

George Starostin

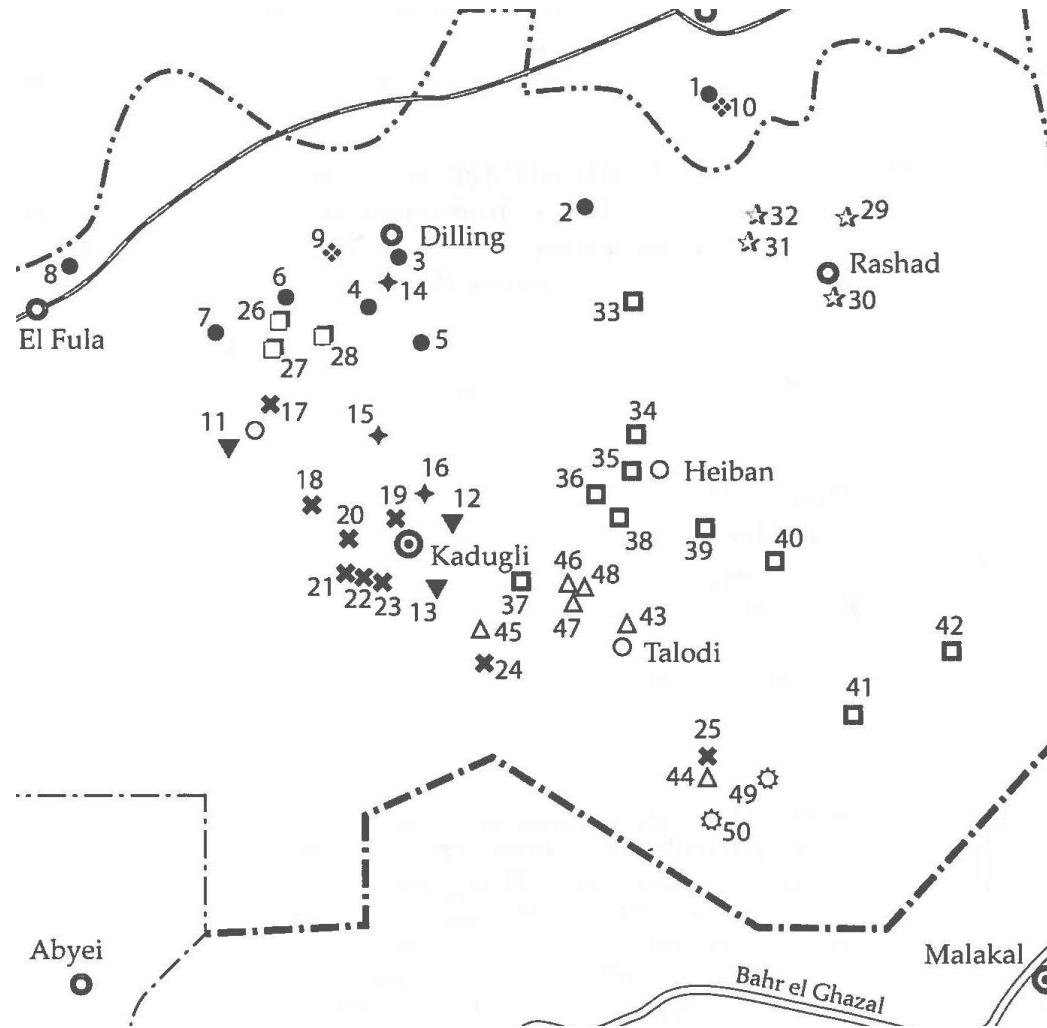
(Russian State University for the Humanities / Higher School of Economics / Russian Presidential Academy)

13th Annual Sergei Starostin Memorial Conference on Comparative-Historical Linguistics (RSUH, March 22-23, 2018)

## **Preliminary lexicostatistical analysis for languages of the Nuba Mountains**



# General scheme of linguistic diversity in the Nuba Mountains



[map from Schadeberg & Blench 2013]

## Ten primary taxa of the Nuba Mountains:

### Legend

- NUBIAN
  - 1 Dair
  - 2 Kadaru
  - 3 Dilling
  - 4 Ghulfan (Uncunwee)
  - 5 Karko
  - 6 Wali
  - 7 Tabag
  - 8 El Higeirat
- ❖ NYIMANG
  - 9 Nyimang (Ama)
  - 10 Afitti
- ▼ DAJU
  - 11 Lagawa
  - 12 Laggori
  - 13 Shatt
- ◆ TEMEIN
  - 14 Temein
  - 15 Keiga Jirru
  - 16 Tese
- ✕ KADU
  - 17 Tulishi
  - 18 Keiga
  - 19 Kadugli
  - 20 Miri
  - 21 Kanga
  - 22 Tumma
  - 23 Katcha
  - 24 Krongo
  - 25 Tumtum
- KATLA
  - 26 Katla
  - 27 Tima
  - 28 Julud
- ☆ RASHAD
  - 29 Tegali
  - 30 Tagom
  - 31 Tagoi
  - 32 Turjok
- HEIBAN
  - 33 Koalib
  - 34 Laro
  - 35 Ebang
  - 36 Shwai
  - 37 Moro
  - 38 Otoro
  - 39 Tira
  - 40 Lukha
  - 41 Werni
  - 42 Kau, Nyaro, Fungor
- △ TALODI
  - 43 Talodi
  - 44 Eliri
  - 45 Masakin, Dagig
  - 46 Acheron
  - 47 Tocho
  - 48 Lumun
- ⚙ LAFOFA
  - 49 Tekeim
  - 50 Al Amira

## Original attempts at classification: Greenberg 1963-66

### (a) Kordofanian:

1. Koalib group: Koalib, Kanderma, Heiban, Laro, Otoro, Kawama,<sup>1</sup> Shwai, Tira, Moro, Fungor.
2. Tegali group: Tegali, Rashad, Tagoi, Tumale.
3. Talodi group: Talodi, Lafofa, Eliri, Masakin, Tacho, Lumun, El Amira.
4. Tumtum group: Tumtum, Tuleshi, Keiga, Karondi, Krongo, Miri, Kadugli, Katcha.
5. Katla group: Katla, Tima.

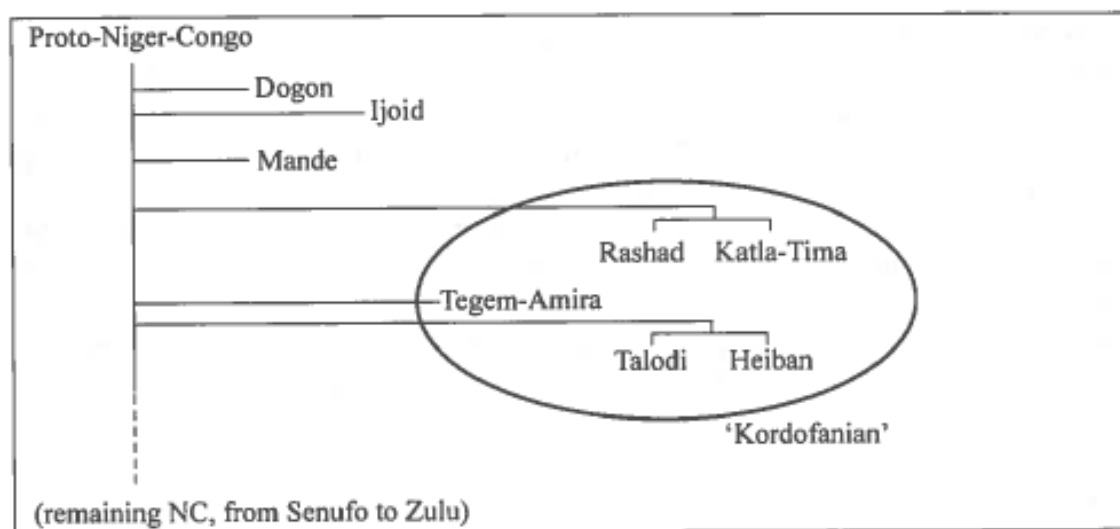
### (b) East Sudanic:

1. Nubian: (a) Nile Nubian (Mahas-Fadidja and Kenuzi-Dongola); (b) Kordofanian Nubian: Dair, Dilling, Gulfan, Garko, Kadero, Kundugr; (c) Midob; (d) Birked.
- ....
5. Nyima, Afitti.
6. Temein, Teis-um-Danab.
- ....
8. Dagu of Darfur, Baygo, Sila, Dagu of Dar Dagu (Wadai), Dagu of Western Kordofan, Njalgulgule, Shatt, Liguri.

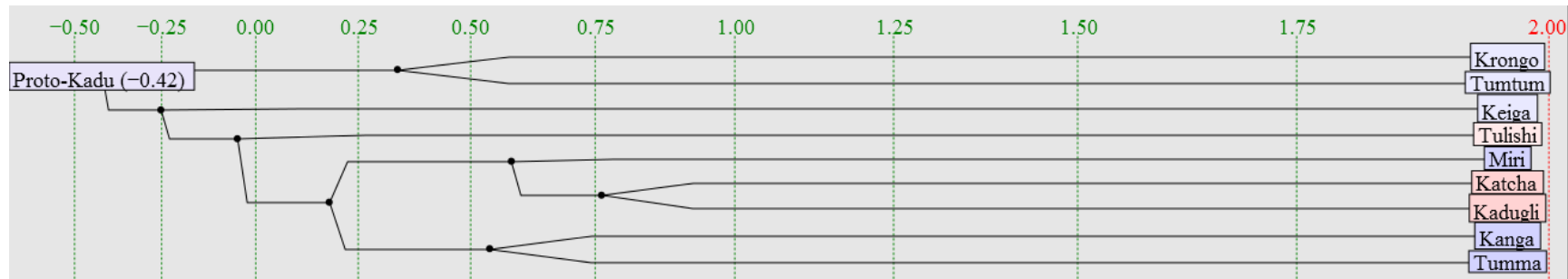
## Reclassification of Kordofanian: Blench 2013

- «Kordofanian 1»: Heiban + Talodi  
«Kordofanian 2»: Rashad + Katla-Tima  
Tegem-Amira (Lafofa): a separate entity, but close to «Kordofanian 2»

*Blench's general scheme of Niger-Congo,  
based on lexical and grammatical (noun classes) evidence*



## Special problem: Kadugli-Krongo languages



Greenberg: part of Kordofanian

Schadeberg 1981: separate branch of Nilo-Saharan

Ehret 1995: no, Niger-Congo!

Blench 2006: no, Nilo-Saharan!

Starostin 2017: not exactly «Nilo-Saharan», but possibly connected to parts of the original «Nilo-Saharan» (most importantly, Central Sudanic)

## The «splitter» view: Glottolog 3.2

Ten top-level families:

Dajuic, Nubian, Nyimang, Temeinic, Kadugli-Krongo, Heibanic, Narrow Talodi, Lafofa, Rashad, Katla-Tima

— no conclusive evidence to assert genetic relationship of *any* two units

E.g.:

*Narrow Talodi (excluding Tegem) is held together by pronouns, lexical cognates, and form-meaning resemblances in the noun class systems. Tegem systematically stands out lexically, pronominally, in the noun class system and alleged sound correspondences are often irregular. The alleged noun class form/meaning correspondences to other Kordofanian groups, namely Rashad and Heibanic, in [Schadeberg, Thilo 1989](#), [Schadeberg, Thilo 1981](#) are not convincing as conclusively genetic; As to form, there is much eclectic selection as well as irregular correspondences, and as to meaning, the methodology of finding non-cognate shared class items is highly dubious. Chance resemblance has not been excluded. The lexical evidence for a Kordofanian unity is weak in the extreme [Blench, Roger M. 2006](#). It follows that the arguments presented for a Niger-Congo affiliation are not compelling either. — Glottolog 3.2 (<http://glottolog.org/resource/languoid/id/narr1279>)*



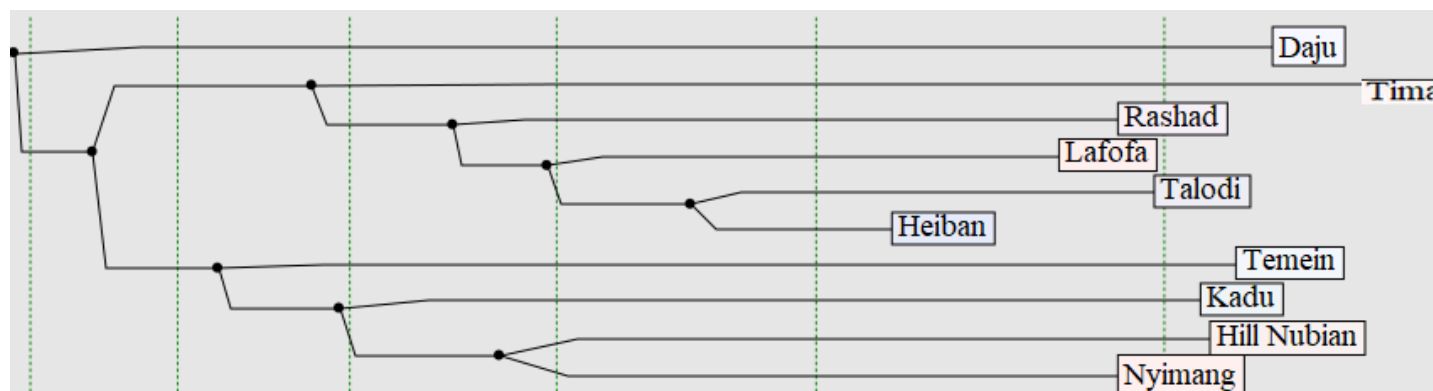
# Preliminary lexicostatistics: Nuba Hill lexical data in the Global Lexicostatistical Database

Number	Word	Rere Koalib	Ebang	Abul	Laru	Utoro	Shirumba	Tiro	Moro	Ko	Warnang
1	all	riðèt <sup>1</sup>	pét <sup>2</sup>	pét <sup>2</sup>	pér <sup>4</sup>	kúrɛɲ <sup>3</sup>	kəɾubɛlən <sup>3</sup>	tárra <sup>4</sup>	prét <sup>2</sup>	=ɓáp <sup>5</sup>	=iðdám <sup>6</sup>
2	ashes	y=ári <sup>1</sup>	árò <sup>1</sup>	arɔ <sup>1</sup>	ɲw=arú <sup>1</sup>	ɟ=aru <sup>1</sup>	árò <sup>1</sup>	iriðà <sup>2</sup>	ùβú:ná <sup>3</sup>	ɛwún <sup>3</sup>	kú:gá <sup>4</sup>
3	bark	k=íaw <sup>1</sup>	g=iú <sup>1</sup>	g=iýú <sup>1</sup>	ɟ=iýú <sup>1</sup>	g=iya <sup>1</sup>	ábúgúr <sup>2</sup>	ð=iyyú <sup>1</sup>	ð=isiá <sup>2</sup>	k=iýé <sup>1</sup>	k=iya <sup>1</sup>
4	belly	k=à:ri <sup>1</sup>	k=á:ɾé <sup>1</sup>	g=à:ré <sup>1</sup>	l=à:ri <sup>1</sup>	g=à:ré <sup>1</sup>	ì:rɪɲ <sup>1</sup>	ári <sup>1</sup>	ára <sup>1</sup>	è=wɪɲ <sup>2</sup>	à:riɲ <sup>1</sup>
5	big	=úppà <sup>1</sup>	=ippà <sup>1</sup>	=ippà <sup>1</sup>	=ppà <sup>1</sup>	=bò:ðò <sup>2</sup>	=bú:qò <sup>2</sup>	=ɾá <sup>3</sup>	=ù:ɾá <sup>3</sup>	=bát <sup>4</sup>	áɽ-ámmàn
6	bird	ð=íaw <sup>1</sup>	ɲ=iɾá <sup>1</sup>	ɲ=iɾá <sup>1</sup>	k=ðɾá <sup>1</sup>	ɲ=iúwá <sup>1</sup>	ɲ=iúvɪá <sup>1</sup>	ù=iúvò <sup>1</sup>	ù=iúfiá <sup>1</sup>	ɲò=iúwé <sup>1</sup>	ɲ=iɾwá <sup>1</sup>
7	bite	ì-ðu <sup>1</sup>	itt-i <sup>1</sup>	itt-i <sup>1</sup>	itt-ú nwàqúné <sup>1</sup>	=èrùm # <sup>2</sup>					
8	black	=únúná <sup>1</sup>	=únúnán <sup>1</sup>	=ámə̀nə̀ɲ <sup>1</sup>	=únna <sup>1</sup>	=únə̀nán <sup>1</sup>	ú:mə̀nà <sup>1</sup>	=ámə̀nə̀ <sup>1</sup>	=múnúá <sup>1</sup>	=ɲə̀nɛ̀túú <sup>1</sup>	=únàn <sup>1</sup>
9	blood	ɲ=in <sup>1</sup>	ɲ=in <sup>1</sup>	ɲ=in <sup>1</sup>	y=in <sup>1</sup>	ɲ=in <sup>1</sup>	ɲ=ə̀vɪn - ɲ=ə̀vɪvɪn <sup>1</sup>	ɲ=ə̀vɪn <sup>1</sup>	ɲ=ə̀fáni <sup>1</sup>	ɲ=ún <sup>1</sup>	ɟ=unú <sup>1</sup>
10	bone	l=iya <sup>1</sup>	l=ú <sup>1</sup>	l=óy <sup>1</sup>	g=úyá <sup>1</sup>	l=úyá <sup>1</sup>	l=iá <sup>1</sup>	l=úðá <sup>1</sup>	l=ðiá <sup>1</sup>	k=wɔɲ <sup>2</sup>	k=úr <sup>2</sup>
11	breast	kw=òrdóm <sup>1</sup>							l=urumia # <sup>1</sup>		
12	burn tr.	w=údu-nnò # <sup>1</sup>	ì-ɟú-nnù # <sup>1</sup>	w=úd-i # <sup>1</sup>	ɲú=m=ú:d-i # <sup>1</sup>	yá=wùd-é # <sup>1</sup>	y=ù:d-ò # <sup>1</sup>	yá=úd-à # <sup>1</sup>	l=ò:d-à # <sup>1</sup>	ɟ=ɔ̀nɔ̀:ɾan # <sup>2</sup>	gi=má=y-á <sup>1</sup>
13	claw(nail)	k=iɾiɾi <sup>1</sup>	k=ámún <sup>2</sup>	ɟ=ámún <sup>2</sup>		g=ámún <sup>2</sup>	úmàr <sup>3</sup>	id=apɾá <sup>3</sup>	id=apɾá <sup>3</sup>	k=əmàn <sup>2</sup>	ábɪɲé <sup>4</sup>
14	cloud	l=áblét <sup>1</sup>	g=ibá:ru <sup>2</sup>	Δβá:ru <sup>2</sup>	k=éré <sup>3</sup>	ábáru <sup>2</sup>	ɲ=iɾɪɲ <sup>3</sup>	ɲ=ɔ̀mè <sup>4</sup>	ib <sup>0</sup> :ɾuá <sup>2</sup>	kúliɾi <sup>5</sup>	
15	cold	=ði <sup>1</sup>	=ɟiɾib <sup>2</sup>	=iɟilò <sup>2</sup>	=iɟilú <sup>2</sup>	=ɪɲidɪɲ <sup>3</sup>	=iɟilb <sup>4</sup>	=mòlò <sup>5</sup>	=mòlì <sup>5</sup>	=ɔ̀tɔ̀ <sup>6</sup>	
16	come	il-á <sup>1</sup>	il-á <sup>1</sup>	il-á <sup>1</sup>	il-á <sup>1</sup>	il-á <sup>1</sup>	è:l-á <sup>1</sup>	il-á <sup>1</sup> / k=ét-ò <sup>4</sup>	è:l-án <sup>1</sup> / g=ét-ú <sup>4</sup>	apá <sup>2</sup> / ú=pinú <sup>5</sup>	kóy <sup>3</sup> / ɟ <sup>6</sup>
17	die	kú=m=áy - kw=áy-ú <sup>1</sup>	gw=áy <sup>1</sup>	gù=m=áy <sup>1</sup>	gwù=m=áy <sup>1</sup>	m=áy <sup>1</sup>	m=áy-é <sup>1</sup>	g=áy-ú <sup>1</sup>	g=à:y-ú <sup>1</sup>	w=àqə̀-ú <sup>1</sup>	ɲ=é:ɟ-i <sup>1</sup>
18	dog	t=iɲən <sup>1</sup>	ɲ=in <sup>1</sup>	ɲ=in <sup>1</sup>	kə̀-ɲən <sup>1</sup>	ɲ=in <sup>1</sup>	ɲ=in <sup>1</sup>	ɲ=é:n - ɲ=é:né <sup>1</sup>	ɲ=inà <sup>1</sup>	t=ón <sup>1</sup>	ɲ=án <sup>1</sup>
19	drink	ì-ðu <sup>1</sup>	iyy-ò <sup>1</sup>	gw=iyy-i <sup>1</sup>	iyy-ú <sup>1</sup>	y-ú <sup>1</sup>	iyy-ú <sup>1</sup>	ɟ-ò <sup>2</sup>	ɟ-ú <sup>2</sup>	iy-ò <sup>1</sup>	idéɟò <sup>1</sup>
20	dry	=ndú <sup>1</sup>	=úqðò <sup>1</sup>	=úndò <sup>1</sup>	=úndú <sup>1</sup>	=úndə̀ɲ <sup>1</sup>	=únd <sup>1</sup>	=úndòttú <sup>1</sup>	=òndòttú <sup>1</sup>	=ɔ̀ɲ <sup>2</sup>	=úwɪɲ <sup>2</sup>
21	ear	k=á:ni <sup>1</sup>	k=é:ni <sup>1</sup>	g=é:ni <sup>1</sup>	l=é:ni <sup>1</sup>	g=é:ni <sup>1</sup>	ə̀ni <sup>1</sup>	áni-ɲò <sup>1</sup>	ánúɲá <sup>1</sup>	k=ə̀ɾi <sup>2</sup>	w=ú:ni <sup>1</sup>
22	earth	wúrɛy <sup>1</sup>	gw=iýán <sup>2</sup>	gw=iýán <sup>2</sup>	ɟ=ián <sup>2</sup>	gw=iýóm <sup>2</sup>	ámúɾá # <sup>3</sup>	ð=ú:ðò <sup>4</sup>	ð=ú:ð <sup>4</sup>	k=ə̀bá <sup>5</sup>	k=ú:ru <sup>6</sup>
23	eat	ì-ði <sup>1</sup>	ì-tti <sup>1</sup>	ì-tti <sup>1</sup>	ì-tti <sup>1</sup>	ì-tti <sup>1</sup>	ì-tti <sup>1</sup>	y-ò <sup>1</sup>	g=às-à <sup>2</sup>	é <sup>1</sup>	yà:-qð <sup>1</sup>
24	egg	l=ɔ̀ɲ <sup>1</sup>	l=ɔ̀ɲ <sup>1</sup>	l=iɲná <sup>1</sup>	k=é:ɲò <sup>1</sup>	l=ɔ̀ɲ <sup>1</sup>	l=é:ɲə̀ɲ <sup>1</sup>	l=é:ɲé <sup>1</sup>	l=é:ɲá <sup>1</sup>	l=ɔ̀ɲ <sup>1</sup>	ɔ̀ɲé <sup>1</sup>
25	eye	l=éy <sup>1</sup>	k=éy <sup>1</sup>	g=éy <sup>1</sup>	l=éy <sup>1</sup>	g=éy <sup>1</sup>	l=éy <sup>1</sup>	l=éy <sup>1</sup>	issi <sup>2</sup>	k=i <sup>1</sup>	k=i <sup>1</sup>
26	fat n.	ɲ=é:lá <sup>1</sup>	ɲ=é:lá <sup>1</sup>	ɲ=é:lá <sup>1</sup>	y=i:lá <sup>1</sup>	ɲ=é:lán <sup>1</sup>	ɲ=é:lá <sup>1</sup>	ɲ=é:lá <sup>1</sup>	ɲ=é:lá <sup>1</sup>	ɲ=é:ɾá <sup>1</sup>	ɟ=é:ɾá <sup>1</sup>

[Part of the 110-item wordlist for Heiban languages. Different colors mark phonetically similar forms, not cognates.]



**StarLing tree based on automated lexicostatistics  
(phonetic similarity according to the consonant class method):**



*Identified:*

- 1) Hill Nubian + Nyimang = (partial) Northeastern Sudanic
- 2) Heiban + Talodi = «Nuclear» Kordofanian
- 3) «Nuclear» Kordofanian + Lafofa, Rashad, Tima = «Macro»-Kordofanian

*Not identified:*

Northeastern Sudanic + Temein + Daju = Eastern Sudanic

## Analyzing the situation with Kordofanian:

### 1) Averaged percentage matches between the various subbranches of Kordofanian

	Talodi	Lafofa	Tima	Rashad	Kadu
Heiban	20%	15%	8%	15%	6%
Talodi		11%	9%	12%	6%
Lafofa			5%	10%	7%
Tima				12%	5%
Rashad					6%

## 2) Some examples of common basic Kordofanian isoglosses:

	'eye'	'bone'	'egg'	'tooth'	'ear'	'tongue'	'head'
Heiban	*=ey	*=uya	*=εɲɔ(η)	*=iŋa-t̥	*=ε:ni	*=ηɛla	*=da
Talodi	*=igi	*=umV	*=uwi(η)	*=iɲi-t	*=ε:nu	*=lɔŋɛ	*=âs
Lafofa	=î	=uami	=uwε-	=ε:ɲi		liɲi	=ay
Katla-Tima	*=igi	*=uga	*=iwɔŋ		*=ɔŋɔ	*=liɲi	*=as
Rashad	*=i(η)gə	*=uh	*iye (?)	*=iɲi-n	*=əni-n	*=ηəla	*=as
Cf. Kadu:	*=yV	*kuba	*sule	*=kini	*=ε:ɔ	*ηadɔ	*=aɬu

### Pronouns:

	'I'	'thou'
Heiban	*(i)ɲ=i	*(u)η=a
Talodi	*η=i	*(u)η=a
Lafofa	ɲε-	ηɔ-
Katla-Tima	ɲɔŋ	ηaŋ
Rashad	*η=i	*η=ɔ
Kadu	*aʔa	*oʔo

## The noun class argument against Kordofanian (Blench 2013):

- Talodi and Heiban have a well-developed system of noun classes
- Lafofa (Tegem-Amira) has a somewhat different system of noun classes
- Katla-Tima has a very limited amount of sg./pl. nominal prefixes
- within Rashad, Tagom (Rashad proper) has no class prefixes whatsoever; the noun class system of Tagoi (Orig) is quite different from the one in Talodi/Heiban

Cf.:

	Rashad		Tagoi	
	sg.	pl.	sg.	pl.
'bird'	uden	uden-e	w=ut	y=ud-εn
'egg'	ye	ye-ndε	=iyε	η=iyε
'eye'	git	gid-e	yi=git	ηə=git
'fire'	ibe		k=ibe	h=ibe
'hair'	aam		k=aam	h=aam
'tooth'	ɲin	ɲin-e	t=iɲin	y=iɲin

Consequently:

Rashad + Katla-Tima viewed as early splits from the Niger-Congo tree (prior to the formation of noun class system).

## Conclusions:

1. There are at least three *very* distinct genetic lineages in the Nuba Mountains: East Sudanic, Kordofanian, and Kadu. This is partially confirmed by automated lexicostatistics and partially by preliminary attempts at reconstructing small subsets of basic lexicon for all involved subgroups.
2. Areal contact, involving occasional borrowing of even the most basic items and typological restructuring of grammatical systems (mainly nominal morphology) is a very common issue in the Nuba mountains, where language shifts must have been quite common.
3. Lexical evidence for the genetic unity of all potential Kordofanian subgroups is currently stronger than grammatical evidence. However, it still remains to be ascertained whether the observed isoglosses are exclusively Kordofanian (and not shared with other branches of Niger-Congo).