

CAMBRIDGE  
APPROACHES  
TO LANGUAGE  
CONTACT

# A Linguistic Geography of Africa

Edited by  
**Bernd Heine and Derek Nurse**



CAMBRIDGE

This page intentionally left blank

# A Linguistic Geography of Africa

More than forty years ago it was demonstrated that the African continent can be divided into four distinct language families. Research on African languages has accordingly been preoccupied with reconstructing and understanding similarities across these families. This has meant that an interest in other kinds of linguistic relationship, such as whether structural similarities and dissimilarities among African languages are the result of contact between these languages, has never been the subject of major research. The aim of this book is to show that such similarities across African languages are more common than is widely believed. It provides a broad perspective on Africa as a linguistic area, as well as an analysis of specific linguistic regions. In order to have a better understanding of African languages, their structures, and their history, more information on these contact-induced relationships is essential in order to understand Africa's linguistic geography, and reconstruct its history and prehistory.

BERND HEINE is Professor Emeritus of African Studies at the University of Köln, Germany.

DEREK NURSE is Henrietta Harvey Research Professor in the Linguistics Department at the Memorial University of Newfoundland, Canada.

## *Cambridge Approaches to Language Contact*

### *General Editor*

SALIKOKO S. MUFWENE

University of Chicago

### *Editorial Board*

Robert Chaudenson, *Université d'Aix-en-Provence*

Braj Kachru, *University of Illinois at Urbana*

Raj Mesthrie, *University of Cape Town*

Lesley Milroy, *University of Michigan*

Shana Poplack, *University of Ottawa*

Michael Silverstein, *University of Chicago*

Cambridge Approaches to Language Contact is an interdisciplinary series bringing together work on language contact from a diverse range of research areas. The series focuses on key topics in the study of contact between languages or dialects, including the development of pidgins and creoles, language evolution and change, world Englishes, code-switching and code-mixing, bilingualism and second language acquisition, borrowing, interference, and convergence phenomena.

### *Published titles*

Salikoko Mufwene, *The Ecology of Language Evolution*

Michael Clyne, *The Dynamics of Language Contact*

Bernd Heine and Tania Kuteva, *Language Contact and Grammatical Change*

Edgar W. Schneider, *Postcolonial English*

Virginia Yip and Stephen Matthews, *The Bilingual Child*

Bernd Heine and Derek Nurse (eds.), *A Linguistic Geography of Africa*

### *Further titles planned for the series*

Guy Bailey and Patricia Cukor-Avila, *The Development of African-American English*

Maarten Mous, *Controlling Language*

Clancy Clements, *The Linguistic Legacy of Spanish and Portuguese*

# A Linguistic Geography of Africa

---

*Edited by*

Bernd Heine and Derek Nurse



**CAMBRIDGE**  
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS

Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press

The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

[www.cambridge.org](http://www.cambridge.org)

Information on this title: [www.cambridge.org/9780521876117](http://www.cambridge.org/9780521876117)

© Cambridge University Press 2008

This publication is in copyright. Subject to statutory exception and to the provision of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published in print format 2007

ISBN-13 978-0-511-36808-0 eBook (Adobe Reader)

ISBN-10 0-511-36808-9 eBook (Adobe Reader)

ISBN-13 978-0-521-87611-7 hardback

ISBN-10 0-521-87611-7 hardback

Cambridge University Press has no responsibility for the persistence or accuracy of urls for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

# Contents

---

<i>List of maps</i>	<i>vii</i>
<i>List of figures</i>	<i>viii</i>
<i>List of tables</i>	<i>ix</i>
<i>List of contributors</i>	<i>xi</i>
<i>Series editor's foreword</i>	<i>xii</i>
<i>Acknowledgments</i>	<i>xiv</i>
<i>Abbreviations</i>	<i>xv</i>
1 Introduction	1
Bernd Heine and Derek Nurse	
2 Is Africa a linguistic area?	15
Bernd Heine and Zelealem Leyew	
3 Africa as a phonological area	36
G. N. Clements and Annie Rialland	
4 Africa as a morphosyntactic area	86
Denis Creissels, Gerrit J. Dimmendaal, Zygmunt Frajzyngier, and Christa König	
5 The Macro-Sudan belt: towards identifying a linguistic area in northern sub-Saharan Africa	151
Tom Güldemann	
6 The Tanzanian Rift Valley area	186
Roland Kießling, Maarten Mous, and Derek Nurse	
7 Ethiopia	228
Joachim Crass and Ronny Meyer	
8 The marked-nominative languages of eastern Africa	251
Christa König	

9	Africa's verb-final languages Gerrit J. Dimmendaal	272
	<i>Notes</i>	309
	<i>References</i>	323
	<i>Index</i>	354

# Maps

---

2.1 A sketch map of northern Nigerian languages: isoglosses of the number of shared typological properties	32
3.1 Six phonological zones in Africa	37
3.2 Northern Bantu languages with labial-velar stops	43
3.3 Distribution of contrastive nasal vowels in a sample of 150 African languages	46
3.4 Distribution of voiced or laryngealized implosives in a sample of 150 African languages	58
3.5 Distribution of emphatic and ejective consonants in a sample of 150 African languages	62
3.6 Distribution of 76 African languages with three or more distinctive tone levels	73
3.7 Distribution of “lax” question prosody markers, which occur in 41 of a sample of 75 African languages for which relevant information was found	79
5.1 Logophoricity across African lineages	156
5.2 Labiovelar consonants across African lineages	158
5.3 ATR vowel harmony across African lineages	160
5.4 Word order S-(AUX)-O-V-X across African lineages	163
5.5 Word order V-O-NEG across African lineages	165
5.6 Labial flaps across African lineages	166
5.7 Minimal-augmented pronoun systems across African lineages	171
6.1 The languages of the Tanzanian Rift Valley area	187
8.1 Case in northern and eastern Africa	266
8.2 Areal distribution of case in southern Ethiopia and adjacent areas	267
9.1 Typological zones (based on Heine 1976)	286
9.2 Languages with converbs in northeastern and north-central Africa	296

## Figures

---

8.1 Definitional characteristics of case systems	254
8.2 Functions covered by the nominative and accusative cases in Turkana	258
8.3 Functions covered by the nominative and accusative cases in Dhaasanac	262
8.4 A schematic overview of the distribution of low tone for nominative encoding	268

## Tables

---

1.1	Past tense markers in Luo	5
2.1	Related Nouns for ‘meat’ and ‘animal’ in Hausa and !Xun	25
2.2	Relative frequency of occurrence of eleven typological properties in African languages	29
2.3	Distribution of eleven typological properties according to major world regions	30
2.4	Distribution of eleven typological properties: sub-Saharan Africa vs. rest of the world	31
2.5	Number of typological properties shared by selected languages of northern Nigeria	33
3.1	Number of languages having each of three consonant types in 150 African languages and 345 non-African languages	40
3.2	Frequency of four types of labial-velar stops in the African database	42
3.3	African languages in our sample with nasal vowels	45
3.4	Languages reported to lack distinctive nasal consonants	47
3.5	Ikwere consonants	47
3.6	Frequency of vowel systems in 150 African languages, classified by number of contrastive vowel heights	51
3.7	Frequency of three characteristic consonant types of the South zone	61
3.8	Some varieties of non-click and click consonants in Khoisan and southern Bantu languages	64
3.9	African and non-African languages lacking P-sounds	66
3.10	Some languages of the eastern Sudan displaying minimal contrasts between dental and alveolar or retroflex consonants	68
3.11	Phonological characteristics of African languages, by zone	81
3A	Composition of the Sudanic database, by genetic groups	83
3B	Composition of the North, East, Rift, Center, and South databases	83
3C	Composition of the tone level database	84

3D	Composition of the question intonation database	85
5.1	Logophoricity across African lineages	155
5.2	Labial-velar consonants across African lineages	157
5.3	ATR vowel harmony across African lineages	159
5.4	Word order S-(AUX)-O-V-X across African lineages	162
5.5	Word order V-O-NEG across African lineages	164
5.6	Labial flaps across African lineages	166
5.7	Distribution of linguistic features across African lineages	168
5.8	Minimal-augmented pronoun systems across African lineages	170
6.1	The languages and their genetic classification	188
6.2	Preverbal clitic cluster of Proto-West Rift	197
6.3	Preverbal clitic cluster and finite verb in Nyaturu	198
6.4	Preverbal clitic cluster and finite verb in Gisamjanga Datooga	204
6.5	Features of the Tanzanian Rift Valley area according to language families	221
6.6	Features of the Tanzanian Rift Valley area and membership index of individual languages	222
7.1	Phonological features	231
7.2	Grammatical features	232
7.3	Possible new features of the ELA	249
8.1	Case terminology in Dhaasanac	260
8.2	Examples of case forms in Dhaasanac	260
9.1	Case marking in Nyimang	287
9.2	Head marking and dependent marking in the Kalahari Basin	291
9.3	Dependent marking in Nilo-Saharan	292
9.4	Inflection of main verbs and converbs in Omotic	294
9.5	Clausal cohesion markers in Wolaitta	297

## Contributors

---

PROFESSOR G. N. (NICK) CLEMENTS, Laboratoire de Phonétique et Phonologie (UMR 7018), CNRS / Sorbonne-Nouvelle, France

DR. JOACHIM CRASS, University of Mainz, Germany

PROFESSOR DENIS CREISSELS, Laboratoire Dynamique du Langage, MRASH, Lyon, France

PROFESSOR GERRIT J. DIMMENDAAL, Institut für Afrikanistik, University of Cologne, Germany

PROFESSOR ZYGMUNT FRAIJZYNGIER, Department of Linguistics, University of Colorado, USA

DR. TOM GÜLDEMANN, Max Planck Institute for Evolutionary Anthropology/Linguistics, Germany

PROFESSOR BERND HEINE, Institut für Afrikanistik, University of Cologne, Germany

PROFESSOR ROLAND KIEBLING, Asien-Afrika-Institut, Abteilung für Afrikanistik und Äthiopistik, University of Hamburg, Germany

DR. CHRISTA KÖNIG, Institut für Afrikanische Sprachwissenschaft, University of Frankfurt, University of Hamburg, Germany

DR. ZELEALEM LEYEW, Department of Linguistics, University of Addis Ababa, Ethiopia

DR. RONNY MEYER, University of Mainz, Germany

PROFESSOR MAARTEN MOUS, Department of Languages and Cultures of Africa, Leiden University, The Netherlands

PROFESSOR DEREK NURSE, Department of Linguistics, Memorial University of Newfoundland, Canada

DR. ANNIE RIALLAND, Laboratoire de Phonétique et Phonologie (UMR 7018), CNRS / Sorbonne-Nouvelle, France

## Series editor's foreword

---

The series Cambridge Approaches to Language Contact (CALC) was set up to publish outstanding monographs on language contact, especially by authors who approach their specific subject matter from a diachronic or developmental perspective. Our goal is to integrate the ever-growing scholarship on language diversification (including the development of creoles, pidgins, and indigenized varieties of colonial European languages), bilingual language development, code-switching, and language endangerment. We hope to provide a select forum to scholars who contribute insightfully to understanding language evolution from an interdisciplinary perspective. We favor approaches that highlight the role of ecology and draw inspiration both from the authors' own fields of specialization and from related research areas in linguistics or other disciplines. Eclecticism is one of our mottoes, as we endeavor to comprehend the complexity of evolutionary processes associated with contact.

*A Linguistic Geography of Africa* diverges from the above tradition only in not being a monograph. We have made this exception because of the rich addition that the volume makes to the scholarly goals of CALC. Bernd Heine and Derek Nurse enrich the series with an outstanding collection of contributions that make evident how the linguistic history of a vast territory is naturally complicated by an intricate entanglement of genetic and areal ties. The backbone of the book consists of a few fundamental questions, including the following: why are the genetic and typological classifications of African languages not coextensive? Why are there so many typological features that cut across well-established genetic classifications of languages? How significant is the role of areal diffusion, therefore of language contact, in the structural affinities observable among so many languages?

From the perspective of language speciation, a number of other questions arise from the chapters of this book. For instance, what does this complex web of structural and genetic affinities suggest about patterns of population dispersal and subsequent inter-group communication over the millennia that have led to the present African linguascape? Can one do a genetic study of languages of any territory without sorting out among the following classic questions: among the features shared by a particular group of languages,

which are the ones that have been inherited from a common ancestor? Which are those that are due to parallel, independent innovations after diversification from the proto-language? Which are those that are due to mutual influence during post-diversification contacts? How can one tell the difference? Are there any correlations between linguistic groupings and geographical characteristics of the relevant territories?

These questions, and many more, are addressed or brought up in complementary and often also overlapping ways by the contributors to *A Linguistic Geography of Africa*. The authors are all veteran field workers, typologists, and students of genetic classifications of African languages. Like the editors, many of them are also students of various forms of human contacts which produced not only creoles and pidgins but also varieties identified by some as “intertwined languages,” those that display the most “global copying” or “grammatical replication.” They are also aware of the danger of attributing too much to contact, as Derek Nurse had previously shown in his study, with other associates, of the evolution of Swahili. All the authors show that it would be too simplistic to stipulate, without qualifications, that Africa is a *Sprachbund* (also known as *convergence area*). Not even the Balkan-like situation of the Ethio-Eritrean area fits this sweeping characterization.

The bottom line is that one must combine techniques from both genetic and areal linguistics to account for the complex ways in which genetic and typological connections are literally intertwined in Africa and perhaps elsewhere too. Population movements, which lead to language contact, are an important dimension of the study of language evolution. The genetic classification of languages to reflect patterns of language diversification in a geographical space as vast as Africa is a much more complex topic than traditionally assumed. As the editors conclude, the contributors to this volume are just scraping the tip of the iceberg. There is much more to learn; and I hope *A Linguistic Geography of Africa* will be as seminal as other previous publications in this series, especially *Language Contact and Grammatical Change* by Bernd Heine and Tania Kuteva (2005). The books show the extent to which studies of language evolution and of language contact at the macro-level are interconnected. In the particular case of *A Linguistic Geography of Africa*, both the areal and genetic linguistics approaches help us make informed distinctions between convergence and diversification with common genetic inheritance, and/or, in some cases, articulate the relevant problems. The book also drives home the fact that language evolution is conditioned by a wide range of ecological factors, including those that pertain to geography, as these bear on population movements and all that ensue from them, economically and linguistically.

## Acknowledgments

---

Many a student of African languages has pointed out that the way grammatical meanings and structures are expressed in a given language is strikingly similar to what is found in some neighboring language or languages, even though the languages concerned may be genetically unrelated or only remotely related. The main goal of this book is to show that such observations are far from coincidental – rather, that such similarities across African languages are more common than is widely believed and that they can be accounted for in terms of areal relationship. In editing this book, we have benefited greatly from discussions with and comments from many colleagues, in particular the following: Sasha Aikhenvald, Felix Ameka, Walter Bisang, Bernard Comrie, Norbert Cyffer, Östen Dahl, Bob Dixon, Paul Newman, Margaret Dunham, Ludwig Gerhardt, Lars Johanson, Tania Kuteva, Yaron Matras, Salikoko Mufwene, Robert Nicolai, Mechthild Reh, Suzanne Romaine, Marie-Claude Simeone-Senelle, Anne Storch, Thomas Stolz, Martine Vanhove, Erhard Voeltz, Rainer Vossen, as well as many others. Special thanks are also due to Monika Feinen and Will Oxford for their invaluable technical assistance, and to two anonymous referees for all the work they did on an earlier version of this volume.

The second-named editor would also like to thank the Henrietta Harvey Foundation for continued and generous support during the period of his work on this book. The first-named editor is deeply indebted to the Center for Advanced Study in the Behavioral Sciences, Stanford, and the Institute for Advanced Study, LaTrobe University, and in particular the Research Centre for Linguistic Typology in Melbourne and its directors Bob Dixon and Sasha Aikhenvald, who offered him academic hospitality to work on this book. We are grateful to these institutions for their generosity and understanding. Our gratitude also extends to our colleagues Lenore Grenoble and Lindsay Whaley, Dartmouth College, who offered the first-named editor academic hospitality and the means to work on this book when he was invited as a visiting professor in 2002. Finally, thanks are also due to the Netherlands Institute for Advanced Study (NIAS) and its rector Wim Blockmans for the outstanding hospitality that the first-named editor was able to enjoy in 2005/6 at the finalizing stage of the book.

## Abbreviations

---

(In Bantu examples, numbers in glosses of nominal forms, or following ‘3:’ in glosses of pronominal markers, indicate noun classes. Elsewhere, numbers indicate persons. For example, O3:2 means ‘object marker, third person, class 2.’)

A	agent, transitive subject function, aspect (Dimmendaal), verb class A, affirmative
AA	Afroasiatic
ABL	ablative
ABS	absolutive
ABS	absolute form of nouns
ACC	accusative, accusative system
AFF	affirmative, affix
AG	antigenitive
ALBU	Alagwa and Burunge
ANAPH	anaphora
ANN	annexion state
ANT	anterior
APPL	applicative
AP	antipassive
ASS	associative
ATR	advanced tongue root
AUX	auxiliary
B	verb class B
CAUS	causative, causative extension
CL	noun class marker
CMPL	complementizer
CND	conditional
CNV	converb
COM	comitative
COMP	comparative

COMPL	complement
CON	construct case
COP	copula, copulative
CSTR	construct state
D	dative object marker
DAT	dative, dative extension
DC	declarative sentence marker
DECAUS	decausative
DECL	declarative
DEF	definite
DEM	demonstrative
DET	determiner
DETR	detransitivization
DIM	diminutive
DIR	directional
DP	declarative particle
DS	dependent subject
DU	dual
E	evidential
E40	Mara Bantu languages
E50	Central Kenyan Bantu languages
EAB	East African Bantu
EE	end of event marker
ELA	Ethiopian Linguistic Area
ERG	ergative, ergative system
EXCL	exclusive
F	feminine
F10, F20, etc.	groups of the Bantu zone F
FAC	factive
FOC	focus
FF	far future
FP	far past
FPL	feminine plural
FSG	feminine singular
FUT	future
GEN	genitive
GO	go
HAB	habitual
IDEO	ideophone
IMP	imperative
INCL	inclusive
INDEF	indefinite

INDIR	indirect mood
INF	infinitive
INIT	initiator of reported speech
INST	instrumental
IO	indirect object
IPF	imperfective
IRR	irrealis
JUS	jussive
LINK	linker introducing a noun modifier
LOC	locative
LOCPRED	locative predicator
LOG	logophoric
M	masculine
MAL	malefactive
MNOM	marked nominative
MPL	masculine plural
MSG	masculine singular
N	neuter, noun
N.PRED	nominal predicate
NEG	negative
NF	near future
NFIN	non-finite
NOM	nominative
NOML	nominalizer
NP	near past, noun phrase
NS	Nilo-Saharan
NSM	non-interrogative sentence marker
NTS	non-topical subject
O	object, transitive object function, direct object marker
OBJOR	object orientation
OBL	oblique case
OPT	optative
PAST	past
PEC	Proto-Eastern Cushitic
PEE	possessee
PERF	perfect
PERS	persistent
PFV	perfective
PIRQ	Proto-Iraqwoid
PL	plural
PNM	predicate nominal phrase marker
POR	possessor

POS	positive (or affirmative)
POSS	possessive
POST	posterior
PRE	preposition
PRES	present
PROG	progressive
PSAM	Proto-Sam
PSN	Proto-Southern Nilotic
P(N,S)WR	Proto-(North, South) West Rift
PSV	passive
PURP	purpose
Q	question
QUOT	quotative
REL	relative
RSM	resumptive marker
S	subject, intransitive subject function
SC	subject concord
SEQ	sequential
SG	singular
SJN	subjunctive
SIM	similative
SPEC	specifier
SUB	subordination
SUBJ	syntactically marked case form
TA	tense–aspect
TAG	tag question
TAM	tense–aspect–modality
TMP	temporal
TR	transitive
V	verb
VEN	venitive (ventive)
VN	verbal noun
WR	West Rift
I	non-past verbal juncture
II	past verbal juncture
1, 2, 3	first, second, third person, or noun class 1, 2, 3, etc.
=	clitic boundary

# 1 Introduction

---

*Bernd Heine and Derek Nurse*

More than forty years ago, Joseph Greenberg (1963) demonstrated that the African continent can be divided into four distinct genetic phyla, or families as he called them, namely Niger-Congo (or Kongo-Kordofanian), Nilo-Saharan, Afroasiatic, and Khoisan. For subsequent generations of Africanists, this classification has served as a reference system to describe the relationship patterns among African languages. In this tradition, scholars doing comparative work on African languages were preoccupied to quite some extent with reconstructing and understanding similarities across languages with reference to genetic parameters. One effect this line of research had was that an interest in other kinds of linguistic relationship was never really pronounced. Especially the question of whether, or to what extent, structural similarities and dissimilarities among African languages are the result of areal, that is contact-induced relationship, has never attracted any major research activities beyond individual studies dealing with lexical borrowing and related subjects. Whether the African continent constitutes an areally defined unit, or whether it can be subclassified into linguistic areas (or sprachbunds, or convergence areas) remained issues that were the subject of casual observations or conjecture, or both, but not really of more detailed research.

Still, once more it was Greenberg who drew attention to the importance of areal relationship in Africa. Not only did he venture to point out major linguistic areas (1959), but he also was the first to come up with important findings on the areal distribution of phonological and morphosyntactic properties across Africa, and with hypotheses on the areal distribution of these properties (1983). The title of chapter 2 of this book echoes that of a paper by Joseph Greenberg (1959), and this choice is deliberate: with this book we wish to build on the foundations laid by Greenberg, demonstrating that in the course of the last decades some headway has been made in areal classification since his paper appeared nearly half a century ago.

A common thread to all the contributions of this volume is that genetic relationship is far from being a parameter for understanding many of the processes characterizing the history of and typological relationship among African languages, and the message implicit in these contributions is that for a

better understanding of African languages, their structures, and their history, more detailed information on the areal relationship patterns is a *sine qua non* – not only for accounting for the relationship patterns among these languages, and for understanding Africa’s linguistic geography, but also for reconstructing Africa’s history and prehistory.

Work on linguistic areas or sprachbunds is not a new research line in Africa (see chapter 2 on the notion “linguistic area”). As early as 1976, an Ethiopian or, perhaps more appropriately, an Ethio-Eritrean area was proposed (Ferguson 1976), and this area is widely believed to constitute the only sprachbund-type unit to be found in Africa. However, doubts have been raised concerning the validity of this unit (Tosco 2000b). Tosco draws attention to the fact that there are a number of smaller, historically more immediately accessible areal groupings that can tell us more about the linguistic history of the macro-region concerned; chapter 7 will review this discussion and provide a summary and new findings on the nature of this sprachbund.

Otherwise, not much headway has been made in the search for linguistic areas within Africa. Some areas have been proposed, but the evidence to support the hypotheses concerned is in most cases not entirely satisfactory. An exception can be seen in Güldemann’s (1998) attempt to define the Kalahari Basin as an areal unit. Based on the methodology developed by Nichols (1992), he argues that it is possible on quantitative grounds to set off the languages of this arid region of Botswana, Namibia, and South Africa from other African languages. The Kalahari Basin area includes a number of – though not all – Khoisan languages plus the Bantu languages Herero and Tswana.

As we will see in the following chapters, genetic relationship does not provide the only parameter for diachronic language classification in Africa; rather, there is reason to maintain that the African continent can equally well be classified in terms of areally defined groupings. Unlike the genetic stocks proposed by Greenberg (1963) these groupings are not really discrete and exhaustive, they exhibit overlapping structures and fuzzy boundaries. However, as we hope to demonstrate in this volume, the areal relationship patterns characterizing these groupings are immediately relevant for understanding structural properties of African languages.

## Language contact

Areal relationship is the result of contact between languages, more precisely between the speakers of these languages. Language contact may have a wide range of implications for the languages involved, and it may affect virtually any component of language structure (see Thomason & Kaufman 1988). Grossly speaking, contact-induced influence manifests itself in the transfer of

linguistic material from one language to another, where linguistic material can be of any of the following kinds:

- (a) Form, that is, sounds or combinations of sounds
- (b) Meanings (including grammatical meanings) or combinations of meanings
- (c) Form-meaning units or combinations of form-meaning units
- (d) Syntactic relations, that is, the order of meaningful elements
- (e) Any combination of (a) through (d)

Language contact may involve simultaneously all kinds of transfer, that is, it may concern what Johanson (1992, 2002) calls global copying (*Globalkopieren*); but it may also involve only one kind of transfer, i.e. what Johanson calls selective copying (*Teilstrukturkopieren*). The data that are provided in this volume relate to both global and selective copying. But, as we will see in a number of chapters, there is one kind of transfer, namely (b), whose significance has been underrated in many previous studies of language contact: the transfer of meanings and combinations of meanings, occasionally discussed under the label “calquing,” is the one that is most difficult to identify, but that is presumably as common as lexical borrowing or other kinds of (c). And perhaps even more importantly, (b) concerns not only the lexicon, but presumably more often the transfer of functional categories, that is, it qualifies as what is technically known as grammatical replication (Heine & Kuteva 2003, 2005, 2006).

While still ill-understood, grammatical replication appears to be a ubiquitous phenomenon in Africa. One of its main effects is that as a result of language contact, a language acquires a new use pattern or grammatical category, or a new way of structuring grammar. The following example may illustrate this effect. The Ilwana, a Bantu-speaking people living along the river Tana south of Garissa in eastern Kenya, have a history of over three centuries of contact with the Orma, who speak a dialect of the East Cushitic Oromo language. Bantu languages have a robust number distinction singular vs. plural, supported by the noun class system, where there is a singular marker regularly corresponding to a plural marker. Orma on the other hand has a prevailing pattern distinguishing three number categories: singulative vs. transnumeral (unmarked) vs. plural/collective. For example, ethnonyms tend to be used in the unmarked transnumeral form and a singular is formed by adding the singulative suffix. Ilwana speakers appear to have replicated this structure with ethnonyms, whereby the Bantu singular (noun class 1) prefix *mo-* was reinterpreted as a singulative prefix while the Bantu plural noun class 2 was replaced by noun class 10, which is unmarked for number – thereby giving rise to an unmarked plural resembling the transnumeral category of Orma (Nurse 2000b: 125; see also Nurse 1994). Thus, a Bantu structure illustrated in (1) was replaced in Ilwana by the structure shown in (2).

- |     |  |                             |                 |
|-----|--|-----------------------------|-----------------|
| (1) | Swahili (Sabaki, Bantu)                  |                             |                 |
|     | M-pokomo                                 | Wa-pokomo ( <i>plural</i> ) | ‘Pokomo person’ |
| (2) | Ilwana (Sabaki, Bantu; Nurse 2000b: 125) |                             |                 |
|     | mo-bokomo                                | bokomo ( <i>plural</i> )    | ‘Pokomo person’ |

### Change in typological profile

Cases such as the one just looked at will surface in a number of the following chapters: they concern the transfer of a structure from one language to another without involving any lexical or other form–meaning units. But we will also look at more dramatic cases of transfer, involving simultaneously a bundle of structural properties and leading to new typological profiles. With the term “new typological profile” we refer to cases where, as a result of grammatical replication, a language experiences a number of structural changes to the effect that that language is structurally clearly different from what it used to be prior to language contact (Heine & Kuteva 2006). Typically, these changes are in the direction of the model language, thus making the two languages structurally more equivalent and more readily inter-translatable – a process that in contact linguistics tends to be described as “convergence.”

We may illustrate this process with the following example from the Kenyan language Luo. As we will see most clearly in chapter 6, East Africa is a region characterized by massive contact between languages belonging to different genetic stocks. Some of the linguistic effects of this contact concern Nilotic languages (belonging to the Nilo-Saharan family) that have been in contact with Bantu languages (belonging to the Niger-Congo family), especially Kalenjin (Southern Nilotic) and Luo (Western Nilotic) of south-central and southwestern Kenya. Nilotic languages may be called aspect-prominent, in that they commonly distinguish e.g. between a perfective and an imperfective aspect in verbs, mainly by way of tonal inflection. Bantu languages on the other hand are well known for their richness in tense distinctions, and the languages with which Kalenjin and Luo came into close contact are no exception to this rule. For example, the Bantu language Luhya (Luyia), which has been in contact with both Kalenjin and Luo, has among others the following tense categories expressed by verbal prefixes (Bukusu dialect of Luhya): Immediate Past, Near Past, Intermediate Past, Remote Past; Immediate Future, Intermediate Future, and Remote Future (Dimmendaal 1995a, 2001a, 2001b: 92; Kuteva 2000). While in Nilotic languages there are hardly any tense categories, the two languages for which there is an attested history of close contact with Bantu languages, viz. Kalenjin and Luo, have an array of tense distinctions comparable to that found among their Bantu neighbors. However, none of the tense markers in Kalenjin and Luo is etymologically related to corresponding

Table 1.1 *Past-tense markers in Luo (Dimmendaal 2001b: 101)*

Adverb of time	Verbal proclitic or prefix	Tense meaning
nénde	née, n-	'today in the past' (hodiernal)
nyóro	nyóo, ny-	'yesterday's past' (hesternal)
nyóca	nyóc(a), nyóc-	'the day before yesterday'
yandé	yand(é), yand-	'a few days ago'

tense markers in any of the Bantu languages concerned. Further, tense markers precede the verbal subject prefix in Kalenjin and Luo but follow the verbal subject prefix in the Bantu languages (Dimmendaal 2001: 93), and they have normally clearly affixal status in the Bantu languages but vary between clitic and affix status in Kalenjin and Luo.

Assuming that these two Nilotic languages replicated their tense categories from Bantu languages, the question arises as to what accounts for the structural difference between the two kinds of languages. Dimmendaal provides a cogent answer: the Nilotic languages received from their Bantu neighbors a range of tense concepts but neither the corresponding forms nor the morphosyntactic structures. Nilotic languages commonly use adverbs of time clause-initially (or clause-finally) to mark distinctions in time, and transfer had the effect that a set of such adverbs were grammaticalized to tense markers in clause-initial position; see table 1.1. Not surprisingly, therefore, these tense markers appear before the subject prefixes; in contrast to the model Bantu languages, which commonly have tense markers after the subject prefixes (Dimmendaal 2001b: 90–1). That this process happened independently in Luo from that to be observed in Kalenjin is suggested, for example, by the fact that the forms used in the two languages are not cognate (nor are they etymologically related to corresponding forms in the Bantu languages). There is one slight difference between the two Nilotic languages: while the grammaticalized tense markers have been adapted to the vowel harmony pattern of the verb stem in Kalenjin, they have not been affected by vowel harmony in Luo (Dimmendaal 2001b: 101).

To conclude, transfer appears to have had the effect that the Nilotic languages Kalenjin and Luo acquired a new functional domain (= tense) via the grammaticalization of adverbs of time.

The case just discussed is not an isolated instance of grammatical transfer from Bantu to Nilotic languages. Bantu languages are known for their rich paradigms of verbal derivational extensions marked by suffixes. There is nothing comparable in the Nilotic language Luo or its closest relatives, the Southern Lwoo languages of Uganda and the Sudan: verbal derivation is limited, mainly involving internal morphology in the verb root. Now, apparently on the model of neighboring Bantu languages, Luo speakers have

developed a set of what look like verbal suffixes, resembling structurally the Bantu verbal suffixes, expressing functions typically encoded by the Bantu derivational applied suffix *\*-id-* ('for, to, with reference to, on behalf of'). Luo speakers used the prepositions *ne* (or *ni*) benefactive, *e* locative, and *gi* instrumental in order to develop verbal enclitics or suffixes; the following example is confined to the benefactive preposition *ne*, where (3a) illustrates the prepositional use and (3b), where *Juma* is topicalized, the use as a verbal suffix (see also Dimmendaal 2001b: 101–2).

(3) Luo (Western Nilotic, Nilo-Saharan; Heine & Reh 1984: 51)

- a.        jon            nego            diel            ne            juma  
           John        is.killing    goat           for          Juma  
           'John is killing a goat for Juma'
- b.        juma        jon            nego-        ne        diel  
           Juma       John        is.killing- for       goat  
           'John is killing a goat for *Juma*'

On the basis of such evidence one may argue that this Nilotic language is on the verge of experiencing a gradual change of profile on the model of its Bantu neighbors. To be sure, Luo is structurally still unambiguously a Nilotic language, but it is typologically no longer exactly as it was prior to language contact with Bantu languages.

### Areal distribution: word order

Areal diffusion, especially when it does not involve lexical borrowing or other kinds of form–meaning units, is not easy to identify. Still, there are ways of developing plausible hypotheses on how linguistic properties spread from one language to another as a result of language contact. One of these ways concerns the probability of linguistic change. For example, Thomason proposes the following definition for contact-induced language change:

In my view, contact between languages (or dialects) is a source of linguistic change whenever a change occurs that would have been unlikely, or at least less likely, to occur outside a specific contact situation. This definition is broad enough to include both the transfer of linguistic features from one language to another and innovations which, though not direct interference features, nevertheless have their origin in a particular contact situation. (Thomason 2003: 688)

Perhaps the most obvious procedure to seek for hypotheses on contact-induced change concerns areal distribution among languages that are genetically unrelated or only remotely related. This procedure has been employed in some way or other by many students of contact-induced transfer (see especially Aikhenvald 2002), and it is used in several of the chapters in this book.<sup>1</sup>

We may illustrate the procedure with the following example, relating to a number of cases discussed in this book. Africa is commonly divided into four distinct language families or phyla. Assuming that languages belonging to different phyla, that is, genetic stocks, do not share any genetic relationship, one can hypothesize that if there is a linguistic property that is found widely in Africa across language phyla, that property is likely to be due to areal diffusion, that is, to language contact. But it is possible to invoke alternative hypotheses. If one finds similarities in form, meaning, or structure between different languages then that can be due to a number of different causes: it may be due to universal principles of linguistic discourse and historical development, to shared genetic relationship, to parallel development or drift, to language contact, or simply to chance. Assuming that we can rule out genetic relationship, drift, and chance, this leaves us with the possibility that universal principles may be responsible for the widespread occurrence of the relevant property. In such a situation, areal distribution once more provides a convenient parameter for testing the hypothesis: if the relevant property is widespread in Africa but uncommon in other parts of the world then a hypothesis based on universal principles can essentially be ruled out.

As we will see in the following chapters, this procedure has been employed extensively to formulate hypotheses on areal relationship across African languages. But the procedure has also been used to propose areal discontinuities within Africa. The areal distribution of word order can be used as an example to illustrate this observation.

In some of the literature on contact linguistics it is claimed or implied that syntax belongs to the most stable parts of grammar, and that it is most resistant to change. As we will see in this book, such a view is in need of revision: syntactic structures are easily transferred from one language to another. With regard to the classic distinction between verb-initial (VSO), verb-medial (SVO), and verb-final languages (SOV), none of the African language families exhibits any consistent word-order behavior: all three word orders are found in the Afroasiatic and the Nilo-Saharan phyla, and the Niger-Congo and Khoisan phyla exhibit two of these orders, namely SVO and SOV.<sup>2</sup>

But word order shows significant correlations with areal distribution. There is a large areal belt extending from Lake Chad to the west to the Horn of Africa to the east, where essentially only SOV languages are found (see chapter 9 concerning the complexity of this word-order type). This belt includes in the same way Nilo-Saharan languages such as Kanuri, Kunama, or Nobiin (Nile Nubian), furthermore all Omotic, Ethio-Semitic and, with one exception, also all Cushitic languages. In view of this areal contiguity and the genetic diversity involved, language contact offers the most plausible explanation to account for this typological similarity (Heine 1976). The areal-diffusion hypothesis receives further support from the fact that there is one Cushitic language that

has basic SVO order. This language, Yaaku, is spoken in central Kenya and is surrounded by languages such as Maasai and Meru that have, respectively, VSO and SVO rather than SOV word order.

Another example of areal patterning concerns what Heine (1975, 1976) calls type B languages. These languages are characterized by head-final word order (*nomen rectum* – *nomen regens*) in genitive (i.e. attributive possessive) and noun–adposition constructions, but otherwise head-initial order prevails, that is, nominal qualifiers such as adjectives and numerals tend to follow the head noun. What distinguishes them from SOV (i.e. type D) languages mainly is the fact that adverbial phrases follow the main verb. Type B languages are crosslinguistically uncommon; it is only in Africa that they are found in significant numbers. While occasional cases are found in various parts of the continent and in all African language families except Afroasiatic, the largest number exists in West Africa: there is a compact area extending from Senegal in the west to Nigeria in the east where virtually only type B languages are found (Heine 1976: 41–2).

One might argue that this concentration of type B languages in West Africa is genetically induced since with one exception all languages belong to the Niger-Congo phylum. But there are arguments against such a hypothesis. First, the area cuts across genetic boundaries, in that all Kwa languages located within this geographical region are type B, while eastern Kwa languages are not. Second, type B languages do not correlate with the genetic relationship patterns within the Niger-Congo phylum, that is, they do not form a genetic unit within Niger-Congo. And third, there is only one Nilo-Saharan language spoken in this West African region, namely Songhai, and it is exactly this Nilo-Saharan language which is type B.

A third example demonstrating that word order in African languages patterns areally rather than genetically is provided by what Heine (1976: 60) calls the Rift Valley (not to be confused with the Tanzanian Rift Valley area discussed in chapter 6). VSO languages form a distinct minority among African languages. Ignoring the Berber languages of northwestern Africa, whose status as VSO languages is not entirely clear, and a few Chadic languages, all African VSO languages are concentrated in a small geographical belt within or close to the East African Rift Valley stretching from southern Ethiopia to central Tanzania. While these languages belong with one exception to the Nilo-Saharan phylum, they consist on the one hand of Eastern Nilotic, Southern Nilotic, and Surmic (Didinga-Murle) languages, and on the other hand of the Kuliak languages Ik, Nyang’i and So, whose genetic position within this family is largely unclear. But perhaps most importantly, the area also includes Hadza (Hadzapi), which some classify as a Khoisan language while others prefer to treat it as a genetic isolate. On account of this areal patterning, the most convincing explanation for this typological clustering again is one in terms of areal relationship.

These are but a few examples showing that it is possible to formulate hypotheses on areal groupings within Africa on the basis of word-order characteristics. Some of these characteristics are also relevant in order to locate Africa typologically vis-à-vis other parts of the world. For example, as has been shown by Dryer (forthcoming), negation markers placed at the end of the clause can be found in a vast area extending from the river Niger in the west to the river Nile in the east, and including a wide range of languages belonging to Niger-Congo, Afroasiatic, and Nilo-Saharan, that is, to three of the four African phyla<sup>3</sup> (see chapter 4, pp. 163–5). The fact that the distribution of this typological property patterns areally and at the same time cuts across genetic boundaries is strongly suggestive of areal relationship. But verb-final negation does not only stand out typologically within the areal landscape of Africa; rather, it is also of worldwide significance: there appear to be only few languages outside Africa that have it.

### Micro-areas

Our focus in this book is on macro-situations, that is, on areal perspectives dealing with Africa as a whole or with significant regions of the continent. In doing so, we are aware that most of the data that are relevant for a better understanding of the mechanisms leading to areal diffusion in Africa have come not from macro-surveys but rather from micro-analyses of contact situations involving a limited number of different speech communities, in many cases only two, where one serves as the donor or model while the other acts as the receiver of linguistic transfers. We are not able to do justice to this rich research that has been carried out in Africa in the course of the last decades; suffice it to draw attention to a couple of studies resulting in fairly well-documented micro-situations of long-term and intense language contact. These studies have been volunteered by Nurse (2000b) on East African contact situations. One of them concerns the Daiso people of northeastern Tanzania, who originate from the central Kenyan highlands and appear to have reached their present territory early in the seventeenth century. By now, they have a history of nearly four centuries of contact with the Tanzanian Bantu languages Shamba(l)a, Bondei, Swahili, and Digo in the course of which their language has been influenced in a number of ways by these languages. The second study deals with the Ilwana, a Bantu-speaking people living along the river Tana south of Garissa in eastern Kenya. They have a history of over three centuries of contact with the Orma, who speak a dialect of the East Cushitic Oromo language (Nurse 2000b), and as a result of Orma influence have experienced a range of grammatical changes.

Intense language contact may result in situations of stable bilingualism, but it can as well lead to language shift, where one language gives way to another.

A number of studies carried out in Africa deal with contact-induced linguistic transfer in this kind of situation. Arguably the most substantial work dealing with such transfers is that by Sommer (1995) on Ngamiland in northern Botswana, where there is a detailed linguistic and sociolinguistic documentation of the process of transition from the minority language Yeyi (Siyeyi) to the national language Tswana (Setswana).

### **The present volume**

All the wealth of information that has been amassed in such studies has been made use of in the chapters to follow, but unlike these studies, the goal of this book is to present a more general perspective of areal relationship in Africa. The contributions are mainly of three kinds. First, there are those that argue that there is reason to consider the African continent as an areal-typological unit that stands out against the rest of the world. This perspective is highlighted in chapters 2, 3, and 4. In the subsequent chapters 5 through 7, specific linguistic regions of Africa are analyzed and evidence is presented to define them as linguistic areas. The remaining chapters 8 and 9 each highlight one particular typological feature with a view to exploring their significance as parameters for areal classification.

That there are a number of properties that are widespread in Africa but uncommon elsewhere has been pointed out by a number of scholars. The authors of chapter 2 go on to look for quantitative information to test this hypothesis, using a catalogue of eleven phonetic, morphological, syntactic, and semantic properties. The conclusion Bernd Heine and Zelealem Leyew reach confirms what has been established in earlier research, namely, that it is not possible to define Africa as an area in terms of a set of properties that are generally found in Africa but nowhere else. Nevertheless, they argue on the basis of their quantitative evidence that it is possible to maintain that areal diffusion must have played some role in shaping Africa's linguistic landscape and to predict with a certain degree of probability whether or not a given language is spoken on the African continent.

Another finding that surfaces in chapter 2 is that the highest concentration of Africa-specific properties is found in the Sudanic belt of west-central Africa, a region that includes languages of three of the four African language phyla, while northeastern and northern Africa are typologically quite different from the rest of the continent, sharing with the languages of western, central, and southern Africa hardly more properties than they share with languages in other parts of the world.

The question of whether Africa can be defined as a distinct area vis-à-vis other language regions of the world is also the central issue of chapter 3. Surveying a range of phonological phenomena and comparing their distribution

with that to be found outside Africa, Nick Clements and Annie Rialland are able to establish that there are in fact significant clusterings of phonological properties in sub-Saharan Africa. They go on to demonstrate that these clusterings concern most of all the Sudanic zone (i.e. roughly what is referred to in chapter 5 as the Macro-Sudan belt), that is, the sub-Saharan region roughly between the rivers Niger and Nile.

A major finding presented in chapter 3 concerns the areal subgrouping of Africa: the authors propose to classify the continent on the basis of typologically salient phonological parameters into six zones, which they call the North, Sudanic, East, Rift, Center, and South zones. These areal groupings cut across genetic boundaries, being suggestive of contact processes that characterize the prehistory of Africa.

While chapter 2 presents a crude template for identifying African languages and for distinguishing them from languages in other parts of the world, chapter 4 provides an extensive analysis of the main morphological and syntactic characteristics of the languages spoken in Africa, thereby building on the foundations laid in works such as Welmers (1974), Meeussen (1975), or Gregersen (1977). But in this chapter, Denis Creissels, Gerrit Dimmendaal, Zygmunt Frajzyngier, and Christa König go far beyond the scope of such works in building on substantive typological information on languages in other parts of the world. In this way, they are able to offer a truly contrastive perspective, demonstrating that there is a range of typological properties that are found extensively within Africa but are rare elsewhere in the world, and vice versa. Accordingly, the authors of this chapter present a balanced profile of African languages and contrast it with that of other linguistic regions of the world. In addition, this chapter also focuses on the internal typological complexity of the continent, suggesting areal groupings of various kinds, and enabling the reader to determine, for example, what structural characteristics to expect from a West African as opposed to an East African language.

The authors conclude chapter 4 with a list of nineteen morphosyntactic properties suggesting that African languages show a distinct areal behavior vis-à-vis other languages, exhibiting either an extraordinarily high or a clearly low rate of frequency of occurrence. This list also includes a number of perhaps surprising areal generalizations, such as the fact that no African language has been found so far where the verb obligatorily agrees with the object,<sup>4</sup> or that a number of African verb-final (SOV) languages exhibit a typologically unusual behavior in that they consistently place adverbial constituents after the verb (SOVX) (see also chapter 9 on this issue).

Ever since Westermann (1911) published his classic on the *Sudansprachen*, the large belt in the northern half of Africa south of the Sahara between the Niger and the Nile valleys has been the subject of hypotheses on the genetic relationship patterns in Africa. While there were scholars who claimed that the

affinities found between the languages spoken in this belt cannot be defined in terms of genetic relationship, no convincing alternative has been presented so far. Based on structural evidence from a wide range of languages, Tom Güldemann is able to demonstrate in chapter 5 that an areal approach offers the most convincing means to account for these affinities. A number of the typological properties studied by him are considered by some to be characteristic of Africa as a whole (see chapter 2), but their clustering in the vast area between West and East Africa – cutting across genetic boundaries, including those of three of the four African language phyla – in fact suggests that the Macro-Sudan hypothesis is a robust one, even if the boundaries of this area are fuzzy.

This chapter discusses a fundamental problem surfacing in some way or other in most contributions to this volume, namely: what do we really know about the genetic relationship patterns in Africa? Güldemann's suggestion that a number of the taxonomic units proposed by Greenberg (1963) may turn out to be more appropriately analyzable in terms of areal, that is, contact-induced relationship rather than in terms of genetic affiliation provides a challenge for future comparative linguistics in Africa.

The region of the Tanzanian Rift Valley is genetically one of the most complex linguistic regions of the world: it includes languages of all four African language phyla and, if one classifies Hadza as a genetic isolate, the region hosts even five different genetic stocks. The region therefore provides an ideal laboratory for the study of language contact, as Roland Kießling, Maarten Mous, and Derek Nurse aptly demonstrate in chapter 6.

The authors use what Campbell et al. (1986) call a historicist approach, that is, their areal description is based on properties that are likely to require an explanation in terms of language contact rather than of general typological similarities. However, the analysis on which this chapter is based is of a different nature from that characterizing Campbell et al.'s (1986) description of Meso-America as a linguistic area: rather than being confined to searching for a catalogue of properties that neatly define the linguistic area, Kießling, Mous, and Nurse go on to reconstruct the historical processes that can be held responsible for the presence of these properties and, hence, for the rise of the Tanzanian Rift Valley as a linguistic area.

The only sprachbund-type area in Africa figuring in textbooks of contact linguistics, on the same level as the Balkans, Meso-America, South Asia, etc., is the Ethiopian linguistic area (also called the Ethio-Eritrean area). More recently, however, some students of African languages have shown that defining this area as a sprachbund is not unproblematic; we have drawn attention to this research above. In chapter 7, Joachim Crass and Ronny Meyer offer a comprehensive appraisal of previous research. Based on their own recent field research, they come to the conclusion that the areal hypothesis is sound, and they add new evidence to further substantiate this hypothesis.

If there are genuinely African typological properties then marked-nominative systems are one of them: languages having a grammaticalized case system where the nominative is the functionally marked category are world-wide extremely rare; they are essentially confined to the African continent and are mainly concentrated in eastern Africa. As Christa König demonstrates in chapter 8, there is reason to assume that this regional patterning is due to some extent to genetic factors, but language contact also must have played some role in the diffusion of such systems across genetic boundaries. What is perhaps noteworthy is the fact that the area covered by marked-nominative languages cuts across the Ethiopian highland area and the lowland region of the Nile valley.

The author defines a number of typological properties characterizing case marking in the Afroasiatic and Nilo-Saharan languages of this area, among them being one according to which there is a generalization to the effect that in marked-nominative languages case is distinguished only after the verb, that is, there are no case distinctions before the verb; for obvious reasons, verb-final languages, such as Cushitic and Omotic languages, are not covered by this generalization.

Gerrit Dimmendaal's survey of verb-final languages in Africa in the final chapter 9 is not strictly on areal linguistics, but it is a demonstration of how areal forces and genetic inheritance interact in shaping the syntax of African languages. To be sure, there are strong correlations between SOV (subject–object–verb) word order and phrasal modifier–head order; as the author convincingly shows, however, labels such as “SVO” and “SOV” are not very helpful for understanding the dynamics underlying the syntax and the discourse-pragmatic structure of the languages concerned. His detailed analysis demonstrates that some of the generalizations proposed for SOV languages are in need of revision, considering the enormous diversity of morphosyntactic structures to be found in the so-called verb-final languages of Africa.

Among the many issues discussed in this chapter there is one that raises general problems for the typology of clause combining. It is widely assumed that the distinction between coordination and subordination is typologically neat; as Dimmendaal shows, however, the situation in Africa – but probably elsewhere as well – is more complex, and a more fine-grained typology of clause combining is required.

The impression one may get when reading the contributions to this volume is that work on the contact-induced patterns of linguistic relationship in Africa is still in its infancy, even though for more than a century, students of African languages have been drawing attention to the fact that neighboring but genetically unrelated or only remotely related languages exhibit a high degree of conceptual and structural intertranslatability. The present volume offers a multitude of examples confirming such observations and proposing significant

areal relationship patterns; still, it can achieve hardly more than revealing the peak of the iceberg. It is hoped that the volume makes it clear that Africa's linguistic geography, and the social dynamics of language contact underlying it, is a research topic that deserves much more attention than it has received in the past.

## 2 Is Africa a linguistic area?

---

*Bernd Heine and Zelealem Leyew*

The question raised in the title of this chapter has been posed by a number of students of African languages (e.g. Greenberg 1983; Meeussen 1975; Gilman 1986), it has figured in the title of a seminal paper by Greenberg (1959), and it is raised in various parts of this work (see especially chapters 3 and 4). In the present chapter it is argued that it is possible, on the basis of a quantitative survey on African languages of all major genetic groupings and geographical regions, to define a catalogue of phonological, morphosyntactic, and semantic properties that can be of help in defining African languages vis-à-vis languages in other parts of the world.<sup>1</sup>

### 2.1 On linguistic areas

Areal linguistics is a much neglected field of comparative African linguistics. While there are a number of studies that have been devoted to contact between individual languages or language groups (e.g. Mutahi 1991; Nurse 1994; 2000b; Sommer 1995; Bechhaus-Gerst 1996; Dimmendaal 1995a; 2001b; Storch 2003), not much reliable information is available on areal relationship across larger groups of languages. The following are among the questions that we consider to be especially important in this field:

- (1) Can Africa be defined as a linguistic area vis-a-vis the rest of the world?
- (2) Are there any clearly definable linguistic macro-areas across genetic boundaries within Africa?
- (3) Are there any linguistic micro-areas?

While the majority of chapters in this book deal with questions (2) and (3), our interest in this chapter is exclusively with question (1). A variety of different terms have been proposed to refer to sprachbunds, such as linguistic area, convergence area, diffusion area, *union linguistique*, *Sprachbund*, etc. (see Campbell et al. 1986: 530). Perhaps the most frequently discussed sprachbunds are the Balkans (for convenient summaries, see e.g. Joseph 1992; Feuillet 2001),<sup>2</sup> Meso-America (Campbell et al. 1986), Ethiopia (Ferguson 1976),<sup>3</sup> South Asia (Masica 1976; Emeneau 1980), the East Arnhem Land (Heath 1978),

the Amerindian Pacific Northwest (Sherzer 1973; Beck 2000), the Vaupés basin of northwest Amazonia (Aikhenvald 1996; 2002), Standard Average European (Haspelmath 1998; 2001), and the Daly River area of Australia (Dixon 2002: 674–9). Furthermore, there are quite a number of less widely recognized sprachbunds, such as the Circum-Baltic (Nau 1996; Koptjevskaja-Tamm & Wälchli 2001), the Middle Volga region (Johanson 2000), or the Circum-Mediterranean area (Stolz 2002).

Substantial work has been done to define sprachbunds, with the result that there are now a few areas in all major parts of the world that can be described in terms of language contact. With regard to defining sprachbunds, two different stances can be distinguished. On the one hand it is argued that a definition of sprachbunds should highlight the fact that they are the result of language contact, that is, of historical processes; the following is representative of this view:<sup>4</sup>

*A linguistic area is defined . . . as an area in which several linguistic traits are shared by languages of the area and furthermore, there is evidence (linguistic and non-linguistic) that contact between speakers of the languages contributed to the spread and/or retention of these traits and thereby to a certain degree of linguistic uniformity with the area. (Sherzer 1973: 760)*

On the other hand, sprachbunds are defined exclusively in terms of linguistic parameters without reference to the historical forces that gave rise to them. Emeneau's classic definition<sup>5</sup> is a paradigm case of such definitions; a more recent version is the following (see also Aikhenvald 2002: 7–8):

*A linguistic area can be recognized when a number of geographically contiguous languages share structural features which cannot be due to retention from a common proto-language and which give these languages a profile that makes them stand out among the surrounding languages. (Haspelmath 2001: 1492)*

In the present chapter, we will be confined to the second kind of definition, and we will assume that there is a sprachbund whenever the following situation obtains:

#### (4) Characterization of linguistic areas

- a. There are a number of languages spoken in one and the same general area.
- b. The languages share a set of linguistic features whose presence can be explained with reference neither to genetic relationship, drift, universal constraints on language structure or language development, nor to chance.
- c. This set of features is not found in languages outside the area.
- d. On account of (b), the presence of these features must be the result of language contact.

This characterization is fairly general, it is not meant to be a definition; rather, it is used as a convenient discovery device for identifying possible instances of

sprachbunds. Note that this characterization does not address crucial problems that have been raised in the relevant literature, for example, how many languages and how many features (or properties or traits) are minimally required, whether these features should be shared by all languages, whether individual features should not occur in languages outside the sprachbund, whether the languages should really be geographically contiguous, whether the languages should belong to different genetic groupings, to what extent isoglosses of features need to bundle, how factors such as the ones just mentioned influence the strength of a sprachbund hypothesis, or whether sprachbunds have any historical reality beyond the linguistic generalizations proposed by the researchers concerned.

## 2.2 Earlier work

Pre-Greenbergian comparative African linguistics suffered from the fact that no systematic distinction between different kinds of historical relationship was made, that is, it remained for the most part unclear whether the linguistic classifications proposed were intended to be genetically, areally or typologically defined or, more commonly, were an amalgamation of all three kinds of relationship. Accordingly, most of the works published prior to 1959 do not offer unambiguous evidence on areal patternings within Africa or between Africa and other parts of the world.

Greenberg's contribution to areal linguistics was of two kinds. First, he proposed a genetic classification of the languages of Africa (1963). A crucial problem associated with many cases of crosslinguistic comparison concerns the fact that it frequently remains unclear whether a given similarity found between languages is due to genetic or to areal relationship. Once it has been established where genetic boundaries are it is possible to propose viable hypotheses on areal diffusion and areal relationship. With his genetic classification therefore, Greenberg made it possible to draw a clear demarcation line between genetic relationship and other kinds of relationship.

Second, Greenberg also made the first substantial contribution to areal relationship in Africa. In an attempt to isolate areal patterns both within Africa and separating Africa from other regions of the world, he proposed a number of what he called "special" features of African languages. The properties listed by Greenberg (1959) include in particular a number of lexical polysemies, such as the use of the same term for 'meat' and 'wild animal,' the use of the same term for 'eat,' 'conquer,' 'capture a piece in a game,' and 'have sexual intercourse,' and the use of a noun for 'child' as a diminutive, or of 'child of tree' to denote 'fruit of tree.'

Another noteworthy contribution to areal relationship within Africa appeared in the same year, 1959: Larochette (1959) presented a catalogue of

linguistic properties characteristic of Congolese Bantu (Kikongo, Luba, Mongo), an Ubangi language (Zande), and a Central Sudanic language (Mangbetu), but a number of the properties proposed can also be found in other regions and genetic groupings of Africa. Another range of properties characterizing many African languages was proposed by Gregersen (1977) and Welmers (1974). Building on the work of Greenberg (1959) and Larochette (1959), Meeussen (1975) presented an impressive list of what he called “Africanisms,” that is, phonological, morphological, syntactic, and lexical properties widely found in African languages across genetic boundaries. Quite a number of the “Africanisms” proposed by Meeussen are in fact promising candidates for status as properties that are diagnostic of Africa as a linguistic area (see section 2.3 below).

Another seminal work on areal relationship was published by Greenberg in 1983. He defined areal properties “as those which are either exclusive to Africa, though not found everywhere within it, or those which are especially common in Africa although not confined to that continent” (Greenberg 1983: 3). As an example of the former he mentioned clicks; as instances of the latter he discussed in some detail the following four properties (“characteristics”; Greenberg 1983: 4): (i) coarticulated labial-velar (or labiovelar) stops, (ii) labial (or labiodental) flaps, (iii) the use of a verb meaning ‘to surpass’ to express comparison, and (iv) a single term meaning both ‘meat’ and ‘animal.’ He demonstrated that these four properties occur across genetic boundaries and, hence, are suggestive of being pan-African traits, especially since they are rarely found outside Africa.

Greenberg (1983) went on to reconstruct the history of these properties by studying their genetic distribution. He hypothesized that (i), (iii), and (iv) are ultimately of Niger-Kordofanian origin even though they are widely found in other African language phyla, in particular in Nilo-Saharan languages. For (ii), however, he did not find conclusive evidence for reconstruction, suggesting that it may not have had a single origin but rather that it arose in the area of the Central Sudanic languages of Nilo-Saharan and the Adamawa-Ubangi languages of Niger-Congo.

Search for areal properties across Africa is associated to some extent with creole linguistics (see e.g. Boretzky 1983). In an attempt to establish whether, or to what extent, the European-based pidgins and creoles on both sides of the Atlantic Ocean have been shaped by African languages, students of creoles pointed out a number of properties that are of wider distribution in Africa. Perhaps the most detailed study is that by Gilman (1986). Arguing that a large number of African-like structures in Atlantic and other pidgins and creoles are best explained by influence of areal properties widely distributed among the languages of Africa, Gilman proposed an impressive catalogue of pan-African areal properties (but see section 2.5).

### 2.3 “Africanisms”

In the works discussed in section 2.2 there are a number of properties that – following Meeussen (1975) – we will call Africanisms. With this term we are referring to properties that satisfy the following set of criteria:

- (a) They are common in Africa but clearly less common elsewhere.
- (b) They are found, at least to some extent, in all major geographical regions of Africa south of the Sahara.
- (c) They are found in two or more of the four African language phyla.

A number of properties that are clearly more widespread in Africa than elsewhere are not considered here, for the following reasons. First, because they appear to be genetically determined. The presence of gender or noun class systems is a case in point. Most instances of such systems to be found in Africa are presumably genetically inherited. This can be assumed to apply on the one hand to the nature-based noun class systems found in Niger-Congo and Khoisan languages, and on the other hand to the sex-based gender systems of Afroasiatic and Central Khoisan languages.<sup>6</sup>

Perhaps surprisingly, we will also not consider the presence or absence of clicks a relevant property, although it appears to be the only property that is confined exclusively to Africa, and although it satisfies all of the criteria proposed above. The reason for doing so is the following: the main goal of this chapter is to find out whether African languages resemble one another more than they resemble other languages and what factors can be held responsible for such resemblances. To be sure, clicks occur in three of the four African language phyla, not only in all Khoisan languages, but also in South African Bantu (Niger-Congo) languages, and in the Cushitic (Afroasiatic) language Dahalo; still, their occurrence is geographically restricted to southern Africa and three East African languages.

Furthermore, the fact that Khoisan languages are among the phonologically most complex languages in the world, some of them distinguishing more than 110 distinct phonemes, is ignored here since it does not appear to be characteristic of Africa as a linguistic area, being restricted to a few North and South Khoisan languages.

In the following we will discuss a catalogue of properties that have been proposed to be characteristic of Africa as a linguistic area (especially Greenberg 1959; 1983; Larochette 1959; Meeussen 1975; Gilman 1986). Our selection is to some extent arbitrary in that we will ignore some properties that have been mentioned by other authors but where we are not entirely convinced that they are possible candidates for status as “Africanisms.”

### 2.3.1 *Grammar*

A general phonological property that has been pointed out by a number of students of African languages is the preponderance of open syllables and an avoidance of consonant clusters and diphthongs (Meeussen 1975: 2; Gilman 1986: 41). Furthermore, tone as a distinctive unit is characteristic of the majority of African languages, in most cases both on the lexical and the grammatical levels (see section 2.4).

Ignoring click consonants, there are a number of consonant types that are widespread in Africa but uncommon elsewhere (see chapter 3 for detailed treatment). This applies among others to coarticulated labial-velar (or labio-velar) stops (Meeussen 1975: 2; Greenberg 1983: 4; Gilman 1986: 41). Labial-velars may be voiceless (*kp*) or voiced (*gb*). There are also corresponding nasals and/or fricatives, but they do not show the wide distribution of stops, and their occurrence is largely predictable on the basis of stops (Greenberg 1983: 4). The distribution of this property is clearly areally constrained: labial-velar stops occur in a broad geographical belt from the western Atlantic to the Nile–Congo divide, and they are also occasionally found outside this belt (see Welmers 1974: 47–8), e.g. in Katla and Giryama. Still, they are found in three of the four African phyla; only Khoisan languages have no labial-velar stops (see chapters 3 and 5). Also, in the Afroasiatic and Nilo-Saharan phyla, their occurrence is restricted essentially to one branch each, namely Chadic and Central Sudanic, respectively (Greenberg 1983: 7). Outside Africa, coarticulated labial-velar stops are found only sporadically, especially in northeastern Papua New Guinea in the Kâte-Ono group of the Indo-Pacific languages,<sup>7</sup> in some languages of Melanesia, and in the Austronesian language Iai (Greenberg 1983: 5; Maddieson 1984: 215–16); see section 3.2.4 for more details.

Perhaps even more characteristic are labial (or labiodental) flaps, where the teeth touch well below the outer eversion of the lip, which is flapped smartly outwards, downwards (see chapter 3). They have been found in all African phyla except Khoisan, e.g. in Chadic of Afroasiatic (Margi, Tera), Niger-Congo (Ngwe, Ngbaka, Ngbaka Mabo, Ndogo-Sere, some Shona dialects), and Nilo-Saharan (Kresh, Mangbetu) (Gregersen 1977: 31; Greenberg 1983: 4, 11). Still, their occurrence is confined to a relatively restricted number of languages, and even there they show restrictions in their use as phonemic units; not infrequently, these sounds are found only in special vocabulary such as ideophones. In their survey of 250 African and 345 non-African languages, Clements and Rialland did not find a single non-African language, but at least 70 African languages having such flaps (see chapter 3).

A third type of consonants that is widespread in Africa can be seen in implosives, which – following Clements and Rialland (chapter 3) – we define as non-obstruent stops. To be sure, these can be found in non-African