

# African Languages An Introduction

---

EDITED BY  
BERND HEINE

*Universität zu Köln*

AND  
DEREK NURSE

*Memorial University of Newfoundland*



**CAMBRIDGE**  
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE  
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS  
The Edinburgh Building, Cambridge CB2 2RU, UK  
40 West 20th Street, New York, NY 10011-4211, USA  
10 Stamford Road, Oakleigh, Melbourne 3166, Australia  
Ruiz de Alarcón 13, 28014 Madrid, Spain

www.cup.cam.ac.uk  
www.cup.org

© Cambridge University Press 2000

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

First published 2000

Printed in the United Kingdom at the University Press, Cambridge

*Typeset* Times 9/13 [GC]

*A catalogue record for this book is available from the British Library*

*Library of Congress Cataloguing in Publication data*

African languages: an introduction / edited by Bernd Heine (Universität zu Köln) and  
Derek Nurse (Memorial University of Newfoundland).

p. cm.

Includes bibliographical references and index.

ISBN 0 521 66178 1 (hardback)

I. African languages. I. Heine, Bernd, 1939– II. Nurse, Derek.

PL8005 .A24 2000

496 – dc21 99-056881

ISBN 0 521 66178 1 hardback

ISBN 0 521 66629 5 paperback

# CONTENTS

---

|  |         |
|--|---------|
| <i>List of maps</i>                            | page vi |
| <i>Notes on contributors</i>                   | vii     |
| 1 Introduction, BERND HEINE and DEREK NURSE    | I       |
| 2 Niger-Congo, KAY WILLIAMSON and ROGER BLENCH | 11      |
| 3 Nilo-Saharan, LIONEL M. BENDER               | 43      |
| 4 Afroasiatic, RICHARD J. HAYWARD              | 74      |
| 5 Khoisan, TOM GÜLDEMANN and RAINER VOSSEN     | 99      |
| 6 Phonology, G. N. CLEMENTS                    | 123     |
| 7 Morphology, GERRIT J. DIMMENDAAL             | 161     |
| 8 Syntax, JOHN R. WATTERS                      | 194     |
| 9 Typology, DENIS CREISSELS                    | 231     |
| 10 Comparative linguistics, PAUL NEWMAN        | 259     |
| 11 Language and history, CHRISTOPHER EHRET     | 272     |
| 12 Language and society, H. EKKEHARD WOLFF     | 298     |
| <i>References</i>                              | 348     |
| <i>Index of authors</i>                        | 374     |
| <i>Index of languages</i>                      | 379     |
| <i>Index of subjects</i>                       | 389     |

# MAPS

---

|      |  |               |
|------|--|---------------|
| 1.1  | African phyla and major languages                      | <i>page 2</i> |
| 2.2  | Niger-Congo  | 12            |
| 3.3  | Nilo-Saharan   | 44            |
| 4.4  | Afroasiatic  | 75            |
| 5.5  | South African Khoisan (pre-colonial situation)         | 100           |
| 9.6  | Basic word order types (according to Heine 1976a)      | 251           |
| 12.7 | The main African language-based <i>linguae francae</i> | 325           |
| 12.8 | Pidgins and creoles (based on Crystal 1997)            | 328           |
| 12.9 | Dominant non-African languages                         | 343           |

---

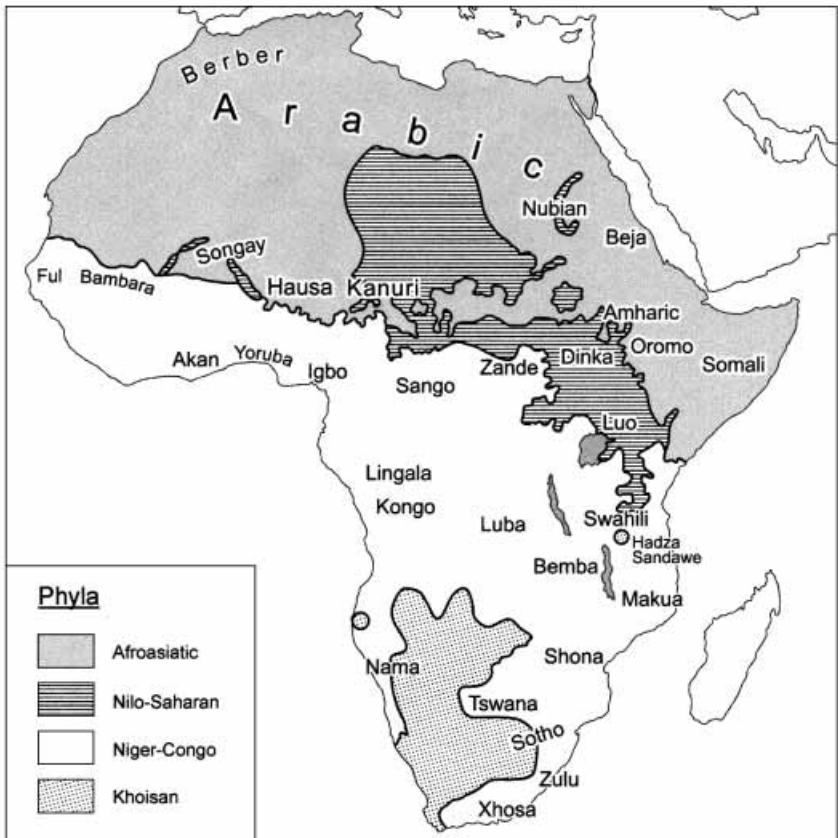
# Introduction

BERND HEINE AND DEREK NURSE

## 1.1 **How many African languages are there?**

We are sometimes asked ‘Do you speak African?’, as if there were but a single African language. A recent authority (Grimes (ed.) 1996) puts the number of African languages at 2,035: this number is not fixed, as some languages are still being ‘discovered’, while others with few speakers are being eliminated. Excluding languages introduced over the past two millennia or so, such as Arabic, Malagasy, Afrikaans, English, French, Spanish, and Portuguese, this figure of just over 2,000 breaks down into four large phyla: Niger-Congo 1,436 languages (including the Bantu family, which itself is often said to have 500 members), Afroasiatic 371, Nilo-Saharan 196 and Khoisan 35. A few Afroasiatic languages are spoken exclusively outside Africa, in the Middle East, which would reduce the figure for Africa somewhat. If we believe this figure of 2,000, then it represents nearly one-third of the world’s languages.

But in fact it is an estimate which should be treated with caution, because it depends crucially on where one draws the line between language and dialect. A language is often defined by some combination of: having national status; being written; being the standard form of a range of speech varieties; not being intelligible to speakers of other ‘languages’; and having a relatively large number of native speakers. By contrast, dialects are said to be local, not written, not the standard form, be mutually intelligible, and to have fewer speakers. In Africa, and also often outside Africa, such definitions frequently fail. There are of course many cases where what are generally considered ‘languages’ or even two or more varieties of what is usually considered the same language are not mutually intelligible, especially across language families. In many cases some agreed standard form is used as a national or official language, is the only officially condoned written



Map 1.1. African phyla and major languages.

form, and is used by millions of people, either as first or second language. In what follows, we sometimes use ‘variety’ to avoid confusion between language and dialect.

These conditions are not met in many other places in Africa. A common situation is a string of similar varieties, in which the speakers of variety A understand those of adjacent B, who in turn understand those of C, and so on, but the speakers of A do not understand speakers of the variety at the other end of the continuum, or even those part way along. Even if we can define ‘understand’, where is the divide between language and dialect in this situation? While some African countries, such as those in North Africa, or Somalia, harbour few languages, many others have many language communities within their boundaries: Nigeria is said to have nearly 500, Cameroon nearly 300, and three other countries over 100. This

proliferation introduces the practical problem of how to communicate across the nation, often solved by using an ex-colonial language, or Arabic, as the national language, which then reduces all other languages, some spoken by millions of people, often with a rich written literature, to the status of local ‘dialects’. In other cases, no variety, or several varieties, are written. In yet other cases, speakers are not agreed on a standard variety, or on what ought to be a or the national language.

Linguists who try to deal with this welter of languages are often referred to as ‘splitters’ or ‘lumpers’. Splitters tend to regard ‘varieties’ as distinct languages, thus boosting the 2,000, while lumpers treat varieties as just dialects, reducing the number.

## 1.2 **The intended audience of this book**

Anyone producing a book about this mass of languages has inevitably to make certain choices. There are older introductions to African languages in English (Berry and Greenberg 1971, Welmers 1973, Gregersen 1977), there are volumes that deal with some but not all African languages (Bender 1997, Bendor-Samuel 1989, Ehret 1995, Heine 1970, 1976a), there are books about African languages in languages other than English (Alexandre 1967, Heine *et al.* 1981). But there is no up-to-date, reasonably comprehensive, basic introduction to African languages in English. It is a gap we have heard mentioned by many, both colleagues and students, and it is the gap which this book aims to fill. We want to produce a book that will above all be accessible to undergraduates worldwide, and especially in Africa, but also elsewhere in the world. At the same time we hope it will be of interest to other audiences, such as general linguists or cognitivists who know little of the African situation, and to Africans or Africanists who are not linguists, but are knowledgeable in the history, culture or anthropology of Africa.

Our primary target, ‘undergraduates worldwide’, is an amorphous bunch of readers. African students are relatively easy to characterise as they acquire one or more African languages well as they grow up, and come to linguistics later, during their university studies. Outside Africa, undergraduates learn African languages as adults but while some learn an African language before studying linguistics, for others it is the other way round. So ‘undergraduates’ approach this book with different mixes of languages and linguistics. They will also come with differing linguistic approaches. Some will come with ‘traditional’ theoretical assumptions, others will come with older versions of ‘modern’ linguistics (often reflecting the unavailability of up-to-date texts in much of Africa), and others again will come equipped with different versions of contemporary theory.

### 1.3 **Contents of the volume**

To deal with this range of background and interest, we have in practice divided the book in three. The first four chapters deal with the four African phyla, Niger-Congo (Williamson and Blench), Nilo-Saharan (Bender), Afroasiatic (Hayward), and Khoisan (Güldemann and Vossen). Readers should understand that space limitations mean that no language receives more than a few words of coverage, and many are not mentioned at all (for a complete list, see Grimes (ed.) 1996). Our aim is the general, not the specific. A further four chapters (chs. 6–9) examine certain linguistic components of African languages. Our aim is to produce a text that is linguistically informed but not about linguistic theory. After discussion, we narrowed our range of topics to: phonology (including phonetics; Clements), morphology (Dimmendaal), syntax (Watters), and typology (Creissels). A third part consists of two chapters (chs. 10 and 11), on comparative linguistics (Newman) and language and history (Ehret). These are included because, while the proportion of professional linguists interested in this topic has diminished over the past quarter-century, we felt our audience would certainly be interested in it, and also because linguistic findings have contributed significantly to our understanding of recent African history. Finally, no book of this type would be complete without reference to language in its social setting, hence the chapter ‘Language and Society’. While we recognise the limitations of this range of topics, we feel it best corresponds to the likely knowledge and interests of our audience. Some of the more important omissions are covered briefly at the end of this introduction. All the chapters are written by scholars who are specialists in their areas. With a couple of exceptions, the authors have been in their fields for at least twenty years, and in some cases, for thirty or more.

The first four chapters share certain design features. Each contains: a survey of the main branches and languages of the phylum; a statement of the current evidence for, and status of, the phylum; a brief history of work on the phylum; and reconstruction of some features of the language assumed to be ancestral to the members of the phylum. These features are above all lexical and morphological, and only in some cases phonological. We use the term ‘phylum’ to refer to a language grouping larger, less well defined, and less widely accepted than a ‘family’, and which typically contains several families. Readers should note that while this distinction of phylum versus family is common usage today, it is not necessarily followed by all the authors in the current volume. The presentation of the kinds of evidence used to support each phylum has been sharpened by Dixon (1997), which casts doubt on the genetic validity of African phyla in general.

Readers will also note differences among these four chapters, differences which reflect partly the different genetic statuses of the phyla, partly the quantity



of scholarship that has gone into them, and partly the authors' personal choices in what they consider the essential characteristic features of the phylum. The least secure of the four is Khoisan. Is Khoisan a language phylum, or is it a collection of languages that have grown together over tens of thousands of years, and thus share certain typological features? The great age of Khoisan has led to a massive loss of shared material, if you believe in the genetic unity of Khoisan, or a massive convergence, if you do not. Nilo-Saharan, in more or less its present shape, was first proposed by Greenberg forty years ago, using evidence which many considered provisional. In the meantime a dedicated group of scientists have worked hard at collecting the kinds of evidence needed to substantiate or modify Greenberg's proposals. Although a grouping akin to Niger-Congo has been acknowledged for longer, since the nineteenth century, and although most Africanists recognise the validity of Niger-Congo, work here has been beset by the same problems as affect the other phyla, and especially by its having more members than any other phylum in the world. Of the four, Afroasiatic is the most widely recognised, the best analysed, and has the longest history of scholarship carried out by the largest number of scholars. However, most of this activity has gone into a few large languages in the Middle East, while the majority, that is, its African members, have suffered from relative neglect.

The search for the truth in all four phyla is slowed down by the same factors. These are: the small size of the group of scholars who worked on them; the many languages involved; the poor documentation for most; the long-standing interaction between adjacent languages; and, in the second half of the twentieth century, the disappearance of some languages. We will not presume to estimate the numbers of scholars involved in work on the various African phyla or even families because discussion suggests that there are too many variables and much disagreement.

The quality and quantity of the documentation for African languages ranges from fairly high to nil. We say 'fairly high' because no African language has been documented or analysed to the extent of the better researched European or Asian languages. If we define 'fairly high' as having a reasonably accurate and comprehensive reference grammar available, then less than a hundred African languages are in this category. For most, the documentation consists of an inadequate grammar, an analysis of part of the language, an article or two. For yet others, all we have is a reliable word list, or less than that.

A particular problem that affects work on comparative linguistics, that is not always acknowledged, results from having many languages adjacent over a long period of time. We now know that most linguistic features, not vocabulary alone, can be transferred from one language to another, or to several. Unique sounds such as the clicks of southern Bantu, word order in some Ethiopian languages,

predominantly suffixing morphology becoming predominantly prefixing, these and many other features can all be diffused across language and phylum boundaries. Areas are recognised elsewhere in the world, where long-standing interaction between settled communities has led to such an areal mixing of features that it is often difficult to distinguish their point of origin. Outside Africa, the Balkans are one such *Sprachbund*, the Indian subcontinent another. In Africa, such areas exist in Ethiopia, where speakers of different branches of Afroasiatic came together, in highland parts of East Africa, where speakers of all four phyla interacted historically, in southern Africa, where Bantu and Khoisan met and mixed, apparently in two discrete areas and times, and the broad region south of Lake Chad, where Central Sudanic, Chadic, and Adamawa-Ubangian speakers mixed. In fact, the vast number of African languages combined with the small size of many and the fact that many current languages, or rather their ancestors, have been in place for millennia means there has been interaction in many other places, too, and much transfer of inherited features. It is often hard to know where some of the features started. Readers will note that in some places in the first four chapters, authors will say that such and such a feature defines a language phylum or family but is also found in some neighbouring group(s).

A final phenomenon that has not affected comparative work much yet but is likely to do so in the twenty-first century is language death. Social, political and economic pressures are already conspiring to eliminate the languages of smaller communities, and the pressures will increase. In Africa and elsewhere these languages are being eliminated because they have lost their function and it is in no one's interest to maintain them. Almost no one's interest, that is, because for comparative and historical linguists, data from Friesian, on the north-west coast of Europe, may be just as valuable for language family work as is data from its much larger and better-known sibling, English. In the same way, data from Dahalo, a dying Cushitic language with clicks in north-east Kenya, from the isolated and small West African languages mentioned in chapter 2, or from Kwadi, a Khoisan language of south-western Angola mentioned in chapter 5, are of great importance to some linguists. Worldwide, there is a growing awareness of the need to document endangered languages. In Africa alone, more than a hundred languages are seriously endangered (see Brenzinger ed. 1992, 1998). We hope that this book will also contribute to stimulate interest in the study of these languages, some of which will no longer be there in a few decades.

Despite all that has just been said, at the start of the twenty-first century, we are very much further ahead than we were a century ago, both in terms of African language data and in terms of the syntheses made from it. The linguistic progress made in the twentieth century can be better seen in the second part of the book,

in the chapters dealing with specific linguistic topics. A century ago such overview chapters would have been impossible since all we had was analyses of a few African languages, mainly done by missionaries, mainly using Latin-based models. We have moved from missionary work through analyses meant to provide teaching materials to sophisticated analyses done by professional linguists. Some of the latter are so technical that it is part of the job of the authors in our second part to explain it to the worldwide undergraduate. At the same time, theory has changed. From the growing recognition in the first half of the twentieth century that the world was full of languages whose design features were not those of Latin or Western European languages, emerged new linguistic theories and insights that transformed the linguistic landscape in the second half.

Each of the authors of the four chapters on phonology, morphology, syntax and typology was faced with an unenviable task: how to reduce to twenty or thirty pages the significant features of two thousand languages and express them in terms which are at once professional, yet easily understood by the worldwide undergraduate? That was perhaps easiest for Clements. As he himself says, most contemporary phonology emerged from the research perspective of generative grammar, and many of the theoretical advances of the last thirty years have been made by phonologists who were also Africanists. For them, theory and African languages are inseparably connected. This was less true of syntax and typology. Chapters 8 and 9 have two roots. On the one hand they owe much to the perspectives of generative grammar, on the other hand they also partly rest on the typological impetus provided by Greenberg in the 1960s. Language typology arose in the nineteenth century and was originally based on the notion that languages could be reduced to a small set of morphologically based types. This was increasingly replaced by Greenberg's (1963b) approach, which was based on the study of word order and syntactic patterns. Syntactic patterns came to subsume much of morphology, and this in turn fitted well with generative perceptions: both the chapters in this book have to do with morphosyntax, how syntax and morphology meld to produce sentences and express meaning.

Despite this relegation of morphology to being a subcomponent of morphosyntax (and of phonology), we think it important to include a separate chapter on morphology. One reason is that the undergraduate audience will be regularly faced with words that make better sense if segmented morphologically. Another is that there are clearly recognised links between phonology and morphology and between syntax and morphology. A third is that African languages show a great range of morphological patterns. They are rich in noun-class and verbal tense-aspect systems. Undergraduates need insight into grammatical categories and how they are expressed.

We also feel it important to include the chapters on historical and comparative linguistics. The tools for building African history are different from those for Europe. African history has no long written tradition and thus resources such as archaeology and linguistics become more important. Linguistics has provided new insights into the history of, especially, the last three millennia in Africa, and Ehret has been associated with many of these. At the same time we feel it is proper procedure to have the methods of historical and comparative linguistics evaluated separately from their results, hence Newman's chapter. These methods are also not always clear to non-linguists and non-Africanists.

The author of the last chapter has to try to deal with the daily interaction of 750 million people speaking some 2,000 languages. Some things do not much change. Thus Africa has, and had long had, many bi- or multi-linguals. It has older koinés, pidgins, and creoles, all used to facilitate communication among different linguistic populations. It has communities giving up their traditional form of speech for others. At the same time the language situation in Africa is changing rapidly. Some languages themselves are changing, some are disappearing, new languages are arising. People are flocking into cities. Countries and leaders are struggling to formulate new language strategies, in situations where languages are often not viewed just as languages but as an integral part of people's culture, which they vigorously defend. This all makes for an exciting situation, but one which is hard to encapsulate adequately.

#### 1.4 **Further issues**

One of the most difficult tasks we have to face is how to reduce the multitude of important topics that have been raised in the history of African linguistics to a manageable set. The various authors in this volume have tried to take care of many of them, and have provided suggestions for readers who want to know more about a particular topic. In some cases these suggestions take the form of a suggested further reading list at the end of the chapter, but more often there are references liberally scattered through the text, which are up-to-date or standard works, which in turn mention older works. Inevitably there remain some subjects that could only be mentioned in passing or could not be covered at all.

We would have liked, for example, to expose the reader to the whole gamut of scholarly discussion on the description and classification of African languages. In a volume that aims at presenting what is widely or commonly accepted, this is possible only within limits. Nevertheless, not infrequently authors express contrasting views on the same subject and this is reflected in this volume. It is perhaps most obvious in the reconstruction of Nilo-Saharan, the subject of controversy ever since Greenberg first proposed it as a genetic grouping. The reader may participate

in this discussion in chapters 3 (Bender) and 11 (Ehret), where the two main alternative hypotheses are presented, each with considerable implications for our understanding of Africa's prehistory.

Work on classifying African languages has focused mainly on problems of genetic relationship. Accordingly, the classification most widely accepted is genetic, as is apparent in the present volume, where chapters 2 to 5 each treat one of the four African phyla proposed by Greenberg (1963a). Compared to that work, other approaches to language classification have received much less scholarly treatment. Areal, that is, contact-induced linguistic relationship in particular, has been out of vogue for a long time and has only very recently started to receive the kind of detailed attention it deserves. Conceivably, some of the classifications proposed could be more profitably analysed with reference to areal rather than genetic relationship. Earlier we drew attention to the problems in establishing Khoisan, the subject of chapter 5, as a phylum. It may well turn out that Khoisan could be more appropriately defined as a convergence area rather than as a genetic unit.

While the study of African-language structures can be divided fairly well into the four domains highlighted in this volume, namely phonology, morphology, syntax and typology, we have not been able to treat other topics in the way they merit. Thus, for instance, some colleagues wanted us to devote space to linguistic aspects of the diaspora, to African languages in the Americas. We settled for asking John McWhorter to write a brief overview for us, as follows, and we are indebted to him for this.

The principal fate of African languages in the New World has been to serve as primary sources for the creoles which slaves developed in plantation colonies. Often speaking closely related languages while having minimal contact with whites, early slaves' transfer-laden approximations of a given European language conventionalised into new languages, African-derived as much as European. The most extreme manifestations are Surinam creoles, whose syntaxes are broad reproductions of Kwa ones: the Saramaccan *di nákináki dágú bi wáka gó a wósu báka* 'the beaten dog walked behind the house' superimposes English lexicon on Kwa features such as a reduplicated attributive adjective, verb serialisation, and a postposed nominal as spatial deictic. Other creoles include Gullah, Haitian, Papiamentu and the extinct Negerhollands.

Lighter African influence can be seen in the speech of many Afro-Hispanics, whose speech diverges slightly from local Spanish varieties in features such as a double negator pattern (*no lo tengo no* 'I don't

have it') found in Kongo. Popular Brazilian Portuguese is similar, in idiom calques such as *o dia ta limpo* 'The day is clean' for 'It's dawn', an expression found also in Yoruba. These New World Iberian varieties, as well as creoles in general, also preserve many West African lexical borrowings.

Evidence suggests that African slaves did not usually transmit their native languages to following generations. A notable exception is in Brazil, where Fon, Kongo, and Yoruba were maintained by communities of blacks, the latter into the twentieth century. More typically, African languages were preserved in fossilised ritual registers often kept today, such as Twi and Gbe in Jamaica; these two and Kongo in Surinam, and Mende in the Sea Islands of South Carolina.

The editors would also have wished to include a separate treatment of semantics. Most of what determines how African languages are used and structured relates to how meanings are expressed. The study of meaning has been approached in different ways, using contrasting theoretical frameworks, and quite a number of semantic characteristics of African languages have been identified. But we lack a more cohesive view of how the different scholarly traditions dealing with meaning and the many details we have on the semantic characteristics of individual African languages can be presented as a book chapter.

A related matter covers discourse structure, text analysis and forms of creative language use. African narrative discourse and conversation structure are fascinating fields of research and they have been approached variously by syntacticians, anthropologists, translators, literary scholars and others.

A final issue concerns new directions in linguistic research paradigms. Grammaticalisation theory, for instance, has yielded fresh insights on how grammatical forms arise and develop, how the boundary between the grammar and the lexicon should be studied, and how to explain why grammar is structured as it is. African languages have figured prominently in formulating principles of grammatical evolution and in proposing new parameters of linguistic explanation (see Heine and Reh 1984, and Heine, Claudi, and Hünemeyer 1991 for details).

It is hardly possible to name all those who have contributed to this volume in some way or another. To you all we say: thank you. Our immediate gratitude is to Erhard Voeltz, for taking the time and trouble to make valuable comments on various chapters, to Monika Feinen for her work on the maps, and to Yvonne Treis for assisting in the editorial work.