

The numeral system of Proto- Niger-Congo

A step-by-step reconstruction

Konstantin Pozdniakov

Niger-Congo Comparative Studies 2



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.

This title can be downloaded at:

<http://langsci-press.org/catalog/book/191>

© 2018, Konstantin Pozdniakov

Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0):

<http://creativecommons.org/licenses/by/4.0/>

ISBN: **no digital ISBN**

no print ISBNs!

no DOI

Source code available from www.github.com/langsci/191

Collaborative reading: paperhive.org/documents/remote?type=langsci&id=191

Cover and concept of design: Ulrike Harbort

Fonts: Linux Libertine, Arimo, DejaVu Sans Mono

Typesetting software: Xe_{La}TeX

Language Science Press

Unter den Linden 6

10099 Berlin, Germany

langsci-press.org

Storage and cataloguing done by FU Berlin

no logo

Ирине Поздняковой

Contents

| | |
|---|-------------|
| Acknowledgments | xv |
| Abbreviations | xvii |
| 1 Preface | 1 |
| 1.1 Introduction | 1 |
| 1.1.1 Niger-Congo: the state of research and the prospects for reconstruction | 1 |
| 1.2 Sources and the monograph structure | 6 |
| 1.2.1 Sources | 6 |
| 1.2.2 Monograph structure | 7 |
| 2 Noun classes in the Niger-Congo numeral systems | 11 |
| 2.1 Noun classes in the counting forms of numerals | 15 |
| 2.1.1 The specific marking of numerals | 17 |
| 2.1.2 The grouping of numerals by noun class | 18 |
| 2.2 Noun classes in derived (reduplicated) numerals | 23 |
| 2.3 Noun class as a tool for the formation of numerals | 32 |
| 3 Analogical changes in numerals | 37 |
| 3.1 Issues pertaining to the detection of alignments by analogy . . . | 37 |
| 3.2 Mande | 40 |
| 3.3 Atlantic | 41 |
| 3.4 Kwa | 43 |
| 3.5 Adamawa | 49 |
| 3.6 Ubangi | 51 |
| 3.7 Gur | 51 |
| 3.8 Dogon | 52 |
| 3.9 Kordofanian | 53 |

| | | |
|----------|---|-----------|
| 4 | Step-by-step reconstruction of numerals in the branches of Niger-Congo | 55 |
| 4.1 | Benue-Congo | 55 |
| 4.1.1 | The Bantoid languages (including Bantu) | 56 |
| 4.1.2 | Benue-Congo (the Bantoid languages excluded) | 73 |
| 4.1.3 | Isolated BC languages | 103 |
| 4.1.4 | Proto-Benue-Congo | 104 |
| 4.2 | Kwa | 119 |
| 4.2.1 | Ga-Dangme | 119 |
| 4.2.2 | Gbe | 120 |
| 4.2.3 | Ka-Togo | 120 |
| 4.2.4 | Na-Togo | 121 |
| 4.2.5 | Nyo | 122 |
| 4.2.6 | Proto-Kwa | 128 |
| 4.3 | Ijo | 140 |
| 4.4 | Kru | 141 |
| 4.4.1 | ‘One’, ‘Two’ and ‘Three’ | 141 |
| 4.4.2 | ‘Four’ and ‘Five’ | 143 |
| 4.4.3 | ‘Six’ to ‘Nine’ | 143 |
| 4.4.4 | ‘Ten’ and ‘Twenty’ | 144 |
| 4.4.5 | ‘Hundred’ and ‘Thousand’ | 144 |
| 4.5 | Kordofanian | 146 |
| 4.6 | Adamawa | 148 |
| 4.6.1 | Fali-Yingilum (G11) | 152 |
| 4.6.2 | Kam (Nyimwom, G8) | 152 |
| 4.6.3 | Leko-Duru-Mumuye (G4, G2, G5) | 153 |
| 4.6.4 | Mbum-Day (G13, G14, G6, Day) | 155 |
| 4.6.5 | Waja-Jen (G9, G10, G1, G7) | 158 |
| 4.6.6 | Laal | 159 |
| 4.6.7 | Proto-Adamawa | 160 |
| 4.7 | Ubangi | 173 |
| 4.7.1 | Banda | 173 |
| 4.7.2 | Gbaya-Manza-Ngbaka | 173 |
| 4.7.3 | Ngbandi | 175 |
| 4.7.4 | Sere-Ngbaka-Mba | 175 |
| 4.7.5 | Proto-Ubangi | 175 |
| 4.8 | Dogon and Bangime | 182 |

| | | |
|---------|---|-----|
| 4.9 | Gur | 184 |
| 4.9.1 | ‘One’ | 186 |
| 4.9.2 | Bariba | 192 |
| 4.9.3 | Central Gur | 192 |
| 4.9.4 | Kulango | 199 |
| 4.9.5 | Lobi-Dyan | 200 |
| 4.9.6 | Senufo | 201 |
| 4.9.7 | Teen | 202 |
| 4.9.8 | Tiefo | 202 |
| 4.9.9 | Tusia | 202 |
| 4.9.10 | Viemo | 203 |
| 4.9.11 | Wara-Natoro | 203 |
| 4.9.12 | Proto-Gur | 205 |
| 4.10 | Mande | 215 |
| 4.10.1 | ‘One’ | 216 |
| 4.10.2 | ‘Two’ | 217 |
| 4.10.3 | ‘Three’ | 217 |
| 4.10.4 | ‘Four’ | 220 |
| 4.10.5 | ‘Five’ | 220 |
| 4.10.6 | ‘Six’ | 221 |
| 4.10.7 | ‘Seven’ | 222 |
| 4.10.8 | ‘Eight’ | 224 |
| 4.10.9 | ‘Nine’ | 225 |
| 4.10.10 | ‘Ten’ | 227 |
| 4.10.11 | ‘Twenty’ | 228 |
| 4.10.12 | ‘Hundred’ | 229 |
| 4.10.13 | ‘Thousand’ | 229 |
| 4.11 | Mel | 231 |
| 4.11.1 | Southern Mel | 231 |
| 4.11.2 | Northern Mel | 232 |
| 4.11.3 | Proto-Mel | 233 |
| 4.12 | Atlantic | 233 |
| 4.12.1 | Northern | 233 |
| 4.12.2 | Bak | 243 |
| 4.12.3 | North Atlantic and Bak Atlantic numerals in the comparative perspective | 254 |
| 4.13 | Isolated languages vs. Atlantic and Mel | 254 |
| 4.13.1 | Sua | 255 |

Contents

| | | |
|----------|--|------------|
| 4.13.2 | Gola | 255 |
| 4.13.3 | Limba | 255 |
| 5 | Reconstruction of Numerals in Niger-Congo | 257 |
| 5.1 | ‘One’ | 257 |
| 5.2 | ‘Two’ | 259 |
| 5.2.1 | ‘Two’ | 259 |
| 5.2.2 | ‘Two’ = ‘one’ PL? | 260 |
| 5.3 | ‘Three’ | 262 |
| 5.4 | ‘Four’ | 269 |
| 5.5 | ‘Five’ | 274 |
| 5.6 | ‘Six’ | 283 |
| 5.7 | ‘Seven’ | 284 |
| 5.8 | ‘Eight’ (‘Four’ and ‘eight’) | 284 |
| 5.9 | ‘Nine’ | 287 |
| 5.10 | ‘Ten’ | 290 |
| 5.11 | Large numbers (‘twenty’, ‘hundred’ and ‘thousand’) | 294 |
| 5.12 | Proto-Niger-Congo | 295 |
| 5 | NC numbers as reflected in particular families, groups and branches | 311 |
| 5.1 | Benue-Congo | 312 |
| 5.2 | Kwa | 314 |
| 5.3 | Ijo | 316 |
| 5.4 | Kru | 317 |
| 5.5 | Kordofanian | 318 |
| 5.6 | Adamawa | 319 |
| 5.7 | Ubangi | 322 |
| 5.8 | Dogon | 324 |
| 5.9 | Gur and Senufo | 325 |
| 5.10 | Mande | 327 |
| 5.11 | Mel | 329 |
| 5.12 | Atlantic | 330 |
| 5.13 | West African NC isolates | 332 |
| 5.14 | Summary | 333 |
| 5.15 | Conclusion | 336 |
| | References | 337 |

| | |
|--------------------------|------------|
| Index | 341 |
| Name index | 341 |
| Language index | 341 |

5 Reconstruction of Numerals in Niger-Congo

5.1 ‘One’

The five stems present in Table 4.1 are the most likely candidates for the reconstruction of ‘one’ in NC (Table 5.1):

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. However, 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

5 Reconstruction of Numerals in Niger-Congo

Table 5.1: Niger-Congo stems for ‘1’

| | | | | |
|-------------------------------|----------------|-------------------|-----------------------|----------------------------------|
| *di | | Dogon | | Kordofan lel/ led? |
| Atlantic di(n)/li(n)/ne(n) | Mande | Gur de/ le/ re | Ubangi le/ ne | Adamawa di |
| Mel -lɛ? | Kru | Kwa di- | Ijo die ? | Benue-Congo (o-)di(n), ni/nye |
| *in | | Dogon | | Kordofan |
| Atlantic -in? | Mande | Gur | Ubangi | Adamawa in? |
| Mel -in | Kru | Kwa | Ijo | Benue-Congo hin/kin/cin/-in |
| *do | | Dogon do | | Kordofan |
| Atlantic don/ lɔŋ | Mande do | Gur | Ubangi | Adamawa do? |
| Mel | Kru do | Kwa | Ijo | Benue-Congo |
| *ti | | Dogon ti(i) | | Kordofan tɛ(k)/ lu(k) |
| Atlantic | Mande | Gur | Ubangi | Adamawa |
| Mel | Kru | Kwa | Ijo | Benue-Congo (o-)ti |
| *gbo, *kpo | | Dogon | | Kordofan |
| Atlantic mbɔ | Mande | Gur (k)po | Ubangi k(p)o(k) | Adamawa *(g)bunu, (mon) |
| Mel bul, mɔ | Kru (g)bolo | Kwa | Ijo gbéri?/n-kèni? | Benue-Congo gbon, m-o? |

also attested in the Laal language isolate (*bìdíl* (*bì-díl?*) ‘1’). 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 *-ò-* 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 labio-velar (or labial in the case of Western NC) that are not necessarily etymologically related. The root *guñŋ* ‘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 *ʔən* (cf. Sua *sən*), Gur *túru/ tumə*, Mande West *kelen*, and Atlantic Bak *-anor, əkon*.

5.2 ‘Two’

5.2.1 ‘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*. However, 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 5.2):

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

Table 5.2: Niger-Congo stems for ‘2’

| | | | | |
|---|---------------------|--|------------------------------|---|
| *di/ni | | Dogon lé(y)/ ló(y)/ né(y)/ nó(y) | | Kordofan |
| | | | | |
| Atlantic di(k), nak | Mande | Gur nyi/ ne(n) | Ubangi | Adamawa du/ru, te/ re/ si |
| Mel díŋ/ tsɪŋ/ tɪŋ/ rəŋ | Kru | Kwa | Ijo | Benue-Congo ba-di / ba-ji |
| isolated roots | | Dogon | | Kordofan kok/kek/cik, (can/ʔan, rak, rɔm) |
| | | | Ubangi | Adamawa |
| Atlantic ɬubəʔ, -təb/-təw, -pugut/pugus | Mande pila/ fila | Gur nyu/ ju, hin/ han | si/ ʃi, (wa/gbwɔ, to/ so) | ra(k)/ ra(p), gba/ gwa |
| Mel | Kru so(n) | Kwa ɲɔ | Ijo mamV | Benue-Congo pa ? ba(i) ? |

below, this term is the least stable in the Niger-Congo languages. However, the NC root **di* is well-attested across the families.

5.2.2 ‘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 pre-

fix) 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* > **mò-òdì* / *mòi* ‘1’ / *ba*(CL2)-*di* > *badi* ‘2’ (the dialectal Proto-Bantu form *jòdè* (zones BH) (< **jò*(CL5 ?)-*dì*?)). 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 5.3):

Table 5.3: 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- |

As stated above, examples of this kind are exceptionally rare. A possible explanation 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 5.4):

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’.

5 Reconstruction of Numerals in Niger-Congo

Table 5.4: 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 | yén, ní | dènnì, yěnde/yènn(do), yòn, *de | den, ni | yen/ yin, dam | hén, nyǎŋ | den/ yen, ni, de? |
| 2 | yè, li | děé(ni), yēdē | le/ dé | yi(?) | li/ réʔ/ *ryaʔ | li/yi |

5.3 ‘Three’

Table 5.5: Niger-Congo stems for ‘3’

| | | | | |
|---|--|-------------------|----------------|-------------------------------------|
| | | Dogon taan | | Kordofan tat/ tǝʔ/ ʔak |
| Atlantic North: ʔat | Mande | Gur tat/ ta(n) | Ubangi taar | Adamawa taat |
| Mel sas/ ra | Kru taa(n) | Kwa ta | Ijo tato | Benue-Congo tat |
| | | Dogon | | Kordofan (ʔitin/ ʔicin, hwɔy) |
| Atlantic Bak: feegir, yant/ jent, habi/yabi | Mande sakpa/sagba/sawa, ʔaākɔ/yaākáʔ | Gur | Ubangi | Adamawa kunuy/ gbunuy |
| Mel | Kru | Kwa | Ijo | Benue-Congo |

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*. However, 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 5.6) provides an overview of the pertinent Bantu reflexes of **tátò* (ABEFGHJKLMNPRS) / **cátò* / **cáçò* (CD) ‘three’ (these reconstructions follow BLR3):

Table 5.6: Reflexes of **tátò* ‘3’ in Bantu

| zone | Language | Form | zone | Language | Form |
|------|----------|--------|------|-------------|--------|
| A | Nyo’o | tá | *PB | *PB (dial.) | cátò |
| 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 | ntfaro | K | Mbwela | hatu |
| C | Bua | salu | L | Kete | sàcw |
| C | So | saso | S | Lozi | talú |
| 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).

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).

5 Reconstruction of Numerals in Niger-Congo

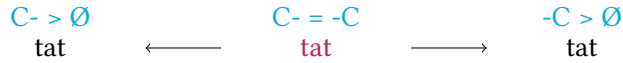


Figure 5.1: Scheme 4.1

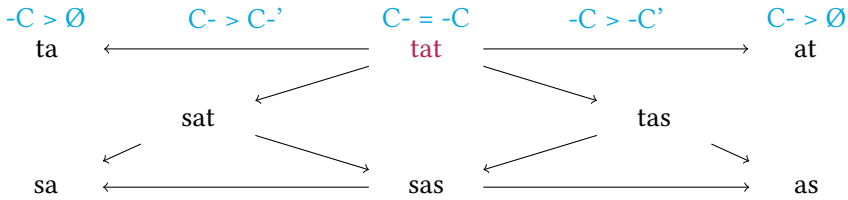


Figure 5.2: Scheme 4.2

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 5.7) 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 5.8), many of which are attested in Bantu, originate (forms without square brackets):

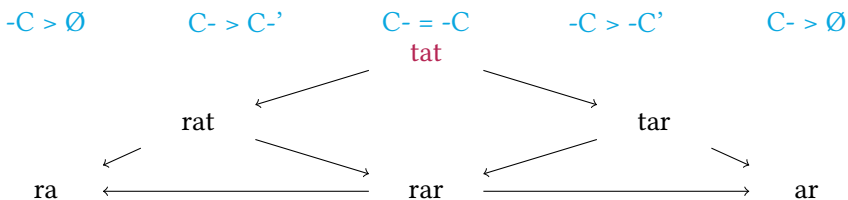


Figure 5.3: Scheme 4.3

Table 5.7: 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 |

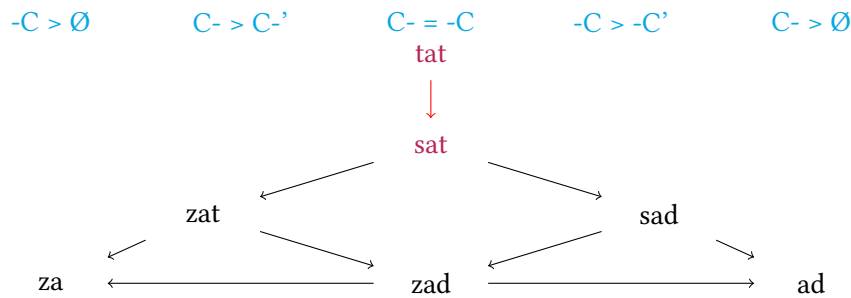


Figure 5.4: Scheme 4.4

Table 5.8: 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 |

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.

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 5.9-4.10):

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 5.11):

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 5.12):

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**.

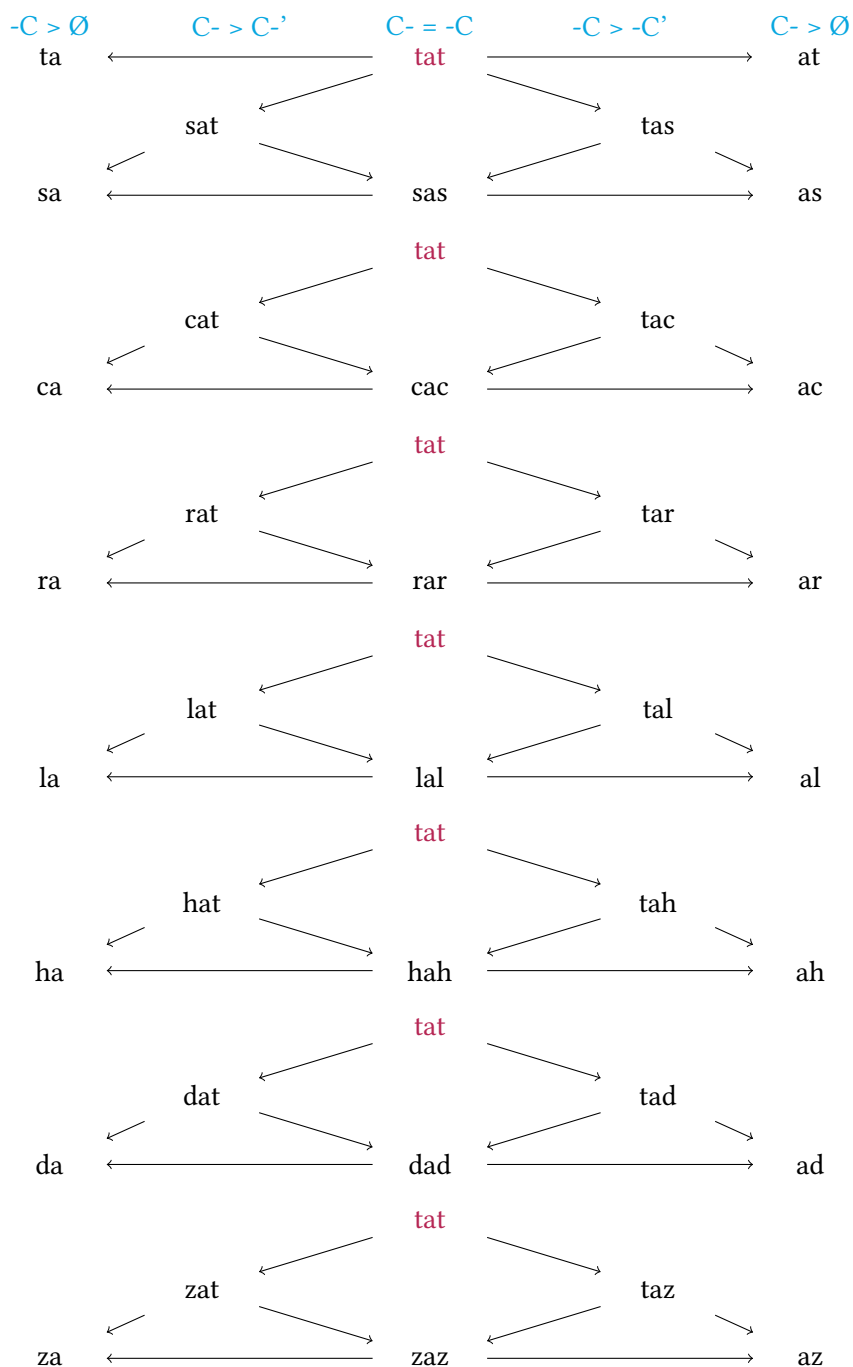


Figure 5.5: Scheme 4.5

5 Reconstruction of Numerals in Niger-Congo

Table 5.9: 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 | tā'l |
| TAD | Mbere | -tadi | | | Sereer | tad-ak |
| TAS | | | Kulaal | tòòs | Bapen | ḡΛ-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 | pě-sās |
| 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 | tjaro | Kam | tshar | | |
| CA | Bubi | -cha | Galke | cha-?a- | | |
| HAT | Nkoya | -hatu | | | Manjak | go-hant |
| DER | | | | | Baga | der |
| | | | | | Mboteni | |

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, giving 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*?, **-t̥*?, **-ts̥*?, **-c*? As we

Table 5.10: Reflexes of **tat-* in Niger-Congo (2)

| | Bantoid | | BC | | Dogon | | Gur | |
|-----|---------|--------|-------------|----------|------------|---------|-----------|------------|
| TAT | Bankala | tát | Biom | be-tat | kolum so | tūāti | Ditammari | -tāāti |
| TAR | Mambila | tar | Jiru | i-tar | bangeri-me | ke-taro | Senari | tāre |
| TAL | Kom | tál | Olulumo | è-tál | toro tegu | taali | Nateni | tālī, tādī |
| TAD | Ngwe | tád | Upper-Cross | *-ttáD | tommo so | tadu | Nateni | tādi, tālī |
| TAS | | | ikaan | tás | | | | |
| TAZ | | | | | | | | |
| TA | Abon | -ta | Ibibio | i-tá | | | Dagbani | -ta |
| SAT | | | Morwa | sat | | | | |
| SAR | Mbe | bé-sár | Kugbo | i-sàr | | | Lorhon | sā:r |
| SAS | | | | | | | Viemo | saasi |
| SA | Ekoi | é-sá | Oloma | e-sa | | | Kulango | sā |
| AT | | | Kohumono | a-átá | | | Hanga | ata |
| AR | | | | | | | | |
| LAL | | | | | | | | |
| RAR | | | Abua | i-rààr | | | | |
| RA | Nkem | í-rá | Ukue | è-rhá | | | | |
| CAR | | | Ufia | kù-tshàr | | | | |
| CA | | | Bandawa | ni-ca | | | | |

tried to show in (PozdniakovSeegerer2007), 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*, *sVf*, 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”, 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.

5.4 ‘Four’

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

5 Reconstruction of Numerals in Niger-Congo

Table 5.11: 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 | | | | | | 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 | 123 |

Table 5.12: 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 |

Table 5.13: Niger-Congo stems for ‘4’

| | | | | |
|-----------------------------------|--------------|-----------------|-------------|---------------------------------|
| | | Dogon nay(n) | | Kordofan |
| Atlantic | Mande | Gur | Ubangi | Adamawa |
| Nord: nak | náání / nãñi | naan | naar | naX, nēn/ niŋ, nda |
| Mel | Kru | Kwa | Ijo | Benue-Congo |
| | na | na | néin | nai |
| | | Dogon keeso | | Kordofan |
| | | | | -ɽɔŋ /-ɽando/ -rɔm?,-(gálɪm) |
| Atlantic | Mande | Gur | Ubangi | Adamawa |
| Bak: baakar/wakar, tasala | | | (syɔ), lu) | |
| Mel | Kru | Kwa | Ijo | Benue-Congo |
| Nord: ’-ŋkile/- nle, Sud: hiol | | | | |

Kordofanian), as well as in the Western NC isolates, cf. Sua *b-nan*, Gola *tii-nàŋ*, Limba *ka-nan*. 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. *ḡī-sa-n* (*ḡī-sa-n?*) ‘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. However, 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.

However, 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 5.14):

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 5.15):

The Adamawa evidence is as follows (Table 5.16):

My suggestion is that the variety of similar forms attested in the majority of

Table 5.14: 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 5.15: Main n- forms for '4' in Kwa

| | |
|---------------|-------|
| Agni (Anyin) | n-na |
| Abron | n-nai |
| Baule | nu-ne |
| Eotile (Beti) | a-ni |

Table 5.16: Main n- forms for '4' in Adamawa

| | |
|---------|-----|
| Tupuri | na |
| Mundang | nai |
| Gula | nay |
| Waja | ni |

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’.

5.5 ‘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 NursePhilipppson1975, 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 5.17):

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*. In addition, the root **dòngò*, which probably goes back to **dòngò* ‘line, row’ (zones: ABCDEFGHJKLMNRS) deserves our attention as well.

The initial consonant in **táànò* is the same as in **tátò* ‘three’, which is proba-

Table 5.17: 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 | limb: arm; leg; thigh | 4: NW SW Ce NE | 8: A B C E H L M R |
| jádà | nail (> finger > ‘hand’) | | > ‘hand’ A D E F G J L N P S |

bly 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 following Bantu evidence (Table 5.18) is illustrative of this admittedly unconventional approach (further BC evidence will be quoted later in this chapter).

Table 5.18: 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 5.19):

The fact that the same consonants are reflected differently may have several

Table 5.19: Different initial consonants in ‘3’ and ‘5’ in Bantu

| | Language | ‘3’ - *tátò | ‘5’ - *táànò |
|---------|----------|-------------|--------------|
| 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 |

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’).

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 5.20): 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 5.21):

The Table 5.21 shows peculiar forms attested in one of the Southern Mel languages (Bom) that are virtually identical to the BC reconstructions. Thus, we

¹Elugbe1987.

Table 5.20: 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 | taij |
| 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 |

5 Reconstruction of Numerals in Niger-Congo

Table 5.21: 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-tɔ́n |
| Kwa | Tuwuli | ɛ-lale | 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 |

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 5.22):

Table 5.22: **tan* ‘5’ in Niger-Congo

| | | | | |
|---|---------------------------|------------|-------------|---------------------------|
| | | Dogon | | Kordofan dinin/ dulin? |
| Atlantic tɔk, tən? | Mande **tan? (> ‘10’?) | Gur tɔ | Ubangi | Adamawa sa? |
| Mel kə-ʔamaʔ (< *kə- ʔa ‘hand’?), tan? | Kru | Kwa ton | Ijo túnó | Benue-Congo tan/ ton |

The attestations of this root in Southern NC (namely in BC, Kwa and Ijo) are more systematic. In Western NC the root is reliably attested as well, despite the fact that the Northern Mel form *kə-ʔamaʔ* allows a two-fold interpretation (i.e. as a derivative of either *ʔam-* or **kə-ʔa* ‘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 5.23):

Table 5.23: **nun* ‘5’ in Niger-Congo

| | | | | |
|----------|-----------|-------------------------------------|--------|------------------|
| | | Dogon núnéé(n)/ nú:(yn) / nùm | | Kordofan |
| Atlantic | Mande | Gur nu(n) | Ubangi | Adamawa nu(n) |
| Mel | Kru mm | Kwa nu(n) | Ijo | Benue-Congo |

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 5.24):

Table 5.24: ‘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 5.25):

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 **nòè* ‘finger, toe’ (Regions

Table 5.25: 'Hand' and potential reflexes of *nun* '5' in Gur

| Group | Source | Language | 'hand' | 5 |
|-----------|-----------------------|------------------|--------------|----------|
| Bariba | Koelle 1963[1854] | Baatonum | nóma | nǒwu |
| Bwamu | BloemartsdeRasby 2010 | Bly 2010 | núumánnu | |
| Grusi | Koelle 1963[1854] | Tem | | nónūa |
| Grusi | CLNK1999 | Kabiye | | naanūwa |
| Grusi | Koelle 1963[1854] | Kiamba | noon / noozi | noonuua |
| Grusi | Koelle 1963[1854] | Sisaala Tumulung | | ínōm |
| Oti-Volta | Koelle 1963[1854] | Mosi | nuro | nu |
| Oti-Volta | Koelle 1963[1854] | Gurma | unu / inui | mu ~ mmu |

5: NW SW Ce NE SE ; Zones 9: ADJKLMPRS), (cf. Bantu, zones MN – Nyiha-Malila-Lambya (NursePhilipppson1975) *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 *nǎ̀nǎ̀ŋ* 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 5.26).

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 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 5.27).

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

Table 5.26: ‘Hand’ > ‘5’ in Atlantic

| Group | Language | ‘hand’ | ‘5’ | ‘10’ |
|----------------|-----------|--------------|-----------|-----------------|
| Atlantic-Bak | Balant | f-cef / k- | cef | f-cef meen |
| | Kentohe | | | (‘whole hands’) |
| Atlantic-Bak | Bijogo | kɔ-ɔkɔ / | nde-ɔkɔ | n-rua-kɔ |
| | Kagbaaga | ŋa-akɔ | | |
| Atlantic-Bak | Bijogo | kɔ-kɔ / | ŋu-duβ-kɔ | ŋɔ-rúŋa-kɔ |
| | Kamona | ŋa-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 | si-lax | ci-lax | haa-lax |
| | Djibonker | | | |
| Atlantic-North | Nyun | ci-lax / xa- | ci-lax | xa-lax |
| | Gujaxer | | | |
| Atlantic-North | Biafada | gə-bəda / | gə-bəda | |
| | | ma-bb- | | |
| Atlantic-North | Jaad | ko-bəda | ko-bəda | |

Table 5.27: ‘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 |

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 *be-tak / be-tuk / be-tik* (cf. also Limba *bi-sohi* ; Sua *sungun*), cf. Table 5.28.

Table 5.28: ‘Hand’ > ‘10’ in Joola (Atlantic: Bak)

| Language | ‘hand’ | ‘5’ | ‘10’ |
|----------------------|----------------------|-----------|---------|
| Joola_Banjál | ga-ɲen / gu-ɲen | fu-tox | gu-ɲen |
| Joola_Fogny | ka-ɲen / u-ɲen | fu-tɔk | u-ɲen |
| Joola_Gusilay | ga-ɲen / u-ɲen | fu-tɔk | u-ɲen |
| Joola_Kasa | ka-ɲen | hu-tɔk | ku-ɲen |
| Joola_Kasa_Esuulaalu | ka-ɲen | hu-tɔk | ku-ɲen |
| Keeraak | ka-ɲen-ak / ʊ-ɲen-aw | hʊ-tɔk | kʊ-ɲen |
| Joola_Kwaatay | ɛ-ɲɔmu | hu-tɔk | si-ɲɔmu |
| Joola_Kwaatay | ɛ-mɔɲo | hu-tɔk | su-muɲo |
| Joola_Mlomp | ɛ-be:s | ɲa:-suwaɲ | se-be: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 5.29).

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 *kɔ́*, Proto-Eastern Mande *gɔn* (?), Dida (Kru) *kɔ́*, etc.

The following Kordofanian terms that attest to the development of ‘hand’ > ‘5’ are also noteworthy: Dagik (Kordofanian) *si-s-ɜl:v* ‘5’ (lit. ‘one hand’): “The *si* in

Table 5.29: ‘hand’ > ‘5’/‘10’ in some Atlantic and Mel languages

| Branch | Language | ‘hand’ | ‘5’ |
|--------------------|-----------------|----------------------------|----------|
| Atl.-Centre-Manjak | Mankanya | ka-nyen | ka-nyɛɛn |
| Atl.-Centre-Manjak | Manjak | ka-ñen | ka-ñen |
| Atl.-Centre-Manjak | Manjak | kádṣāg | kányan |
| Atl.-Centre-Manjak | Mankanya | úlōl | kányɛn |
| Atl.-Centre-Manjak | Manjak Bassarel | pëndānd | kañan |
| Atl.-Centre-Manjak | Manjak Tame | wúepalōl, pl. n·gípalōl | kényān |
| Temne-Baga-Landuma | Temne | kə-ta / mə- | ta-maṭ |
| Temne-Baga-Landuma | Temne | a-loŋk (i), ma- | ṭamaṭ |
| Temne-Baga-Landuma | Temne | à.loŋk | -tàmath |

5 comes from the word ‘hand’. So 5 is ‘one hand’², Acheron *zəguŋ zulluk* (lit: ‘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/ ?i:p*, Mel *wan/wen*, Mande *dúuru/ sóóru*, Kru *gbə / gbo*, Gur *mwan/ bwa*, Ubangi *du(w)/ lu(w)*, Kordofanian *ŋer-/ per-*. As a rule, these innovations (not quoted here exhaustively) are only attested in particular branches of the families under study.

5.6 ‘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, cie* (‘7=6+1’), Adamawa *jup, gu*, Ubangi *zala/ zya*, Dogon *kuro/ kule*, Gur *do(b)*, Mande *t(s)um?* (the examples

²John Vanderelst, <https://mpi-lingweb.shh.mpg.de/numeral/Dagik.htm>

³Russell Norton, <https://mpi-lingweb.shh.mpg.de/numeral/Acheron.htm>

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.

5.7 ‘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-ránèfi* (‘6+1’)), Adamawa (*bir/ bil, rɪɪ, nbutu*), Ubangi (*sílànā, lè-rèzi*), 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.

5.8 ‘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 5.30).

The reduplication can also be ‘partial’ (as a rule the reduction of the first syllable is involved), cf. Table 5.31.

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 5.32.

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-ddú* ‘4’ > *ka-ʔa-nda-ddú* ‘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-jo 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

Table 5.30: ‘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 | je | je-ri-je |
| BC-Cross | Kukele | na | i-na-mi-na |
| BC-Bantoid | Esimbi | mō-jī | mō-jī-ō-jī |
| BC-Jukunoid | Mbembe | nyε | é-nyεnyε |
| Bc-Ikaan | Ikaan | nā ¹ / nā | nà:ná ¹ / nà:ná |
| Adamawa-Fali | Fali | ná:n | nàn nán |
| Adamawa-Duru | Gəunəm | náárək | náárək àp náárək |
| Gur-Southern | Lamba | nasa | nasì-nasa |
| Gur-Southern | Lyele | na | nana |
| Laal | Laal | ḃīsān | ḃīsān.ḃīsān |

Table 5.31: '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 5.32: '8' = 4PL

| Branch | Language | '4' | '8' |
|------------------------|----------|-----------|-------------|
| Kwa-Nyo | Lelemi | í-né | máá-né |
| Kordofanian Heiban | Warnang | ɣɛ̀lamlàn | ɣelamlaan-ɔ |
| BC Platoid | Ikulu | ín-nāā | nín-nāā |
| Adamawa Leko-Nimbari | Yendang | nâ:t | ḃṵ-lá-nâ:t |
| Adamawa Mbum-Day | Niellim | ɲɛ́ní | twā:-ɲɛ́ní |
| Adamawa Waja-Jen | Waja | nu | wu-nii |
| Ubangi Sere-Ngbaka-Mba | Gbanzili | ḃṵ-nā | sá-nā |
| Gur Grusi | Delo | a-naara | gya-naara |
| Gur Grusi | Tampulma | a-naasi | ɣmɛ-naasa |

vowel present in ‘four’ or has [a] in the first syllable. This mechanism is confirmed at least in the case of Bantu (Table 5.33).

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 5.34).

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 5.35) or with an innovation (Table 5.36):

The evidence presented above strongly suggests that the pattern ‘8=4 redupl.’ 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 5.37):

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, tɔŋi-*, Mande *seki/ segi*, Dogon *sele/ sagi* (< Mande ?), *gá(a)rà*, Atlantic Bak **ʋAs-*. These forms (as well as some additional ones) are interpreted as local innovations.

5.9 ‘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. Ban-

5 *Reconstruction of Numerals in Niger-Congo*

Table 5.33: 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 5.34: '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 5.35: '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 | penet | jurom-ŋeta |
| Gur | Birifor (dial.) | anan | anu-ni-ata |
| Gur | Teen | nan | to sanr |
| Mande | Vai | nani | sog sakpa |
| Adamawa | Karang | niŋ | tòŋ ndók sé'de ('10-2') |

Table 5.36: '8=4+4' > '8' innovated

| Family | Languages | '4' | '8' |
|------------|--------------------|---------|---------------|
| Bantu-A | Bafo | benin | wam |
| Bantu-A | Bankon | bi-nan | mwam |
| Bantu-A | Fang | ɲɪɲ | mwom |
| Bantu-A | Ndambomo | li-naɲi | li-mwabi |
| Bantu-B | Kota | naɲi | mwabi |
| Bc-Platoid | Mabo | nen | hur |
| Dogon | Tene Kan | nani | silá |
| Dogon | Tene Kan | nani | sira |
| Kwa | Abron | nain | ɲocie |
| Kwa | Akan (Akuapem Twi) | anan | awotcye /tw/ |
| Kwa | Baule (Baoulé) | nan | nmocue |
| Kwa | Foodo | naɲ | dukwe / dukoi |
| Kwa | Mbato | ne-ni | o-gbi |
| Mande | Mandinka | náani | segi |
| Mande | Looma | náanĩ | dosawa |

toid *bukV* (if indeed primary), Akpes ɔ-kpɔlɔf(i), Defoid *sá(n), dà (cf. Edoid *cien/sin*), Igboid *totu/tolu*, Ubangi *kùsì*, *me-newá*, Laal *yàɲjǎɲ*, Dogon *túwó*, Mande *kòntonɔ/kòɲɔndɔ(n)* (historically perhaps '10-1').

5.10 '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 5.38).

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*, *fəɲən*, *tiəɾum*, 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 *zìyà* '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).

Table 5.37: ‘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-G?E? | 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 | já-ciaat | bé- <u>ni</u> | jí-rii- <u>ni</u> |
| BC-Cross-River | Alege | é-cɛ | é-ce-e-ce | é-ne | ee-né-ne |

Table 5.38: **pu/ fu* ‘10’ in Niger-Congo

| | | | | |
|----------------------|-----------------|----------------|--------------------|-----------------------|
| | | Dogon | | Kordofan |
| Atlantic pok | Mande pu/ fu | Gur fu/ po | Ubangi bú / fu? | Adamawa boo/ fu? |
| Mel pu/ tɔ-f-ɔt ? | Kru pu | Kwa fo / wo | Ijo | Benue-Congo pu/ fu |

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 related to the 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 5.39).

Table 5.39: Labial suffix in Tula numerals

| | | | |
|---|-----|----|----------|
| 1 | -iñ | 6 | nukuñ |
| 2 | rəp | 7 | nibiñ |
| 3 | táa | 8 | na'á-rəp |
| 4 | naa | 9 | túrukup |
| 5 | nu | 10 | kwəp |

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* (Dimmendaal 1978 **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 *kɔgba* ‘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 **kɔ́* ‘hand, arm’). In Kru, this root is attested not only in the Eastern group (Dida *kɔ́* ‘hand’), but in the Western group as well (Glio-Oubi *hō*, Krumen *hō*). Finally, it is (admittedly only marginally) attested in Bantoid (as an alternative to the wide-spread root *kómi* ‘10’): according to Larry Hyman (in Paulin 1995) this root is distinguishable in Kom (*ɔ̄-kɔ̄*) and Narrow Bantu, e.g. in zones B (Mpur *kɔ*, Yansi *kwɔ*) and E (Mashami *oko*, Meru *uko*, Nurse/Philippson 1975). 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 (*‘spécilisation de “ligne” dòng?’*). However, it has parallels in other BC branches, namely in Cross River (Connell1991) and probably Idomoid (Table 5.40).

Table 5.40: 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 | dòn· & ndòn· (Koelle 1963[1854]) |

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 *kómì/ 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 *kóm* ‘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 as others discussed above.

Other isolated forms for ‘ten’ include Atlantic (*n*)*taaj*, *taim*, *-suwan*, Mel *wi-tfɔ?*, Western Mande *tan* (< *5’?), Gur *ke(n)*, Kwa *bula* (cf. Ubangi *bale*), Ubangi *busa*, *sui*, Kordofanian *tu(l)*, *di*, *rakpac*, *fəŋən*, *tiərɪm*, Adamawa *kutu(n)* (< **kutu(n)*), cf. Laal *tūū*, Kordofanian *ṭṭṭ*, Sua *teŋi* etc.

5.11 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.

5.12 Proto-Niger-Congo

The reconstruction of the Proto-Niger-Congo number system may be summarized as follows (Table 5.41):

Table 5.41: 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.

References

- 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.
- 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>.
- 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.

References

- 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 Études 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 natio-ro*. 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 natio-ro*. Electronic document.
- 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 *Aspekty komparativistiki-2. Orientalia et classica XI: Trudy instituta vos-tochnykh kul’tur i antichnosti*, 409–498. Moscow: Izdatel’sтво 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 Perekhval’skaya. 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.
- Winkelmann, Kerstin. 2007a. Akasele. In Gudrun Mieke, 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 Mieke, Brigitte Reineke & Kerstin Winkelmann (eds.), *Noun class systems in Gur languages. Southwestern Gur languages (without Gurunsi)*, vol. 2007, 181–193. Köln: Rüdiger Köppe.

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.

