OXFORD

The Oxford Guide to the Atlantic Languages of West Africa

EDITED BY

FRIEDERIKE LÜPKE

OXFORD GUIDES TO THE WORLD'S LANGUAGES

THE OXFORD GUIDE TO THE ATLANTIC LANGUAGES OF WEST AFRICA

OXFORD GUIDES TO THE WORLD'S LANGUAGES

PUBLISHED

The Oxford Guide to the Atlantic Languages of West Africa
Edited by Friederike Lüpke

The Oxford Guide to Australian Languages
Edited by Claire Bowern

The Oxford Guide to the Languages of the Central Andes
Edited by Matthias Urban

The Oxford Guide to the Malayo-Polynesian Languages of Southeast Asia
Edited by Alexander Adelaar and Antoinette Schapper

The Oxford Guide to the Romance Languages
Edited by Adam Ledgeway and Martin Maiden

The Oxford Guide to the Transeurasian Languages
Edited by Martine Robbeets and Alexander Savelyev

The Oxford Guide to the Uralic Languages
Edited by Marianne Bakró-Nagy, Johanna Laakso, and Elena Skribnik

IN PREPARATION

The Oxford Guide to the Bantu Languages
Edited by Ellen Hurst, Nancy Kula, Lutz Marten, and Jochen Zeller

The Oxford Guide to the Papuan Languages
Edited by Nicholas Evans and Sebastian Fedden

The Oxford Guide to the Slavonic Languages
Edited by Jan Fellerer and Neil Bermel

The Oxford Guide to the Tibeto-Burman Languages
Edited by Kristine Hildebrandt, Yankee Modi, David Peterson, and Hiroyuki Suzuki

THE OXFORD GUIDE TO THE

Atlantic Languages of West Africa

EDITED BY

Friederike Lüpke

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2024910095

ISBN 9780198736516

DOI: 10.1093/oso/9780198736516.001.0001

Printed and bound by
CPI Group (UK) Ltd, Croydon, CR0 4YY

To the memory of G. Tucker Childs (1948–2021)
and
Alpha Naby Mané (1964–2021)

Short contents

Detailed contents

Series preface

We know that the close study of individual language families and linguistic areas is vital both to the synchronic and diachronic study of language and to cognitive science more widely. Comparative investigations of this type stimulate exciting synergies between different subdisciplines of linguistics, such as language change, contact linguistics, sociolinguistics, linguistic typology, textual philology, and microvariation in grammar, sound, and meaning within and across languages. Besides reflecting and encouraging the links between these subdomains, the fundamental goal of the series is to publish high-quality, substantial reference works which represent a set of theoretically informed and systematic guides to what is known about the world's languages.

Each *Guide* focuses on a particular language family, subfamily, or areal grouping, and is edited by leading authorities, who bring together contributions from the best international scholars in the field. The *Guides* aim to show the more general theoretical significance of the languages' history, linguistic and sociolinguistic characteristics, and overall to provide an indispensable reference tool both to specialist scholars and students and to professional linguists. The approach adopted in all the *Guides* is systematic and comparative, informed by the latest research and theoretical and methodological perspectives, and, where appropriate, the authors draw on relevant work in such fields as anthropology, archaeology, and cognitive science.

Adam Ledgeway and Martin Maiden
University of Cambridge and University of Oxford

List of abbreviations

.	marks syllable boundaries when inserted into Nalu words (Chapter 17)
:	segmentable morpheme when the text is not segmented
_	stem that never appears without suffixes
<	from
I, II, III	consonant mutation grade
Ø	zero marker
1, 2, 3	first, second, third person
1, 2, 3 . . .	agreement classes (Chapter 5)
1, 2, 3 . . .	noun classes (Chapter 30)
A	agent
ABC	Aruba-Bonaire-Curaçao
ACC/Acc	accusative
ACT	active; actualizer (Chapter 9)
ad	Adamawa
ADJ	adjective
ADP	adposition
ADV	adverb
ADVCOP	adverbial copula
ADV.PART	partitive adverb (*ci*)
AF	assertive focus
AG	agent
AGNM	agent nominalization
AGR/AGR	agreement
AGR/AGR	agreement class (Chapter 5)
AGR/AGR	agreement marker (Chapters 14, 17, 18)
AGR0	agreement marker consisting of a single consonant
AGR1	agreement marker (*-ʊ, -ɪ, ɛ-* series)
AGR2	agreement marker (*-a* series)
AGT	agentive
ALTLOC	altilocal
ALV	alveolar
AM	Associated Motion
ANAPH	anaphoric
ANC	action nominal construction
ANT/ANT	anterior
ANT	anteriority (Chapter 24)
ANTCP	anticipative
ANTER	anterior suffix
ANTIP/ANTIP	antipassive
AOR/AOR	aorist
APPL/APPL	applicative
APPL1	applicative verb suffix for role of recipient, beneficiary, or companion
APPL2	applicative verb suffix for role of instrument, manner, or location
APTH	apenthesis
ar	Arabic
ART	article
ASAL	argument salience
ASP/Asp	aspect
ASS	associative
ATR	Advanced Tongue Root
ATTR	attributive
AUX/Aux/AUX	auxiliary
AW	archaic Wolof (as attested in nineteenth-century documents)
B	Bühnen
BAL	Balanta Ganja
BALf	Balanta Fraase
BALj	Balanta Jabada
BALk	Balanta Kentohe
Bamb.	Bambara
BAPPL	benefactive applicative
Ben/BEN/BEN	benefactive
BF	Baga Fore
BGR	backgrounding marker
BK	Baga Koba
BLS	Bliss
BNJ	Banjal
BW	Banjul Wolof
C/C	consonant
CAUS/CAUS	causative
CAUS1	causative verb suffix for direct causation (*-al*)
CAUS2	alternate causative verb suffix for direct causation (*-e*)
CAUS3	causative verb suffix for sociative–assistive causation (*-le*)
CAUS4	causative verb suffix for omitted causee (*-lu*)
CAUS5	causative verb suffix for indirect causation (*-loo*)

CC	Casamance Creole
CDS	child directed speech
CEL	celerative
CEN	Central-Eastern Niger
CENT	central varieties
CFG	centrifugal verb suffix
CIR/CIR	circumstantial
CL	class (Chapter 18)
CL	class (cell in the inflectional paradigm of agreement targets) (Chapter 15)
CL	class agreement marker (Chapter 13)
CL	class membership marker (Chapter 17)
CL/CL/CL	noun class (Chapters 5, 9, 12, 22, 24)
CL	noun class prefix (Chapters 10, 11, 14, 23)
CLPF	class prefix
clX/CLX	class marker of class x
CM1/CM2	
CMM	class membership marker
CMP	completive
COLL	collective
COME&DO	prior centripetal associated motion marker
COMP/COMP	complement(izer)
COMP	completive (Chapter 17)
COMPL	complementizer (Chapters 6, 13)
COMPL	completive (Chapter 11)
COND	conditional
COND	conditional or temporal verbal marker (Chapter 11)
conj/Conj/ CONJ/CONJ	conjunction
CONN/CONN	connector/connective
COORD	coordination
COORD.V	verbal coordinator
COP/COP	copula
COP.LOC	locative copula
CP	centripetal
CPL/CPL	completive
CPLT	complement
CS/CS	causative
CTF	centrifugal
CTP	centripetal
CTP.RES	resultative centripetal
CTRP	centripetal
CurTR	current time region
CV	Cape Verdean (Chapter 6)
CV	consonant–vowel
CVB	converb
CvP-model	Creolization vs. Pidginization-Model
D	default determiner (Chapter 7)
D	definite (Chapter 21)
DAT	dative
DD	Deictic Directional
DEF/DEF	definite
DEF	definite suffix (Chapter 8)
DEM/DEM	demonstrative
DEM	demonstrative determiner (Chapter 8)
DEMI	identificational demonstrative
DEMm	demonstrative of manner
DER	derivation
DER	derivational suffix
DESCR	descriptive
DET	default determiner
DET	determiner
DETR	detransitivizer
DetTR	detached time region
DF	default
DIM	diminutive
DIR	directional
DIST/DIST	distal
DIST	distal spatial suffix (Chapter 4)
DIST	distributive (Chapter 17)
DisTR	distal time region
DISTR/DISTR	distributive
DJ	disjoint
DKB	Documenting Kim and Bom (project)
DP	discourse particle (Chapter 17)
DP	distant past (Chapter 18)
DVN	deverbal noun
DW	Dakar Wolof
DX	deictic marker
EAST	eastern varieties
EJM	Ejamat
ELDP	Endangered Languages Documentation Programme
EMPH/EMPH	emphatic
EP	epenthetic vowel
EPENT	epenthetic consonant
EPENTH	epenthesis/epenthetic
ESL	Esuulaalu
EvN	event noun
EXC	exclusive
EXCL/EXCL	exclusive
exp	phrase; expression
EXPECT	expectative
F	feminine

FACTC	factice
FNY	Fogny
FOC	focalizer (Chapter 21)
FOC/Foc/FOC	focus
FOC/FOC	focalization (Chapters 7, 24)
FOC	subject-focus copula (Chapter 4)
FOCCOMP	Complement-focusing conjugation
FOCS	Subject-focusing conjugation
FOCV	Verb-focusing conjugation
fr/Fr./FR/FR	French
FUB	Fula Bible
FUT/FUT	future
FUT	Future conjugation (Chapter 4)
FUT1	immediate future
FUT2	delayed future
FUT3	predictive future
FV	final vowel
GA	Gaawoore
GBC	Guinea-Bissau Creole
GC	genitive constructions
Gen/GEN	general (Chapters 19, 20)
GEN	genitival morpheme (Chapter 11)
GEN/GEN	genitive
GEN(.SG)	genitival suffix (*-u* SG, *-i* PL)
GER/GER	gerund(ive)
Gl	Gola
GLOT	glottalic
GO	*Gombe*
GO&DO	prior centrifugal associated motion marker
GSL	Gusilay
H	high (tone)
Hab/HAB/HAB	habitual
hau	Hausa
h/d	hours per day
HES	hesitation
HPX	Hapax
HUM	human
HYP	hypothetical
IAPPL	instrumental applicative
ICA	initial consonant alternation
ICESCO	Islamic Educational, Scientific and Cultural Organization
ICPL/ICPL	incompletive
ID	identification (marker)
IDEF	indefinite
IDEN	identificational
IDEOPH	ideophone
IFV	imperfective
IMP/IMP	imperative
Imperf/IMPERF	imperfective
IMPERS	impersonal
IMPFV	imperfective
IMPS	impersonal suffix
IMPV	imperfective
INACT/INACT	inactual(is)
INC	inclusive
INCHO	inchoative
INCL/INCL	inclusive
INCOMP	incompleteness
INCPL	incompletive
INDEF	indefinite
INDEP	independent
INDET	indetermined
INDF	indefinite article
INEV	inevitability marker
INF/INF	infinitive
INJ	injonctive
INSNM	instrument nominalization
int/INT	interrogative
INTER	interrogative
INTERR/INTERR	interrogative
INTJ	interjection
INTR	intransitive
INTRG	interrogative
INTSV	intensive
INV	reversive
IP	impersonal (Chapters 17, 18)
IP	interrogative pronoun (Chapter 5)
IPA	International Phonetic Alphabet
IPFV/IPFV	imperfective
IPFVCOP	imperfective copula (*di* ~ *d-*)
IRR/IRR	irreal(is)
IT	iterative
ITER	iterative/iteration
ITR	intransitive
JBK	Djibonker Nyun
JBL	Djibelor Nyun
JFG	Djifanghor Nyun
JLD	Jean-Léonce Doneux
JUS	jussive
Kik.	Kikongo
Km	Kim
KRK	Keeraak
KRN	Karon
Ks	Kisi

KSA	Kasa
KWT	Kwaatay
L	Lespinay (Chapter 13)
L	low (tone)
LAB	labial
Lab-Vel/ LAB-VEL	labial-velar
LE	Linking element
LEX	stem
LINK	verbal linker
LIP	*Liptaako*
lit.	literally
LK	linker
LNK	verbal linker
loc	localization word (Chapter 16)
LOC/Loc/loc	locative
LOC/loc	locative preposition (Chapters 4, 11)
LOC.PART	locative/partitive pronoun (*ci*)
LOL	laugh out loud
LZ	linguistic zone
M	masculine
MA	*Maasina*
MCI	*merci*
MDK	Mandinka
Mdka(.)	Mandinka
MDR	*mort de rire*
MED	medial
Mid/MID/MID	middle
MIDD	middle
MK	marked
MLP	Mlomp
Mm	Mani
MNR	deverbative nominal suffix for manner
MO	*mɔ* particle
MOD/MOD	mode/modality
MOVE&DO	prior associated motion marker
n.	n-form (for bound personal subjects) (Chapter 11)
n/n./(n)/N/N	noun
N	nasal (Chapters 17, 18)
NA	non-attested
NAR	narrative
NARR	narrative
NA/U	non-attested or unrelated form
NC	Niger-Congo
NC	noun class (Chapter 22)
Nc	noun phrase or compound noun (Chapter 16)
NCA	nominal class agreement
NCM	noun class marker
NCP	noun class pronoun
NE	Northeastern Nyun
Neg/NEG/ NEG /NEG	negative/negation
NEG(.3SG)	Negative perfective conjugation (3sg)
NEGEMPH	Emphatic negative conjugation
NEG.HAB	negation of habitual
NEG.REL	negation in relative clause
NEUT	neutral
NEUTR	neutral
NF	noun form
NFIN	non-finite
n.l.	not listed
NL	Dutch
NLS	Noon-Laala-Saafi
NN	number neutral
NO	*Noŋaare*
NOMCOP	nominal copula
NOMR	nominalizer
NOTLOC	spatial suffix (*-u*) indicating absence of localization
NP/NP	noun phrase
NP	proper noun (Chapter 16)
NTM	New Tribes Mission
NULL	Null tense conjugation
num/NUM	numeral
NYM	Niamone Nyun
O/O	object
O(.3SG)	object pronoun
o, ɓe, ngel…	class markers
O, BE, NGEL…	noun form classes
OBJ/OBJ/OBJ	object
OBL	oblique (Chapter 6)
OBL	obligative mood (Chapter 8)
OBLG	obligative
OBLIG	obligative conjugation
oI	object index
OM	object marker
ONM	obligatory noun marker
OP/O_{PRO}	pronominal object
ORD	ordinal (marker/number)
os.	oneself
OST	ostensive
OUT&DO	prior associated motion marker
P	patient (Chapter 7)
P	plural (Chapter 10)

P	Pukur
PAL	palatal
Pap.	Papiamentu
PART	particle
PAS	passive voice
PASS/Pass/ PASS	passive
PAST	past
PASV	passive
PERF/perf	perfect(ive)
PFV/pfv	perfective (aspect)
PHB	prohibitive
pl/PL/Pl/pl	plural
PLA	pluractional
PLC	calculable plural
PLD	plural dual
PLE	plural exclusive
PLI	incalculable plural (Chapter 13)
PLI	plural inclusive (Chapter 8)
PLUR	pluractional
PM	personal marker (Chapter 13)
PM	predicative marker (Chapter 7)
PN	pronoun
POLY	single morpheme with more than two functions
POS	positive
POSS/poss	possession/possessive
POSS. V	possessive verb suffix
POSS(.1SG)	(1SG) possessive determiner
POSTP	postposition (The generic gloss POSTP is used for multifunctional postpositions with a range of uses that cannot be analysed straightforwardly as extensions of an identifiable basic meaning)
POT	potential
PP	personal pronoun
PPI/D	plural dual or inclusive
PPRO	past progressive
pr	personal pronoun (Chapter 16)
PR	pronoun
PRED	predicative
PREF/pref	prefix
PREF	noun class prefix (Chapter 8)
PREP/prep	preposition
PRES	presentative (Chapters 10, 24)
PRES	present tense (Chapters 17, 18)
PRESTL	presentational
PRF/prf	perfect
PRG	progressive
Pro/PRO/ Pro/pro	pronoun
PRO	independent personal pronoun (Chapter 4)
PROG/prog	progressive
PROH	prohibitive
PROH(.2SG)	prohibitive conjugation
PRO.HU	*hu*-class pronoun
PRON/pron	pronoun
PROP	property
PROSP	prospective
PROX/ Prox/prox	proximal
PROX	proximal spatial suffix (*-i*) (Chapter 4)
PROX	proximate (Chapter 24)
PRS	present
PRST	presentative conjugation (discontinuous morpheme)
PRT	particle
PRTCL	particle
PSPH	'in the personal sphere of'
PST/pst	past (suffix)
PST.REM	remote past suffix
PT	Classical Portuguese
PTCL	particle
PTCP	participle
PU	Pular
PUL	Pulaar
PUNC	punctual
PUNCT	punctual
PURP	purposive
Q/q	interrogative/question particle (Chapters 5, 15)
Q	interrogative (Chapter 7)
Q	question marker/word (Chapters 4, 6, 13)
qlt	quality word
qnt	quantity word
QNT	quantizer
QNT.ALT	alterity quantifier
QNT.TOT	totalizing quantifier
Q.PLR	question particle for polar questions
QUAL	qualifier/qualifying modifier
QUANT	quantifier
QUOT/quot	quotative
R	recipient; rural (chapter 6)
RBT	*ree ba tas*
RDP	reduplication

RDPL	reduplication
REC/REC	reciprocal
RECIP/RECIP	reciprocal
RED	reduplicated form
REDUP	reduplication
REFL/REFL	reflexive
REL/REL/REL	relative/relativizer
REL	relative linker (Chapter 9)
REM.FUT	remote future
REPET	repetition
RESULT/RESULT	resultative
REV/REV	reversive
RP	recent past
RT	reference time
RUF	Revolution United Front
RVRS	reversive
S	singular (Chapters 10, 23)
S/s	subject
s	subject index (Chapter 8)
Ṡb	Sherbro
SBJ	subject
SBJV/SBJV	subjunctive
SD	situational dependency
SEQ	sequential
sg/SG/SG/SG	singular
sI	subject index
SIL	Summer Institute of Linguistics
SIM	simulative (Chapter 5)
SIM	verb suffix for simultaneous action (Chapter 4)
SM	subject marker
SNS	standard numeral system
s.o.	someone
SOC	sociative/associative
SODÉCOTON	*Société de Développement de Coton*
SP	shared inside perspective (Chapter 18)
SP	Spanish (Chapter 6)
Sp	speaker (Chapter 18)
SP, sp.	species (Chapter 4)
sth.	something
SUB	subjunctive (Chapter 12)
Sub/SUB	subordinating/subordinator
SUBJ/SUBJ	subject (Chapters 12, 22)
Subj/SUBJ/SUBJ	subjunctive (Chapters 14, 17, 19, 20)
SUBORD	subordinator
SUF	suffix(al element)
SUFF	suffix
SVI	subject–verb inversion
SVO	subject–verb–object
syll.	syllable
T	term
TAM	tense–aspect–mood/modality
TAMP	tense–aspect–mood–polarity
TBR	Tobor Nyun
Tem.	Temne
temp	temporal word
TH	theme
Tm	Temne
TMA	tense–mood–aspect
TMAP	tense–mood–aspect–polarity
TN	transnumeral
t.o.	type of
TOP	Topic
TOT	
TR/TR	transitive
TRM	terminative
UGPC	Upper Guinea Portuguese Creole
UNINT	Unintentional
UPC	*Unions des Populations du Cameroon*
V/v/v.	verb
V/v	vowel
(v)	verb
var.	variable
VBLZ	verbalizing noun suffix (*-e*)
Vc	verb phrase or compound verb
VEL	velar
VEN/VEN/VEN	venitive
VEN	venitive verbal suffix (Chapter 4)
VFOC	verb focus/focalization
VH	vowel harmony
VIR	virtual
VP	verb phrase
VSAL	verb salience
VSI	verb subject index
VT	thematic vowel
WAAW	W. A. A. Wilson
WEST	western varieties
WN	Western Niger
X	oblique
Z	Ziguinchor

The contributors

Alain-Christian Bassène is professor at the Department of Linguistics and Language Sciences of the university of Dakar, Senegal (Université Cheikh Anta DIOP). He obtained his PhD in 2006 at the university of Lyon, France (Université Lumière Lyon 2 - Laboratoire Dynamique du Langage).

Noël Bernard Biagui is a junior researcher in linguistics at CLAD (Université Cheikh Anta Diop, Dakar, Senegal). A specialist and native speaker of Casamance Creole, he is the author of the first book-length grammar of this language.

Jean Pierre Boutché is a senior lecturer of African linguistics and German studies at the University of Maroua (Northern Cameroon). His activities of research and teaching include Fula grammar, sociolinguistics, contrastive analysis, and second language instruction.

† **G. Tucker Childs** was professor of applied and West African linguistics at Portland State University. He conducted pioneering research on ideophones and was a world-leading specialist on Mel languages. He was a longtime editor of the journal Studies in African Linguistics and recipient of the Kenneth L. Hale Award of the Linguistic Society of America in 2018.

Ibrahima Abdoul Hayou Cissé a lecturer in the Faculty of Languages and Linguistics at Université des Lettres et des Sciences Humaines de Bamako (Mali). His research interests include language, education and human rights, and language acquisition and socialization in multilingual environments.

Alexander Y. Cobbinah is professor at the Linguistic Department of the University of São Paulo, teaching and researching in general and African linguistics. Research interests include language contact, transatlantic linguistics, discourse structure, anthropological linguistics, Bantu languages, and Creolistics.

Jules-Jacques Coly obtained his PhD on the morphosyntax of Joola Kuwaatay at the University of Cologne. His research includes the structure of Joola languages, multimodality, and the expression of emotions in Atlantic languages.

Denis Creissels retired in 2008 after teaching general linguistics at the universities of Grenoble (1971–1996) and Lyon (1996–2008). His research interests centre on linguistic diversity, the description of less-studied languages, and morphosyntactic typology. He has been engaged in fieldwork on a variety of languages, among which Mandinka (Mande) and Balant Ganja (Atlantic) and is currently working on Jóola Fóoñi in collaboration with Alain-Christian Bassène.

Samantha Goodchild is guest researcher at the University of Oslo. She has held postdoctoral positions at the University of Galway and at the Center for Multilingualism in Society across the Lifespan, University of Oslo. Her research investigates the relationship between language, ideologies, and space. She is co-founder and co-director of Language Landscape, a not-for-profit organization which raises awareness of linguistic diversity through academic and public outreach events.

Jacqueline Knörr is a social anthropologist and professor at the Max Planck Institute for Social Anthropology in Halle/S., Germany. She works on connections between creolization and decolonization in postcolonial contexts, on (re-)migration, gender and generation, on (trans-)ethnic and national identity, and on issues of identity and integration. Regionally, her work focuses on Atlantic West Africa, Indonesia, and Western Central Europe.

Raija Kramer is a former junior professor at the University of Hamburg, specializing in the description and documentation of Adamawa languages and the Fulfulde varieties of northern Cameroon. A particular focus of her research lies in the dynamics of language use and language contact and their impact on linguistic description.

Kristin Vold Lexander, PhD is a researcher at Inland Norway University of Applied Sciences and former postdoctoral fellow at MultiLing Center for Multilingualism in Society Across the Lifespan, University of Oslo. Her research interests centre on multilingualism in digitally mediated interaction with a particular focus on Senegal.

Daniel Alcón López works at the Department of Romance Languages at the University of Freiburg. His main interest lies in the search for methodologies of management, representation, and analysis in textual, audiovisual, and sensory data.

Friederike Lüpke is professor of African studies at the University of Helsinki and a Core Fellow at the Helsinki Institute for Advanved Studies (HCAS). She works on Mande and Atlantic languages of West Africa in their societal contexts, focusing in particular on language documentation and description, multilingualism, language in education, and language and writing.

Fiona Mc Laughlin is professor of linguistics and African languages at the University of Florida. She does research on the sociolinguistics of African cities, the sociolinguistics of writing, and the formal phonological and morphological properties of Atlantic languages, including language contact effects.

Joseph Jean-François Nunez is a Wolof teacher at INALCO (Paris, France) and a post-doctoral researcher at SEDYL-UMR 8202. A specialist of contact linguistics and native speaker of Casamance Creole, he has researched in depth the influence exerted by French and Wolof on the actual language practices of today's Casamance Creole speakers.

Loïc Perrin is lecturer at INALCO (the French National Institute for Oriental Languages and Civilizations) in Paris where he teaches the Wolof language. He is associated with LLACAN, the French laboratory of the CNRS specializing in African linguistics. After a PhD in cognitive semantics, which dealt with the representations of time in Wolof, he undertook a description of the Bassari language.

Konstantin Pozdniakov is emeritus professor at INALCO (Paris; African and comparative linguistics, Wolof) and a researcher at LLACAN (CNRS). He was a member of Institut Universitaire de France (IUF). He has authored various publications on Atlantic reconstruction and Niger-Congo languages in general.

Nicolas Quint is a senior researcher in linguistics at LLACAN-UMR8135 (CNRS/EPHE/INALCO, Paris, France). He is mainly involved with the description of languages belonging to the three following groups: Afro-Portuguese Creoles (Capeverdean, Casamance and Guinea-Bissau Creoles, and Papiamentu), Niger-Congo (Kordofanian and Nyun), and Romance (mainly Occitan).

Stéphane Robert is senior researcher in linguistics at the French National Center for Scientific Research (CNRS) and a member of the LLACAN research unit specialized in African languages. Her research interests cover Atlantic languages and typology, general linguistics (semantics, grammar, information structure, and discourse), cognitive linguistics, and cognition.

Guillaume Segerer is a researcher at LLACAN (CNRS, France). He has been working for the past twenty-five years on the description, comparison, and classification of Atlantic languages, especially their Bak branch. He has also created RefLex, a database containing more than 1.5 million words of more than 1,000 African languages.

Frank Seidel studied at the University of Cologne and University of San Diego, achieved his doctoral studies at the Special Research Center ACACIA, University of Cologne and carried out his post-doctoral research at the University of Florida. He is currently employed as an analyst at a logistics company.

Anne Storch is professor of African linguistics at the University of Cologne. Her principal research has been on the various languages of Nigeria, on the Atlantic language region, on Western Nilotic, with a current interest in Digo (Kenya). Her work combines contributions on cultural and social contexts of languages, the semiotics of linguistic practices, epistemes and ontologies of colonial linguistics, as well as linguistic description.

Sylvie Voisin is professor at Aix-Marseille University and member of the DDL research unit of the CNRS. A specialist of field linguistics, she works on the description of Wolof and on the verbal extension system of Atlantic languages. With Stéphane Robert, she co-directed the ANR project Sénélangues (2009–2014) which resulted in the production of sixteen new grammar sketches and collective work on nominal classification (2015).

Sophie Wade is originally from Senegal. She obtained her PhD in African languages and cultures at the University of Cologne. Interested in the empowerment and inclusion of migrants, she coordinated the community project Space2groW. Sophie is also involved in initiatives for self-determination of refugees and migrants and in development cooperation.

Rachel Watson is a research associate at SOAS, University of London. Her research focuses on the Kujireray variety of Joola, spoken in Brin (Jire), nominal classification systems, the epistemology of language, and the use of Joola in European diaspora contexts.

PART I

Background and classification of the Atlantic languages

CHAPTER 1

Language, land, and languaging in the Atlantic space

Friederike Lüpke

1.1 Introducing the dialectic relationship between language and languaging

> "Language is *an ongoing process that only exists as languaging, not as isolated items of behaviour*". (Maturana and Varela 1984: 210)

Linguistic practice is not primordially language-based. The categorization of speech forms is a sociocognitive process that results in the creation of entities that can be named and referred to, by speakers and outsiders, including linguists. The creation of a named language by linguists through the selection of particular linguistic features to represent it has huge ramifications. Many studies emphasize the relationship between descriptive activities and colonialism. In fact, many of the named languages known in Africa today are a direct product of colonial language creation, certainly related to a desire to govern through regulation but also intimately linked to the language experiences of Europeans in the long nineteenth century, an era of ethnonationalist romanticism that had great repercussions on imaginations of language, including in the budding discipline of linguistics (Bonfiglio 2010; Makoni 2013; Schneider 2018). These events were contemporaneous with the colonization of most of Africa and hence provided the lens through which African linguistic ecologies were explored by colonial actors. Cysouw and Good (2013) address the artificiality of glossonyms in the practice of descriptive linguists and typologists and propose practical solutions to manage it, but the ontologies underpinning language names and the categorization of linguistic features deserve to be investigated in more detail because they reveal language ideologies of different actors at work. In contrast to the regulated languages of nation states in the global north, including the official languages of colonial origin, enregisterment of named languages on the Upper Guinea Coast is weak. Where standards have been proposed, they remain symbolic and are not enforced by institutions (Lüpke 2021b). This means that what constitutes a language remains an important empirical question which cannot be sidestepped by recurring to a standard established through description or codification serving as the baseline of comparison, but can only be answered by investigating research participants' indexical fields, i.e. categorial options and their grounding (Silverstein 2003; Eckert 2008). As the following paragraph argues, this endeavour is of prime importance for the Atlantic space, a world area in which linguistic diversity has deeply rooted social motivations.

The Atlantic space roughly comprises an area reaching from the present-day states of Senegal in the north to Liberia in the south and their hinterlands, characterized by a sunken coastline, a network of tidal rivers, mangrove swamps, and areas of thick tropical vegetation including forests.[1] The title photo shows the sunken coastline of this area, criss-crossed by tidal rivers and mangrove swamps. These topographic conditions have had repercussions on the type of communities inhabiting it: mainly small frontier communities which maintained intense ties with their neigbhours. In large parts unsuitable because of its topographic and climatic conditions to the formation of larger polities and to direct conquest of its coasts, it is a region of deeply entrenched linguistic diversity, which has already been recorded in the earliest outsider testimonies, such as those of Christopher Columbus who visited the area in 1498 (Hair 1990: 119). It is therefore not surprising that this area has been included in a fragmentation zone first described by Dalby (1970) as the sub-Saharan fragmentation belt, reaching from present-day Senegal to the Ethiopian highlands. Intense contact and widespread multilingualism have been upheld despite the availability of languages of wider communication such as Mandinka (see Chapter 7, this volume) for hundreds of years and regardless of the great linguistic affinity between many (but not all) languages of the area. Yet, new identities have constantly emerged during the

[1] The geographical extension of the Atlantic space provided here is only meant to give a rough orientation. The contours of the space are subject to constant change induced by changing climatic conditions (Brooks 1993) and political factors, for instance the advent of the transatlantic slave trade (Barry 1998; Hawthorne 2003; Green 2012, 2019) and European colonization (Roche 2000).

Friederike Lüpke, *Language, land, and languaging in the Atlantic space.* In: *The Oxford Guide to the Atlantic Languages of West Africa.* Edited by: Friederike Lüpke, Oxford University Press. DOI: 10.1093/oso/9780198736516.003.0001

past millennium for which records are available. Particularly visible among them—because of the great linguistic and cultural contrast involved in the creation of new linguistic and cultural identities—are those that involve languages of European origin. To these belongs Kriolu, a Portuguese-based creole that was created on the Cape Verde archipelago in the sixteenth century and integrated into language ecologies on the Upper Guinea Coast mainland from the seventeenth century onwards (see Chapter 6, this volume). The linguistic and cultural processes related to the formation of Creole in Sierra Leone from the eighteenth century onwards (see Chapter 26, this volume) also fall within this category, as do new registers often categorized as mixed, featuring colonial languages such as French in combination with Wolof and other local languages in Senegal (see Chapter 31, this volume).

In the light of the continuity of diversity going hand in hand with changes in recorded language names, their intension and extension, and the creation of new languages, we can assume that societies in the area are built on the premise of maintaining internal diversity while adapting to constant change. These processes can be observed both at a very fine-grained local level and at regional and transnational scales. Upholding diversity and multilingualism must consequently be understood as a socially motivated habitus of maintaining multiple identity articulations through linguistic difference. This socially motivated inclination perpetually recreates diversity and thus counteracts tendencies of wholesale language shift and of massive convergence of linguistic features that might otherwise occur, because of social as well as cognitive pressures. This means that the social meaning contained in language names, and how labels are correlated with specific linguistic forms, deserve special attention. Additionally, languages must be understood not as a fixed set of linguistic features but as abstractions bound to particular sociopolitical chronotopes and social stereotypes. This is particularly so because the area is characterized by a continuous reshuffling of spatial organization and identities through small-scale migration rather than through large-scale migration events (Hair 1967; Wright 1985, 1999), leading Hair to ask the following question:

> 'The overworked question "Where did the people come from?" might be profitably given a rest in favour of the question "When and how did the people living in this place come to be the people they are today".' (Hair 1967: 268)

For languages, we can paraphrase parts of Hair's dictum as follows: 'When, how and for whom did the linguistic features used in this space come to be the languages they are now?' And because the processes of identity indexing through language are ongoing, we can add 'And how are new languages being forged to respond to new conceptualizations of identity?'. In any multilingual area where languages are not acquired as monolingual solitudes (Cummins 2008), this has repercussions on the reach of named languages and on their unequivocal enregisterment, i.e. on the association of particular linguistic features with a named language or register (Agha 2005). The situation can be sketched in the broadest brushstrokes as one where there are *more languages* present in a given area and in individuals' repertoires, but there is *less of language*. This entails a smaller scope for a named language in terms of domains of use, less distinctive matter in terms of lexicon and grammar in an individual's repertoire, and more creative linguistic practice focused on performing one or several named languages than in regulated monolingual environments, as will be explored in more depth in Section 1.3.

As a result, although languages are kept distinct in name and in parts of their linguistic features, the traces of intense contact and multilingualism are omnipresent, and their conceptualization of convergences as contact-induced vs. genetically inherited is, in many cases, contentious (see Chapter 2, this volume). Doubt is also cast on the usefulness of distinguishing these processes as distinct operators of language change (see Chapter 3, this volume). It is certainly the case that some linguistic features are shared over large areas, which lends them great temporal stability, as discussed in Section 1.3 below. However, whether it can be established beyond doubt that they can be traced back to proto-languages, and whether such presumed inheritance is qualitatively different from speech forms diffused through active or receptive multilingualism remains debated. As argued by Mufwene (2002), once conventionalized in a group of speakers and used frequently enough, features are transmitted to language learners regardless of being inherited or acquired through contact, making the distinction obsolete. In fact, we can even argue that not all linguistic features are necessarily acquired with a strict association with one language. There are many reasons favouring such a perspective. Firstly, language socialization and acquisition are not focused on one language only (see Chapters 29, 30, and 32, this volume). Monolingual language acquisition (i.e. acquisition in a context pragmatically intended to focus on one named language) is limited to early toddlerhood and to a very limited selection among children's caregivers and community contexts. As soon as children become mobile, they are exposed to linguistic diversity through multiple caregivers not part of the nuclear family, in courtyards, neighbourhoods, through visits, and in schools. As adults, they continue their linguistic journeys through professional mobility, family visits, and long-term migration, resulting in the continuous adjustment of repertoires and concomitant recalibration of the categorical values of linguistic features.

Mobile individuals adapt their repertoires over their entire lives and encounter many linguistic features that are potentially ambivalent (Woolard 1998b), that is, features that can be seen as shared between a large number of languages alongside features that can be unequivocally associated with one named language only. The more widely shared a feature is, the less it is associated with named languages in a multilingual speaker's mind, and the lower its indexical potential (Watson 2019). Features from all areas of language have been reported as convergent through contact, as discussed in Section 1.4 below. Applying a different model of language, we can conceptualize these features as linguistically potentially polyvalent. This entails that we can study how they are mapped onto named languages and how multiple categorization occurs based on the perspectives of different speakers or observers (Lüpke and Watson 2020; Lüpke 2021a), rather than departing from the notion that they belong to a particular named language and are spread through language contact.

This dialectic relationship between fluid language use and its reified representations demands new theorizations of language in multilingual areas such as the Atlantic space and beyond, since linguistic practice is always fluid, variable, and open to multiple categorizations, while reifications are bounded and have less or no room for variability and polyvalency. It is therefore useful to explicitly investigate the relationships between fluid practice or *languaging* and their categorization or *language*.

Languaging (Maturana and Varela 1984; Mignolo 1996; Canut 2010; García and Wei 2014), as synthesized below by Li Wei, transcends named languages, registers, and modalities. Multilingual individuals make use of semiotic resources as required by the pragmatic context of interaction:

> '1. Multilinguals do not think unilingually in a politically named linguistic entity, even when they are in a "monolingual mode" and producing one nameable language only for a specific stretch of speech or text.
>
> 2. Human beings think beyond language, and thinking requires the use of a variety of cognitive, semiotic, and modal resources of which language in its conventional sense of speech and writing is only one.' (Wei 2018: 18)[2]

Two questions arise from these observations: Firstly, variability, fluidity, and multimodality characterize the entirety of human communication. Yet, we do describe some of this languaging in terms of languages, by calling individuals or settings multilingual, whereas in other contexts we do perceive this fluidity as operating within the boundaries of a named language, for instance in the case of special registers or dialects. Thus, since some, but not all, languaging is categorized as multilingual, how do we construct the discrete objects that languages are from heteroglossic speech or categorize even languaging as involving discrete languages, by conceptualizing it as *trans*languaging (García and Wei 2014), i.e. as overcoming discrete languages?

The answers to these questions are social in nature. Languages, as a famous dictum by Wilhelm von Humboldt reminds us, draw circles about their speakers. In ethnolinguistically imagined nation states, these circles have been drawn in permanent ink, and been given absolute distinctive power. Individuals are expected to be situated firmly within one circle, exceptionally in two. In the Atlantic space, these circles have only been pencilled in and are constantly being redrawn, and individuals can place themselves or be positioned by others in multiple and different circles. Here, social processes of identity formation through language are ongoing and not (yet) constrained by language institutions. Because no independently set standard imposes a particular perspective, we can study how categorization is achieved linguistically. This means that this area is a testing ground *par excellence* for examining the changing relationships between languaging and language. We can investigate how particular speech forms become associated with particular named registers or languages, and how speech forms can be categorized differently based on scale, perspective, and fractal recursivity (Silverstein 2003; Gal and Irvine 2019). We can analyse which value potentially polyvalent speech forms receive, and which emblematic forms become invested with social meaning based on repertoires, social spaces, and language ideologies governing them (Woolard 1998a, b; Lüpke 2021a). We can thus move away from the heuristic measure of taking a priori established languages as the baseline from which to calibrate variation and distance towards an investigation of how these languages are constructed.

It follows that we need to treat the products of descriptive linguistic efforts with caution not regarding the linguistic features they present, but in terms of reifying fluid and ambivalent speech into one named language. Walter Mignolo describes the process of language creation as follows:

> '[W]hen you write grammars and vocabularies of a complex set of languaging processes among a given population, you convert the process into an object and you own, you possess that process that you call *language*. *Language* becomes then an object, with a grammar and vocabulary that you have and regulate. It also becomes the point of reference to measure and rank *languaging practices* that do not comply with the regulatory force of *language*.' (Delgado and Romero 2000: 16f.)

[2] Language, in its conventional sense of course also includes sign languages, although they often remain at the perceptual margins of linguistics.

The case studies in this volume offer an ideal opportunity to explore in detail how both language and languaging manifest themselves in an area of known linguistic diversity, contact, and multilingualism throughout the past millennium. This endeavour provides a further step towards the holistic turn in African language description and documentation, historical linguistics and multilingualism, and contact research asked for by Lüpke (2010b), Childs et al. (2014), and Di Carlo et al. (2021), to name but a few—a turn that incorporates socio- and anthropological-linguistic vantage points into linguistic description.

1.2 Named languages from expressions of sociopolitical organization to expressions of ethnicity

> 'Les noms ont une histoire; il y des conjonctures et des situations locales qui font qu'un nom est adopté par les uns, refusé par les autres. La lecture en termes ethniques de la réalité sociale est une sorte de mise à plat, de mise sur carte, dont le principe est l'oubli obstiné de cette histoire.' (Bazin 1985: 105)
>
> [Names have a history; there are junctures and local situations which mean that a name is adopted by some, refused by others. A reading of social reality in ethnic terms is a kind of flattening, of mapping, the principle of which is the obstinate oblivion of this history.]

One of the areas where a combination of perspectives sheds new light on contentious issues is that of glossonyms; their social meaning and linguistic significance. Language names are windows into ideas of identity and their change over time. In the Atlantic space, we find different patterns for naming languages that illustrate which aspects of sociopolitical organization, language ideology, or linguistic pattern motivates the glossonyms. While naming and thus fixing languages is often an outcome of colonial practice, there is also evidence that linguistic practice and sociopolitical identity categories were nameable already in the earliest known written sources for the Upper Guinea Coast, accounts provided by Iberian traders (Hair 1967). In many cases, the exact reference or even referentiality of these terms remains opaque. They certainly are mostly based on outsider categorizations, often not drawing on linguistic criteria. This means that they constitute larger-scale identity constructs that hide many more localized levels of diversity, linguistic and otherwise. For instance, concepts such as Baïnounk—an ethnolinguistic label in the Casamance going back to an exonym used by Mandinka—can split into several contrastive entities, among them Gujaher and Gubëeher. These entities in turn can be broken down into more localized and fine-grained lects that then stand in opposition to each other and to their seeming hyponyms (Cobbinah 2010, 2019; Lüpke 2010a, 2016a).

Even if groupings involve linguistic resemblance, this does not necessarily entail that much importance is accorded to the unequivocal categorization of (all of) speech—a desire to do so may well be a reflection of nationalist language ideologies active in later periods. Yet, the recorded names reveal areas of both continuity and ruptures in sociopolitical organization and concomitant possible categorization of people (and sometimes their speech). They are not exclusively based on the perception of European outsiders but also reflect the perspectives of local neighbours and of insider identity concepts of social groups and political entities. We can trace this through the designations used today, for instance through the areally widespread pattern of naming languages after the places with which they are associated, as exemplified below.

Isomorphy between place names and language names is frequent. In this naming pattern, a place name, for instance that of the village of Jire (Brin in French) is lexically related to the language name. For Jire, this name is Kujireray, featuring the noun class prefix *ku-* and the suffix *-ay*, an abstract nominalizer with the meaning 'something having to do with' (see Chapter 10, this volume). In some instances, the relationship between place name and language name is equipollent, that is, both are morphologically marked, for instance in the place name Kaboi (Cobiana in Portuguese), with the noun class marker *ka-* and the corresponding glossonym Guboi, featuring the noun class marker *gu-*. Extending the pattern, designations for speakers occur with the same root in the noun class paradigm for human/agent nouns. Table 1.1 illustrates this pattern for topo- and glossonyms.

Even where place names and language names are not phonologically related, the reification of lects is often based on political units, regardless of linguistic affinity to other lects. Childs (Chapter 20, this volume) discusses the case of Kim and Bom, which are close to identical regarding their linguistic features but kept apart by speakers because of being associated with different chiefdoms (see also Di Carlo and Good 2014 on similar groundings for language territorialization in northwestern Cameroon). Different scales of political organization can be mirrored in fractal recursivity of language names—thus, lects of the Banjal peninsula in the Lower Casamance can be named based on the smallest political units, wards or villages (cf. Gusiilay, the language of Essyl, or Gubanjalay as the language of Banjal), or based on the overarching polity, the kingdom of Mof Avvi. In the latter case, Gubanjalay is often used, but in this

Table 1.1 Related topo- and glossonyms on the Upper Guinea Coast[a]

Toponym	Glossonym	Colonial placename	Source
Jibëeher	Gubëeher	Djibonker	Cobbinah 2013a, Chapter 12, this volume
Jire	Kujireray	Brin	Watson 2014, Chapter 10, this volume[a]
Eer	Keerak	Kabrousse (from Port. Cabo Roxo)	Robert and Segerer, Chapter 11, this volume
Kaboi	Guboi, Gubuy	Cobiana	Voisin 2015b
Ñaamol	Guñaamolo	Niamone	
Tobor	Gutobor	Tobor	
Banjal	Gubanjalay	Banjal	Goodchild 2018/2019, Chapter 32, this volume
Gusiil	Gusiilay	Essyl	
Jibelor	Gubelor	Djibelor	
Mansoa	Sua	Mansoa	Segerer, Chapter 16, this volume
Gubaabo	Gubaabo	Ziguinchor[b]	Lüpke 2021a

[a] The toponyms and glossonyms are given in the Wolof-based orthographies for national languages of Senegal. The spelling for colonial place names follows the conventions of the respective colonial language, in this area mainly Portuguese and French. Where no information source is indicated, the names are based on my own field notes.
[b] Gubaabo is an interesting case—it is the name for the Portuguese-based Creole of the area in Gujaher and Gubëeher. It is the result of a morphological backformation of the word *tubaab* 'European', whose initial syllable was reanalysed as a human prefix and replaced with the prefix *gu-* whose function includes the derivation of language names in Nyun languages, and whose functions in Bak languages comprise the derivation of place names. Thus, Ziguinchor, a city founded by the Portuguese in 1645, is named 'place of the Europeans', and Kriolu, associated with Portuguese settlers and a local class of intermediaries, as the 'language of the Europeans'. See also Lüpke 2021a.

case metonymically extending the designation of a village to the entire area. Languages can also be named after geographical and/or political areas, as is the case for many intermediate-scale varieties of the Joola cluster, for instance Fooñi/Fogny (Bassène, Chapter 9, this volume), Kaasa (Sambou 1979; Bassène 2007), or Karon (Sambou 2007). Analogical patterns are attested for Gujaher and Kasanga, which are each and in some cases both associated with a group of villages, in all likelihood reflecting historical patterns of political organization into kingdoms/sacred chiefdoms (Bühnen 1994).

Glossonyms that are not associated with particular settlements or smaller polities and regions typically lack the derivational relationship between location and language described above. These are languages associated with larger precolonial political structures, for instance Wolof. Or they denote polycentric languages such as Fula, whose different lects are connected to particular geographic and/or political areas, but where widespread nomadic pastoralism and long-range migration also result in great mobility and the cohabitation of Fula-speaking cattle herders with more sedentary people and political structures (see Chapters 5 and 30, this volume). Mandinka is a language that is both associated with state formations in the area (the kingdom of Kaabu) and with Islam, through Mandinka traders and clerics who spread Islam and with it, the Mandinka language (Wright 1985; Green 2012). For many of the inhabitants of the Lower Casamance who witnessed the Jihad led by Fode Kaba in the nineteenth century, Mandinka has become homonymous with Muslim, and converts to Islam often declare themselves Mandinka, regardless of their linguistic repertoire.

People, and by extension the languages they are identified with, have been named by outsiders. Many of the labels used have been recorded in travellers' accounts from the fifteenth century onwards (Hair 1967). That these terms were not only figments of the imagination of Portuguese travellers can be shown through the fact that many of these exonyms are still used today by people in the Atlantic space to designate their neighbours. For instance, the ethnonym Bassari has been used by Mandinka to refer to people who call themselves *alëyan* (plural *bëlëyan*) and call their (internally diverse) language *onëyan* (see Chapter 8, this volume). Ariatta, a term recorded in Portuguese sources to refer to particular groups or areas seen synchronically as Joola, is in current use. Guriat, with a stem related to Ariatta, is the term

used by speakers of Gujaher to refer to Joola. It is important to note that these terms depend on the scale at which the differentiation is socially meaningful. Speakers of Gujaher, 40km away from the kingdom of Mof Avvi, use Guriat indistinctively for this kingdom's Joola language(s) and any other Joola language. Close neighbours, for instance those from Jire, refer to it as Kubanjalay and differentiate it from other Joola varieties, demonstrating the attention they pay to local patterns of sociolinguistic organization. Which exact linguistic or social features are associated with these denominations is not absolute but hinges on which contrasts are seen as contextually meaningful and which features index this meaning.

In addition to political dimensions, languages can be named to signal ethnicity. This reading, as Irvine reminds us, is prevalent in linguistics today:

> '[M]uch of the African linguistic literature is written as if the only differentiation among Africans that matters is ethnic, and as if ethnic difference and linguistic difference always coincide.' (Irvine 2008: 331)

Many current ethnonyms can be traced back to exonyms already recorded in traveller reports, but this does not mean that they have the 'ethnic' meaning now entailed by them. For example, Banyun, Baïnounk, and related terms, lumping together people speaking related languages but exhibiting different sociopolitical organization (Bühnen 1994) were grouped together under one umbrella term because of their similarity to outsiders, which overrode the internal differences (Lüpke 2010a). Likewise, Balant(a) is a term not used by the people so designated themselves (see Chapter 15, this volume)—it simply means 'people' in different languages grouped together as Balant, whose speakers themselves use other, more differentiated endonyms.

In the late nineteenth century, a noticeable shift in denotation towards ethnicity occurs: colonial actors catalogue populations in the wake of the creation of colonial territories at this time; and they perceive West African realities through ethnic lenses freshly ground to fit the ethnonationalist feelings en vogue in Europe (Hobsbawm and Ranger 1983; Amselle 1990; Lentz 1995). Thus, either old terms with vague and changeable meanings receive ethnic interpretations, henceforth meant to designate a people united by a shared language and culture, or new terms, such as Joola, are minted to fulfil the same ethnic purpose (Roche 2000; Nugent 2010). Ethnic identities have become powerful social representations in contemporary societies, and this is reflected also in language naming strategies. A common practice is the creation of double-barrelled glossonyms: Joola Fooñi, Balant Ganja, Joola Keerak (or, in its 'modernized' version, Joola Kabrousse), Baïnounk Gujaher, Baga Mandori, and so forth. Ethnonyms can also stand alone, then being comparable in scope to the older outsider terms out of which they sometimes develop. Sometimes these ethnonyms have been used by linguists to label intermediate-level groupings of lects, for instance Joola or Balant. Outside this specialist usage, the extension of these terms is not codified and remains highly variable, depending on the kind and extent of group identity desired. Thus, Baïnounk can be used to refer to the speaker's own smaller-scale local identity and associated lect (i.e. as a stand-in for Gujaher, Gutobor, and so forth), or as a hypernym encompassing not only closely related languages, but also the only remotely related Kasanga, just as Joola can not only comprise this language cluster but sometimes also include the various lects united under the level Bayot. Sometimes, ethnic (re)categorizations are still traceable. For instance, Creissels (Chapter 15, this volume) reports that Ganja, today identifying as Balant, see themselves as former Baïnounk, without a linguistic trace to corroborate this earlier affiliation. It is an empirical question whether the shift was one of language or merely of ethnic affiliation. Many present-day Joola and/or Baïnounk see themselves as members of both groups (Cobbinah 2020; Watson, Chapter 10, this volume). In the case of Baga varieties, the mismatch between ethnic identity and linguistic classification is especially stark: while speakers assume an overarching Baga identity (see Chapter 18, this volume), Baga Mandori is classified as a Mel language, together with Temne, Landuma, Baga Koba, Baga Sitemu, and other Baga varieties, and outside the Atlantic language family. Nalu, Baga Fore, and Baga Mboteni in contrast are classified as belonging to the North branch of Atlantic (see Chapters 2 and 3, this volume). Baga ethnicity can encompass all these identities or be conceptualized as people having multiple identities.

Thus, ethnic identity, although widely adopted in contemporary Upper Guinea Coast societies, is more performative and multiple, and much less based on sharing one and the same language than an essentialist reading entails, as captured here for Mandinka identity:

> 'I went into the field to study the history of a Mandinka state. Everything I had read pointed to its Mandinka identity [. . .]. British colonial rulers and The Gambian government noted that the country's majority population was Mandinka. (Such authoritative sources as the CIA's World Factbook still give us The Gambia's population in percentages of ethnic representation—it is forty-two percent Mandinka, eighteen percent Fula, sixteen percent Wolof, ten percent Jola, nine percent Serahuli, and four percent "other", with non-Gambians making up the final one percent.) So one can imagine my surprise when I met "Mandinka" people residing in Niumi who spoke five languages and who sought

> marital partners across what I thought were clear ethnic boundaries. I met people who spoke Mandinka and resided in what people (and no end of ethnographic maps) said was Mandinka territory, yet had been born of parents who considered themselves Wolof; I met people who were "Serer" but who had children who were "Mandinka"; I encountered a number of people whose ethnicity depended as much on lifestyle as on parentage.' (Wright 1999: 416f.)

Finally, linguists have grouped together lects based on shared linguistic features. Sagna (2008) for instance, has coined the term Joola Eegimaa for the varieties spoken in the Mof Avvi because of the word *eegimaa* 'here is what I am telling you' used in conversations, and not shared with other Joola varieties. The term Mel stands for a grouping based on the shared isogloss for 'tongue'; Bak is based on the *bVk*-prefix for plural forms of nouns and pronouns denoting humans in languages so grouped together (Sapir 1971).

Denominations for languages, clusters, and subgroupings can thus be based on a combination of local glossonyms and insiders', various outsiders', and linguists' perceptions of similarity, shared (and multiple) ethnicity, and selected isoglosses or isogloss bundles at different points in time. This means, as also forcefully argued by Childs (Chapter 3, this volume), that language names cannot be codified and captured in fixed ISO codes or standardized glossonyms. How linguistic features are categorized as reflecting reified sociocultural identities or selected linguistic features is a matter of constant recategorization and contingent on variable and time-bound vantage points of political actors. Speakers' views should not be replaced by categories intended to reflect only linguistic criteria, something that cannot be achieved in any case because of the multiple ways in which linguistic features align and overlap. Crucially, how languages are conceptualized—associated with a territory in the minds of their speakers—and how they are used in society influences their substance, as the following section argues.

1.3 Language territorialization and sociolinguistic spaces for language and languaging

> 'Cultural centers, discrete regions and territories, do not exist prior to contacts but are sustained through them, appropriating and disciplining the restless movements of people and things.' (Clifford 1997: 3)

Spaces are not created as language communities, i.e. as regrouping only people united by the desire to form 'a social group, generally a primary reference group, the members of which are, by degrees, oriented to a denotational norm, however much within its compass they recognize situated variation' (Silverstein 2015: 8). The creation of a language community is a language-ideological act, fuelled by conscious acts of identity (Le Page and Tabouret-Keller 1985). It is thus useful and important to distinguish between language community and speech community, although they are often conflated in sociolinguistic and descriptive linguistic parlance. The original definition of speech community proposed by Gumperz (1968) applies this distinction, making it close to later theoretical notions such as the community of practice (Wenger 1998). Gumperz designates as speech community '[a]ny human aggregate characterized by regular and frequent interaction by means of a shared body of verbal signs and set off from similar aggregates by significant differences in language usage' (Gumperz 1968: 381). His emphasis is on shared social conventions for language use, **not** on shared linguistic signs, as made explicit in this statement:

> 'Regardless of the linguistic differences among them, the speech varieties employed within a speech community form a system because they are related to a shared set of local norms.' (Gumperz 1968: 381)

The difference between language community (close in meaning to the ethnic group) and speech community is perhaps most obvious for polycentric languages such as Fula (Chapter 5, this volume), which is spoken in a vast area from the shores of the Atlantic in the west to Sudan in the east. Fula spread through nomadic pastoralists and conquest and sedentary state formations. While the lifestyles and spoken lects of people identifying as Fula differ considerably, they are united by their identification with an overarching Fula language community. At the same time, they integrate speech communities that range from relatively homogeneous (for instance the Middle Casamance) to very multilingual (northern Cameroon, see Chapter 30, this volume), depending on who they cohabit with, and what wider multilingualism patterns they adapt to. The difference between language communities and speech communities or communities of practice is of utmost importance for heterogeneous, small-scale multilingual places like much of the Atlantic space, both in terms of linguistic interactions and of ideas of language. As have many adjacent areas of West Africa, this area has been described by historians and anthropologists as a frontier, a place where family-based groups claim (seemingly) uninhabited or unclaimed spaces of which they become the firstcomers or landlords (Kopytoff 1987a; Brooks 1993; Barry 1998; Lentz 2013; Nyamnjoh 2017). In these places, the original settlers receive and settle latecomers or strangers with whom they develop client

relationships. People therefore always juggle language lives in a multilingual speech community, consisting of founders and linguistically heterogenous strangers and in an abstract language community linking them with all those who speak languages with which they focally identify, wherever they are located. A multilingual speaker of (Baïnounk) Gujaher living in the village of Agnack Grand is surrounded by speakers of Joola languages, Kriolu, Balant, Manjaku, Fula, and Wolof, which all make up its speech community. Gujaher connects this person to all villages with Gujaher as the founders' language; and Baïnounk to the overarching language community, an abstract 'imagined community' (Anderson 1983). The Frontier process with its creation of multiple communities and multiple links constantly perpetuates itself: in order to avoid conflict or unrest, people move to establish new homesteads, either as landlords or by joining established settlements as strangers. I have sketched this process as follows:

> 'The Frontier as a region is characterized by the fictional vacuum it presents to its first settlers, who turn themselves into the autochthones even though they often really are not. Firstcomers are the ones that lay claims to the land and, through their descendance and linguistic identity, determine its patrimonial language.' (Lüpke 2018a: 187)

What makes a space empty and unclaimed depends on its categorization: it is always possible to carve a piece out of an already owned space and declare it as independent and thereby turn its inhabitants from strangers into landlords. Unlike ethnolinguistic nation states which like to see themselves as finished projects, with every change to the established order being a challenge that needs to be regulated and controlled, frontier societies are, by nature, unfinished and ongoing, since individuals live in many different multilingual speech communities and maintain large-scale ties of belonging to language communities at different scales. This results in mallable and relational ideities, as explained by Nyamnjoh, echoing (Kopytoff 1987a):

> '[F]rontier Africans are able to navigate and negotiate myriad margins of identity and belonging. Their capacity to straddle physical and cultural geographies enables them to point attention to the possibility and reality of a world beyond neat dichotomies. Their world is characterised by flexibility in mobility, identity, citizenship and belonging. Myriad interconnections, inextricable entanglements and creative interdependencies, despite persistent hierarchies at global and local levels, afford Africans the opportunities to explore the fullness of their potentialities without unduly confining themselves with exclusionary identities.' (Nyamnjoh 2017: 259)

So how can the language-ideological process of territorializing a language (Blommaert 2010)—i.e. the creation of a deictic link between a place and (a) language(s) rudimentarily sketched out above—happen in societies in constant flux? In earlier work, I have focused on two particular processes: patrimonial and ancestral deixis (Lüpke 2016a, b, 2018a). Through patrimonial deixis, a connection is made between a focal language of the perceived founders with the place. Patrimonial languages thus index firstcomer status, which is often associated with land rights and control over important land-related rituals, such as sacred groves, sacred trees, some spirit shrines, etc. A patrimonial language is not seen as reflecting the linguistic profile of the totality of the place's dwellers—it iconizes founders and their privileged relation to language and land, the patrimony. This deictic process acknowledges their special status without erasing the presence of newcomers, who are welcome as clients as long as they do not threaten the founders' authority. Although introduced in colonial times, village chiefs are still often issued from the founder's family and thus reinforce this social category. It follows that in these language ecologies, landlords are not interested in linguistically integrating strangers, because mastering the patrimonial language would give them access to rituals and hence, founders' powers. It also follows that the patrimonial language of a place is not necessarily its most widely spoken language, nor one that occupies a privileged place in founders' and their descendants' daily language use. While the patrimonial language has the monopoly in a number of ritual practices related to land and ancestors, other sociolinguistic spaces give room to all inhabitants' linguistic repertoires. Patrimonially organized language ecologies are thus an instantiation of a multilingual speech community in Gumperz's sense. Its members may see themselves or be seen by others as members of different language communities, depending on the scale at which wider language-based identity claims are made.

Ancestral deixis conceptualizes present-day inhabitants of a place as the descendants of ancestors perceived as having the same ethnolinguistic identity. Linguistically, this identity model is associated with seeking to preserve a language as an 'ancestral code' (Woodbury 2005; Childs et al. 2014) of a community seen as one where language community and speech community are desired to coincide, and where any deviation from their overlap is interpreted as a sign of unwelcome change. Inhabitants not conforming to the linguistic profile are erased from its representation, and multilingual practices not conforming to the monolingual ideal are discounted and rendered invisible or illegitimate (Chapters 29 and 32, this volume). The ancestral model of language territorialization is connected to (ethno)nationalist imaginations of identity—although these

unite and homogenize historically heterogeneous places, they are often post hoc extended into the past.

At the current time, both models of language territorialization co-exist, often with conflicting claims over the same space. These tensions are as much an outcome of ongoing political changes favouring ethnonationalist identity models as a reflex of conflicts inherent in the landlord–stranger model. From the nineteenth century onwards, we find a break with the firstcomer–newcomer cohabitation model and a tendency to create places according to ethnic identity (Baum 1999). Individual landownership is also on the rise and perturbs older patterns of land stewardship, which often remain reflected only in rituals. However, as described in detail by Lentz (2013), inhabitants who were originally in a situation of clients or strangers have strong interests to overthrow their dependency on landlords and reject this dualistic model altogether, thus being the agents of far-reaching changes in political organization that also have repercussions on which languages are seen as having a right of existence. It is thus perhaps not surprising that it is among Joola-identifying elites that ancestral models of language territorialization and nationalism (Foucher 2005a; Evans 2013) are particularly widespread, while among many Baïnounk, pride in being multilingual and accommodating strangers is still very widespread (Ducos 1979; Lüpke 2016a; Cobbinah 2020; Chapter 13, this volume)—archaeological research points to a later immigration of populations now identified as Joola (Linares de Sapir 1971). In oral history and public imagination, Baïnounk, regardless of actual migration history, are generically seen as the autochthones (Lespinay 1987a). Thus the commitment of Baïnounk-firstcomers to diversity and to patrimonial deixis can be seen as continuing adherence to a political patrimonial model that grants them special rights, and the nationalist turn of many Joola-latecomers to an ancestral model as motivated by changing a status quo that politically disfavours them.

In addition to these two language-ideological patterns, it remains to be researched what the effect of the widespread spirit shrines is on imaginations of language, individual and communal identity, and language use. Many shrines, such as the powerful shrine of Mama Jombo in Kaboi, attract thousands of followers and have satellite shrines all over the region (Crowley 1990). In patrilineal societies, founders and their descendants are male. The extents to which mobile women, children, and labourers are represented, what their communities of practice are and how these intersect with language and speech communities remains a topic for future research, as are studies on the language territorialization and identity among matrilinear societies in the Atlantic space, for instance in the Bijogos archipelago (Segerer 2002).

Based on the sociopolitical significance of language in different pragmatic contexts, we can differentiate contexts for languages and contexts for languaging in the speech communities of the Atlantic space. The former contexts are associated with situations in which the language community is referenced. Depending on the prevailing language ideologies, these may be mainly confined to ritual contexts and situations related to land rights (Cobbinah et al. 2017; Lüpke 2016a, 2018a, 2021a; Cobbinah 2020) in patrimonial language ecologies. Or they may be extended to more sociolinguistic spaces within a settlement (Goodchild 2018, Chapter 32, this volume). Language-based contexts, here meant as contexts characterized by the pragmatically motivated desire to remain within the confines of a named languages, may also comprise multiple languages kept pragmatically separate, for instance in contexts of translation.

Language-based contexts are flanked by situations in which the emphasis is not placed on performing a language, but on being convivial and sharing communication. Such contexts call for languaging, 'the process of using language to gain knowledge, to make sense, to articulate one's thought and to communicate about using language' (Makalela 2015: 203). In these contexts, it is not the relationship between language and land or political rights that is foregrounded, but relationships between people. Relational identities, not based on projecting absolute social and linguistic identities but on creating and maintaining multiple relationships have been revealed as central for West African ideas of identity (Nyamnjoh 2017; Di Carlo et al. 2020). It is crucial to stress that these contexts not only give rise to thickly multilingual speech, but that they blur the distinctions made in language contexts, yielding speech forms that can be seen as code-switched, ambivalent, fused, or mixed depending on analysts' repertoires and assumptions.

It is because of the social importance of language contexts that different named languages survive as bundles of divergent linguistic features. Regarding convergences, I propose that different contexts generate qualitatively different types of mergers. In this context, it is important to stress the great agency of speakers to control different socially governed language contexts and adapt cognitively and socially to them (see also Green and Abutalebi 2013). Language contact situation are often theorized based on dominance relations between languages, both in terms of prestige and of frequency (Winford 2020). While these hold without doubt, they are counteracted by the pragmatic intentions of versatile multilingual speakers in particular discourse contexts (see also Backus 2020).

In language-based monolingual interactional contexts, the pragmatic intent is to produce speech in only one named

language. Language-based contexts can also occur in more than one language. Such multilingual language-based contexts are, for instance, situations in which the intention is to produce equivalent speech in several named languages, for instance in religious and federal contexts requiring translations. I postulate that the main contact effect of this context is to create semantic and syntactic equivalences in order to facilitate translatability, through priming and conscious stylistic choices. Languaging contexts stand in contrast to these two types. In them, individuals use those parts of their repertoire deemed appropriate in the situation (based, for instance on knowledge about the linguistic regime of the space, on knowledge about speech participants' status and repertoires) in order to communicate and create a community of practice.

The assumption of qualitatively different contexts based on their proclivity for language vs. for languaging resolves one of the most striking paradoxes of language contact: that it appears to be at the same time highly constrained **and** omnipresent and unrestricted; and that communities can be reported to exhibit both strict language separation and purist language ideologies **and** fluid multilingual speech (Stenzel and Khoo 2016; Lüpke et al. 2020). In particular, basing scenarios not on entire languages but on discourse contexts allows us to model where convergent linguistic features originate and how divergent features are maintained or created.

Language-based contexts are characterized by the pragmatic intent of speakers to produce speech in a particular named language. Cognitively, this means that multilingual speakers have to deactivate large parts of their repertoires (Green and Abutalebi 2013) and select only those features they deem appropriate for the discourse context (Lüpke and Watson 2020). Their speech will never be completely monolingual in hermetically sealed fashion; however, their intent, coupled with knowledge of their repertoire, makes it possible to analyse some of the linguistic forms that seem to violate the monolingual constraint originating as code alternations—mainly in terms of lexical insertions into what is otherwise a clear matrix language in pragmatic terms. These code alternations are qualitatively equivalent to borrowings, the distinction being one of synchrony vs. diachrony and innovation vs. conventionalization (Backus 2020).

Language contexts may be multiple, particularly in situations where translation is required for successful communication or political representation and equity, for instance in village assemblies, in parliament, or in religious domains (see Chapters 27 and 30, this volume). In these contexts, the constraints for language-based contexts apply, but in addition, fuelled by the motivation to make texts equivalent or even parallel, far-reaching convergences in semantic organization and syntax originate here.

Both language-based contexts are also sources of divergence—the social motivation for keeping named languages distinct as markers of identity counteracts the tendencies towards convergence. This results in the maintenance or creation of differential linguistic features.

When not constrained by the nature of sociolinguistic spaces, large parts of speakers' repertoires are active, although in thickly multilingual settings it will never be the case that repertories completely match, and considerations of creating common ground will constrain choices. Here, very little deselection occurs cognitively or socially, and the results are far-reaching priming effects from a structural perspective, and playfulness and creativity from a social perspective. This context is one of languaging that can be analysed as languaging and/or translanguaging—transcending and recombining features of several named languages—based on participants' and observers' language experience and concomitant categorial options. Language use in social media is one of the prime examples for this context in the written domain (see Chapter 31, this volume).

The bulk of linguistic description is inscribed in a monolingual paradigm with the intention to describe or document a single named language. Therefore, the data collected in this framework are mostly in mono- or bilingual language modes—they consist of texts produced in a monolingual language mode and elicitations and translations in mono- or bilingual language modes. These products of linguistic research constitute perfectly admissible and possible samples of language use. However, they have two side effects. First, through the artefactualization that characterizes the process, they unequivocally enregister forms as belonging to a named language that may be weakly or multiply enregistered for speakers. Second, because of the text types collected, they favour a particular look at multilingualism: one of contact, i.e. of the static outcomes of language interaction in the most **lingual** setting, instead of investigating dynamic processes of feature recombination and feature creation in interaction.

While only an investigation of languaging contexts allows us to witness the dynamicity of linguistic settings and the actuation of linguistic change, many of the features that crosscut the boundaries between named languages enable us to see the lasting effects of multilingual communication in different discourse contexts. They will be examined in the following section.

1.4 Relationships between named languages and linguistic features

> '[S]tudies on bounded languages reinforce participants' meta-awareness of languages as separate bounded systems. The aims of language documentation and people's meta-awareness of languages [. . .] reinforce the perception of this language existing as a bounded code. The challenge is for the sociolinguist to accept these bounded systems as existing in people's perceptions, whilst simultaneously studying actual multilingual practices where these boundaries may not necessarily be evident, and furthermore to problematise language as a system or construct.' (Goodchild 2016: 85)

It has become evident that most glossonyms originate from social actors reifying linguistic features or feature bundles because of their significance for identity. This process can have radically different outcomes: it can lead to the differentiation in name of languages that are near-identical in their linguistic characteristics (such as Kim and Bom, Chapter 20, this volume). Or it can result in heterogeneous and highly variable features being labelled as belonging to the same language because this internally diverse feature pool (Mufwene 2001b) references a language community or because the speaker is identified as belonging to a particular language community in the speech context. Recall that where speech forms are categorized according to linguistic criteria, both the maintenance of minimal differences at a local level and the pooling together to create entities at a larger scale coexist, depending on the way in which the contrast is framed.

It follows that it is of particular interest which features and feature bundles are widely shared, and which are not, since the indexical potential of widely diffused features, and indeed their enregisterment and unambigous reification are likely to be very weak. Many linguistic features are shared cross the entire Atlantic area and beyond, cross-cutting genetic groupings. Typological studies highlight linguistic convergences of particular linguistic features. Clements and Rialland (2008) identify a Sudanic belt based on a number of statistically significant phonological features: the presence of labial flaps, labial-velar stops, implosives, nasal vowels, vowel systems distinguishing two height levels and with vowel harmony, multiple tone levels, and 'lax' question prosodies. They conclude that Atlantic languages do not fit this pattern very well—for instance, there are only few tonal languages in this family, and many of the phonological features are not attested or only present in a small subset of languages. Güldemann (2008) employs the criteria of logophoricity, labial-velars, labial flap consonants, ATR vowel harmony, and S-AUX-O-V-X and V-O-NEG word order in order to identify what he calls the macro-Sudan belt, but grants that Atlantic languages do not show a clear picture regarding the occurrence of these features and lack many of them. More in-depth studies of feature distribution in the Atlantic space are thus needed in order to identify micro-convergence zones within it in order to re-appraise its relationships with other linguistic areas based on them. At the present moment, we can observe a number of features that robustly occur in the Atlantic area, such as widespread vowel harmony, although the exact types and their significance based on cross-linguistic frequency of occurrence remain to be determined. The widely noted tendency in Niger-Congo for dynamic verbs to have past reference when zero-marked for TAM and for stative verbs to have present reference when zero-marked for TAM (Welmers 1971) applies to the Atlantic area, including languages not genetically Atlantic, such as Kriolu (Chapter 6, this volume) and Mandinka (Chapter 7). Verbal extensions (Chapter 24, this volume) as a morphological means to code changes in argument and participant structure are likewise common in the Niger-Congo phylum as a whole.

Night-time taboos have been reported by Biagui et al. (Chapter 6, this volume). In Kriolu, *karboŋ* 'charcoal' cannot be uttered at night but needs to replaced with *wuru pretu* 'black gold'; likewise, *sal* 'salt' becomes *wuru braŋku* 'white gold' and *kobra* 'snake' turns into *liña di coŋ* 'thread from ground'. Similar lexical avoidance strategies of items associated with witchcraft are spread not only through the entire Atlantic space and adjacent areas—in Joola Fogny, Baïnounk Gujaher, Bambara (Lüpke, fieldnotes), but also in Songhay (Minkailou 2016) and Kinubi (Luffin 2002), thus spread out from the Atlantic to the Indian ocean.

Within the noun class systems, a gender reserved for humans and a noun class prefix for naturally occurring liquids with the phonological shape **m-** as well as augmentative and diminutive genders are not only widespread in Atlantic languages, but link these languages to Niger-Congo (Chapter 22, this volume). Note, however, that the latter class is absent from Balant Ganja (Chapter 15, this volume). Creissels notes that, overall, languages of the Bak branch of Atlantic are closer to the Niger-Congo prototype than those of the North branch.

Candidates for features that might, especially in particular feature clusterings, have areal significance within a more locally Atlantic space, are the following: inclusive/exclusive distinctions in first person plural pronouns, which are

attested in many Atlantic languages, and Kriolu (Chapter 6, this volume), although not in Mandinka (Chapter 7).

Shape-based genders probably originating in the botanical domain are very characteristic for the semantics of Atlantic noun class systems overall (Chapter 22, this volume), while triadic genders distinguishing mass/collective, singular, and plural are mainly attested in Bak and Nyun languages. Non-finite verb forms that are deeply integrated into the noun class system and occur in different genders based on semantic distinctions appear to be a defining trait of Atlantic noun class systems (Chapter 23, this volume). Number marking outside the gender system with suffixes of the shape -Vn are common in Buy (Voisin 2015b), Nyun (Cobbinah and Lüpke 2014; Cobbinah 2017), and Sua (Chapter 16, this volume), and a number marker not indexing gender also exists in Nalu (Chapter 17, this volume) and Biafada (Bassène 2015b).

A closer look at language-individual and shared lexicalization patterns reveals characteristic traits. Patterns are often similar in templatic fashion but also exhibit fine-grained differences, different alignments of isoglosses for different lexical items, and multiple source languages for items identifiable as loanwords. I will illustrate shared general patterns through data on numeral systems with anthropomorphic components attested, for instance in Bainounk Gubëeher (Chapter 12, this volume), Baïnounk Gujaher (Chapter 14), Joola Kujireray (Chapter 10), and Joola Banjal (Bassène 2007). Three of these languages, Gubëeher, Kujireray, and Banjal, are the patrimonial languages of neighbouring villages, with different patterns of multilingualism in all three languages through recent history. Gujaher is closely related to Gubëeher but not in direct contact with it. In all four languages, numbers up to twenty are based on the human body, with the basic units 'five' and 'ten' related to hands, 'fifteen' expressed through an added foot, and 'twenty' designated with a word that means 'king'—standing in for a person and all the digits of their hand and feet (*unam* in the two Baïnounk languages, *ávvi*/*ëvi* in the two Joola languages. In Gubëeher, the word for 'five' is *cilax* 'hand'. In Kujireray, spoken 200 metres down the road in Brin, 'five' is not expressed with the word for hand, but with the word for 'fist', *futox*, cognate to *futok*, the form used in Banjal. But 'ten' is based on the word for 'hands' in all three languages—*halax* in Gubëeher and Gujaher, *kuñen* in Kujireray, and *guñen* in Banjal. From twenty to hundred, the numeral system is organized around hands, feet, and multiples of kings. Hundreds are counted decimally (with multiples of ten). All three languages share the lexeme for 'hundred': *teemeer* (Gubëeher and Gujaher) or *eteemir*/*etémel* (Kujireary and Banjal), originally from Wolof. 'One thousand' is expressed with a word originating in Mandinka, another lingua franca of the wider area: it is *wuli* in Gubëeher and Gujaher, and *euli*/*éuli* in Kujireray and Banjal. Phone numbers are given in French, and currency has its own counting system based on five as the basic unit, with larger sums are given in French. This pattern for counting money is attested throughout former French Colonial Africa. When expressing a currency amount, the base number needs to be divided by five in order to arrive at the denominational amount. For instance, if a price is 100 Francs CFA, in Wolof, Bambara, Jalonke, Gujaher, etc., it would be expressed with the number twenty. If the expression of the equivalent of 5,000 in monetary terms is desired, this would be the number 1,000. The probable reason for this parallel counting system is that a five-francs piece has been the smallest coin in circulation since colonial times, so this became equivalent with one (unit of currency).

It is likely that areal convergences originate in small-scale multilingual ecologies which are not confined to this area alone but attested throughout rural and urban Africa (Di Carlo et al. 2019). The discourse contexts in which speakers communicate(d) and index(ed) identities through language are of prime importance for a nuanced understanding of their spread.

1.5 Presentation of the volume and outlook

> 'Le terrain africain représente pour les linguistes un réel défi, du fait de la grande hétérogénéité des groupes et des moyens de s'exprimer. Lorsqu'on laisse émerger cette hétérogénéité linguistique à partir des interactions elles-mêmes, sans trop projeter de catégories préconstruites pour la faire apparaitre et l'analyser, les concepts, les théories et les techniques de recueil des données sont mis à l'épreuve.' (Juillard 2010)
>
> [The African field presents a real challenge to linguists, because of the great heterogeneity of groups and their means of expression. Once one lets this heterogeneity emerge from the interactions themselves, without much projecting of preconstructed categories in order to make them appear and to analyse them, the concepts, theories and techniques of data collection are put to the test.]

The present volume presents new data on one of the most linguistically diverse and complex areas in Africa, the continent itself being a hotspot of linguistic diversity and multilingual practices. Through focusing on four different

perspectives, the contributions to this volume allow us to approach the language ecologies in this area from different vantage points. The first part of the volume focuses on two different ways to conceptualize relatedness and feature spread, through the lens of genealogical relatedness (Chapter 2) and challenges to this method of capturing relatedness (Chapter 3). The sixteen chapters contained in Part II present grammatical overviews of individual languages in the Atlantic space. In addition to languages classified as genetically Atlantic in the latest classification are also included Kriolu (Chapter 6) and Mandinka (Chapter 7), the most important non-Atlantic contact languages as well as Kisi (Chapter 19), Bom-Kim (Chapter 20), and Baga Mandori (Chapter 18), Mel languages formerly classified as Atlantic. Many of these vignettes constitute the very first descriptions of languages (Chapters 6, 11, 13, 16, 17, 18, and 20), the first published research data on them (Chapters 8, 10, 12, and 14), or the first up-to-date overview articles available in English (Chapters 4, 5, 9, 15, and 19), thus constituting precious empirical data on languages whose potential to inform linguistic typology is largely untapped. Part III focuses on areas of linguistic structure of particular interest for morphosyntactic and semantic typology, with chapters on formal (Chapter 21) and semantic aspects of nominal classification systems (Chapter 22), nominalization (Chapter 23), verbal extensions (Chapter 24), and initial consonant mutation (Chapter 25). The final Part IV sheds light on the social circumstances in which languages are spoken. This part contains contributions on creolization (Chapter 26), the long tradition of Ajami writing (Chapter 27), secret languages and registers (Chapter 28), multilingual children's language socialization (Chapter 29), Fula as a lingua franca (Chapter 30), digital multilingual repertoires (Chapter 31), and the relationship between monolingually conceptualized places and multilingual people (Chapter 32).

The outcome is a kaleidoscopic account of diachronic and synchronic linguistic diversity in the Atlantic space. The volume provides answers to what languages are from social, typological, genealogical, and areal perspectives, how linguistic features emphasize languagehood or blur boundaries through diffusion of their features. It offers case studies on how identities based on overcoming difference are built, how multilingualism is played out in speech and writing, how children grow up in multilingual societies and how adults use languages in different sociolinguistic spaces. In particular the attention paid to the multilingual language ecologies in which Atlantic speakers live sets the volume apart from many earlier accounts of Atlantic. It is hoped that this holistic outlook will inspire future research to further advance the integration of descriptive and sociolinguistic research.

Acknowledgements

My heartfelt thanks go to an anonymous reviewer and to Alexander Cobbinah for their insightful feedback. As always, I'm indebted to all research participants and researchers of the DoBeS project 'Pots, plants and people—an investigation of Baïnounk knowledge systems' and of the Leverhulme Research Leadership Project 'At the Crossroads—investigating the unexplored side of multilingualism' which I had the privilege to lead, and on whose findings the present chapter is based. I began planning this volume while at SOAS, University at London, continued working at it after moving to the Department of Languages at the University of Helsinki and finished it while at the Helsinki Collegium of Advanced Studies at the University of Helsinki. I am grateful to these institutions for granting me research leave and fellowships to bring this project to fruition.

CHAPTER 2

A genealogical classification of Atlantic languages

Konstantin Pozdniakov and Guillaume Segerer

2.1 Introduction

The most widely cited classification of Atlantic languages is that of David Sapir (1971), which is reproduced in Figure 2.1.

Sapir's work was undoubtedly the most comprehensive of the time on these languages. It is based on lexicostatistics but it also presents a detailed account of nominal classification and consonant mutation.

These results have often been republished in various scholarly works as well as well-known media such as *Ethnologue* (Lewis et al. 2015), sometimes with some minor errors: for example, in Wilson (1989: 92), Fula is grouped together with Wolof as opposed to Sereer, whereas Sapir correctly stated that Fula and Sereer form a valid subgroup.

Wilson thus apparently used Sapir's figures, but whereas Sapir carefully set five independent groups in the Northern branch, Wilson proposed a more complex internal classification but without any convincing evidence. His 1989

I. Northern Branch
  A. Sénégal Languages
  1. a. Fula
  b. Serer
  2. Wolof
  B. Cangin Languages
  Lehar, Safen, Non; Ndut, (Falor)
  C. Bak Languages (Lower Casamance)
  1. Diola Group: Fogny, Huluf, (other Diola dialects); Gusilay; Karon; Kwaatay; (Bayot)
  2. Manjaku; Papel; (Mankanya)
  3. Balanta; (Ganja)
  D. Eastern Sénégal-Portuguese Guinea Languages
  1. a. Tenda Group: Tanda, Basari, Bedik, [Bapɛŋ]
  b. Konyagi
  2. Biafada; Pajade
  3. a. Kobiana; (Kasanga)
  b. Banhum
  E. a. Nalu
  b. [Mbulungish (otherwise known as Baga Foré)]
  c. [Baga Mboteni]
II. Bijago (isolated)
III. Southern Branch
  A. Sua
  B. Mel Languages
  1. Temne, Baga Koba, (Banta, Landuma, Tyapi, other Baga languages (excluding 'Foré' and 'Mboteni'))
  2. Sherbro, Mmani, (other Bullom dialects), Krim, Kisi
  3. Gola
  C. Limba.

Figure 2.1 David Sapir's classification (Sapir 1971: 48–9)

Konstantin Pozdniakov and Guillaume Segerer, *A genealogical classification of Atlantic languages*. In: *The Oxford Guide to the Atlantic Languages of West Africa*. Edited by: Friederike Lüpke, Oxford University Press. DOI: 10.1093/oso/9780198736516.003.0002

paper as well as other publications still represent a crucial contribution to Atlantic comparative linguistics. While Sapir had worked with materials collected from other scholars (except for Joola Fogny for which he had done personal fieldwork), Wilson had collected many first-hand data on all the languages spoken between the Gambia and Northern Guinea, materials that he eventually presented in his 2007 book (Wilson 2007) but which was available long before in manuscript form to any scholar asking for it.

The next step in Atlantic comparative studies was made by J. L. Doneux. Based on sound correspondences, grammatical and lexical innovations, and typological features, he formulated several interesting hypotheses, as for example a Wolof–Buy subgrouping (1978). He was also the first to present lexical and grammatical isoglosses, as well as a preliminary reconstruction of the proto-North Atlantic noun class system (Doneux 1975a). It was much later that Doneux (1991: 142) presented a general classification of Sapir's Northern branch, with a few differences. Doneux's contribution has been largely ignored, and since Sapir's genetic tree is still used as a standard by most scholars, the latter must be updated, all the more since we have access to a much richer documentation:

- Many grammatical descriptions of languages that were virtually unknown in the 1970s or 1980s have appeared since then: Bijogo (Segerer 2002); Baynunk Gubaher (Cobbinah 2013a); Baynunk Gunyaamolo (Bao-Diop 2013); Buy (Doneux 1991); Tenda (Ferry 1991; Wade 2015); Jaad (Meyer 2001a, b; Cover 2010); Noon (Soukka 2000); Laala (Dieye 2010); Bayot (Diagne 2009); Joola Banjal (Bassène 2007; Tendeng 2007; Sagna 2008); Joola Kujireray (Watson 2015); Joola Karon (Sambou 2007); Joola Kwaatay (Coly 2012); Pepel (Ndao 2011); Sereer (Renaudier 2012); Manjaku (Buis 1990; Doneux 1993); Balant Kentohe (Doneux 1984a); Balant Ganja (Creissels and Biaye 2016); and Gola (Koroma-Fachner 1994).
- Studies in various fields of theoretical or comparative studies: noun classes (Doneux 1991; Pozdniakov 1993; Creissels and Pozdniakov 2015); consonant mutation (Pozdniakov 1987; Storch 1995; Pozdniakov and Segerer 2006); verbal extensions (Becher 2000; Nouguier Voisin 2002); pronominal systems (Pozdniakov and Segerer 2004a, b); and language contact (Vydrin 1999b; Childs 2003b; Lüpke 2010a, b; Pozdniakov, Segerer, and Vydrin 2019).
- A considerable amount of lexical data has been digitized: today we can rely on more than 200,000 lexical records representing all the Atlantic languages. These records are available through the RefLex website (Segerer and Flavier 2011–2023 - www.reflex.cnrs.fr), which provides a set of tools for cognate search and reconstruction.

The authors of the present chapter are currently working on the reconstruction of proto-Atlantic: we have postulated hundreds of cognates, many of which give convincing evidence for subgrouping.

Before presenting the classification, we wish to emphasize the methodological and theoretical aspects that distinguish our work from the previous ones (in addition to the new database, which was just unavailable half a century ago).

1) Today we can rely on regular sound correspondences between Atlantic languages rather than on surface similarities, as used by both Sapir and Wilson. Without regular correspondences, how could one be sure that Joola *yuut* 'star' and Kobiana *a-woolu* 'star' are cognates? The regular regular reflexes of Proto-Atlantic *-ʈ in Table 2.1 give us the piece of evidence we need.

Table 2.1 A regular sound correspondence in Atlantic

Gloss	'star'	'to fly'	'to die'	'to rot'	'three'	
	*kʷʊʈ	*yiiʈ	*keʈ	*pʊʈ	*taʈ	*-ʈ
Joola Kasa	*yuut*	*it*	*kɛt*	*fut-en*		t
Bayot	*doos*			*-ɸus*		s
Manjaku Bassarel		*iʈ*	*cʌʈ*		*w-anʈ*	ʈ
Manjaku Cur		*iih*	*keeh*			h
Pepel	*wʊʂʊ*	*yɪʂ*	*kɛʂ*		*a-jenʂa*	ʂ
Mankanya		*p-yiʈ*	*p-keʈ*	*p-puʈ*	*wa-jenʈ*	ʈ
Balante-Fca	*f-gbóóθè*	*yiθ*				θ
Baynunk Gunyaamolo	*huur*	*yiir*	*cir*	*pʊr*		r
Kobiana	*a-woolu*	*k-il*		*wul*		l
Basari			*a-ʃʌ̀s*	*a-wàr*	*sàs/tàs*	r/s
Bedik	*ɛ-kór*		*ɔ-ʃʌ́s*	*u-buír*	*sás/tás*	r/s
Biafada	*wweela*			*bwəl*		l
Jaad	*puoor*			*pər*		r
Sereer		*yet*		*fot*		t
Nalu	*hooθ*	*yeeθ-en*				θ

Let us take another example concerning the two subgroups of the Cangin group (Table 2.2). Without etymological analysis it is difficult to relate *kinil* 'four' (Ndut-Palor) and *nikis* 'four' (Laala-Noon-Saafi). However, we can see in Table 2.2 that the correspondence -l ~ -s is regular.

Table 2.2 Cangin correspondences

*-ɬ	'eye'	'black'	'road'	'year'	'four'
Ndut	*ʔil*	*suul*	*wal*		*kinil*
Palor	*ʔil*	*suul*	*waal*	*kiil*	*kinil, enil*
Noon	*kwas*	**sujus*	*waz*		*nigis*
Laala	*kɔs*	**susus*	*was*	*kiis*	*nikis*
Saafi	*xas*	**suzus*	*was*	*kiis*	*nikis*

In these languages most numerals are recorded with a noun class prefix (Palor *ka-nak* 'two', *ke-ye* 'three', *ki-nil* 'four', *kip* 'five'), but in the NLS (Noon-Laala-Saafi) subgroup some of them have a suffix *-Vs* (Noon *jet-us* 'five'). The prefixed and the suffixed forms can be shown to be related, as illustrated with the words for 'four' in Table 2.2 which are reflexes of *Cangin *nik-iɭ (from Proto-Atlantic *nak(i) 'four'; Table 2.3)

Table 2.3 Cangin reconstruction

Proto-Cangin	**nik-Vɬ*		
Ndut/Palor	**ki-nik-Vl* >	*ki-nik-il* >	*kinil*
Noon/Laala/Safin	**nik-Vs* >		*nikis*

2) It is important to formulate our attitude towards lexicostatistics—the quantitative comparison of lexical cognates—when applied to a family as divergent as Atlantic. Lexicostatistics is a technique that allows us to obtain genealogical trees from the percentage of shared basic vocabulary within a group of languages. It is still in use although it has been regularly criticized for two major reasons: the cognacy judgements are mostly based on formal resemblances, which may well signal borrowing or shared retentions instead of real cognacy; and it will output a tree whatever the actual data look like. Nevertheless, it has also showed some efficiency (i.e. gave similar results to those obtained by more conventional methods) in the case of clear-cut subgroupings.

We consider that the lexicostatistical counts used by Sapir and his followers are too low to be useful for subgrouping at the intermediate levels, i.e. they may be used to identify obvious groups such as Tenda or Cangin but they cannot capture the relationship between, say, Nyun and Tenda. In fact, figures are so low that any slight difference in cognacy judgement may considerably affect the shape of the tree. Actually, Sapir's method was to count as cognates roots that do not diverge by more than one phonological feature (Sapir 1971: 49, footnote 8). In doing this, he was sure to count all the borrowings as cognates, which is precisely what lexicostatistics would like to avoid. But our own counts obtained with more data and more elaborated judgements yield no better results.

We have therefore decided NOT to rely on lexicostatistics at all for our subgroupings. However, statistics can provide indications on different phenomena: for example, an analysis of contradictions in lexicostatistical data may be used to assess the importance of contact in Atlantic languages and to detect large-scale internal borrowings within the family. We shall illustrate this approach below.

3) Last, we build our classification on shared innovations rather than on the study of similar retentions. This is especially important when we consider one striking feature of the Niger-Congo lexicon: some fifty lexical roots show an impressive stability throughout the various branches of the Niger-Congo phylum, while in many individual languages (independently of their grouping) they are replaced by borrowings or innovations. This means that the retention of such lexical roots in some Atlantic languages and their absence in other languages give no evidence of their grouping. The fact that the Niger-Congo root *tɪ 'tree' (Proto-Bantu *tí, Proto-Gbaya *tè, Proto-Kwa *tí, Nembe (ijo) *tị̀*, Proto-Oti-Volta *ti etc.) can be found in Bijogo (*ŋu-te*) and in Konyagi (*æ̀-tɔ́x*) but not in Wolof (*garab*), or in Manjaku (Manjaku Bok *-ko*; Manjaku Cur *bo-mol*), does not prove that Bijogo is more closely related to Konyagi than to Manjaku. Nevertheless, considering the very low figures of lexicostatistical scores, these kinds of roots may have a big influence on the configuration of the genealogical tree of Atlantic languages. On the contrary, the detection of common innovations within the Atlantic family is the most effective criterion for their classification. Our conclusions

will therefore be based on four kinds of innovations: phonetical/phonological, morphological, lexical, and semantic.

The structure of this chapter will be as follows: first, we will address the question of the relationship between Atlantic and Mel languages. Then we will present the updated genealogical tree of Atlantic languages and we will discuss its major splits. In the following step every node of the proposed tree will be argued for. Finally, we will consider internal borrowings within the Atlantic family, and address the question of isolate languages.

2.2 Outline and limits of the Atlantic family

One important point of this chapter is to put emphasis on the fact that there is, to date, no convincing evidence that the Mel languages should be put together with the other Atlantic languages in a single Niger-Congo branch. The first scholar who clearly formulated this idea was Dalby (1965), who even recommended to drop the label 'West-Atlantic' except as a mere referential term. Nevertheless, a few years later, Sapir presented fifteen lexical cognates between Sua, Limba, Mel, and 'Northern' Atlantic languages (Sapir 1971: 54–5). Sapir himself noted that seven of them are retentions from Niger-Congo, which means that they cannot be used to support the Mel–Atlantic hypothesis. This is the only attempt to prove a connection between Mel and Atlantic, a hypothesis that has been systematically repeated since,[1] without any additional evidence. Moreover, Sapir's lexical counts show that Mel and the rest of Atlantic are only very remotely connected (never more than 10% similarities in the Swadesh-100 list between any Mel language and any North Atlantic language, down to 3% between Temne and Wolof, which is less than chance resemblance). There is little doubt that Fula, Basari, and Balant on one side, Temne, Kisi, and Sherbro on the other side are genetically related, but only because both sets belong to the Niger-Congo phylum.

So far, and in the light of many published and unpublished data at our disposal, there is not enough evidence to support the fact that Mel is more closely related to Atlantic than to any other branch of Niger-Congo. We wish to insist that we are not claiming that there is no such thing as a Mel–Atlantic family, but rather that there is still no basis for such a grouping. From now on we will use the term 'Atlantic' to refer to those languages previously labelled as 'Northern Atlantic' plus Bijogo, whose position in the Atlantic family will be discussed below.

Hence, we do not give details on the classification of the Mel languages, since they are excluded from the scope of this chapter. Those interested in Mel comparative study may refer to Dalby (1965) and Chapter 3, this volume). Let us just add two comments about Mel:

- There are two uncontroversial subgroups: Northern Mel, with Temne, Landuma, Baga Koba, Baga Maduri (see Chapter 18, this volume), Baga Sitemu; and Southern Mel, with Kisi (see Chapter 19, this volume), Sherbro, Mani, Kim, and Bom (see Chapter 20, this volume).
- Three languages had been included with Mel in a 'Southern' branch of Atlantic by both Sapir and Wilson: Sua, Limba, and Gola (the latter being even included in Mel proper as a third subgroup). We have not found any solid evidence for this grouping. We have therefore included these three languages in the scope of our study in order to assess their possible inclusion within Atlantic.

Overall, the situation of the Atlantic languages is as follows: along the coast of Africa between Senegal and Guinea one finds 50–60 languages that can be grouped together in what we call the Atlantic family of Niger-Congo. In addition, there are a few languages whose affiliation is more disputable (Limba, Gola, Sua). Finally, some languages are very likely to be Atlantic but their position inside the group is yet to be defined (Nalu, Baga Fore, Pukur).

2.3 The internal classification of Atlantic languages

Figure 2.2 shows the internal classification of Atlantic languages.

In our view, the Atlantic family consists of two main branches called North and Bak (this avoids the term 'South' or 'Southern' which would evoke the branch postulated by Sapir). In addition, there are a few isolated languages or clusters (namely Gola, Limba, and Sua), for which we could find no convincing evidence for subgrouping.

[1] Principally by Wilson (1989), cited in Williamson and Blench (2000), repeated in Wilson (2007) and still used in *Ethnologue* (Lewis et al. 2015).

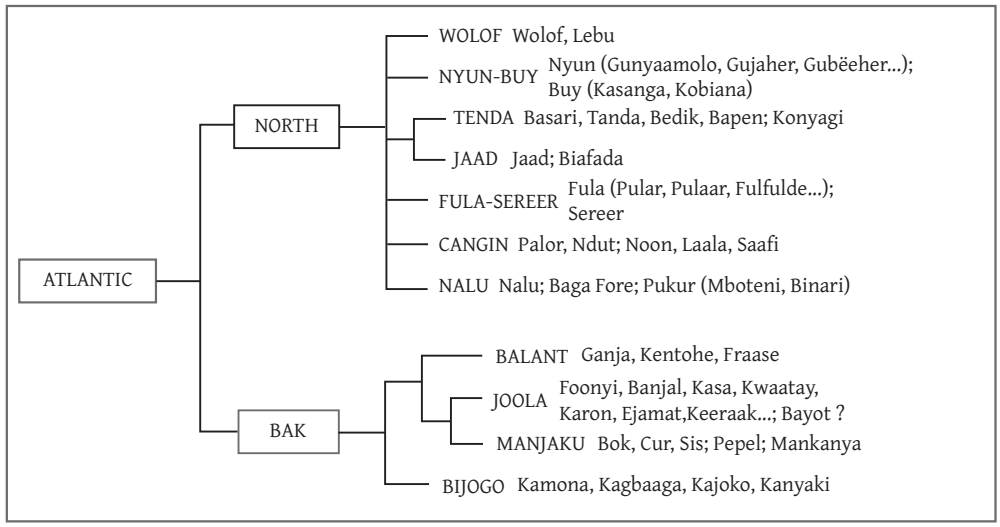

Figure 2.2 Classification of Atlantic languages

2.4 The North/Bak split

The earliest split of proto-Atlantic led to two branches, namely North and Bak. The North branch includes the following groups:

2.4.1 North Atlantic languages

A. Wolof; Lebu.
B. Nyun-Buy: Nyun (*aka* Baynunk): Gutobor, Gunyaamoolo, Gubəher, Gufangor, Gubelor, Northeastern, Gujaher; Buy (Kobiana; Kasanga) (Cobbinah 2013a: 31).
C. Tenda: Basari, Tanda, Bedik, Bapen, Konyagi; Jaad *aka* Badiaranke; Biafada.
D. Fula (Pular, Pulaar, Adamawa Fulfulde, Maasina Fulfulde...); Sereer (Sine, Petite Côte, Nyominka, Fadiout).
E. Cangin (Ndut, Palor, Laala *aka* Lehar, Noon, Saafi *aka* Safen).
F. Nalu (Nalu, Baga Fore, Pukur *aka* Baga Mboteni).

The most striking innovation that unites North Atlantic languages is the development of initial consonant mutation. There are three mutation grades, and every noun class is attached to a specific grade. One can find this consonant mutation mechanism in four North Atlantic subgroups among five. Only the Cangin languages seem to keep no trace of it. The Baynunk subgroup does not show any consonant mutation either, but the closely related Buy subgroup does. As for Baynunk, it is very likely that consonant mutation has been lost due to the heavy influence of Joola (Bak branch) during centuries of contact. This strong influence is, of course, also reflected in the lexicon. As for Cangin, the situation is less clear. There are two possible scenarios: 1) If Proto-Cangin represents the first split from Proto-North Atlantic, then the emergence of consonant mutation took place after this split; 2) if not, we have to conclude that consonant mutation has disappeared from Proto-Cangin,[2] maybe in conjunction with the degradation of the noun class system. Consonant mutation in North Atlantic languages can be reconstructed as in Table 2.4 (Pozdniakov 1987).

In Proto-North Atlantic, only nasal sonorants (*m, *n, *ɲ, *ŋ) and glottalized stops (*ɓ, *ɗ, * ʃ, *ɠ ?) did not undergo mutation. The diachronic aspects of Atlantic mutation

[2] Actually, there is a handful of words that could possibly illustrate an earliest stage when consonant mutation was at work: Laala *ka-gud* 'to cut'/*ngud* 'stump' (Dièye 2010); Palor *wət* 'to surpass'/*mbət* 'to be numerous' (d'Alton 1987); Saafi *nuhun* 'to point at'/*ndukun* 'finger' (M'bodj 1983).

Table 2.4 Consonant mutation in North Atlantic

degree	*voiceless						*voiced				
III	(m)p	(n)t	(n)c	(ŋ)k	(ŋ)kw	(ŋ)q	mb	nd	nj	ng	ngw
II	p	t	c	k	kw	q	b	d	j	g	gw
I	f	r	s	h,x	xw	x	w	l	y	ɣ	ɣw

are discussed in various publications (Storch 1995, 1996; Mc Laughlin 2000, Chapter 25, this volume; Pozdniakov 2015, etc.).

Interestingly, the distribution of mutation systems in North Atlantic languages nicely correlates with the (complementary) distribution of glottalized and geminate consonants (Table 2.5).

Table 2.5 Glottalized and geminate consonants in North Atlantic

Subgroup	Languages	ɓ, ɗ, ƒ	bb, dd, jj
Nyun-Buy	Nyun (Baynunk)	–	–
	Buy	–	+
Wolof	Wolof	–	+
Jaad-Biafada	Jaad, Biafada	–	+
Tenda	Basari, Bedik, Konyagi	+	–
Fula-Sereer	Fula, Sereer	+	–
Cangin	Ndut, Palor, Laala, Noon, Safin	+	–

All Northern languages except Nyun have either glottalized or geminate consonants and no language has both. In Bak languages these phonemes are not attested. The question whether this distribution and consonant mutation are two independent phenomena (thus making two isoglosses for North Atlantic) is beyond the scope of this chapter and remains open.

North Atlantic languages show a number of lexical isoglosses that separate them from Bak languages. Some examples are given in Table 2.6.

2.4.2 Bak languages

A very important grammatical innovation of Bak languages was used by Sapir for the naming of this group. For the Joola, Manjaku, and Balant clusters, Sapir proposed the name 'Bak', based on a particular feature concerning noun classes for human beings. In Proto-Atlantic (as in Niger-Congo in general) the plural of nouns is obtained through substitution of the noun class marker: in particular, in plural forms for human beings, CM2 **ba-* replaces CM1 **ku-*. However, in Proto-Bak this common strategy was modified: CM2 began to be used as a preprefix: SG **ku*-STEM ~ PL **ba-ku*-STEM. This atypical strategy caused various morphological changes in dependent markers and, more specifically, in pronouns. Sapir describes these changes in a note:

> 'In general the *BAK* form appears in demonstrative dependents beginning with a vowel and plural personal pronouns: Balanta: *bək-ɔɔbɔ* 'these'; Manjaku: *ba-ŋan bik-i, buk-un* 'these, those people'; Papel: *(b)ɔk-i, (b)ɔk-un* 'these, those'; Fogny (where its use is optional): *(b)uk-ɛ* 'these'; Kasa Diola: *a-kum-buk-e* 'these'; Gusilay: *bug-ɛ* 'these'; and Kwaatay: *bɔk-ɔn-u* 'those there'. For the disjunctive plural pronouns, third person: Manjaku: *bukul*; Papel (object): *bɔcɔ, bɔkɔ*; Joola: **bukɔ* (in Kasa it also appears in the second person plural *bukurul*); Gusilay: *bugɔ* and second person plural *bugul*; and Kwaatay: *bɔwɔkɔ*. Otherwise it appears in the Manjaku and Papel plural for 'woman': M. *ŋaats/bakaats*; P. *ŋaars/ba'ars*, and the Diola (all dialects) and Gusilay word for 'people' *an/bukan*. In the Diola of Samatit (Esulau) it turns up as the definite article, e.g. *ku-ɲil-a-buxu* 'children' (in other dialects this is simply *ku-ñil-a-k(u)*). Kwaatay makes more extensive use of the *BAK* marker. Besides appearing in demonstratives and as a pronoun it also concords with the genitive marker *-ta*: *ɛ-waatɛ yu baka-ta ku juwʌt* 'the people of Juwat' (note class shift in the dependant *ɛ* → *baka*); marks numerals *haarɛ bʌgʌ-suba* 'two women'; and, more prominently, appears as the regular

Table 2.6 Lexical innovations in North Atlantic

Gloss	TENDA-JAAD	FULA-SER[a]	CANGIN	WOLOF	NYUN-BUY
'to bear child'	**ɗəm*	**dim-*	*lim*	*doom*	*dim*
'to bite'	**ŋat*	**ŋat*		*matt*	*ŋar*
'old'	**ɗæf*	**rap*		*ràpp*	*def*
'fool' ~ 'stupid'		**ɗof*	*dəf*	*dof*	*dof*
'stranger'		**gen-*	*xan-eel*	*gan*	*gənaal*
'pestle'		**kun*	*kuɗ*	*kuur*	*huur*
'to hit'	**laɓ*		*laɓ*		*law-n*
'to last'	**ɓiŋ*	**miŋ-*	*mañ*		*miñ*
'easy'		**yɔɔɓ*	*yoɓ*	*yomb*	*yomb*
'calabash'	**gwom*	**gamb*	*gum*	*yomb*	
'lightning'	**beleŋ*	**meleŋ*	*ɓ/meleŋ*	*melax*	
'to inherit'	**dəng*	**don-*	*dɔ̄n*	*donn*	
'bow'	**kala-*	**qali*	*xalə*	*xala*	
'to wound, hurt'	**gaŋ*	**gaaŋ*	*gaŋ-*	*gaañu*	
'luck'	**mal?*	**mal*	*mal*		
'to dance'	**mbɛs-*	**mbec*	*mbəs*	*fecc*	
'day'	*pis*	**bes*	*bes*	*bés*	
'early'	**tii*	**teel*	*teel*	*teel*	

[a] The Proto–Fula–Sereer reconstructions come from Pozdniakov (2022).

> regent marker for personal plurals (especially kin terms): *baka-ti-am* 'my same sex siblings', *bʌkʌ-i* 'kings'.' (Sapir 1971: 81, italics are his).

Sapir did not include Bijogo in his Bak group but some traces of the *bak* forms do seem to exist there. Table 2.7[3] shows all the forms that exhibit an initial *bVk-*. We have included Bijogo as well.

In Bijogo, class pronouns are characterized by a final *-g*. Our hypothesis here is that this *-g*, exactly like the *-k* of the *bVk-* forms, originates in the former class 1 noun prefix **ku*. But Bijogo, unlike Joola, Manjaku, and Balant languages, expanded this element to all noun class pronouns, therefore choosing a 'paradigmatic' levelling instead of the 'syntagmatic' one typical of the continental Bak languages. There is not much evidence for this hypothesis, since the traces of this former class 1 **ku* are now very scarce:

Joola: the definite marker for class 1 (a-) is *akʊ* in Banjal and Gusilay (the regular forms are *aCʊ* for all CV-type classes). In other Joola varieties, *akʊ* is the (regular) definite marker for class 2 (*kʊ-*). As a consequence of this (probable) archaism, in Banjal and Gusilay the class 2 definite is *bʊkakʊ* ~ *bʊgagʊ*.

Balant: in the Ganja variety, class 1 is labelled as HA-class in the recent description by Biaye and Creissels (2015).

[3] A blank cell means that the corresponding form could not be found, but makes no assumption about its possible existence in the language. '—' means that the corresponding form is absent (i.e. another form exists for this function/meaning).

Table 2.7 Reflexes of *bak in the Bak branch

	3pl Pronoun	Object	Interr.	Possessive	Definite	Demonstr.	'people'
JOOLA							
Fogny[a]	*bʊkɔ*	—	*bukay*				*bʊkan*
Banjal	*bʊgɔ*	—	*bʊgai*	*bʊg-*	*bʊgagʊ*	*ʊbʊg-V*	*bʊgan*
Gusilay	*bʊkɔ*	—			*bʊkakʊ*	*ʊbʊk-V*	*bʊkan*
Kasa	*bʊkɔ*	—	*(bʊ)kai*			*akumbuk-V*[b]	*bʊkan*
Keeraak	*bʊkɔ*	—		—	—	—	*bʊkan*
Kwaatay	*bɔkɔ ~ bɔkun*	—	—	*bakatɛ (kʊ)*	—	*bɔkɔnd-V*	*bakan*[c]
Karon	*pak-o*	—	*pak-*			*pakeŋk-V*	*pakan*
Bayot	*bʊkɔ ~ bʊgɔ*	—	—	—	—	*buki-*	—
BALANT							
Ganja	—	—	*bìg-ílà*	*bìg-íná*	—	*bVg-*	—
Kentohe[d]	—	—	*bakala*	*bakina*	—	*bVk-*	—
MANJAKU							
Bok	*bukul*	*bukul*	*beke*	*bukul*		*bVk-*	—
Bassarel	*bukul*	*bukul*	*bikiyen*	*bukul*		*bVk-*	—
Mankanya	*bukal*	*baka*		*baka*		*bVk-*	—
Pepel	*bɔkɔ*	*baka*	—	*bɔkɔ*	—	—	—
BIJOGO							
Kagbaga[e]	*yag*	—	—	—	—	—	—
Kanyaki	*ag*	—	—	—	—	—	—

[a] Says Sapir (1965: 62): '*buka* is an anachronic plural person marker used with only one noun, *buka-n* "people", but frequently used as an optional prefix in a number of the concord forms: *kuŋil bukɔkɛ* or *kuŋil kɔkɛ* "the children are here".'. Unfortunately, the exact forms are not listed in the rest of the text.
[b] Apparently only in the 'emphatic' series (Wintz 1909: 19).
[c] Unlike in other Joola languages, there are (many?) other nouns with this plural prefix.
[d] The words for 'other' and 'only' are *bakolo* and *bakudi* respectively.
[e] In Bijogo, all the class pronouns have the shape *CL-g* with CL being the class exponent.

In fact, three dependent forms show an initial *h-* that could be a reflex of **k*: concord marker before V (*h-*), class pronoun (*hí*), and object index in certain conditions (*-hí*). In the Kentohe variety, the noun class prefix is *hə-*. However, there is no regular correspondence between Balant *h* and Joola *k*.

Manjaku: there is apparently no trace of **k* in class 1 concord paradigm.

Bijogo: the noun prefix of class 1 is *o-* which could possibly originate in **ko*, but there is no clear evidence for this hypothesis.

Another argument for considering Bijogo *g* as a reflex of Proto-Bak **k* can be found in a few lexical series (Table 2.8). Some Joola lects, as well as Balant Ganja, also seem to show *g* there but this can be argued to be a recent change.

Table 2.8 Sound correspondences in the Bak branch

	*K	'four'	'hippo potamus'	'crocodile'	'to dry'
JOOLA					
Fogny	***k/c***	*-bacɪr*		*cɪkɔn*	
Banjal	***g/x***	*sɪ-bbagɪr*			*ɛ-xay-ɛn*
Gusilay	***g/h***	*sɪ-baagɪr*			*ga-hay-ɛn*
Kasa	***k***	*ku-bakir*			
Keeraak	***k***	*-bakɪr*			
Kwaatay	***k***	*-baakɪr*			
Karon	***k***	*paakɪl*			
Bayot	***Ø***	*i-ßei*			
BALANT					
Ganja	***g***		*gɔ́mɔ̀nɛ̀*		
Kentohe	***k***		*nkom*		
MANJAKU					
Bok	***k***	*bakər*	*komʌl*	*tɪʌk*	*kʌy-an*
Bassarel	***k***	*kə-bakər*	*u-kɔmʌl*	*u-tɪʌk*	*u-kʌi*
Mankanya	***k***	*baakr*	*u-komal*	*u-teek*	*u-kay*
Pepel	***k***	*u-wakər*		*tɛkɛ*	*kay-ɔ*
BIJOGO					
Kagbaga	***g***	*ɛ-aganɛk*	*ɛ-gɔmɔr*	*ɛʧɛɛga*	*-ga-ɛ*
Kanyaki	***g***	*ŋɔ-agɛn*			*ŋɔ-gay*

Bak languages show other lexical isoglosses which allow to group them. Table 2.9 gives some examples.

Table 2.9 Bak exclusive lexical isoglosses

	Bijogo	Balant	Joola	Manjak
'soft'	*-bod(ɛ)*	*bud*	**bol*	*P: bʊrɪ*
'to break'	*-ɔm*	**ham*	**fɔm*	**fom*
'two'	*-sumb*	*sɪbɪ*	**łubə*	*təb*
'oil'	*-kidiŋ*	*ngɪlɪ*	**kir*	*-kər*
'leech'	*-tot*	*ntɔ́ntɔ́*		*ntuntu*
'to play'	*-rokom*		**lok*	*lok*
'to lean'	*-sink(ɛ)*		**tinko*	*cik*
'ant sp.'	*-taβɛʧ*		**taf-a*	*tapeʧ*
'to go down'	*wal*	**wal-o*		**wal-an*
'to choose'	*gì-ñjɔ̀gí*	*Bayot: -zoku*		*jok*
'to carry on ahead'	*yɪnd.ɛ*	*JF: -jont*		**yunʧ*

Another important innovation in Bak concerns personal markers. The *-ŋj-/-ɲ-* forms for 1sg. are limited to the Bak languages. Table 2.10 shows all the forms of emphatic personal pronouns in the Bak languages.

Thus, the genealogical unity of the Bak languages (which now include Bijogo), although resting on a small number of features, is posited here. Four distinct groups can be easily recognized: Joola, Manjaku, Balant, and Bijogo. From the distribution of *bVk-* forms (see Table 2.7 above), we can posit that Bijogo was the first to emerge from Proto-Bak. This is consistent with the fact that Bijogo is lexically marginal, but also geographically isolated.

In the two next sections we will examine internal subgroupings inside the two main branches of Atlantic—respectively Bak and North.

2.5 Bak languages: subgroupings

2.5.1 Primary groupings

2.5.1.1 *The Joola cluster*

The Joola languages form the largest cluster in the Atlantic family. The exact number of languages depends on how one conceives the language/dialect distinction. Segerer (p.c.) has elaborated a list of more than 1,200 cognates and the sound correspondences allow to identify the following main languages: Fogny, Gusilay, Banjal, Kasa, Keerak, Kwaatay, Karon, Bayot. There are a few additional lects, for which the documentation is too small to decide about their exact position within the cluster: Bliss and Gulompay in Senegal, Fulup in Guinea-Bissau.

Bayot, Kwaatay, and Karon are the most divergent lects (Bayot showing as few as 12% of similarity in the Swadesh-100 list with Gusilay; Carlton and Rand 1993: 15), but yet they are closer to other Joola varieties than to any other language. Table 2.11 shows lexical innovations shared by Joola languages.

It is worth noting that while the rates of lexical similarity vary continuously from 12% to 85% (Carlton and Rand 1993), the grammatical features of the Joola languages are generally very stable, as are the basic phonological features. A

Table 2.10 Personal pronouns in the Bak branch

INDEPENDENT ~ EMPHATIC							
	1sg	2sg	3sg	1pl incl	1pl excl	2pl	3pl
JOOLA							
Fogny	***inje***	*aw*	*akila*	*walal*	*ʊlɪ*	*mʊyʊl*	*bʊkɔ*
Banjal	***inje***	*au*	*acila*	*wɔlal*	*woli*	*buru*	*bʊgɔ*
Gusilay	***inje***	*au*	*acila*	*wɔlal*	*woli*	*buru*	*bʊkɔ*
Kasa	***inje***	*au*	*akila ~ yoo*	*ɔlal*	*oli*	*bukurul ~ miul*	*bʊkɔ*
Keeraak	***inje***	*aw*	*ɔɔ*	*ʊlaal*	*ʊlɪ*	*mʊyʊl*	*bʊkɔ*
Kwaatay	***inje***	*awe*	*akina ~ oo*	*une*	*ʊnɪ*	*aruun*	*bɔkɔ ~ bɔkɪn*
Karon	***iɲci***	*awe*	*a-kina*	*ʊna*	*ʊnɪ*	*alu*	*pak-o*
Bayot	***inje***	*ab*	*o*	*(y)i*	*wa*	*ka*	*bʊkɔ ~ bʊgɔ*
MANJAKU							
Bok	***ɪnji***	*wi*	*(n)ul*	*ɪnja*	*wənd*	*ɪnd*	*bukul*
Bassarel	***ɪnji***	*wi*	*nul*	*ɪnja*	*uənd*	*ɪnd*	*bukul*
Mankanya	***nji***	*iui*	*ul ~ uŋ*	*nja*		*an*	*bukal*
Pepel	***nji***	*wuy*	*ɔlɔ*	*ndɔ*	*nda*	*nde*	*bɔkɔ*
BALANT							
Ganja	***ɲí***	*hú*	*hí*	*bân*	*báa*	*bǎa*	*bá*
Kentohe	***ɲi***	*hu*	*hɔn*	*bah*		*bàa*	*baan*
BIJOGO							
Bijogo	***ɛŋɔ***	*amɔ*	*ɔg*	*atɛ*		*anɛ*	*yag*

part of the lexical divergence may have arisen from cultural practices very briefly described by Tendeng (2007: 6):

> '(. . .) chez les jóola, avec la mort, l'individu rentre dans l'oubli total. On ne dit plus son nom et tous ceux qui portent le même prénom que lui doivent automatiquement changer de prénom. Peu à peu, tout ce qui a fait sa personne rentre également dans l'oubli total.'
>
> [. . . among the Joolas, with death, a person falls into complete oblivion. One cannot say his name, and all those who wear the same name have to change it. Gradually, everything that has built his personality also falls into complete oblivion.]

This kind of taboo may have had consequences for the lexical inventory. Since '[m]ost traditional Eegimaa names are meaningful' (Sagna and Bassène 2016: 49), avoiding a meaningful name might lead to avoiding lexical roots which are part of it. This practice is rapidly disappearing today, and although we cannot evaluate its importance in the past, the hypothesis of its effects on the core lexicon over a long period of time is at least plausible.

About Bayot, the very low rates of similarity with other Joola lects, along with a few phonological and grammatical peculiarities have led Segerer (2016) to postulate a pre-Atlantic substrate largely replaced by Joola material after centuries of permanent contact. If we set apart this question, which requires much more work, it is safe to say that the Joola cluster is uncontroversial. The internal classification of Joola is obscured by the fact that these languages have never ceased to be in contact, and therefore may have exchanged linguistic material at the same time that they

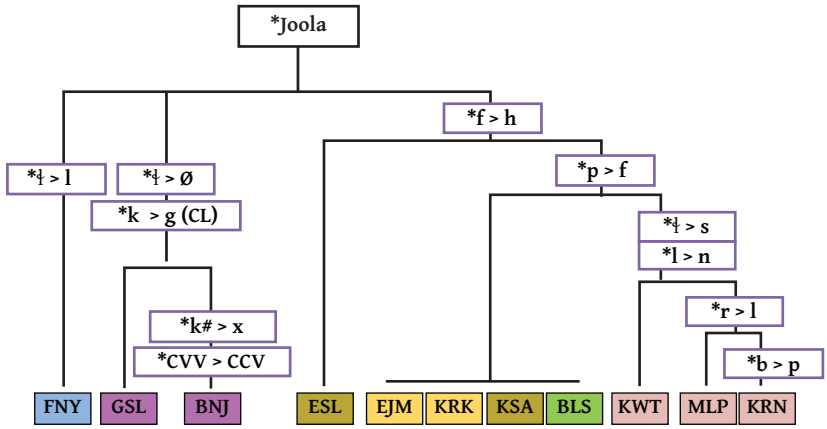

Figure 2.3 Phonological innovations in the Joola cluster

Key: FNY: Fogny, GSL: Gusilay, BNJ: Banjal, ESL: Esuulaalu, EJM: Ejamat, KRK: Keeraak, KSA: Kasa, BLS: Bliss, KWT: Kwaatay, MLP: Mlomp, KRN: Karon.

Table 2.11 Joola lexical innovations

	'to tie'	'to hold'	'salt'	'embers'	'ground'	'to float'
Fogny	*kɔk*		*mʊ-sɪs*	*bʊ-jɛk*	*ɛ-ntaam*	*-rɛnd*
Gusilay	*kɔk*	*jɔk*		*ba-jɛk*	*ɛ-taam*	
Banjal	*xɔx*	*jɔx*	*mɪ-sɪs*	*bɪ-jɛx*	*ɛ-ttam*	*-rend*
Kasa	*kok*	*jok*	*mu-sis*	*hu-jɛk*	*ɛ-taama*	*-rend*
Keerak		*jɔk*	*mʊ-sɪs-am*	*ba-jɛk*	*ɛ-taam-ay*	
Kwaatay		*jɔk*	*mɪ-sɪs*	*hi-jeek*	*ɛ-taam*	*-rand*
Karon		*cok*		*hi-ceek*	*ɛ-taam*	
Bayot		*?*	*múzíó*		*e-tam*	

underwent particular patterns of evolution. Thus, sound correspondences may appear irregular due to the fact that borrowed forms are very similar to inherited ones.

Today, the Fogny variety of Joola (see Chapter 9, this volume), spoken north of the Casamance river, is widely considered as a standard since it has the largest number of speakers. It is, or has been, taught in schools even in non-Fogny areas, and it is at least passively known by all the people who claim to be Joola. Kasa is the largest variety south of the river. Other varieties are often spoken in one village only (Keeraak in Kabrousse (see Chapter 11, this volume), Kujireray in Brin (see Chapter 10, this volume), Kwaatay in Diembering), or in very small areas, like Banjal, Karon, or Gusilay.

Figure 2.3 illustrates how regular sound shifts may be used to obtain an internal classification of the Joola cluster.

2.5.1.2 *The Manjaku cluster: Manjaku–Mankanya–Pepel*

The situation of the Manjaku cluster is not very different from that of Joola, except that there are three main divisions, namely Manjaku proper, Mankanya, and Pepel, each of these being in turn divided into local varieties. The most complex division is apparently Manjaku, in which the Bok variety is considered the standard. In the introduction to his 'Lexique manjaku', where a Proto-Manjaku lexicon is given, J. L. Doneux (1975b: 2) proposes the internal structure for the cluster shown in Figure 2.4.

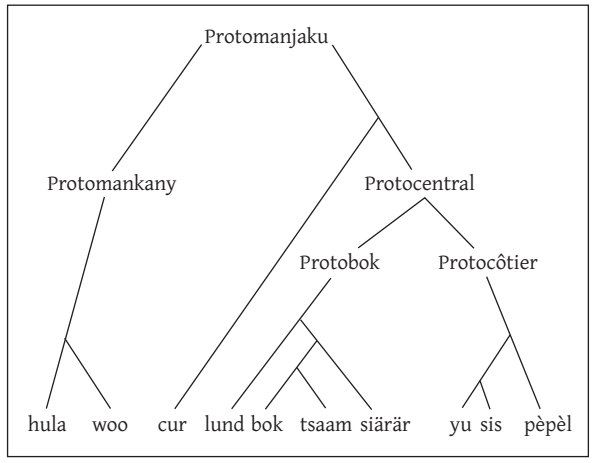

Figure 2.4 Internal structure of the Manjaku cluster

Table 2.12 Manjaku lexical innovations

	'baobab'	'to sleep'	'black'	'wing'	'to spoil'	'embers'	'ear'
Bok	*-dʊngʌl*	*-ŋɔy*	*-jin*	*-bakəla*	*-cɔk*	*-jʊʌs*	*-bats*
Bassarel	*-rungal*	*-ŋɔy*	*-jin*	*-bakəla*	*-cɔk*	*-jʊʌs*	*-baaʈ*
Ulund	*-dungal*	*-ŋɔay*	*-jin*	*-baklʊ*	*?*	*-joasal*	*-baʈ*
Cur	*?*	*-ŋɔy*	*-ʐə̃*	*-bakəla*	*-cɔk*	*-ʐous*	*-bah*
Mankanya	*duŋal*	*-ŋɔy*	*-jən*	*-baaklu*	*-θɔk*	*-joʂ*	*-baʈ*
Pepel	*rungal*	*-ŋʊy*	*-jin*	*-akʊlʊ*	*-sɔk*	*-jɔl*	*-waʂ*
***Manjaku**	—	***-ŋɔy**	***-jin**	***-bakəl**	***-cɔk**	***-joos**	***-bats**

Table 2.13 Balant lexical innovations

	'broom'	'baobab'	'tomorrow'	'hand'	'to buy'	'warm'	'dirty'
Ganja	*tààsí*	*f-lààθé*	*b-lùθí*	*f-jíf*	*-wus-*	*-ɲú*	*lɔ́rɛ̀*
Fraasɛ	*taasn*		*kə-luθn*	*f-cef*	*-wus-*	*-ɲuh*	*lər*
Kəntohe	*f-taastɪ*	*f-laaθe*	*b-lʊθ*	*f-cɪf*	*-wʊs-*	*-ɲʊh-*	*lora*

There are many lexical innovations for the Manjaku cluster. Table 2.12 gives a few examples with their corresponding proto-forms according to Doneux (*1975b.*).[4]

2.5.1.3 *The Balant cluster*

Balant is often considered a single language in the literature, but it consists of at least two distinct languages, Ganja in Northern Guinea-Bissau and Southern Senegal and Kəntohe in Central Guinea-Bissau. In addition, there are some local varieties about which very little is known: Fora, Mane ~ Manoj, Soofa, Naga, Fraasɛ, and maybe more. The Ganja variety is the best documented one, with one grammar sketch including a short lexicon (Ndiaye-Corréard 1970) and a more comprehensive grammatical description accompanied by a larger lexicon and texts (Biaye and Creissels 2015), see also (Chapter 15, this volume). The differences between the two main varieties are much more grammatical than lexical. Table 2.13 gives a few examples of lexical innovations.[5]

Balant languages are undoubtedly tonal. However, in the available literature, tones are marked for the Ganja lect only, and they are analysed differently by both Ndiaye-Corréard and Biaye and Creissels. In the Fraasɛ and Kəntohe data tones are left unmarked, although Doneux (1984a: 23) posits their existence. The question of whether tones emerged in Balant only or disappeared in all three other Bak branches is still problematic. Wilson, in his unpublished outline of Bijogo (where he deals with the Kajooko lect, see below) identifies two tones. However, Segerer (2002) sees no trace of lexical tone but admits that prosody plays an important role in grammar. Tone is synchronically absent from the Joola-Manjaku group.

2.5.1.4 *The Bijogo cluster*

The Bijagos archipelago, situated about 40 miles off the coast of Guinea-Bissau, is home to the Bijogo language cluster. Bijogo is still considered a single language in the major sources for classification, including *Ethnologue* (Lewis et al. 2015). Yet, phonological, grammatical, and lexical peculiarities allow us to identify at least four main varieties, corresponding to the major populated islands: Kagbaaga (island of Bubaque), Kanyaki (island of Canhabaque), Kamona (islands of Caravela and Caraxe), Kajooko (islands of Uno and

[4] Sources: Bok, Proto-Manjaku: Doneux 1975a, b; Bassarel: Buis 1990; Ulund: Segerer (personal fieldnotes); Cur: Wilson p.c.; Mankanya: Gaved and Stammers 2004; Pepel: Ndao 2011.

[5] Sources: Ganja: Biaye and Creissels 2015; Fraasɛ: Wilson p.c.; Kəntohe: Doneux 1984a.

Table 2.14 Bijogo lexical innovations

	'animal'	'three'[a]	'egg'	'to buy'	'rope'	'leg, foot'	'nose'
Kagbaaga	*ŋɔ-kɔma*	*ɲ-ɲɔɔkɔ*	*ne-keke*	*ŋɔ-rɛs*	*nɛmɛda*	*e-be*	*ŋɔ-mɔ*
Kamona	*ŋo-koma*	*ŋuɲ-ɲɔkɔ*	*nɛ-kɛkɛ*	*ŋɔ-rɛc*	*nɛmɛnda*	*e-be*	*ŋɔ-mɔ*
Kanyaki	*ŋ-koma*	*ɲ-ɲɔɔkɔ*	*nɛ-kɛkɛ*	*ŋo-leʃ*	*nemeƅa*[b]	*e-be*	*ŋɔ-mɔ*
Kajooko	*ŋo-kooma*	*-ɲɔɔkɔ*	*ne-keeke*		*nemeƅa*	*e-bə*	*ŋɔ-mɔ*

[a] The word for 'three' in Bijogo literally means 'fingers', which may seem bizarre. But this makes sense if one considers that Bijogo kings never go out without their sceptre, which consists of a wooden hand with. . . three fingers.
[b] The symbol; 'ƅ' stands for a linguolabial voiced stop, a consonant that is found in only a handful of the world's languages. This consonant is attested in Kanyaki and Kajooko, but is absent from Kagbaaga and Kamona.

Table 2.15 Bijogo internal lexical variation

	'smoke'	'man'	'dream'	'seat'	'to sing'	'bee'
Kagbaaga	*maaʈukpa*	*o-gude*	*ŋo-soʈe*	*sammaj*	*ŋɔ-raɛ*	*nɛ-mɛ*
Kamona	*mumbənk*	*ɔ-mpɛndɛ*	*mɔ-ədə*	*na-paʈ*	*ŋo-rom*	*na-tamiɔ*
Kanyaki	*maʈiko*	*o-wanto*	*ŋo-ʃoʈoko*	*ŋo-kota*	*ŋu-lin*	*e-to*
Kajooko	*mamƅu*	*ka-suka*	*ŋɔ-ƅɔnatɛ*	*ŋo-ɲoata*		*nɛ-mɛ*

Orango). There might also be local varieties on the islands of Formosa and Galinhas.

Since there is only one published source for grammatical description and lexicon (Segerer 2002), it is still premature to estimate the exact status of the Bijogo lects. Let us just say that the Kamona variety is not understood by speakers of other varieties.

Although Greenberg had placed Bijogo in his Northern Atlantic branch (Greenberg 1963: 8), both Sapir (1971) and Wilson (1989) made it an isolate within Atlantic. In the section devoted to the installation of the Bak group, we presented evidence for the inclusion of Bijogo, one of the main innovations proposed in the present chapter.

With lexical percentages being not higher than 10% with any other Atlantic branch, it is not hard to find Bijogo lexical innovations (Table 2.14).[6]

But Bijogo can also show great internal variation (Table 2.15).

[6] Sources: Segerer 2002 (Kagbaaga); Segerer 1998b, unpublished (Kamona); Scantamburlo p. c. (Kanyaki); Wilson p. c. (Kajooko).

2.5.2 Secondary groupings

2.5.2.1 *The Joola–Manjaku node*

The Joola and Manjaku groups show the most common features in the distribution of their *bVK-* forms (Table 2.7) as well as in their personal pronouns paradigms (Table 2.10). In addition, these are the two Bak groups that apparently display the most lexical exclusive cognates, i.e. probable innovations, although those do not number much (a few dozens). Table 2.16 presents some innovations shared by the two languages.

2.5.2.2 *Higher Bak nodes*

Once Joola–Manjaku has been installed as a valid subgrouping, and since we have shown that Bijogo is the result of the first Bak split, we are only left with the Balant group, which then fits in between the Bijogo split and the Joola–Manjaku split. Thus, we get the picture in Figure 2.5 for Bak (excerpt from Figure 2.2).

Table 2.16 Joola–Manjaku lexical innovations

	'bee'	'mat'	'to dress'	'slippery'	'hand'	'fist'	'woman'
JOOLA							
Fogny	*f-aaj*	*ka-pɛnk*		*-lʊʊlɪ*	*ka-ɲɛn*	*e-tuk*	*anaare*
Gusilay	*y-aj*	*ga-pɛɛk*		*-luuli*	*ga-ɲɛn*		
Banjal	*y-aaj*	*ga-ppɛx*	*ɛ-ssɪmɔ*		*ga-ɲɛn*		*aar(e)*
Kasa	*h-aaj*	*-fenk*			*ka-ŋɛn*	*ka-tʊk*	*anaara*
Keerak	*ɛ-aj-ay*			*-lʊʊlɪ*	*ka-ŋɛn-ak*		*annaaraw*
Kwaatay	*ɛyaj*	*nɪ-fɛnk*	*bɪsɪumɔ*			*e-tuka*	
Karon							*aal*
Bayot			*bɔsɪemʊ*				
MANJAKU							'wife'
Manjaku Bok	*uwaaj*	*ka-mpenk*	*tiəman*	*-lʊəl*	*-ɲʌn*	*ʊ-sʊk*	*ar*
Manjaku Cur	*u-haaz̨*	*ka-mpɛnk*		*-lul*	*ka-ŋeẽ*		*ahaar*
Pepel					*-ɲenɛ*	*ɔ-lʊkʊ*	*aarʊ*
Mankanya				*-luul*	*ka-ɲen*	*uʂuuki*	*ahar*

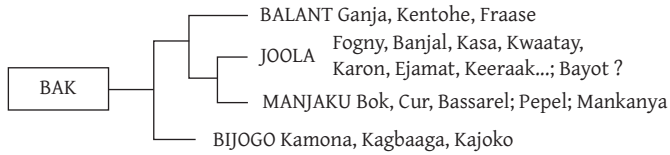

Figure 2.5 Internal structure of the Bak branch

2.6 North Atlantic languages: subgroupings

2.6.1 Primary groupings

2.6.1.1 *The Tenda cluster*

The uncontroversial cluster called Tenda groups together five languages, one of which—Bapen—is most probably extinct since the end of the last century. Three languages—Bedik (*aka* Mənik), Basari (*aka* Oniyan), and Konyagi (*aka* Mey or Wamey)—are spoken in southeastern Senegal and across the Guinean border (Konyagi). Tanda is spoken in and around the village of Cacine in southwestern Guinea-Bissau.

For the three Senegalese languages we have very complete lexical data at our disposal (Ferry 1991; Santos 1996), enough to undertake phonological and lexical reconstruction (Pozdniakov 2022). The noun class system of Proto-Tenda has already been published (Ferry and Pozdniakov 2001). The internal structure of the Tenda cluster puts Konyagi aside the others, Bapen and Tanda being grouped with Bedik and Basari respectively (see Chapter 8, this volume for a grammatical overview of Basari). Figure 2.6 shows the internal structure of the Tenda group.

The Tanda language (not to be mistaken for Tenda as a group) is an interesting case of recent migration: the language is linguistically very close to Basari, but is located some 350 km away from its supposed homeland.

The items in Table 2.17 are Tenda shared lexical innovations. They also exemplify regular sound correspondences.[7]

[7] Sources: Basari, Bedik: Ferry 1991; Konyagi: Sachot (Santos) 1996.

Table 2.17 Tenda lexical innovations

	'green, ripe'	'belly'	'to winnow'	'fist'	'to work'	'fan palm'
Basari	*xàf /k-/ng-*	*càl*	*a-fèl*	*ɛ-màl*	*a-ndĭyèn*	*a-péỹ*
Bedik	*gāf /ng-*	*ga-céɗ*	*o-fɤ́ɗ*	*u-mɤ́ɗ*	*gi-ndìyél*	*ga-pèŋ*
Konyagi	*xáf/k-/k-*	*æ̀-cèry*	*ì-pæ̀ry*	*ì-mùryá*	*ì-díyèl̰ɔ́*	*ì-pêŋ*

Table 2.18 Jaad–Biafada lexical innovations

	'three'	'goat'	'palm-tree'	'water'	'to eat'	'to begin'
Jaad	*-caw*	*wan-daafe*	*bə-nic*	*ma-mbe*	*jaar*	*fet*
Biafada	*-njo /j/j*	*w-ndaf*	*bu-ninci*	*ma-mbiya*	*jal*	*fər*

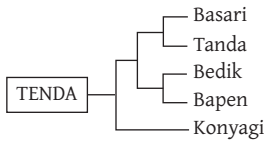

Figure 2.6 Internal structure of the Tenda group

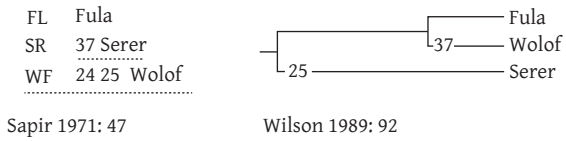

Figure 2.7 Wilson's mismatch in citing Sapir

2.6.1.2 *The Jaad–Biafada cluster*

This cluster groups only two languages: Jaad (*aka* Pajade, Badyara, or Badyaranke) and Biafada (*aka* Biafar or Bidyola). No information is available about possible dialectal variation.

The clustering of these two languages is quite obvious if the lexicon only is considered. More puzzling is the fact that these languages are spoken very far from each other, which makes it very unlikely that the lexical resemblances are due to borrowing (although some of them are, but from other languages such as Fula or Mandinka). While Jaad is spoken on both parts of the Senegal–Guinea border, Biafada is located near the coast of Guinea-Bissau in front of the Bijagos archipelago.

From a grammatical point of view, there are very important differences between these languages: Biafada has a complex noun class system and a very developed system of consonant mutation whereas Jaad is in the process of losing entirely its noun classes and keeps only traces of its former consonant mutation (Wilson 1965).

Some lexical innovations are given in Table 2.18.[8]

[8] Sources: Jaad: Meyer 2001a; Biafada: Wilson (p.c.)

2.6.1.3 *The Fula–Sereer cluster*

First, we wish to insist here that comparative data do not allow us to include Wolof in this cluster, despite repeated allusions to a Senegambian or Senegal language group in the literature (Sapir 1971; Wilson 1989). It must be emphasized that there is no solid argument to attach Wolof to the Fula–Sereer branch. Numerous 'cognates' between Wolof and Sereer or between Wolof and Fula or even between Wolof and Fula–Sereer represent borrowings that are generally easy to detect. Moreover, when Wilson (1989) included Sapir's (1971) chart of Atlantic languages with their percentages of lexical similarity, he made a mistake in that he switched Wolof and Sereer, thus putting Wolof closer to Fula, as can be seen in the excerpts in Figure 2.7.

This mistake has been repeated as such in many works where a brief presentation of Atlantic was needed.

Wolof and Sereer have, for a long time, been spoken in adjacent areas, and it is no surprise that there are a fair number of apparent Wolof–Sereer exclusive lexical isoglosses. However in many cases it is possible to spot borrowings. In Table 2.19 we list a few examples of probable borrowings from Wolof into Sereer. For these notions, Sereer has its own lexemes in addition to the Wolof ones. The reverse is not

Table 2.19 Probable Wolof borrowings in Sereer

	Wolof	Sereer < Wolof	Sereer
'lazy'	*tayal*	*tail*	*poos*
'vein'	*siddiit*	*sidit*	*pamb*
'last'	*mujj*	*muj*	*fakit, falak, falakit*
'friend'	*xarit*	*xarit*	*sil*
'armpit'	*poqtaan*	*poqotaan*	*naaplan, naapand. . .*
'to covet'	*xemmem*	*xemem*	*bug, yaaw, yoxum*
'fat'	*duuf*	*duuf*	*ɓuɓel, dakel, neeɓ*
'to be cross-eyed'	*jell*	*jel, jol*	*taal, xaɗ*

true; our Wolof data show no other lexemes than those for these notions.

It should also be noted that among all our apparently exclusive Wolof–Sereer isoglosses, there is no notion from the basic vocabulary (i.e. the Swadesh 100-list). In comparison, Fula–Sereer exclusive isoglosses contain seven items from the basic vocabulary (Table 2.20).

It is quite enough to posit a Fula–Sereer group. But there are also striking Fula–Sereer parallels in grammar. Let us mention only one emblematic grammatical/morphological isogloss, namely the reorganization of the noun class system in Proto-Fula–Sereer. These are the only two languages in the Atlantic family that have 'augments' in the Bantu way, that is, a preprefixing vowel on noun prefixes in derivated noun class markers (*di- > e-de > el, *du- > o-du > ol). Details are given in Pozdniakov (1988, 2015). Doneux (1975a), who first presented the augment hypothesis, suggested that it could be traced not only in Fula and Sereer but also in Tenda, Wolof, and even Joola. As for Tenda, Ferry and Pozdniakov (2001) show that the noun class system can be accounted for without any augment. Doneux's arguments for the augment in Wolof and Joola are not convincing: they are only based on the fact that both languages have a definite concord marker with an initial *a-*, whereas the Bantu-type augment should be a vowel of the same quality as that of the noun class marker. See Chapter 5, this volume for an overview of Fula.

2.6.1.4 *The Cangin cluster*

The Cangin cluster consists of five languages spoken around the town of Thies in Senegal, not very far from Dakar. They are surrounded by Wolof and, to a lesser extent, Sereer. Therefore we can expect to find many common words in Wolof (resp. Sereer) and Cangin.

The Cangin languages were long regarded as dialects of Sereer until Pichl (1966) showed that they form a distinct linguistic group. He called this group Cangin, after one of the local names of Thies. The five languages are Palor (~ Falor), Ndut, Saafi (~ Safen), Noon, and Laala (~ Lehar).

Linguistically, this cluster is quite compact and homogeneous, but some of the sound correspondences are not trivial (for more detail, see Drolc 2006). It is divided into two

Table 2.20 Fula–Sereer lexical innovation within basic lexicon

	'bone'	'come'	'drink'	'knee'	'kill'	'white'	'path'
Proto-Fula–Sereer	**giɣ/ʔ*	**gar*	**yɛr*	**kʋɓ/f*	**war*	**d/tan*	**ɗat*
Fula	*giʔ-*	*war/g-*	*yar-*	*hof-/k-*	*war*	*ran/d-*	*ɗat-*
Sereer	*hiiɣ/k-*	*gar*	*yer*	*nguɓ/k*	*war/b-*	*ran/t-*	*ɗat-*

Table 2.21 The first Cangin split

	'to die'	'mouth'	'nail'	'elephant'	'to do'	'axe'	'tomorrow'
Ndut	*hʊl*	*ɓʊk*	*cigilan*	*caʔ*	*paŋ*	*yun*	*kɛy*
Palor	*xul*	*ɓuk*	*cɪgɪlaan*	*caʔ*	*paŋ*	*yiin*	*key*
Saafi	*kaan*	*nguɓ*	*cakəna*	*cɔx*	*tum*	*kiw*	*kim*
Noon	*kan*	*ku*	*cogəna*	*cox*	*tum*		
Laala	*kaan*	*kuu*	*cɔgɔna*	*cɔh*	*ka-tum*	*kɛw*	*kim*

main branches: Ndut-Palor and Saafi-Noon-Laala. Table 2.21 gives some isoglosses that distinguish these branches.[9]

Table 2.22 gives some lexical innovations for the entire group. They also illustrate regular sound correspondences.

Table 2.22 Cangin lexical innovations

	'to drink'	'black'	'water'	'name'	'arm, hand'	'blood'
Ndut	*xan*	*suul*	*mulub*	*tii*	*yaʔ*	*ɲɪf*
Palor	*han*	*sʊʊlə*	*mʊlʊb*	*tii*	*yaʔ*	*ɲɪf*
Saafi	*ʔan*	*susus*	*masuɓ*	*tiik*	*yax*	*ɲif*
Noon	*ʔan*	*sujus*	*mᵘoju?*	*tek*	*jah*	*ɲif*
Laala	*an*	*susus*	*məsu*	*tɛɛk*	*yah*	*ɲif*

2.6.1.5 *The Nyun–Buy cluster*

This cluster is composed of two well-separated subgroups—Nyun and Buy. The Nyun languages are often called Baynouk (or Baynunk, Banhum, Banyun, and similar). The Buy subgroup comprises two nearly extinct languages spoken in Northern Guinea-Bissau: Kobiana (or Buy, Kabuy, Guboy) and Kasanga (or Guhaca). For an up-to-date list of Nyun languages with existing sources and subclassification, see Cobbinah (2013a).

The last decades have seen many publications on various Nyun–Buy languages: Cobbinah (2013a, Chapter 12, this volume) and Biagui (2012a) on Baynunk Gubëher; Bao Diop (2013) on Baynunk Gunyaamolo; Lüpke (Chapter 14, this volume) on Baynunk Gujaher; Quint (2015, Chapter 13, this volume) on Djifanghor Nyun; Doneux (1991) and Voisin (2015b) on Kobiana.

There are not many exclusive lexical innovations that characterize the Nyun cluster, partly because of the paucity of Buy data. Some are given in Table 2.23.[10]

As far as grammar is concerned, the Nyun–Buy languages share a unique noun class system, with the following two features:

- The classical singular/plural opposition is complexified in that many nouns may have two different plural forms (i.e. classes): a countable plural and a collective plural. This feature is also attested, although more sporadically, in at least some Joola languages, where it might have arisen through ancient contact.

  Gubaher (Cobbinah 2013a):
  gu-lihan 'stick' (singular)
  ha-lihan 'sticks' (count plural)
  ja-lihan 'wood (substance)/sticks' (unlimited plural)

- While plural forms are typically obtained by a change in the prefix, in many cases a suffix is used, be it with or without prefix substitution.

  Gubaher (Cobbinah 2013a):
  u-diin 'friend'/*in-diin-eŋ* 'friends'
  ba-xon 'African fan palm'/*ba-xon-oŋ* 'African fan palms'

These two features may combine in various ways. For instance, the suffixed plural may be used as collective (*u-dikaam* 'woman'/*in-dikaam* 'women'/*in-dikaam-aŋ* 'women' (collective)). Both features are not unheard of in Atlantic, but it is only in this cluster that their use is so frequent.

[9] Sources: Ndut: Doneux (ms); Palor: D'Alton (1987); Saafi: M'Bodj (1983); Noon: Williams and Williams (1993); Soukka (2000); Laala: Pichl (1981); Deye (2010).

[10] Sources: Gubaher: Cobbinah 2013a; Gunyaamolo: Bao Diop 2013; Sonkodou: Lespinay 1992; Kobiana, Kasanga: Wilson (ms).

Table 2.23 Nyun–Buy lexical innovations

	'woman'	'man'	'fire'	'hair'	'to laugh'	'rope'	'dry season'
Gubaher	*ʊ-dikaam*	*ʊ-digeen*	*kʊʊl*	*gʊjənd*	*ɲanciil*	*sincind*	*buun*
Gunyaamolo	*-dɪkaam*	*-diigeen*	*kʊʊl*	*gu-gend*	*bʊ-kʊl*	*sɪn-kɪnd*	*buun*
Sonkodou	*udikam*	*udigen*	*kulo*	*gugeno*	*ñankilo*	*sinkindo*	*buhuno*
Kobiana	*u-likaam*	*u-ligen*	*kooh*	*jɛgɛn*	*keeh*	*si-kind*	*gamboon*
Kasanga	*u-likaam*	*u-lien*	*hoor*	*jɛgɛn*	*keer*	*si-kind*	*gomboon*

The position of Nyun–Buy within Atlantic is problematic. These languages have been in close contact with Manjaku and Joola for centuries—Nyun being closer to Joola and Buy to Manjaku. These contacts have led to massive borrowing, probably in both directions. This question will be addressed in Section 2.7 below.

2.6.1.6 *Wolof*

Wolof (including the Lebu dialect) is the only language in North Atlantic that cannot be grouped with other languages. The attempts to link it to Fula and Sereer are mentioned in Section 2.6.1.3 above. Doneux (1978) suggested that Wolof might be closer to Nyun than to any other Atlantic cluster, on the basis of various arguments including the following features of their noun class systems (Doneux 1978: 43–6):

- The singular class prefixes are very similar in Wolof and Nyun.
- Both languages have *ɲ-* and *y-* as their regular plural class exponents.
- Both languages lack a true 'human' class.
- The use of a *j-* class as a collective is regular in both languages.

Among the main innovations that may characterize Wolof is the loss of productive noun class prefixation. In fact, initial consonants of nouns in Wolof may often be shown to be traces of older prefixes, but today Wolof needs not (and is even difficult to) be described as having noun class prefixes. See Chapter 4, this volume, for an overview of Wolof.

2.6.1.7 *The Nalu cluster: Nalu–Baga Fore–Pukur)*

These three languages, about which very little is known,[11] are spoken close to each other on both sides of the western part of the border between Guinea and Guinea-Bissau. The number of speakers may be estimated to a few thousands for Nalu (see Chapter 17, this volume), a few hundred for Baga Fore, and is unknown for Pukur. These languages have been grouped together by both Sapir and Wilson, and although the arguments for this grouping are not explicit, it appears as the most reasonable position. Below we will look at the available evidence both for the internal unity of the group and for its inclusion in the North branch of Atlantic.

Among the features that may justify this grouping, the particular morphology of noun classes is promising, even though we don't have that much data at our disposal. A recent fieldtrip on Baga Fore by E. Golovko has brought more precisions about its noun class system which, with its mandatory confixes, is not typical of Atlantic (Creissels 2015a: 21).

In Golovko's data, more than 50% of nouns show a mandatory plural suffix **-lVŋ* or **-il*, and some of them even require a combination of these class suffixes. Finally, for a few nouns denoting human beings (or more generally animate beings), the plural suffix *-y* is mandatory:

Ø-dok/pl. n-dok-il 'duck', *ɨ-tim/pl. ɛ-tim-il* 'old person', *n-puŋ/pl. o-puŋ-ul* 'fish', *Ø-bac/pl. ci-mbac-il* 'stick';
i-tiya /pl. ɛ-tiya-laŋ 'crocodile', *ki-len/ɛ-len-diŋ* (< **ɛ-len-liŋ*) 'horn', *Ø-fako/ci-fako-loŋ* 'knife';
Ø-su/pl. n-su-luŋ-ul 'horse', *Ø-kabi/pl. n-kabi-liŋ-il* 'blacksmith;
i-landuma/pl. a-landuma-y 'Landuma', *Ø-sofale/pl. n-sofale-y* 'donkey'.

This feature is apparently shared by Pukur and Nalu. In Pukur, the rare nouns showing a suffix in the plural don't

[11] A recent ELDP grant allowed F. Seidel to gather important documentation on Nalu, of which we could access only the lexicon (~3,000 words). Baga Fore is being studied by E. Golovko but this work is just beginning at the moment this chapter is written. For Pukur, we only have a short wordlist collected by M. P. Ferry some twenty years ago. Swadesh-like wordlists are given for these three languages in Wilson (2007).

seem to belong to a single semantic category, but one of the suffixes (*-VlVŋ*) looks like a compound of suffixes that resemble those of Baga Fore:

céx/céx-àlòŋ 'neck'
nʌ́k/nʌ́g-àlʊŋ 'medicine'
ndʌ̀mb/ndʌ̀mb-àlóŋ 'basket'
a-vàr/a-vàr-ʌl 'dog'
nsèl̀/nsèl̀-ʊŋʊn 'canoe'

In Nalu, a part of the nouns (mostly animates) show a plural suffix *-ɛ, either in addition to the prefix change or as the only plural marker:

m-kemet/bɛŋ-kemt-ɛ 'spitting cobra'
mmaafeek/baafeekɛ 'crocodile'
kɔŋ/kɔŋ-ɛ 'fruit bat'

Some parallels may be found in the Nyun–Buy cluster (North branch), but not in the Bak branch.

The other reason for grouping together these three languages is of course lexical cognacy. In a paper by Fields (2004), an attempt was made to select common isoglosses, but this list included retentions which may be found in other Atlantic languages. Let us have a look at those that seem innovative (Table 2.24).

Table 2.24 Some Nalu–Baga Fore–Pukur probable cognates

Gloss	Nalu	Baga Fore	Pukur
'mouth'	*-sool*	*túl/cì-ntúl*	*sùl*
'fire'	*met*	*m-bùt/o-but-uluŋ*	*múr*
'leg'		*ì-sí/è-sí*	*i-sù/o-sù*
'hand'	*tɛ*	*itɛ*	
'moon'	*-bilaŋ*	*ìbìl/ɛbel-liŋ*	*pʌ́l*
'to hunt'	*mil*	*mel*	
'inheritance'	*-kɛɛ*	*ké/cì-ké*	
'mangrove'	*-yɔɔf*	*kí-yòp/yòp, yop-il*	
'fly (n.)'		*mpàl/à-pál-àŋ*	*mfʌ́lã*
'ashes'	*ferfer*	*ki-pocpoci/pocpoci*	*fʌ̃s*
'witch'	*-bes*	*bíc/è-bíc-íllìŋ*	
'elephant'	*-rɛɛp*		*lʌ̀p'*

In addition, a number of cognates (including Atlantic or even NC retentions) show regular sound correspondences. Here is an illustration with the f/p/f correspondence (Table 2.25).

Table 2.25 Regular sound correspondence within the Nalu group

	Nalu: f	Baga Fore: p	Pukur: f
'ashes'	*ferfer*	*ki-pocpoci*	*fʌ̃s*
'goat'	*miifi*	*ɛpi/ɛpi-lliŋ*	*í-fí*
'strength'	*fanka*	*panka*	*fángà*
'dead'	*ref*	*nrʌ́p*	*i-félʌ̀f*
'ear'	*(nɛɛw)*	*í-nɔ́p/ɔ́-nɔ́p*	*ɛ-nʌ́f/sʌ̀-nʌ́f*

Thus, despite the paucity of data for Baga Fore and Pukur, it is quite safe to postulate that these three languages are genetically related. However, their position with respect to the language families of the region is not easy to assess. Many of the cognates are retentions, whether Atlantic or Niger-Congo. Here are a few examples of such ancient roots (Table 2.26).

Table 2.26 Shared ancient roots in Atlantic, Mel, and Nalu

	ATL-Bak	ATL-North	Mel	Nalu
'dog'	Manjaku *bus*	Fula *ɓoos-*	Sherbro *gbos*	BF *bùt*
'tongue'	Balant **dɛm-*	Bedik *dém*	**d-dem-el*	BMb *lʊm*
'hill'	Bijogo *dingi*	Saafi *taŋgi*	Sherbro *thɔk*	BF *ronk*
'tree'	Balant *tah*	Basari *tʌ̃x*	Temne *tək*	Nalu *-ti*
'to tell'	Balant *tum*	Nyun *tum-*	Kim *tum*	Nalu *rɔɔm*
'ancestor'	Mankanya *maam*	Wolof *maam*	Kisi *-màmà*	Nalu *maama*
'star'	Joola **-oot*	Fula **hoor-*	Baga Mad *k-ɔɔs*	Nalu *hooθ*

Thus, the Nalu group could be classified either with Atlantic, with Mel, or as an isolate.