundred' and 'thousand')

It is better to treat large numbers together for the following reasons:

First, these terms were probably lacking in Niger-Congo, so it comes as no surprise that they are often borrowed from European languages, Arabic, Hausa, Lingala or other "languages of influence".

Secondly, these roots are often identical, i.e. the root that means 'thousand' in one language may mean 'hundred' or even 'ten' in another. Some of the forms simply denote 'a large number'. The well-known migrating root *keme* that has the meaning 'hundred' in the majority of the Mande languages may be used with the meaning 'eighty' or even 'sixty' in other Mande languages.

However, each of the roots has its own characteristics.

In the majority of the NC languages, the term for 'twenty' goes back to lexical roots that mean 'person', 'leader', 'body', 'head', 'grain', 'sack' and 'large number'. Numerous examples of this kind are discussed in Chapter 3. The etymology of those terms for 'twenty' that seem to be primary at the synchronic level should be sought with this in mind.

It can be safely stated that the terms for 'hundred and 'thousand' were absent in Proto-Niger-Congo. Thus, the pattern 'twenty' = 'person' remains the only reconstruction possibility for large numbers in Proto-Niger-Congo.

## 4.12 Proto-Niger-Congo

The reconstruction of the Proto-Niger-Congo number system may be summarized as follows (Table 4.41):

Table 4.41: Proto-Niger-Congo numeral system

| | | | |
|---|---|---|---|
| 1 | ku-(n)-di (> ni/ -in), do, gbo/ kpo | 7 | 5+2 |
| 2 | ba-di | 8 | na(i)nai (< 4 redupl.) |
| 3 | tat/ tath | 9 | 5+4 |
| 4 | na(h)i | 10 | pu/fu, |
| 5 | tan, nu(n) | 20 | < 'person' |
| 6 | 5+1 | | |

This table summarizes our discussion. However, it is tempting to apply our conclusions to the evidence pertaining to particular families in order to identify the most archaic families, groups and branches within NC. Such a review of data within a wider NC context could also help, enhancing the intermediate reconstructions suggested in Chapter 3.

# 5 NC numbers as reflected in particular families, groups and branches

No new reconstructions are presented in this chapter that offer the alignment of intermediate reconstructions on the basis of wider Niger-Congo evidence and conclusions based on the reconstruction suggested earlier. Hopefully, these results will enable an evaluation of each of the families (or a group/branch when possible) with regard to the inventory of NC roots preserved in them. In addition, this may enhance our understanding of the NC linguistic taxonomy. We will begin our analysis with the Benue-Congo evidence (Table 5.1).

## 5.1 Benue-Congo

Table 5.1: NC numerals reflected in Benue-Congo (+)

|  | 1 | 2 | 3 | 4 | 5 | 8 | 10 | Total |
|---|---|---|---|---|---|---|---|---|
| Nupoid | + | ba | + | + | + | 5+3 | + ? | 4 |
| Defoid | + | + | + | + | + | jo/ ro | gwá | 5 |
| Edoid | kpa/ gwo | va | + | + | + ? | + | gbe | 4 |
| Igboid | tù? | bɔ́ | + | + | + ? | 5+3 | ɗì/ ri/ li | 3 |
| Idomoid | + | pa | + | + | do/ lo/ ro/ ho | 5+3 | gwo | 3 |
| Kainji | + | + ? | + | + | + | ro, 5+3 | + | 6 |
| Platoid | + | + | + | + | + | + | kop | 6 |
| Cross | + | + ? | + | + | + | + | + ? | 7 |
| Jukunoid | jun, ʃíʃe, tə́ŋ | pa(n) | + | + | + | + | + ? | 5 |
| Bantoid (-B) | + ? | pa/ba/fe | + | + | + | + | + | 6 |
| Bantu | + | + | + | + | + | + | kʊ́mì | 6 |
| Oko | -ɔ́rɛ, -jɛ́rɛ | -bɔ̀rɛ̀ | + | + | -pi | + | + | 4 |
| Akpes | + ? | + | + | + | + ? | + | -yōf(ì), *t-ēfī | 6 |
| Ikaan | ʃí | wa | + | + | + | + | + | 5 |
| Lufu | + ? | máhà | + | + | + | 5+3 | + ? | 5 |

Commentary:

- Reflexes of the reconstructed NC forms are marked with /+/ in the table above.

- It should be repeatedly stressed that some of the etymologies accepted here are in need of further investigation and evaluation by experts. In case it is not clear whether the form is indeed a NC reflex, /+-?/ is used henceforward.

- Since the Bantu evidence is of great importance to our reconstruction, it is

treated separately, i.e. the Bantoid (-B) section only includes forms attested in these languages except for those found in Bantu.

- The terms for 'six', 'seven' and 'twenty' are not present in the tables. The assumed NC patterns that are employed for them are typologically widespread, which means that the evidence pertaining to their reflexes will only mar the overall distribution picture.

- If a reflex is supposedly lacking, a selection of basic forms (interpreted as innovations) is provided.

- The total number of Proto-Niger-Congo roots that have reflexes in each of the BC branches (out of the seven numbers represented in the table) is quoted in the rightmost column.

Table 5.1 demonstrates the following: If we accept this reconstruction, it appears that in only Cross-River do all seven terms discussed above directly reflect their NC prototypes, which makes this branch the most archaic within BC. Six terms out of seven represent NC reflexes in Kainji, Platoid, Bantoid, Bantu and Akpes. In other words, the Proto-NC numerical terms are better preserved in Eastern BC than they are in Western BC. It should be noted that only three terms out of seven have their reflexes in Idomoid and Igboid, i.e. they are the most distant from Proto-Niger-Congo among the languages under study.

Reflexes of 'three' and 'four' have been preserved in all BC branches. The reflection of 'five' is consistent as well. The same can be applied to 'eight' (the replacement of the pattern '8' = '4 redupl.' with '8' = '5+3' may have occurred independently in some of the branches).

Why the assumed reflexes of the Proto-terms for 'two' and 'ten' underwent a massive replacement is more difficult to explain. In the case of 'ten' a Proto-Western-BC innovation may be assumed, i.e. the replacement of *pu/ fu with *gbV/gwV. This is applicable to the Nupoid form *wo* (represented as /+?/ in the table above) as it probably reflects the Western innovation *gwo rather than *pu/fu. This raises doubts as to whether our interpretation of the forms attested in Cross (*kpo), Jukunoid (*wo*) and Lufu (*wo*) is correct (these forms were explained above as NC).

The reflexes of the Proto-NC term for 'two' are limited to 4-6 branches (out of the fifteen branches under study). At the same time, the forms that do not go back to *di are phonetically quite homogeneous in both main groups of BC (*pa/ ba/ wa/ va*). This suggests that the by-form of 'two' with the initial labial may have already existed at the Proto-BC level.

## 5.2 Kwa

Table 5.2: NC numerals reflected in Kwa (+)

|  |  | 1 | 2 | 3 | 4 | 5 | 8 | 10 | Total |
|---|---|---|---|---|---|---|---|---|---|
| 1. | Ga-Dangme | -kē, *go/ wo | - ɲɔ̀(n) | + | -ɟwὲ | -nùɔ̃ | 6+2 | ɲɔ̀ŋmaˊ | 1 |
| 2. | Gbe | + | -wὲ | + | + | + | -ɲí, 'hand'+3 | + | 5 |
| 3. | Ka-Togo | + | + | + | + | + | + | + | 7 |
| 4. | Na-Togo | + | -nyɔ | + | + | -no(N) | + | + | 5 |
| 5.1. | Nyo-Agneby | + | -ɲʊ̯ | + | + | -ne | -pyὲ, wo(n) | diw, 5PL | 3 |
| 5.2. | Nyo-Attié | kə(n) | mwə(n) | + | dʒí(n) | bə(n) | + | kεŋ | 2 |
| 5.3. | Nyo-Awikam | -tɔ̰́ | -ɲɔ̰́ | + | + | -ɲú | -tyέ | -jú | 2 |
| 5.4. | Nyo-Alladian | -tò̰ | -yrὲ | + ? | -zɔ̀ | -nrì | -ɰrì | + ? | 2 |
| 5.5.1. | Nyo-Potou | *ce, bὲ̰ | -noɔ̰́ | ja/je | + | na̰ | ɓyá/ gɓī | + | 2 |
| 5.5.2. | Nyo-Tano | ko(n) | -ɲɔ/ - ɲu(n) | + | + | nu(n) | - kwέ/ - cué | bulu, du | 2 |

Interestingly, Table 5.2 shows that some of the Kwa branches are exceptionally variable with regard to the reflection of Proto-NC terms. All seven Proto-terms under study have their reflexes in Ka-Togo, i.e. the Ka-Togo reconstruction is virtually identical to that of NC. At the same time, Gan-Dangme has only the reflex of 'three' (assuming that -tὲ̃ '3' reflects NC *tath*). In Nyo, the majority of terms are replaced as well: it seems that only the terms for 'three' and 'four' have been preserved in Proto-Nyo, whereas the preservation of 'ten' (not speaking of 'one' and 'eight', let alone the terms for 'two' and 'five', since the reflexes of *di* '2' and *tan* '5' are not traceable in any of the Nyo branches) is questionable. This means (assuming Ka-Togo, Na-Togo and Gbe indeed belong to Kwa) we should assume that: 1) the innovations presented in the table above postdate the division of Proto-Kwa; 2) Proto-Ka-Togo was the first language to separate from Kwa, since many of these innovations are homogeneous. This line of reasoning is more difficult to follow in the case of Na-Togo, since Na-Togo shares its innovations for 'two' (*nyɔ*) and 'five' (*nu*) with Nyo and Ga-Dangme. In other words, the Kwa numbers provide valuable data for the alignment of the internal genealogy

of the Kwa languages.

One important point that I would like to stress here is that if the Ka-Togo languages indeed belong to Kwa, we may state that our reconstruction of the NC number system is fully supported by the Kwa evidence.

It should be remarked that in a number of the Kwa branches the forms of 'five' interpreted as innovations in the table above could go back to an alternative NC prototype *$nu(n)$ '5' with its reflexes attested in Dogon, Gur and Adamawa.

Finally, I'd like to note that such a large-scale replacement of Proto-terms as in Nyo and Gan-Dangme (apparently etymologically related innovations) is a promising subject for both special investigation and discussion within the framework of a NC linguistics conference.

## 5.3 Ijo

Table 5.3: NC numerals reflected in Ijo (+)

|          | 1                  | 2     | 3 | 4 | 5 | 8   | 10        |
|----------|--------------------|-------|---|---|---|-----|-----------|
| Defaka   | gbérí              | mààmà | + | + | + | 5+3 | + ? (wóì) |
| East     | *+, gbérí, ŋ̀gèi   | màmì  | + | + | + | +   | ójí /àtìé |
| West     | *+, kènɪ           | maamʊ | + | + | + | +   | ójí       |

The Ijo languages are closely related, hence they do not differ much in the reflection of Proto-NC numbers. An apparent innovation of Ijo is the term for 'two' (mààmV). As for the term for 'one', the reflexes of the NC prototype are distinguishable in the Ijo compounds die/ zie/ ie. In the case of 'ten' it is, however, unclear whether this form is an innovation or not, since it can also be reconstructed as *wo-(i) based on *pu/ fu. The reconstruction *(w)oji < **ji is an alternative possibility that implies an innovation in Ijo.

In any case, the majority of the Proto-Ijo numbers can be traced to their NC prototypes.

## 5.4 Kru

Table 5.4: NC numerals reflected in Kru (+)

|  | 1 | 2 | 3 | 4 | 5 | 8 | 10 |
|---|---|---|---|---|---|---|---|
| Aizi | mu̲mɔ̲, yre | -ʃɩ | + | yeɓi | -gbo | patɛ | bɔ |
| Eastern | + | sɔ | + | + | gbu / gbi | 5+3 | +, kɔ́gba |
| Kuwa | + | sõr | + | + ? | wàyɔ̀ɔ | 5+3 | kowaa |
| Seme | dyuɔ̃ | nĩ | + | yur | kwɛ̃l | kprɛ̄n̂ | + |
| Western | + | sɔn | + | + | -mm | + | + |

The Proto-Niger-Congo forms are well-preserved in Western Kru (Bassa, Grebo, Klao, Wee). In other branches they are less well represented (especially in Aizi and Seme, where they are nearly completely replaced with innovations (except for the term for 'three') with reflexes attested in all the branches).

## 5.5 Kordofanian

Table 5.5: NC numerals reflected in Kordofanian (+)

|  | 1 | 2 | 3 | 4 | 5 | 8 | 10 |
|---|---|---|---|---|---|---|---|
| Heiban | -(ʈ)ʈɛ(k) | -can / -ʈan, rɔm | + | -ʈɔŋɔ/ -ɬʈʊ | -dìní, ɲer- | + ? | di/ dî/ ri |
| Katla | -ʈʌk | cik/ heek | + | -gʌlʌm | -duliin, -gbəlɪn | ʈʌ́ŋgìl | *ʈʌʌ, -rɔ |
| Rashad | -tta | (k)ko(k | + | -rʊm | *ɲer-, -ram | dubba | 5PL |
| Talodi | +? (lu(k)/ li(k)) | -ʈʌk/ -tta | + | -ʈandɔ, kekka | hand'- '1', -liəgum | + ? | -tu(l), tiəʈum |

This evidence leads to the conclusion that the number systems of the Kordofanian languages are hardly reconcilable with each other. Moreover, none of them seems to have inherited the NC system (with the exception of 'three' that apparenly goes back to its NC prototype, cf. e.g. Katla *ʌ̀-ʈʌ́ʈ̠* '3').

The NC root for 'eight' (< '4') is not represented in the Kordofanian languages. The use of /+?/ for Heiban and Talodi is only due to the fact that the Proto-NC pattern (8 = 4 redupl.) is traceable in them (rather than the form itself), cf. e.g. Warnang (Heiban) *ŋè-làmlàŋ* '4' > *ŋe-lamlaaŋ-ɔ* '8', Lumun (Talodi) *mɔ́ʲɔ̀ʈìn* '4' > *má-mɔ̀ɾmɔ̀ɾ* '8'. This resemblance, however, may be due to typological (rather than etymological) reasons.

## 5.6 Adamawa

It is important to note that Adamawa is one of the most divergent families within NC, hence the remarks below.

First, despite the diversity of forms, reflexes of the NC prototypes are well represented in many of the branches, e.g. five terms out of the total seven are probably reflected in Mbum Bua, Waja Jen, Waja Waja and Waja Yungur. Like in other families, the terms for 'three' and 'four' are the best-preserved.

The table above may create an impression that the term for 'one' is well-preserved in Adamawa as well. This impression is, however, misleading, since multiple forms are reconstructible for 'one'. Moreover, numerical terms attested in particular Adamawa branches go back to a variety of forms (rather than one particular form) that may be unrelated to each other. Thus NC *di* '1' finds parallels in the following branches: Duru *dɔ́ə*, Bua *\*lɛ* and possibly Laal *ɓɨ̀-dɨ́lʔ*. Its reconstructed allomorph *\*n-di* (with further evolution to*\*ni/ -in*) may be reflected in Kam *(-ii)*, Jen *-ín*, Waja *-in*, Mumuye ( ?) *-n*i, Yungur ( ?) *-ni*. The terms reflected in Falo *\*-lo*, Bua *dʊ(ŋ* and Kim *dú* may go back to the reconstructed NC form *\*do* '1'.

The forms observable in these two groups cannot be coalesced on the basis of the presently available evidence. Moreover, it bears reminding that the morphological analysis of the majority of the Adamawa numbers is uncertain. This problem cannot be solved at the moment since any firm criteria for distinguishing noun class affixes (or their traces) from the base are lacking.

The same is applied to the forms of 'two'. The set of reflexes for the NC term *\*di* '2' quoted in the table above is represented by the following isolated forms: Bua *di-di/ ri*, Kim *zí / tʃí-rí*, Day *dīí*, Jen *\*re / rá-b*, Waja *rɔ́-b*, Yungur *raa-p*. Regardless of whether the final -**b** goes back to a suffix or is the result of alignment by analogy (both possibilities are discussed above), it is clear that the relationship of these forms deserves careful examination in the diachronic perspective.

'Four'. This section of Table 5.6 is a result of our cautious treatment of the potentially related forms: the possibility that the forms of Kim-Day *nda* may go back to NC *\*na-* cannot be excluded.

The NC base *\*tan / ton* '5' has not been preserved in any of the Adamawa languages (apart from the doubtful Laal form). On the contrary, reflexes of the alternative NC form *\*nu(n)* are clearly distinguishable in the majority of the mid-range NC families such as Dogon, Gur and Kwa, so they should have probably been marked with the plus sign in the table above.

As for the reflexes of 'ten' (NC *\*pu/ fu*), it should be noted that all forms marked

Table 5.6: NC numerals reflected in Adamawa (+)

|  | 1 | 2 | 3 | 4 | 5 | 8 | 10 |
|---|---|---|---|---|---|---|---|
| Fali | + | gbara, cuk | + | + | kɛ̃rɛw | + | ra |
| Kam | + ? (-ii) | -raak | + | + | ŋwún | sâl | + ? |
| Leko Duru | +, ŋá | du/ ru, to, te/re | + | + | nún- | 5+3 | + ?, kob |
| Leko Leko | *ŋa | ra, in, nu | + | + | núún- | 5+3 | kob |
| Leko Mumuye | + ? | ye, ti, ni | + | + | nɔng | 5+3 | kob |
| Mbum Bua | + | + | + | + | *lu, tɛ, *kɔn, tiso | + | do, kùtù |
| Mbum Kim | + | + ? | + | ndà(y) | nūwē̞y | + | wàl |
| Mbum Mbum | bɔɔ̃ŋ/ búónó | ti, seɗe, gwa | + | + | ndiɓi | 10-2 | + ?, dùɔ, -wàl |
| Mbum Day | ngɔ̃ŋ̀, *mon | + | + | ndà, -yām | sɛ̄rì | + | + ? |
| Waja Jen | + | + ? | + | + | nóob/ *na, *hwĩ | + | ʃóób |
| Waja Longuda | khal, twɛ̀ | shir, kwɛ̃́ | + | + | nyɔ́ | nyíthìn | koo/ kù |
| Waja Waja | + | + ? | + | + | nu(ŋ) | + | kob |
| Waja Yungur | + | + | + | kurun | -nun | + | + ?, kutun |
| Laal | + | (ʔī-sī?) | māā | ɓī-sān | + ?? (saˉb, *swa-) | + | tūū |

with the plus sign in the table originally had a voiced labial as their initial consonant: Adamawa *\*buu/ buu.* The forms of Adamawa *\*ko-b* probably go back to NC *\*ko* 'hand'.

## 5.7 Ubangi

Table 5.7: NC numerals reflected in Ubangi (+)

|     |                          | 1              | 2             | 3 | 4    | 5             | 8       | 10          |
|-----|--------------------------|----------------|---------------|---|------|---------------|---------|-------------|
| 1.  | Banda                    | +              | -ʃi           | + | +    | -ndū          | 5+3     | +           |
| 2.  | Gbaya-Nanza-Ngbaka       | kpɔ́(k)/ ndáŋ   | wáʔ, -too     | + | +    | -(k)ɔ́         | +       | + ? (ɓú)    |
| 3.  | Ngbandi                  | kɔ(i)          | sɛ            | + | siɔ  | kɔ̃/ kū̱        | miambe  | sui, bàlé    |
| 4.1.| Ngbaka-Mba               | +, kpó-        | -ʃì/ -si      | + | +    | ve/ vue       | 5+3     | <'hand'     |
| 4.2.| Sere                     | njẽe           | so            | + | +    | vo            | 5+3     | <'hand'     |
| 5.  | Zande                    | +              | -jō/ -yō      | + | lu ? | -sìbē/ -sùè    | 5+3     | ŋgbɔ̃        |

Here, NC numbers are well-preserved in Banda and Gbaya-Nanza-Ngbaka (each of these branches has four reflexes out of seven) whereas in Ngbandi they have been totally replaced (except for *ta* '3').

The following problematic forms that have been taken as NC reflexes can be reinterpreted as follows (with due attention to their morphological structure and phonetics):

NC *\*di* '1': Banda *bà-lēʔ,* Ngbaka-Mba *ɓī-nì / bì-rì,* Zande *kí-lī*;

NC *\*pu/ fu* '10': Banda *bu-fu,* Gbaya *ɓú/ ɓù-kɔ̀*. Whether the latter form is indeed a NC reflex is not clear (not only due to its phonetics but also because a lexical etymology is suggested for *ɓù*), e.g. Edouard Koya states that *ɓu`* means 'person' in Bokoto (Central Gbaya-Manza-Ngbaka), where *ɓù-kɔ̀* '10' (https://mpi-lingweb.shh.mpg.de/numeral/Bokoto.htm). Moniño suggests an alternative etymology (Moñino 1995: 656): *« \*ɓú 'dix' est en relation avec \*ɓú 'façonner, faire un cercle, joindre les mains' ; la série partielle \*ɓú-kɔ́ 'joindre-mains' est encore plus explicite, et décrit le geste qui accompagne l'énonciation du chiffre 10 chez tous les locuteurs».* The following meanings of *ɓu´* in Gbaya are provided in (**BlanchardNoss1982**: 51):

*ɓú* «joindre les deux extrémités d'une même chose ; faire de la poterie»,

*ɓú* «dix, s'exprime en joignant les doigts de chaque main et en faisant toucher l'une de l'autre».

It is entirely possible that we are dealing with an innovation that follows the pattern described by Moniño. However, similar forms attested in other families may suggest that as finger counting developed, the secondary merger of homonyms occurred.

Finally, the Proto-Ubangi terms for 'two' (*\*se/ so*) and 'five' (*\*ko/ vo*, possibly a derivative from 'hand') should be mentioned as possible shared innovations.

## 5.8 Dogon

Table 5.8: NC numerals reflected in Dogon

|         | 1                      | 2    | 3   | 4   | 5      | 8            | 10           |
|---------|------------------------|------|-----|-----|--------|--------------|--------------|
| Dogon   | túrú/ tumɔ, ti(i)      | + ?  | +   | +   | núnɛ́ɛ́ | gá(a)rà      | pérú/ pɛ́lú   |
| Bangime | tòré / tĭyɛ́           | + ?  | +   | +   | nŭndí  | (borrowed)   | kúrɛ́        |

The Dogon numbers are quite homogeneous, so there is probably no need to treat them by branch. Instead, they will be compared to the numerical terms attested in the Bangime language that is considered a NC isolate.

**Dogon**. The forms *lɛ́(y)/ nɛ́(y)* (with their allomorphs *lɔ́(y)/ nɔ́(y)*) may be viewed as reflexes of NC *\*di* '2'. The reflex of NC *\*tan/ ton* '5' is lacking in Dogon, but the basic form quoted in the table above corresponds to the alternative NC root *\*nu(n)* widely attested in a number of NC families. The term for 'ten' can be compared to *\*pu/ fu*, but this comparison should be substantiated. As previously stated, the reflexes of 'three' (Dogon *\*taan*) and 'four' (Dogon *\*nay(n)*) appear to be the most consistent, which clearly identifies Dogon as a member of the NC family.

**Bangime**. The Bangime numbers are virtually identical to those of Dogon as far as their etymology is concerned. The form *jíndò* '2' may be a palatalized reflex of *\*di*. The term for 'eight' (*sàágín*) is a borrowing from Mande (just as in Dogon where a by-form of this primary term (*sagi*) is widely attested). The only Bangime term that is markedly different from the one found in Dogon is 'ten'.

## 5.9  Gur and Senufo

Evidence of the ten Gur branches is treated in Table 5.9 (cf. the discussion pertaining to the division of Gur into 16 branches in Chapter 3).

The Southern branch of Central Gur (Dogoso-Khe, Gan-Dogose, Grusi, Kirma-Tyrama) has preserved most of the NC terms (six out of the total seven), whereas its Northern branch (Bwamu, Kurumfe, Oti-Volta) preserved five. The NC numbers are well-represented in Teen and Wara-Natioro as well. Nearly the entire inventory of NC terms was replaced in Senufo (except for 'three' – Senufo *tã̀ã̀/ taàr), Bariba (except for *i-ta* 'three' and *ǹ-nɛ* 'four') and Kulango (except for *na* 'four and *tɔ* 'five'). At the same time, Kulango and Teen seem to be the only languages that have a reflex of NC *tan / ton* '5'.

As we have seen, the NC numbers are well-preserved in Gur, the more so that an alternative root for 'five' *(\*nu(n))* is distinguishable in at least four NC families. Its reflexes are attested in Bariba, Central, and Senufo. In view of this, it can be stated that all seven Proto-NC terms are reflected in Southern Central.

The term for 'one' is marked with the plus sign in reference to the reflexes of NC *do* (Central, Lobi-Dyan, Viemo) or NC *di* (Central, Tiefo).

Proto-Oti-Volta (Northern Central) *li/yi* and Proto-Grusi (Southern Central) *lɛ/ le* forms are considered to be reflexes of NC *di* '2'. Other forms of 'two' listed in the table represent a common (Proto-Gur ?) innovation *nyo/ jo /(ni* ?).

The Kulango term for 'three' (*sã̀ã̀be*) must be a borrowing from Mande.

The innovations for '4' are isolates that are irrelevant to the grouping of branches within the Gur family.

Some innovations for 'five' may go back to the lexical root for 'hand' (< *ko*).

The pattern for 'eight' (= '4 redupl.') is preserved in three of the branches.

In the case of 'ten', the similarity between the Senufo and Tiefo innovative forms is noteworthy.

Table 5.9: NC numerals reflected in Gur and Senufo (+)

|  |  | 1 | 2 | 3 | 4 | 5 | 8 | 10 |
|---|---|---|---|---|---|---|---|---|
| 1. | Bariba | tiā | Ru | + | + | nɔ̀ɔbù | 5+3 | -kuru |
| 2.1. | Central Northern | + | +, ɲū | + | + | nu | 5+3, tɔɔ, ni | + |
| 2.2. | Central Southern | + | +, ɲɔ̱/ jɔ̱ | + | + | nʊ(n) | + | + |
| 3. | Kulango | *ta̱à̱ | nyʊ̀ | (borrowed) | + | + | 5+3 | 5PL, *ji |
| 4. | Lobi-Dyan | + | nyɔ̀(n) | + | + | mɔ̀ì/ *mà, dìèmà | 5+3 | -kpo, nyɔ̀ɔ́r |
| 5. | Senufo | ŋgbe, nìkĩ̀ | sin /sun | + | tésyàr, … | -no | 5+3 | kɛ |
| 6. | Teen | tani | Nyor | + | + | + | 5+3 | + ? |
| 7. | Tiefo | + | jɔ̃ | + | ʔuʔɔ̃́ / ŋɔɔ | kã̀ | 5+3 | kɛ̃ |
| 8. | Tusia | nɔ́nkì | nĩ̃ŋ | + | + | k(w)lɔ́ | 5+3 | *gbɔ̃/ bwɔ̀ |
| 9. | Viemo | + ? | Niinĩ | + | jumĩ | *kɔ̱ | + | kwɔmũ |
| 10. | Wara-Natioro-Paleni | pɔ | nɩ́nté, bɔ̌ | + | + | sùsú, sV | + | + |

## 5.10 Mande

This is no doubt the most isolated family in what pertains to the reflection of NC numbers (Table 5.10). The maximum number of reflexes attested in particular branches does not exceed three (out of the total seven). In some of the branches, only two terms have been preserved. At the same time, the branches are quite compact, which enables us to discuss shared innovations within the Proto-Mande number system. The question as to whether these Proto-Mande innovations are of a lexical or morphological nature remains.

The most 'radical' etymological scenario is as follows:

The term *keden* '1' could be explained as going back to *\*ku-den*, which correlates well with the Proto-NC form *\*ku-di(n)* (with **ku**- being the most likely Proto-NC noun class prefix (class 1)).

The term *do* '1' is in line with the alternative NC root *\*do* '1' (without a noun class marker).

The Mande term *\*fida / fide* could be interpreted as going back to *\*fi-de* (assuming the first syllable reflects a noun class, e.g. CL 19).

The term for 'three' could be interpreted as a compound, one that has a reflex of *\*ta* '3' (< *\*tath*) as its first component (the second component remains unidentified).

The Mande term for 'ten' (*\*tan*) as found in Western Mande may be a reflex of the Proto-NC form *tan 'five' with a semantic shift *'5' > '5PL' (='10'). Moreover, its original form may have been preserved in Jowulu.

Any of these bold assumptions may prove true, but presently none of them is substantiated enough, so they are better left for future discussion in the hope that over time more pertinent evidence will become available. In this respect, the study of Samogo and Jowulu looks promising, the more so that the lack of an up-to-date linguistic investigation of these languages, as far as I know, has been a sore gap in present day comparative-historical studies of the Mande languages. In addition, these languages are the only ones that seem to preserve reflexes of both NC terms for 'five' (NC *tan/ ton* and *\*nu(n)*). Moreover, the Jowulu terms that have [p-] ~ [b-] allomorphs may reflect a noun class prefix (the choice between **p**- and **b**- depends on the following consonant, i.e. [p-] appears before a voiceless consonant (cf. *p-ʃɪrɛ* '4') whereas [b-] appears before a voiced consonant (*b-zei* '3', *b-ʒĩĩ* '10').

Table 5.10: NC numerals reflected in Mande (+)

| | 1 | 2 | 3 | 4 | 5 | 8 | 10 |
|---|---|---|---|---|---|---|---|
| Manding | +, *kélen | *fìlá | *sàbá | + | dúuru, *wo | séegi | + ?, *tán |
| Jogo-Jeri | +, *kɛlɛ ? | *fàlá | sègbá | + | sóólò | 5+3 | táà(n), ta |
| Mokole | +, *kéle | *fìla | saba | + | lɔɔlu, *wo | sɛ́ɛn/ seyi | + ? , tán |
| Vai-Kono | +, *N-kélen | *fèLá | sagba | + | dúʔu/ sóó(ʔ)ú | séin, 5+3 | tâŋ |
| Susu | *kédén, ǹdá/ nde | *fìdíń | sawa/ sàxán | + | sùlù, *fò | 5+3 | +, *tòngó |
| SWM | *gìláaŋ, *tà | *fèelé | sāaɓā | + | dɔ́ɔlú, *wɔ/ ngɔ̀ | 5+3 | + |
| Bozo-Soninke | ke̱/ kuɔn, sana̱, … | pẽ̀:ndé/ fíllò | sike/ síkkò | + | kɔ́lɔ́-/kárá- | segi | tan |
| Bobo | tàlá/ tèlé | pálà | sàà | + | kóò | sɛ́ki/ tʃèkí | +, m̥ḿ |
| Samogo | *ké, *so | fíː(kí) | ʒìːgī, ʃwɛ̀, ɣei | + | + | +, kàà | t(s)eu/ ceũ |
| Jowulu | tẽẽna/ tenn | fuuli | bʒei < *jɔnn/i ? | p-ʃɪrɛ | + | 2+'to lose' | bʒĩ / byìnn |
| SE East | +, gôon | *pela | cɔ́w, ʔàà-kɔ̃ | sì/ síirí | *sodu | + ?, 5+3, síɲe | +, kwi |
| SE South | + | *pìì-lāŋ | *yààká | *yìì-si̱i̱yá̱ | sɔ́ɔɗú | 5+3 | +, ko, sɔ́jɔlú |

## 5.11 Mel

Table 5.11: NC numerals reflected in Mel (+)

|        | 1   | 2         | 3 | 4           | 5                          | 8   | 10  |
|--------|-----|-----------|---|-------------|----------------------------|-----|-----|
| North  | +   | -rəŋ      | + | -ŋkɨlɛ/ -nlɛ | < 'hand' ? ~ (-mV- ?)      | 5+3 | +   |
| South  | + ? | tsiŋ / tiŋ | + | hiɔl        | wan/ wen                   | 5+3 | 5PL |

/~/ in the section dealing with the Northern Mel term for 'five' indicates that it allows for a two-fold morphological analysis, namely *kə-ṭa-maṭ* (< *\*kə-ṭa+suffix* < root *ṭa* 'hand'?) or (< *kə-ṭa-m-aṭ* < root *mV*).

The numeral system of the proto-language is generally poorly preserved in both of the Mel groups. At the same time, it should be noted that the most apparent innovations ('four' and 'two') are found in both groups, thus being important isoglosses useful to the assessment of Proto-Mel.

In the Northern group, as well as in a number of other NC families, the term for 'one' is reconstructible as CL-*in* '1' (< NC *\*n-di*). The forms reconstructed for the Southern group include *\*lɛ, \*lɔ* '1' (< *\*di, \*do*). Languages of the Northern group preserve the basic form of 'ten', cf. Landuma *pù* '10', Temne *tɔ-f-ʌt* '10'.

## 5.12 Atlantic

Table 5.12: NC numerals reflected in Atlantic

|  | 1 | 2 | 3 | 4 | 5 | 8 | 10 |
|---|---|---|---|---|---|---|---|
| Cangin | no | nak | haj | + | <'hand', ʔiːp | 5+3 | daːŋkah |
| Nyun | + ? | nak | + | + ? | <'hand' | 5+3 | <'hands' |
| Buy | nɔʔ, tee- | naŋ | + | + ? | + ? | 5+3 | ntaajã |
| Jaad-Biafada | + | ke | jo/ caw | + | <'hand' | 5+3 | + |
| Tenda | +, mbɔ | ki | + | + | <'hand' | 5+3 | + |
| Fula-sereer | + | + | + | + | + ?, jo(w)i | 5+3 | sapp-o, xarɓ- |
| Wolof | + | X-aar | + ? | + | jurom | 5+3 | + |
| Nalu-BF-BMb | +, mbɔ́ | + | + | + | <'hand', ribə(l) | 5+3 | *a-lafaŋ? |
| Joola | + | *-łubəʔ | -feegir | -bääkiɽ | + | 5+3 | -ntaaja |
| Manjak | + | -təb, puguɽ | -jenɽ | baakər | <'hand' | + | (n)taaja |
| Balant | -ɔdaʔ | *-łubəʔ | (borrowed) | tasala | jìíf | + | jímmín |
| Bijogo | +? | *-łubəʔ | <'fingers' | -agɛnɛk | <'hand' | 5+3 | <'hands' |

The Atlantic languages comprise two major groups, namely Northern and Bak (the members of the latter are highlighted in grey in the table above).

The Proto-NC numbers are generally better represented in Northern rather than in Bak (cf. the distribution of data pertaining to 'three', 'four' (generally the most persistent terms) and 'ten' in the table above). The only Northern sub-group where the Proto-NC numbers are poorly preserved is Cangin, while Fula-Sereer, Tenda, Wolof and Nalu are the most conservative.

The distribution of reflexes and innovations presented in the table above suggests the following historical development:

Reflexes of all major Proto-NC terms were present in Proto-Atlantic. The distribution of the terms for '1' may point to the existence of two dialect zones. A form

that goes back to NC *(n)-di* '1' became predominant in the ancestral dialect of Proto-Northern, whereas in the ancestral dialect of Proto-Bak the main form was NC *do* '1'. A specific phonetic (or morphological?) innovation of Proto-Atlantic (in contrast to NC) is the presence of the final **-k** in its numerical terms.

Proto-Northern inherited all basic Proto-Atlantic terms that go back to NC prototypes.

The term for '2' has been preserved in Peul-Sereer (**di-k* '2') and in Nalu (in all three languages). A (shared?) innovation developed in Cangin and Nyun-Buy (**na-k* '2'). Another innovation is characteristic of Tenda-Jaad-Biafada (**ki* '2').

The terms for 'three' and 'four' have been preserved in the majority of the Northern Atlantic languages (cf. e.g. Proto-Fula-Sereer **tati-k* '3', **na(y)i-k* '4').

The NC root **tak/ tok* '5' is probably reflected only in Fula-Sereer *(*ɓe-tV-k)* and Buy (*ju-roo-g*, cf. Wolof **ju-rom* ?). In the majority of the Northern languages the original form was replaced with the pattern '5' < 'hand', which may have influenced the replacement of the pattern *'8' = '4 redupl.' with '8' = '5' (hand) + 3.

The term for '10' has been preserved in three sub-groups (Wolof **fu-kk*, Tenda **pǝ-xw*, Jaad-Biafada **po*). In the remaining sub-groups it is replaced with isolated innovations.

The Proto-Bak numeral system underwent dramatic changes.

The original term for 'two' was replaced with the innovation *-ɫubǝʔ* '2', with its reflexes being traceable in three out of four sub-groups.

The reflexes of the Proto-NC terms for 'three' and 'four' are lacking. Moreover, a shared innovation *baakǝr* '4' is observable in Joola-Manjak.

The original term for 'five' has been preserved in numerous Joola dialects, including Bayot (Proto-Joola **fu-tɔ-k* '5').

The Proto-pattern '8' < '4' has been preserved in Manjak (Mankanya *ŋi-bakɨr* '4' > *bakɾ-êŋ* '8', Pepel *ŋ-uakr* '4' > *bakar-i* '8') and Balant (despite the fact that the original term for 'four' was replaced with an innovation in this language, cf. Balant Ganja *tàllá* '4' > *táhtállà ~ tántállà ~ táttállà* '8' as recorded by Denis Creissels).

The term for '10' was replaced with innovations. Here (just as in the case of '4') we have another shared Joola-Manjak innovation (*ntaaja*). This seems to be another solid argument in favor of grouping these languages together.

## 5.13 West African NC isolates

We will conclude with an overview of the number systems attested in three NC isolates. These languages are traditionally grouped together with Mel or Atlantic (for seemingly no substantial reason, see **PozdniakovSegerer2007**).

Table 5.13: NC numerals reflected in Sua (+)

| 1 | 2 | 3 | 4 | 5 | 8 | 10 |
|---|---|---|---|---|---|---|
| sɔn | cen | + | + | sɔŋgun | 5+3 | tɛŋi |

The reflexes of 'three' and 'four' have been preserved in Sua (*b-rar* and *b-nan* respectively). It should be noted that the innovation for 'two' is comparable to that found in Mel.

The term for 'ten' is possibly a borrowing from Mande *tan* '10'.

Table 5.14: NC numerals reflected in Gola (+)

| | 1 | 2 | 3 | 4 | 5 | 8 | 10 |
|---|---|---|---|---|---|---|---|
| Gola | guù`ŋ | tì-yèe(l) | + | + | + | 5+3 | zììyà |

The term for 'five' may reflect the alternative NC root *\*nu(n)* '5' (Gola *nɔ̀ɔ̀nɔ̀ŋ*).

Table 5.15: NC numerals reflected in Limba (+)

| | 1 | 2 | 3 | 4 | 5 | 8 | 10 |
|---|---|---|---|---|---|---|---|
| Limba | -nthe | + | + | + | bi-sɔhi | 5+3 | kɔhi |

The forms for 'five' and 'ten' in the Koelle records include [-f]: *ta-sóóf ~ ka-sóóf* '5', *koof* '10'.

The form *bi-le* 'two' is noteworthy in that it may be interpreted as a direct reflex of NC *\*be-di* '2'.

## 5.14 Summary

The results of our reconstruction of the basic numeral terms are presented in Table 5.16.

Table 5.16: Niger-Congo numerals reflected in various families (+)

| PROTO-NC | 1 *(n)-di | 1 *do | 2 *di | 3 *tath | 4 *na(h)i | 5 *tan | 5 *nu(n) | 8 < '4' | 10 *pu | Total |
|---|---|---|---|---|---|---|---|---|---|---|
| Bantu | + | | + | + | + | + | | + | | 6 |
| Bantoid (-Bantu) | + ? | | + ? | + | + | + | | + | + | 7 |
| BC (-Bantoid) | + | | + ? | + | + | + | | + | + | 7 |
| Kwa | + | | + ? | + | + | + | + | + | + | 8 |
| Ijo | + | | | + | + | + | | + | | 5 |
| Kru | | + | | + | + | | + ? | | + | 5 |
| Kordofanian | | | | + | | | | + | | 2 |
| Adamawa | + | + | + | + | + | | + | + | + ? | 8 |
| Laal | + | | | | | + ? | | + | | 3 |
| Ubangi | + | | | + | + | | | + ? | + ? | 5 |
| Dogon | | | + | + | + | | + | | | 4 |
| Gur | + | + | + | + | + | + ? | + | + ? | + | 9 |
| Mande | | + | | | + | + ? | +?? | | + | 5 |
| Mel | +? | +? | | + | | | | | + | 4 |
| Atlantic North | + | | + | + | + | + | | | + | 6 |
| Atlantic Bak | | + | | | | + | | + | | 3 |
| Sua | | | | + | + | + ? | | | | 3 |
| Gola | | | | + | + | | + | | | 3 |
| Limba | | | + | + | + | + ? | | | | 4 |
| Total | 11 | 6 | 9 | 16 | 15 | 12 | 7 | 11 | 9 | |

- The lack of a NC reflex in a particular family or branch is highlighted in grey.

- The number of Proto-NC terms (out of nine listed in the table) with reflexes in a particular family or branch is given in the rightmost column.

- The number of branches (out of the total nineteen) with a reflex of a particular proto-form is provided in the lowermost row.

Our step-by-step reconstruction has yielded the following results.

The terms for 'three' and 'four' (*tath* '3' and *na(h)i* '4' respectively) are, as expected, the most stable within the NC number system. Their reflexes are rarely absent.

Surprisingly, the term for '2' appears to be the least persistent (the more so that this is the only numerical term on the Swadesh list). The reconstructed root for 'two' (*di* '2') is traceable in nine (out of nineteen) branches only. This may raise doubts as to whether the proposed reconstruction is correct. However, as we have tried to demonstrate above, no alternative reconstruction suggests itself on the basis of available evidence. The term for '2' shows a great variety of forms, at the same time being surprisingly persistent in particular branches (and other times rather divergent). Thus, the apparent Mande innovation *pila/ fila* '2' is present in all Mande languages.

The most conservative NC branches in terms of the reflection of Proto-NC numbers are Gur, Adamawa and Kwa. All bases/patterns listed in the table have been preserved in Gur, including the alternative bases for 'one' and 'five'. The only reflex that is missing in Adamawa (as well as in Ubangi) is *tan / ton* '5'. All Proto-terms have their reflexes in Kwa (except for the alternative base for 'one', i.e. *do*).

The inventory of the Proto-NC terms is well-preserved in the Bantoid languages, with only two alternative bases lacking (*do* '1' и *nu(n)* '5'). These reflexes are missing in other BC branches outside the Bantoid languages as well. The reflex of *pu* '10' is not present in Bantu as it was replaced with the Bantoid innovation *kum/ kam/ ɣam* (Proto-Bantu *kŏmì/ kámá* '10').

It would seem improper to define the branches with the lowest number of NC reflexes as the most distant from Proto-NC. The probability of finding a reflex of a NC-prototype in an isolate (e.g. Gola or Laal) is much less than, say, in the huge Benue-Congo family. At the same time, the massive replacement of numerical terms in the small West African branches such as Bak (Atlantic), Mel and Dogon is noteworthy.

The Kordofanian languages are the most remote from Proto-NC, as the only term with a NC prototype attested in them is *tath* '3'. The term for '8' is based on '4', which may be seen as another bond between Kordofanian and Proto-NC. However, this pattern may have developed in Kordofanian independently.

## 5.15 Conclusion

In conclusion, I would like to highlight the thesis that I personally consider to be the most important. For me, the current study is an experimental project that aspires to demonstrate what can be done (if anything) in terms of the NC reconstruction, given that a step-by-step reconstruction is not available for all the families and branches of this macro-family.

In this experiment, the emphasis was placed on providing an exhaustive account of the distribution of forms by families, groups and branches. Quasi-reconstructions of Proto-NC numbers that resulted in the process should be viewed as mere possibilities. My intention was to present evidence that the reconstructions offered in this book are more probable than any others.

The author sees his major goal as providing a substantial discussion of the most likely reconstructions of Proto-NC numbers, in the hope that linguists specializing in particular NC families (as well as those who provide speculative 'etymologies') will finally join the debate. Chapter 3, which is the lengthiest and the most important chapter of the book, contains 'technical proposals' regarding the reconstruction of numbers within each of the numerous branches of the macro-family. I would like to thank the specialists who kindly joined the discussion while the book was still in preparation and whose opinions were duly accounted for. I would be grateful if other specialists critically examined the evidence presented in this book and gave their evaluation of data that lies within their competence. Hopefully, this will give way to the real reconstruction of the NC number system. Today it is evident that plausible reconstructions in terms of a macro-family that comprises one and a half thousand languages can only result from the cooperation of dozens of specialists. This book aims at providing data for such an effort.

I hope that the methodology tested in this book will be of use for the reconstruction of the NC lexicon in general. In any case, the author sees no other way of approaching this objective of utmost importance in the coming decades.

# References

Bendor-Samuel, John T. (ed.). 1989. *The Niger-Congo languages: A classification and description of africa's largest language family.* Lanham MD, New York & London: University Press of America. by arrangement with the Summer Institute of Linguistics (SIL).

Blench, Roger. 2013. *Niger-Congo: An alternative view.* http://www.rogerblench.info/Language/Niger-Congo/General/Niger-Congo%20an%20alternative%20view.pdf. Ms.

Boyd, Raymond. 1989. Number systems in the adamawa branch of Niger-Congo. *African Languages and Cultures* 2(2). 149–173.

Boyeldieu, Pascal. 1982. *Deux études laal (Moyen-Chari, tchad).* Berlin: Dietrich Reimer Verlag.

Carlson, Robert J. 1993. A sketch of jo: A mande language with a feminine pronoun. *Mandenkan* 25. 1–109.

Creissels, Denis & Séckou Biaye. 2015. *Le balant ganja: Phonologie, morphosyntaxe, liste lexicale, textes.* Dakar: IFAN.

d'Avezac, Armand. 1845. Notice sur le pays et le peuple des Yébous en Afrique. *Mémoires de la Société Ethnologique* 2. 1–196.

Djilla, Mama, Bart Eenkhoorn & Jacqueline Eenkhoorn-Pilon. 2004. *Phonologie du jôwulu ("samogho"): Langue mandé du Mali et du burkina faso.* Köln: Rüdiger Köppe Verlag.

Elders, Stefan. 2007. Kulango. In Gudrun Miehe, Brigitte Reineke & Kerstin Winkelmann (eds.), *Noun class systems in Gur languages. Southwestern Gur languages (without Gurunsi)*, vol. 2007, 292–330. Köln: Rüdiger Köppe.

Greenberg, Joseph Harold. 1966. *The languages of Africa.* 2nd edn. The Hague, Paris & Bloomington: Mouton & Co. for the Indiana University.

Hochstetler, Lee. 1996. Enquête linguistique sur le duungoma: Une langue samogo parlée au burkina faso et au mali. *Mandenkan* 31. 1–57.

Kaliai, M. H. I. 1964. *Nembe-English dictionary.* (R. Blench ed. 2008).

Kastenholz, Raimund & Ulrich Kleinewillinghöfer. 2012. *Nimbari as a language name.* https://www.blogs.uni-mainz.de/fb07-adamawa/files/2012/06/Nimbari.pdf.

# References

Koelle, Sigismund Wilhelm. 1963[1854]. *Polyglotta africana, or a comparative vocabulary of nearly three hundred words and phrases in more than one hundred distinct African languages.* London: Church Missionary House.

Miehe, Gudrun, Brigitte Reineke & Kerstin Winkelmann (eds.). 2007. *Noun class systems in Gur languages. Southwestern Gur languages (without Gurunsi).* Vol. I. Köln: Rüdiger Köppe.

Miehe, Gudrun & Florian Tham. 2007. Lobiri. I. 211–237.

Moñino, Yves. 1995. *Le proto-gbaya: Essai de linguistique comparative historique sur vingt-et-une langues d'afrique centrale.* Paris: Editions Peeters pour l'Société des Etudes Linguistiques et Anthropologiques de France (SELAF) avec le concours du Centre National de la Recherche Scientifique (CNRS).

Naden, Anthony Joshua. 1989. Gur. In John T. Bendor-Samuel (ed.), *The Niger-Congo languages: A classification and description of africa's largest language family*, 140–168. Lanham MD, New York & London: University Press of America. by arrangement with the Summer Institute of Linguistics (SIL).

Pozdniakov, Konstantin. 2010. *Laal: an isolate language?* Paper presented at the Workshop 'Language Isolates in Africa', Lyon, December 3-4, 2010.

Pozdniakov, Konstantin. 2015. Diachronie des classes nominales atlantiques: Morphophonologie, morphologie, sémantique. In Denis Creissels & Konstantin Pozdniakov (eds.), *Les classes nominales dans les langues atlantiques*, 54–102. Köln: Rüdiger Köppe Verlag.

Pozdniakov, Konstantin & Guillaume Segerer. 2017. Genealogical classification of Atlantic languages. In Friederike Lüpke (ed.), *Oxford guide to the world's languages: Atlantic.* Oxford: Oxford University Press.

Prost, André. 1958. Quelques notes sur le don (samogho). *Bulletin de l'Institut Fondamental d'Afrique Noire* 20-3/4. 612–623.

Prost, André. 1968. *Deux langues voltaïques en voie de disparition: Le wara et le natioro.* Dakar: Université de Dakar.

Sapir, J. David. 1971. West atlantic: An inventory of the languages, their noun class systems and consonant alternation. In Thomas A. Sebeok (ed.), *Linguistics in sub-Saharan Africa* (Current trends in linguistics 7), 45–112. The Hague & Paris: Mouton & Co.

Sawadogo, Tasséré. 2002. *Rapport d'enquête sur le natioro.* Electronic document.

Sebeok, Thomas A. (ed.). 1971. *Current trends in linguistics, 7: Linguistics in sub-Saharan africa.* The Hague & Paris: Mouton & Co.

Segerer, Guillaume. 2002. *La langue bijogo de bubaque (guinée bissau).* Louvain, Paris: Editions Peeters.

Vogler, Pierre. 2015. *Le sèmè / siamou n'est pas kru.* <hal-01182225>.

Vydrin, Valentin. 2007. South mande reconstruction: Initial consonants. In Aa Bb (ed.), *Aspekty komparativistiki-2. Orientalia et classica XI: Trudy instituta vostochnykh kul'tur i antichnosti*, 409–498. Moscow: Izdatel'stvo Rossijskogo Gosudarstvennogo Gumanitarnogo Universiteta.

Vydrin, Valentin. 2016. Toward a Proto-Mande reconstruction and an etymological dictionary. *Faits de langues* 48. 109–123.

Vydrin, Valentin & Elena Perekhvalskaya. 2015. Sistemy schislenija v jazykakh mande (системы счисления в языках манде) [counting systems in mande languages]. In Alexander Zheltov (ed.), *Afrikanskij sbornik - 2015 (африканский сборник - 2015) [African collection − 2015]*. 356–377. St. Petersburg: Muzej antropologii i etnografii RAN (Kunstkamera).

Vydrine, Valentin. 2009. Esquisse de la langue lélé (groupe mokolé). *Mandenkan* 45. 29–104.

Westermann, Diedrich. 1927. *Die westlichen sudansprachen und ihre beziehungen zum Bantu*. Berlin: Walter de Gruyter & Co.

Williamson, Kay [Ruth] [Margaret]. 1988. Linguistic evidence for the prehistory of the Niger delta. In Ebiegberi Joe Alagoa & Nwanna Nzewunwa Anozie (eds.), *The prehistory of the Niger delta*, 65–119. Hamburg: Helmut Buske Verlag.

Williamson, Kay [Ruth] [Margaret]. 1989a. Benue-Congo overview. In John T. Bendor-Samuel (ed.), *The Niger-Congo languages: A classification and description of africa's largest language family*, 247–274. Lanham MD, New York & London: University Press of America. by arrangement with the Summer Institute of Linguistics (SIL).

Williamson, Kay [Ruth] [Margaret]. 1989b. Niger-congo overview. In John T. Bendor-Samuel (ed.), *The Niger-congo languages: A classification and description of africa's largest language family*. 3–45. Lanham MD, New York & London: University Press of America. by arrangement with the Summer Institute of Linguistics (SIL).

Winkelmann, Kerstin. 2007a. Akasele. In Gudrun Miehe, Brigitte Reineke & Kerstin Winkelmann (eds.), *Noun class systems in Gur languages. Southwestern Gur languages (without Gurunsi)*, vol. 2012, 407–424. Köln: Rüdiger Köppe.

Winkelmann, Kerstin. 2007b. Khe. In Gudrun Miehe, Brigitte Reineke & Kerstin Winkelmann (eds.), *Noun class systems in Gur languages. Southwestern Gur languages (without Gurunsi)*, vol. 2007, 181–193. Köln: Rüdiger Köppe.

# Did you like this book?

This book was brought to you for free

language
science
press
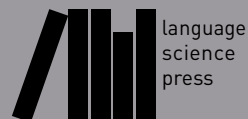
# The numeral system of Proto-Niger-Congo

This book proposes the reconstruction of the Proto-Niger-Congo numeral system. The emphasis is placed on providing an exhaustive account of the distribution of forms by families, groups, and branches. The big data bases used for this purpose open prospects for both working with the distribution of words that do exist and with the distribution of gaps in postulated cognates. The distribution of filled cells and gaps is a useful tool for reconstruction.

The first chapter of this book is devoted to the study of various uses of noun class markers in numeral terms. The second chapter deals with the alignment by analogy in numeral systems. Chapter 3 offers a step-by-step reconstruction of number systems of the proto-languages underlying each of the twelve major NC families, on the basis of the step-by-step-reconstruction of numerals within each family. Chapter 4 deals with the reconstruction of the Proto-Niger-Congo numeral system on the basis of the step-by-step-reconstructions offered in Chapter 3. Chapter 5 traces the history of the numerals of Proto-Niger-Congo, reconstructed in Chapter 4, in each individual family of languages.