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 been 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 (*Global-kopieren*); 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 (*plural*)	'Pokomo person'
(2)	Ilwana (Sabaki, Bantu; Nurse 2000b: 125)		
	mo-bokomo	bokomo (*plural*)	'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'
yandɛ́	yand(ɛ́), 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 *nɪ*) benefactive, *e* locative, and *gɪ* 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.[1]

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 the these orders, namely SVO and SOV.[2]

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[3] (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