

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’	271
5.5	‘Five’	274
5.6	‘Six’	283
5.7	‘Seven’	284
5.8	‘Eight’ (‘Four’ and ‘eight’)	284
5.9	‘Nine’	290
5.10	‘Ten’	291
5.11	Large numbers (‘twenty’, ‘hundred’ and ‘thousand’)	294
5.12	Proto-Niger-Congo	295
6	NC numbers as reflected in particular families, groups and branches	297
6.1	Benue-Congo	297
6.2	Kwa	299
6.3	Ijo	300
6.4	Kru	301
6.5	Kordofanian	301
6.6	Adamawa	302
6.7	Ubangi	304
6.8	Dogon	306
6.9	Gur and Senufo	306
6.10	Mande	308
6.11	Mel	310
6.12	Atlantic	310
6.13	West African NC isolates	312
6.14	Summary	313
6.15	Conclusion	315
	References	337

Index	339
Name index	339
Language index	339

6 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 6.1).

6.1 Benue-Congo

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.

Table 6.1: 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?	bó	+	+	+	5+3	dí/ 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, jífe, táj	pa(n)	+	+	+	+	+ ?	5
Bantoid (-B)	+ ?	pa/ba/fe	+	+	+	+	+	6
Bantu	+	+	+	+	+	+	kómì	6
Oko	-óre, -jére	-bòrè	+	+	-pi	+	+	4
Akpes	+ ?	+	+	+	+	+	-yōf(i), *t-ēfī	6
Ikaan	ǰí	wa	+	+	+	+	+	5
Lufu	+ ?	máhà	+	+	+	5+3	+ ?	5

- 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 6.1 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.

6.2 Kwa

Interestingly, Table 6.2 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. However, Gan-Dangme has only the reflex of ‘three’ (assuming that *-tē* ‘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.

Table 6.2: NC numerals reflected in Kwa (+)

		1	2	3	4	5	8	10	Total
1.	Ga-Dangme	-kē, *go/ wo	- ɲɔ̃(n)	+	-ɣwè	-nùḥ	6+2	ɲòŋma'	1
2.	Gbe	+	-wè	+	+	+	-ɲí, 'hand'+3	+	5
3.	Ka-Togo	+	+	+	+	+	+	+	7
4.	Na-Togo	+	-nyɔ	+	+	-no(N)	+	+	5
5.1.	Nyo-Agneby	+	-ɲɔ̃	+	+	-ne	-pyè, wo(n)	diw, 5PL	3
5.2.	Nyo-Attié	kə(n)	mwə(n)	+	dʒí(n)	bə(n)	+	kɛŋ	2
5.3.	Nyo-Awikam	-tɔ̃	-ɲɔ̃	+	+	-ɲú	-tyé	-jú	2
5.4.	Nyo-Alladian	-tò	-yrè	+	-zò	-nrì	-ɸrì	+	2
5.5.1.	Nyo-Potou	*ce, bɛ̃	-nõ	ja/je	+	na	ɸyá/ gɸi	+	2
5.5.2.	Nyo-Tano	ko(n)	-ɲɔ/ - ɲu(n)	+	+	nu(n)	- kwé/ - cué	bulu, du	2

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.

6.3 Ijo

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, un-

Table 6.3: NC numerals reflected in Ijo (+)

	1	2	3	4	5	8	10
Defaka	gbérí	mààmà	+	+	+	5+3	+ ? (wóì)
East	*+, gbérí, ògèi	màmì	+	+	+	+	ójí /àtié
West	*+, kènɪ	maamɔ	+	+	+	+	ójí

clear 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.

6.4 Kru

Table 6.4: NC numerals reflected in Kru (+)

	1	2	3	4	5	8	10
Aizi	mumɔ, yre	-fɪ	+	yeɓi	-gbo	pate	bɔ
Eastern	+	sɔ	+	+	gbu / gbi	5+3	+, kógba
Kuwa	+	sɔr	+	+	wàyɔɔ	5+3	kowaa
Seme	dyuɔ	nĩ	+	yur	kwɛl	kprɛn	+
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).

6.5 Kordofanian

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

Table 6.5: NC numerals reflected in Kordofanian (+)

	1	2	3	4	5	8	10
Heiban	-(t)ʈɛ(k)	-can / -ʈan, rɔm	+	-ʈɔŋɔ/ -ʈʈɔ	-dini, ɲer-	+ ?	di/ di/ ri
Katla	-ʈʌk	cik/ heek	+	-gʌlʌm	-duliin, -gbəlɪn	ʈʌŋgil	*ʈʌʌ, -rɔ
Rashad	-tta	(k)ko(k	+	-rɔm	*ɲer-, -ram	dubba	5PL
Talodi	+? (lu(k)/ li(k))	-ʈʌk/ -tta	+	-ʈandɔ, kekka	hand'- '1', -liəgum	+ ?	-tu(l), tiərum

to have inherited the NC system (with the exception of ‘three’ that apparently goes back to its NC prototype, cf. e.g. Katla $\lambda\text{-}\text{t}\text{ʈ}\text{ʌ}\text{t}$ ‘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) $\eta\text{è-làmlàn}$ ‘4’ > $\eta\text{e-lamlaa}\eta\text{-}\text{ɔ}$ ‘8’, Lumun (Talodi) $m\acute{s}ʔ\acute{r}\text{in}$ ‘4’ > $m\acute{a}\text{-}m\acute{ə}rm\acute{ə}r$ ‘8’. This resemblance, however, may be due to typological (rather than etymological) reasons.

6.6 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

Table 6.6: NC numerals reflected in Adamawa (+)

	1	2	3	4	5	8	10
Fali	+	gbara, cuk	+	+	kêrɛw	+	ra
Kam	+ ? (-ii)	-raak	+	+	ɲwún	sâl	+ ?
Leko Duru	+, ɲá	du/ ru, to, te/re	+	+	nún-	5+3	+ ?, kob
Leko Leko	*ɲa	ra, in, nu	+	+	núún-	5+3	kob
Leko Mumuye	+ ?	ye, ti, ni	+	+	nɔŋg	5+3	kob
Mbum Bua	+	+	+	+	*lu, tɛ, *kɔn, tiso	+	do, kùtù
Mbum Kim	+	+ ?	+	ndà(y)	nūwɛ̃y	+	wàl
Mbum Mbum	böŋɲ/ búónó	ti, sede, gwa	+	+	ndifi	10-2	+ ?, dùɔ, -wàl
Mbum Day	ngɔ́ɲ, *mon	+	+	ndà, -yām	sɛ̀rì	+	+ ?
Waja Jen	+	+ ?	+	+	nóob/ *na, *hwĩ	+	ʃóób
Waja Longuda	khal, twè	shir, kwé	+	+	nyó	nyíthìn	koo/ kù
Waja Waja	+	+ ?	+	+	nu(ɲ)	+	kob
Waja Yungur	+	+	+	kurun	-nun	+	+ ?, kutun
Laal	+	(ʔi-sĩ?)	māā	ḃi-sān	+ ?? (sā̃b, *swa-)	+	tūū

in the following branches: Duru *dáa*, Bua **le* and possibly Laal *bī-dīl?* Its reconstructed allomorph **n-di* (with further evolution to **ni/ -in*) may be reflected in Kam (*-ii*), Jen *-in*, Waja *-in*, Mumuye (?) *-ni*, Yungur (?) *-ni*. The terms reflected in Falo **-lo*, Bua *dɔ(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 / tʃi-rí*, Day *díi*, Jen **re / rá-b*, Waja *ró-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 6.6 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 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’.

6.7 Ubangi

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-lē?*, Ngbaka-Mba *bī-ni / bi-rì*, Zande *kí-lī*;

Table 6.7: NC numerals reflected in Ubangi (+)

		1	2	3	4	5	8	10
1.	Banda	+	-fi	+	+	-ndũ	5+3	+
2.	Gbaya- Nanza- Ngbaka	kpó(k)/ ndánj	wá?, -too	+	+	-(k)ó	+	+ ? (bú)
3.	Ngbandi	kə(i)	sɛ	+	siə	kɔ̃/ kũ	miambe	sui, bàlé
4.1.	Ngbaka- Mba	+, kpó-	-fi/ -si	+	+	ve/ vue	5+3	<'hand'
4.2.	Sere	njěe	so	+	+	vo	5+3	<'hand'
5.	Zande	+	-jō/ -yō	+	lu ?	-sibē/ -sùè	5+3	ngbɔ̃

NC **pu/ fu* '10': Banda *bu-fu*, Gbaya *bú/ bù-kə*. 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 *bù*), e.g. Edouard Koya states that *bu* means 'person' in Bokoto (Central Gbaya-Manza-Ngbaka), where *bù-kə* '10' (<https://mpi-lingweb.shh.mpg.de/numeral/Bokoto.htm>). Moniño suggests an alternative etymology (Moñino1995): «**bú* 'dix' est en relation avec **bú* 'façonner, faire un cercle, joindre les mains'; la série partielle **bú-kɔ̃* '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 *bú* in Gbaya are provided in (BlanchardNoss1982: 51):

bú «joindre les deux extrémités d'une même chose ; faire de la poterie»,

bú «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.

6.8 Dogon

Table 6.8: NC numerals reflected in Dogon

	1	2	3	4	5	8	10
Dogon	túru/ tumɔ, ti(i)	+ ?	+	+	núnéé	gá(a)rà	péru/ pélú
Bangime	tòré / tǝyɛ	+ ?	+	+	nǝndí	(borrowed)	kúré

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 *lé(y)/ né(y)* (with their allomorphs *lɔ(y)/ nɔ(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’.

6.9 Gur and Senufo

Evidence of the ten Gur branches is treated in Table 6.9 (cf. the discussion pertaining to the division of Gur into 16 branches in Chapter 4).

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-Natioero as well. Nearly the entire inventory of NC terms was replaced in Senufo (except for ‘three’ – Senufo **tǎã/ taàr*), Bariba (except for *i-ta* ‘three’ and *n-ne* ‘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’.

Table 6.9: NC numerals reflected in Gur and Senufo (+)

		1	2	3	4	5	8	10
1.	Bariba	tīā	Ru	+	+	nòòbù	5+3	-kuru
2.1.	Central North- ern	+	+, pū	+	+	nu	5+3, tɔɔ, ni	+
2.2.	Central South- ern	+	+, pɔ/ jɔ	+	+	nɔ(n)	+	+
3.	Kulango	*taà	nyɔ̀ò	(borrowed)	+	+	5+3	5PL, *ji
4.	Lobi- Dyan	+	nyɔ̀(n)	+	+	mòì/ *mà, dièmà	5+3	-kpo, nyòòr
5.	Senufo	ngbe, nikĩ	sin /sun	+	tésyàr, ...	-no	5+3	kɛ
6.	Teen	tani	Nyor	+	+	+	5+3	+ ?
7.	Tiefo	+	jɔ̃	+	ʔuʔɛ́ / ɲɔɔ	kǎ	5+3	kɛ́
8.	Tusia	nónki	nîŋ	+	+	k(w)lɔ	5+3	*gbɔ/ bwɔ
9.	Viemo	+ ?	Niini	+	jumĩ	*kɔ	+	kwɔmũ
10.	Wara- Natorio- Paleni	pɔ	nínté, bó	+	+	sùsù, sV	+	+

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) **le/ 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’ (*sāābe*) 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.

6.10 Mande

This is no doubt the most isolated family in what pertains to the reflection of NC numbers (Table 6.10). 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 *kedən* ‘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-firɛ* ‘4’) whereas [b-] appears before a voiced consonant (*b-zei* ‘3’, *b-ʒiɪ* ‘10’).

Table 6.10: NC numerals reflected in Mande (+)

	1	2	3	4	5	8	10
Manding	+, *kelen	*filá	*sábá	+	dúuru, *wo	séegi	+ ?, *tán
Jogo-Jeri	+, *kelɛ ?	*fàlá	sègbá	+	sóólò	5+3	táà(n), ta
Mokole	+, *kéle	*fila	saba	+	lóolu, *wo	séen/ seyi	+ ? , tán
Vai-Kono	+, *N- kelen	*fèLá	sagba	+	dúʔu/ sóó(?)ú	séin, 5+3	tân tân
Susu	*kédén, ndá/ nde	*fidín	sawa/ sàxán	+	sùlù, *fò	5+3	+, *tòngó
SWM	*giláaŋ, *tà	*fèelé	sāabā	+	dóólú, *wo/ ngò	5+3	+
Bozo-Soninke	ke/ kuɔn, sana, ...	pě:ndé/ fillò	sike/ síkkò	+	kóló- /kára-	segi	tan
Bobo	tàlá/ télé	pálà	sàà	+	kóò	séki/ tʃèkí	+, m̥m̥
Samogo	*ké, *so	fí:(kí)	zì:gī, f̥wè, yei	+	+	+, kàà	t(s)eu/ ceũ
Jowulu	tēēna/ tenn	fuuli	bzei < *jɔnn/i ?	p-f̥ɪrɛ	+	2+'to lose'	bʒi / byinn
SE East	+, gòon	*pela	ców, ʔàà- kɔ̃	sì/ síirí	*sodu	+ ?, 5+3, sípe	+, kwi
SE South	+	*piì- lāŋ	*yààká	*yìi- sìyá	sóódú	5+3	+, ko, sójólú

6.11 Mel

Table 6.11: NC numerals reflected in Mel (+)

	1	2	3	4	5	8	10
North	+	-rəŋ	+	-ŋkile/ -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-ta-mat* (< **kə-ta+suffix* < root *ta* ‘hand’?) or (< *kə-ta-m-at* < root *mV*).

The numeral system of the proto-language is generally poorly preserved in both of the Mel groups. However, 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-at* ‘10’.

6.12 Atlantic

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.

Table 6.12: NC numerals reflected in Atlantic

	1	2	3	4	5	8	10
Cangin	no	nak	haj	+	<'hand', ʔi:p	5+3	da:ŋkah
Nyun	+ ?	nak	+	+ ?	<'hand'	5+3	<'hands'
Buy	nɔʔ, tee-	naŋ	+	+ ?	+ ?	5+3	ntaajä
Jaad- Biafada	+	ke	jo/ caw	+	<'hand'	5+3	+
Tenda	+, mbɔ	ki	+	+	<'hand'	5+3	+
Fula- sereer	+	+	+	+	+ ʔ, jo(w)i	5+3	sapp- o, xarɓ-
Wolof	+	X-aar	+ ?	+	jurom	5+3	+
Nalu- BF- BMb	+, mbɔ	+	+	+	<'hand', ribə(l)	5+3	*a- lafan?
Joola	+	*- ɭubəʔ	-feegir	-bääkiŋ	+	5+3	- ntaaja
Manjak	+	-təb, puguɗ	-jenɗ	baakər	<'hand'	+	(n)taaja
Balant	-ɔdaʔ	*- ɭubəʔ	(borrowed)	tasala	jüf	+	jimmín
Bijogo	+?	*- ɭubəʔ	<'fingers'	-agenək	<'hand'	5+3	<'hands'

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 (**be-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 **pə-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 *-tubə?* ‘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 *baakər* ‘4’ is observable in Joola-Manjak.

The original term for ‘five’ has been preserved in numerous Joola dialects, including Bayot (Proto-Joola **fu-tə-k* ‘5’).

The Proto-pattern ‘8’ < ‘4’ has been preserved in Manjak (Mankanya *ŋi-bakir* ‘4’ > *bakr-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 *tállá* ‘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.

6.13 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 Pozdniakov & Segerer 2007).

Table 6.13: 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’.

The term for ‘five’ may reflect the alternative NC root **nu(n)* ‘5’ (Gola *nò̀̀nò̀̀*).

The forms for ‘five’ and ‘ten’ in the Koelle records include [-f]: *ta-sóóf* ~ *ka-sóóf* ‘5’, *koof* ‘10’.

Table 6.14: NC numerals reflected in Gola (+)

	1	2	3	4	5	8	10
Gola	guu`ŋ	tì-yèe(l)	+	+	+	5+3	zìyà

Table 6.15: NC numerals reflected in Limba (+)

	1	2	3	4	5	8	10
Limba	-nthe	+	+	+	bi-sōhi	5+3	kōhi

The form *bi-le* ‘two’ is noteworthy in that it may be interpreted as a direct reflex of NC **be-di* ‘2’.

6.14 Summary

The results of our reconstruction of the basic numeral terms are presented in Table 6.16.

- 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,

6 NC numbers as reflected in particular families, groups and branches

Table 6.16: 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	

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/ yam* (Proto-Bantu **kómì/ 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.

6.15 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 4, 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

- Bloemarts, Maarten & Père de Rasilly. 2012. *Dictionnaire boore*. Electronic Document.
- Boyd, Raymond. 1989. Number systems in the adamawa branch of Niger-Congo. *African Languages and Cultures* 2(2). 149–173.
- CLNK. 1999. *Tɔm kpou kabiyɛ-fransu - dictionnaire kabiyè-français (avec lexique français-kabiyè et esquisse de grammaire kabiyè)*. Kara, Lomé (Togo): Comité de Langue Nationale Kabiye (CLNK) & SIL.
- Connell, Bruce A. 1991. *Phonetic aspects of the Lower-Cross languages and their implications for sound change*. University of Edinburgh dissertation.
- Dimmendaal, Gerrit Jan. 1978. *The consonants of Proto-Upper Cross and their implications for the classification of the Upper Cross languages*. Department of African Languages, State University of Leiden dissertation.
- Elugbe, Ben Ohi[omambe. 1987. *Comparative edoid: Phonology and lexicon*. Port Harcourt: University of Port Harcourt Press.
- Ganong, Tina Weller. 1998. *Features of бага morphology, syntax and narrative discourse*. Arlington: University of Texas. (M.A. thesis).
- Hochstetler, Lee, Jude Durieux & Evelin Durieux-Boon (eds.). 2004. *Sociolinguistic survey of the dogon language area*. <http://hal.archives-ouvertes.fr/hal-00323718/fr/>.
- 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.
- Nurse, Derek & Gérard Philippon. 1975/1999. *The Tanzanian language survey*. <http://www.cbold.ish-lyon.cnrs.fr/>. Field work material, adapted as a searchable net-database.
- Paulin, Pascale. 1995. *Étude comparative des langues du groupe ring - langues grassfieldss de l'ouest, cameroun*. Lyon: Université Lumière - Lyon 2. (Mémoire de DEA).
- Pozdniakov, Konstantin & Guillaume Segerer. 2007. Similar place avoidance: A statistical universal. *Linguistic Typology* 11(2). 307–348.

References

- Stewart, John Massie. 1983. *Nasality in Tano-Congo*. Colloquium on African Linguistics, Leiden.
- Wilson, William André Auquier. 2007. *Guinea languages of the atlantic group*. Frankfurt: Peter Lang.

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.

