14

Convergence and Divergence in the Development of African Languages

Bernd Heine and Tania Kuteva

In the present chapter a bird's-eye view of some problems associated with contributions to the linguistic history (or prehistory, as some would say) of Africa is offered. It has been stimulated by recent contributions on linguistic methodology, especially by attempts to relate linguistic findings to more general observations on the evolution of the human species. The conclusion reached is that contact-induced language change and the implications it has for language classification in Africa are still largely a *terra incognita*.

1. Introduction

For roughly half a century, work on the reconstruction of African languages and their relationship has been based on the work of Joseph Greenberg (1949, 1955, 1963). What this work has established in particular are findings such as the following:

(a) The most easily accessible way of describing the historical relationship of these languages is by reconstructing their genetic relationship patterns.
(b) The multitude of African languages can be reduced to four genetically defined units, called families by Greenberg and phyla by others. These units are Niger-Congo (or Niger-Kordofanian), Nilo-Saharan, Afroasiatic, and Khoisan.
(c) There are various methods available to the linguist for historical reconstruction. The task of the linguist is to choose that method that appears to be best suited to solve a particular problem. Which method is most suitable in a given

The present paper is based on research that has been generously funded by the Deutsche Forschungsgemeinschaft (German Research Society) to which we want to express our deeply felt gratitude. The paper has benefited greatly from discussions with and critical comments by colleagues at the Institut für Afrikanistik of the University of Cologne. We wish to express our gratitude most of all to Erhard Voeltz, Ulrike Claudi, Marianne Bechhaus-Gerst, and Friederike Lüpke for many suggestions. We are also grateful to Sasha Aikhenvald, Lionel Bender, Hilary Chappell, Gerrit Dimmendaal, Bob Dixon, Geoffrey Haig, Malcolm Ross, and other contributors to the present volume for valuable comments.

situation depends primarily on the time factor involved. Historical processes that happened within the last century require different methods of analysis from processes dating back a thousand or two thousand years.

(d) Most of the work to establish language phyla in Africa, including that of Greenberg (1963), has used the method of resemblances, which is based on the assumption that in order to establish that two or more languages are genetically related, or to determine the degree to which they are related, one simply needs to demonstrate that these languages share a sufficient number of lexical (and/or grammatical) items that are similar in form and meaning. The main problems associated with this method concern the question of how the notions 'sufficient number' and 'similarity in form and meaning' can be defined.

(e) On account of such problems, many students of African linguistics consider this method to be of doubtful value, and some would reject it altogether, arguing that reliable reconstructions of genetic-relationship patterns can only be achieved by means of the comparative method. According to Nichols (1992: 2), this method works reliably only up to a time depth of roughly 8,000 years. So far, it has not been possible to apply the comparative method appropriately to any of the four African language phyla.

Greenberg's genetic classification of African languages is by now widely accepted, but it also leaves many questions on the prehistory of Africa unanswered. Reconstructing family trees is helpful to define one kind of historical process, but it contributes little to our understanding of what has happened in Africa for example in terms of linguistic interaction across languages. With the present chapter we wish to draw attention to the need that exists to study language contact and the ways it may be relevant to linguistic classification in Africa.

2. Language contact

While there are a number of studies on how African languages influence one another, we know little about how this affects linguistic relationship. Still, there are a few studies that suggest that areal forces and linguistic relationship based on contact between languages may cut across genetic boundaries, and a number of convergence areas (or areal groups) have been identified.

2.1. AREAL LINGUISTICS

To start with, there is some evidence to suggest that the African continent forms a convergence area of its own. There are very few linguistic properties that are found almost only in Africa, like click consonants, which occur in all southern African and East African Khoisan languages, in many Bantu languages of southern Africa, and in Dahalo, a Cushitic language of eastern Kenya. But there are a number of features that are widespread in Africa but less common, or uncommon, outside

Africa. Phonetic features are in particular the presence of labiovelar stops (*kp* and *gb*), of implosive stops (*ɓ*, *ɗ*, *ʄ*), of prenasalized stops (*mb*, *nd*, *ŋg*), or of vowel harmony based on the tongue root (+/– advanced tongue root position). Among morphological characteristics one finds the widespread occurrence of a set of verbal derivative extensions expressing grammatical functions such as passive, causative, applicative/benefactive, and reciprocal. One might also mention the presence of noun-class systems, based on the distinction human vs. non-human or animate vs. inanimate (rather than masculine vs. feminine) and distinguishing a larger number of classes, but the occurrence of such systems appears to be genetically determined: they are common in the Niger-Congo family but essentially absent elsewhere.[1] A negative areal feature can be seen in the nearly complete absence of ergative languages in Africa. Semantic features characterizing the African continent are certain polysemies of nouns and verbs. For example, the noun for 'wild animal' also denotes 'meat', and the verb for 'eat' has 'conquer' and 'have sexual intercourse with' as additional meanings in many African languages; see Greenberg (1959: 23), Gilman (1986) for details.[2]

But also within Africa, some convergence areas have been identified (see e.g. Greenberg 1959: 24–5). The most frequently mentioned example is north-eastern Africa. Within roughly the last two millennia, the highlands of Ethiopia appear to have favoured cultural and linguistic exchange on a massive scale, with the effect that the languages of this region now share a number of linguistic properties (Ferguson 1976).

The Kalahari basin of southern Africa appears to form another convergence area; it provides an instance of a refuge area where people have been living over centuries and probably millennia without much interference from outside. It is the homeland of the Khoisan-speaking Bushmen or San peoples. Lewis-Williams (1984) suggests that there has been ideological continuity in San culture for at least two millennia and possibly for as long as 26,000 years, where ideological continuity implies some degree of continuity in social relations.[3] As Güldemann (1997) argues, the Kalahari basin convergence area is not confined to languages conventionally classified as belonging to the Khoisan phylum; rather, it also includes a Bantu language, Tswana (Güldemann 1997).

One linguistic domain that appears to be particularly prone to contact-induced change is word order, more precisely the arrangement of main clause

[1] An example of such a noun-class system outside Niger-Kordofanian can be found in !Xun (Ju|'hoasi), a North Khoisan language of Namibia, Botswana, and southern Angola. What distinguishes this system from canonical Niger-Congo systems is in particular that in !Xun, nouns are not overtly marked for gender (cf. Heine 1981: 210).

[2] Concerning more areal properties of African languages, see Gilman (1972, 1986), Creissels (2000).

[3] Lewis-Williams (1984) suggests in particular that the San trance dance strengthens kinship relationships, which in turn structure the aggregation and dispersal necessary to distribute people over available resources. Whatever the significance of such suggestions may be, we have to be aware that they are not based on any historically relevant data and therefore have to be approached with care.

constituents. Based on a survey of the order of meaningful elements in African languages, Heine (1976) concludes that there are a number of linguistically defined areas cutting across boundaries of language families. One such area consists of a large part of West Africa where Mande, Gur (Voltaic), and western Kwa languages are spoken. In addition to these languages, which are traditionally classified as Niger-Congo, this area also includes Songhai, a language usually classified as belonging to the Nilo-Saharan phylum. What characterizes this area most of all is the presence of a possessor-possessee word-order syntax which is not confined to the noun phrase but has also affected the structure of the clause (see Claudi 1993). Another area, called the Rift Valley Convergence Area, is defined by the presence of verb-initial (VSO) syntax, very rarely encountered elsewhere in Africa.[4] The languages of this East African area belong to Greenberg's (1963) Nilo-Saharan (Surma, Kuliak, Eastern Nilotic, and Southern Nilotic) and Khoisan families (Hadza).

What these studies suggest is, first, that previous research, like that summarized in §1, has relied too heavily on discovering genetic-relationship patterns, ignoring the fact that investigating areal relationship provides a complementary—and equally rewarding—approach to reconstructing Africa's linguistic history. Second, they also suggest that what constitutes areal relationship is still largely unclear. Terms such as 'linguistic area', 'areal group', or 'convergence area' are notoriously fuzzy; as a rule, using them is tantamount to claiming that there is a set of linguistic properties exhibiting an areal distribution that cannot be reconciled with what we know about the genetic relationship of the languages concerned and that the most reasonable explanation therefore is contact-induced relationship.Third, these studies also suggest that we still know very little about the overall situation of areal relationship in Africa and, perhaps more importantly, that we still lack adequate methods and models for describing this kind of relationship. Still, whether, or to what extent, existing models capture salient characteristics of convergence areas remains unclear considering our as yet largely inadequate empirical knowledge of language contact and its implications for language classification. Fourth, and consequently, what we need most urgently is a more detailed account of what happens when languages, or more exactly, when speakers of different languages, are in contact.

The most likely consequence of such situations is lexical borrowing. There are quite a number of studies that describe how one African language has borrowed part of its vocabulary from another language. As a rule, nouns account for by far the largest part of borrowed material, followed by verbs, interjections, and conjunctions, with affixal morphology being much less likely to be affected by language contact. While this is the expected case, there are nevertheless examples

[4] The only other VSO-languages reported so far are the Berber languages of north-western Africa, a few Chadic languages, and Krongo, a Kordofan language nowadays considered to belong to the Kadu branch of Nilo-Saharan.

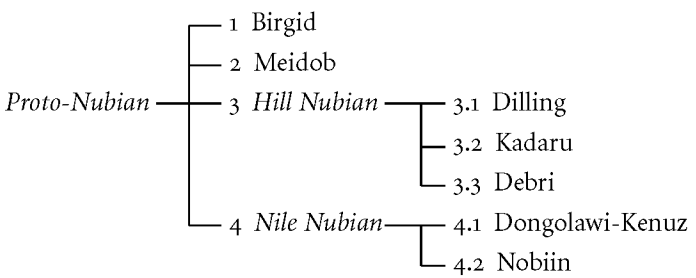
to suggest that the lexicon need not be the primary domain of contact-induced change. The following case study involving the Nile Nubian languages of Egypt and Sudan provides such an example.

2.2. NILE NUBIAN

Prior to the construction of the Aswan dam, the Nile Nubian languages were spoken by some 200,000 to 400,000 people along the Nile River between Aswan in the north and Old Dongola in the south in southern Egypt and the northern end of the Republic of Sudan (see Werner 1987: 29–30). Four Nile Nubian dialects tend to be distinguished in the relevant literature: Kenuz (Kenzi, Kunuzi), Fadijja (Fadidja, Fadičča), Mahasi (Mahas), and Dongolawi (Dongola).[5] Following Bechhaus-Gerst (1984), on which the present account is based, Kenuz and Dongolawi are treated as one group,[6] referred to as Dongolawi-Kenuz, and Mahasi and Fadijja are also grouped together under the term Nobiin. Both groups show a close relationship: their phonological systems are said to be identical, and their morphological and lexical inventories are very similar; a lexicostatistic count yielded 70% of cognates between the two groups.

Nile Nubian shows genetic relationship with the following language groups: the Hill Nubian languages and dialects of Kordofan, such as Debri, Kadaru, and Dilling, and the Birgid and Meidob languages spoken in Darfur, over five hundred kilometres away from Nile Nubian (see Map). On the basis of lexicostatistic counts, Nubian has been classified as described in (1).

(1) A lexicostatistic classification of Nubian (Bechhaus-Gerst 1984: 17; groups are italicized)



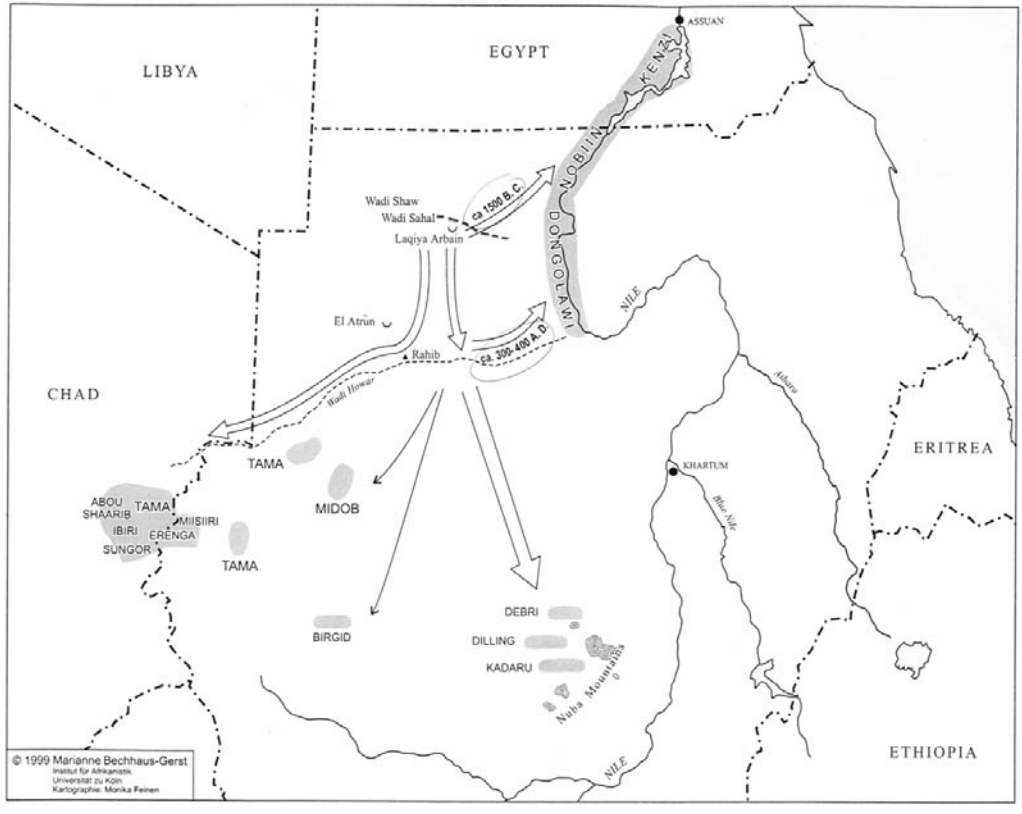
On the basis of such lexicostatistic data, Thelwall (1982) concludes that Dongolawi and Nobiin[7] represent a most recent genetic split within Nubian.[8] More detailed

[5] This four-fold distinction is not shared in every detail by all authors who have written on the subject; virtually every author has come up with a different description of Nile Nubian dialects or languages.

[6] The split of Kenuz and Dongolawi into different dialects appears to be a very recent one (cf. Werner 1987: 28).

[7] Thelwall (1982) does not consider Kenuz in his calculations.

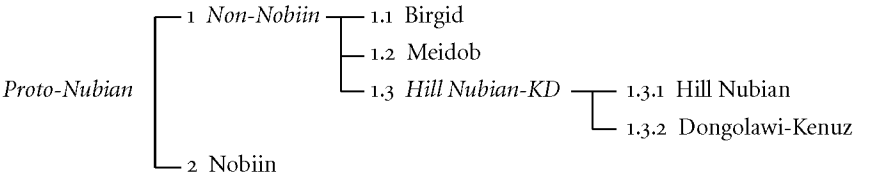
[8] Thelwall's (1982) classification differs also in few other details from that of Bechhaus-Gerst, e.g. in his claim that the first split of Proto-Nubian led to a separation of Meidob from the rest of Nubian.



MAP. Reconstructed migrations of the Nubian and Tama peoples from their presumed homeland in the Wadi Shaw/ Laqiya region

work based on linguistic, archaeological, and Egyptological evidence suggests, however, that (1) does not reflect the actual genetic relationship pattern holding between these languages; rather (2) is a much more appropriate tree diagram:

(2) The genetic classification of Nubian (Bechhaus-Gerst 1984: 121; groups are italicized)



The lexicostatistic tree is at variance with the 'genetic' tree in two ways in particular:

(a) It suggests that Nile Nubian is a genetic unit, while a more comprehensive analysis shows that the two subgroups of Nile Nubian (Kenuz-Dongolawi and Nobiin) belong to different primary branchings of Proto-Nubian.
(b) It does not represent Nobiin as splitting off from Proto-Nubian before all other languages did.

In general, lexicostatistics has turned out to be a fairly reliable tool for establishing first hypotheses on genetic relationship in Africa: in most cases where the comparative method and lexicostatistics have been employed, they yielded similar results (see below). The question is: how is it possible that in the case of Nubian there is such a divergence between the two tree diagrams?

On the basis of combined linguistic, archaeological, and other evidence, Bechhaus-Gerst (1984) volunteers the following answer (see also Thelwall 1982: 32). The Nubian languages were 'originally' spoken in the dry regions of Kordofan and Darfur nearly five hundred kilometres west of the Nile. Roughly three millennia ago, speakers of what is nowadays referred to as Nobiin migrated to the Nile and settled there, adopting a riverine economy and culture.

About one thousand years later, another group of Nubians, now represented by speakers of Kenuz-Dongolawi, also migrated to the Nile via the Bayuda Steppe where they met the Nobiin. With the collapse of the Meroitic empire in the fourth century AD at the latest, Nobiin speakers became the dominant power along that part of the Nile. Their language became a written language used in church and in trade, commonly known as Old Nobiin (Bechhaus-Gerst 1996: 298, 304). The result was heavy borrowing, which was largely though not entirely unilateral: the high-prestige language Nobiin was the main donor, but there were also Kenuz-Dongolawi loans that entered Nobiin. While having stayed separated for no less than a thousand years, Kenuz-Dongolawi and Nobiin became more and more similar, to the extent that they have almost become dialects.[9]

[9] Werner (1987: 24) insists that Kenuz-Dongolawi and Nobiin are not mutually intelligible and hence proposes to consider them as different languages.

The separation of the Nobiin-speaking and the Dongolawi-Kenuz-speaking people was so long that present-day descendants of both groups do not share a common Nubian identity; traditionally, the Nubians (Nobiin) call Dongolawi-Kenuz *oʃkiriin baŋŋid* 'language of the slaves' or *bideriin baŋŋid* 'language of the poor'. Both groups also differ in their traditions about their origin. The Nobiin claim to be the only genuine Nubians of African origin, while the Dongolawi-Kenuz believe they are descendants of immigrants from the Arabian peninsula (Bechhaus-Gerst 1996: 298).

The case of Nile Nubian is also of interest with regard to principles of language classification. Genetic relationship is widely assumed to be based on the family-tree model, even if there are a few examples that seem to challenge such an assumption (Thomason and Kaufman 1988). Nile Nubian offers another case of a challenge. There is some evidence to suggest that Dongolawi-Kenuz is a 'hybrid' language between Old Nobiin and pre-contact Dongolawi. This evidence is of the following kind (see Bechhaus-Gerst 1996: 305 ff.):

(a) PHONOLOGY. Dongolawi-Kenuz has borrowed almost its entire phonological system from Nobiin, even if there was also some borrowing in the reverse direction.[10]

(b) MORPHOLOGY. Dongolawi-Kenuz has borrowed much of its morphology from Nobiin, in particular the following items:[11]
- (i) the postpositions *bokon* 'until' and *takki* 'when'
- (ii) demonstrative pronouns
- (iii) interrogative pronouns
- (iv) the plural suffix *-gu* with pronouns
- (v) the plural suffix *-ri* (in loanwords only)
- (vi) the suffix *-ke(n)* marking habitual aspect (Kenuz only)
- (vii) the plural object suffixes with the verb *den-/tir-* 'give'
- (viii) verbal suffixes for the resultative/perfective
- (ix) verbal suffixes for the stative
- (x) verbal suffixes for the durative
- (xi) verbal prefixes for the durative/habitual
- (xii) verbal prefixes for the intentional/ingressive (future).

Conversely, the influence of Dongolawi-Kenuz on Nobiin was rather restricted. Bechhaus-Gerst (1996: 306) finds the following elements of

[10] Nobiin has borrowed the phoneme /b/ through late loanwords (Bechhaus-Gerst 1996: 306). Concerning the techniques used to determine directions of borrowing, see Bechhaus-Gerst (1996).

[11] Bechhaus-Gerst (1996) gives several types of arguments why these morphological forms are borrowed rather than retained. That morphemes might be adopted into one language from another language, provided that the two languages are significantly similar typologically, has already been observed in other language-contact situations in other geographical areas (cf. the borrowing of Bulgarian inflectional verb endings into Meglenite Romanian; Sandfeld 1938: 59). It remains unclear, however, whether, or to what extent, the borrowed morphology has replaced previously existing grammatical categories.

Dongolawi or Kenuz origin in Nobiin: the formative suffix for ordinal numerals, a second set of personal pronouns based on plural pronouns, the suffix *-ndi* (> *-ni*) with possessive pronouns, and the nominal plural suffixes *-ii* and *-nci* in loanwords.

(c) LEXICON. Even though loanwords and loan translations from Nobiin into Dongolawi-Kenuz can be found, the latter appears to have retained much of its own vocabulary.[12] There is no evidence of massive lexical borrowing from Nobiin in Dongolawi-Kenuz.[13]

Modern Dongolawi-Kenuz is not just a later historical state of one language, rather it is the 'daughter' of both pre-contact Dongolawi-Kenuz and Nobiin. Whether the process underlying this situation requires specific circumstances to happen, like the presence of a close, or even a common, genetic link between the languages concerned (Jeffrey Heath, p.c.), requires further investigation. What is obvious from the data available is that we are dealing with an instance of convergence, not towards a common prototype (see Dixon 1997), but rather of one language towards another, i.e. Dongolawi-Kenuz to Nobiin. To conclude, we are faced

(a) with the emergence of a new language, modern Dongolawi-Kenuz, whose genetic position can no longer be described unambiguously in terms of a tree-diagram model, and
(b) with the continuation of another language, Nobiin, which is slightly 'changed' due to borrowing.

3. Grammaticalizing metatypy

Cases like Nile Nubian do not seem to be very common in Africa. What this case suggests however is, first, that until now such situations of intensive language contact have not received the kind of attention they deserve. Second, that we still know very little about the various kinds of sociolinguistic settings that may be present in situations of intensive language contact, and how each setting affects language structure. And third, that previous research on contact-induced language change has focused primarily on lexical, phonological, and morphological interference. What we lack most of all is more information about how language contact affects meaning and the arrangement of meaningful elements in

[12] That areal influence can strongly affect grammar but spare the lexicon is noteworthy but not entirely uncommon. Aikhenvald (this volume) observes that the Vaupés linguistic area of north-west Amazonia is characterized by the presence of a number of common grammatical features while lexical borrowing is absent.

[13] The discovery of lexical items turns out to be a difficult exercise since Dongolawi-Kenuz has been assimilated phonologically to Nobiin, which means that there are hardly any phonological clues which would be of help when deciding whether a given lexical correspondence is due to borrowing or to common inheritance (cf. Bechhaus-Gerst 1996: 186).

discourse. Apart from a few, largely impressionistic observations, not much is known about how meaning and meaningful structures behave in language-contact situations.

Examples of meaning-transfer from language to language, irrespective of the form this meaning may take in a given language, are more common in Africa than is commonly believed; they are instances of what Ross (1996, 1997, Chapter 6) calls *metatypy*. Note that metatypy is not confined to lexical semantics. It may involve entire constructions or predications and, since it has to do with the combination of meaningful elements, it also has a syntactic component. Underlying metatypy there appears to be a strategy whereby speakers aim to adapt their ways of saying things to those of the target language by reorganizing their expressions of meaning (semantics) and the way meaningful elements are arranged (syntax). Adaptation affects in particular the following domains of language structure:

(a) the range of meanings expressed by a given word or phrase,
(b) the patterns of syntactic encoding, and
(c) the nature of idiomatic expression.

Metatypy has been treated traditionally as calquing. The difference between calquing and metatypy is actually one of degree rather than kind. While the term *calquing* tends to be used for the 'translating' of lexical items, referring to what happens to individual words or groups of words, *metatypy* captures 'loan translation' on a larger scale, relating to more general patterns of linguistic expression; it leads essentially to a change in structural type. No attempt is made here to trace a boundary between the two (for a discussion of the differences between calquing and metatypy, see Ross, this volume). Following Ross (1997: 241) we will assume that underlying metatypy there is a strategy employed by speakers to reduce their cognitive and linguistic-processing burden by bringing their construal of reality into line with that of speakers of another language, what we may term the target language. The result is that the languages concerned become more readily intertranslatable. Note that in metatypy the form, i.e. the phonological substance used to encode meaning, remains unaffected.

Metatypy can be held responsible for a variety of new structures of language use, for example new conventionalized expressions, phrases, idioms, proverbs, and patterns of syntactic encoding. In addition, there appears to be one type of metatypy that leads to the emergence of new grammatical categories: we will refer to this type as grammaticalizing metatypy.

Grammaticalization has been described as a process leading from lexical to grammatical and from grammatical to even more grammatical forms. While such a description has turned out to be useful, it tends to ignore that, more often than not, the relevant process is not confined to individual units such as morphemes or words; rather it involves the reinterpretation of more complex semantic structures as structures serving the expression of grammatical functions. The following is a sketchy treatment of grammaticalizing metatypy, meant to illustrate the

potential this notion may have for understanding certain patterns of contact-induced language change.

There is a kind of semantic structure that tends to be used cross-linguistically for expressing grammatical functions. The term used for this structure is *event schema* (see Heine 1993, 1997a, 1997b for details). Event schemas present stereotyped situations with which we are constantly confronted; they are propositional in structure and take the form of simple predications describing what one does (Action), where one is (Location), who one is accompanied by (Companion), what exists (Existence), etc. There is only a limited pool of such schemas recruited for the expression of grammatical functions. The way event schemas can affect grammatical encoding may be illustrated with two examples. In §3.1 we will look at comparative constructions, while §3.2 will be devoted to reflexive markers.

3.1. COMPARATIVES

Our concern here is more narrowly with the way the standard of comparison is encoded in comparative constructions of inequality (also called superior comparatives). These are constructions having the form *X is Y-er than Z*, where *X* is the comparee (or item compared), *is Y-er* is the predicate, and *than Z* is the standard of comparison. There is only a handful of event schemas that tend to be recruited time and again in the languages of the world to express and grammaticalize this notion. Perhaps the most widespread schema is one in which the standard of comparison is presented by means of an ablative or locative source morphology, as in the following example:

(3) Yaaku (Eastern Cushitic; Afroasiatic)
kédén ké céin òù àì
tree COP big from house
'The tree is bigger than the house.'

Quite a different way of expressing the notion of a comparison of inequality is to establish a polar contrast between the comparee (*X*) and the standard of comparison (*Z*). Polarity may involve either antonymy (= presence of property p vs. presence of property q), as illustrated in (4a), or a negative–positive contrast (= presence vs. absence of property p), as in (4b).

(4) (a) Sika (Moluccan; Austronesian; Stassen 1985: 44)
dzarang tica gahar, dzarang rei kesik
horse that big horse this small
'That horse is bigger than this horse.'
(b) Hixkaryana (Carib; Stassen 1985: 44)
kaw-ohra naha Waraka, kaw naha Kaywerye
tall-not he.is Waraka tall he.is Kaywere
'Kaywerye is taller than Waraka.'

TABLE 1. The main event schemas used for encoding comparative constructions (see Heine 1997b: 112)

Form of schema	Label of schema
X is Y surpasses Z	Action
X is Y at Z	Location
X is Y from Z	Source
X is Y to X	Goal
X is Y, Z is not Y	Polarity[14]

TABLE 2. Event schemas serving as sources for the grammaticalization of comparatives of inequality (Sample: 109 languages of world-wide distribution; Stassen 1985, Heine 1997b: 128)[15]

Source Schema	Europe	Asia	Africa	The Americas	Indian/ Pacific Ocean	Total
Action	0	4	13	1	2	20
Location	0	4	3	4	1	12
Source	0	18	4	9	1	32
Goal	1	0	3	3	3	7
Polarity	0	0	0	10	10	20
TOTAL	14	26	23	28	18	109

These are two schemas commonly recruited to express the notion of a comparative of inequality; we will refer to them as the Source and the Polarity Schemas, respectively. But these are not the only schemas; the whole range of schemas most commonly employed cross-linguistically is summarized in Table 1.

In principle, speakers of a given language may select any of these schemas to develop a new comparative construction, and in many languages, more than one schema has been grammaticalized. It would seem, however, that there is one important factor that influences the choice of schemas, and this factor has to do with geography: neighbouring peoples are more likely to draw on the same schema for a specific purpose than peoples living at some distance from one another. The result is that there are geographically defined regions where a preference for a specific kind of grammaticalizing metatypy can be observed. Table 2

[14] There is reason to assume that the Polarity Schema differs from the other schemas in that it tends to be only weakly grammaticalized, if at all (Geoffrey Haig, p.c.)—to the extent that in some languages where comparative notions are expressed by means of polarity it remains unclear whether there is any justification for assuming that such expressions really have the status of a grammatical category. More research is needed on this issue.

[15] This table differs slightly from that presented in Heine (1997b: 128) in that Classical Arabic and Biblical Hebrew are treated here as 'Asian languages', which means that, instead of a category 'Africa and Middle East' we now have a category 'Africa', which includes only languages that have been spoken natively in Africa for at least one millennium.

summarizes the results of a cross-linguistic survey of these constructions. Note that the sample of 109 languages has been established on what Stassen (1985) argues is a genetically and areally balanced selection of the world's languages (see Stassen 1985 for details).

What the figures in Table 2 suggest is that the macro-areas distinguished are each characteristic of a particular choice of event schemas. In the European languages of the sample, the vast majority are characterized by what Stassen (1985) calls 'particle comparatives', that is, by constructions whose etymological source is opaque: 92% of all European languages of the sample, including English, have grammaticalized their major comparative construction to the extent that it is no longer possible to determine unambiguously the schema from which it is historically derived.[16] In Asian languages there is a clear preference for the Source Schema: more than two thirds (69%) of all sample languages spoken on the Asian continent make use of the Source Schema. What unites the Americas and the region of the Indian and Pacific Oceans again is the widespread grammaticalization of the Polarity Schema to a comparative construction; no sample language outside this general area has been found to have drawn on this schema.

Africa as a macro-area also exhibits a clear preference pattern: more than half of all African sample languages (57%) have grammaticalized the Action Schema to comparative constructions. But perhaps more significantly, almost two thirds (65%) of all languages of our world-wide sample having made use of this schema are spoken in Africa. There is some variation in the exact shape this schema may take, the main ones being either [*X is Y surpasses Z*], as in (5a), or [*X surpasses Z (at) Y-ness*], illustrated in (5b). What is common to all of them is that the standard of comparison is presented by means of a verb meaning 'surpass, defeat, exceed', and the like, that is, the comparee (X) surpasses the standard (Z) with reference to the quality in question (= the predicate Y).

(5) (a) Swahili (Bantu; Niger-Congo)
Nyumba yako ni kubwa kushinda yangu
house your be big to.defeat mine
'Your house is bigger than mine.'

(b) Hausa (Chadic; Afroasiatic; Wolff 1993: 221)
naa fi Muusaa wàayoo
I surpass Moses cleverness
'I am cleverer than Moses.'

This areal distribution sets the African continent apart from the rest of the world: one can predict with a certain degree of probablity that if one finds a language that expresses the notion of a comparative of inequality by means of the Action Schema then that is likely to be an African language.[17] Note that the Action

[16] But see Heine (1997b: 117–18) for possible etymologies.

[17] Geoffrey Haig (p.c.) points out, however, that the Action Schema is also very common in Papua New Guinea.

Schema is not confined to some particular language phylum or phyla in Africa; rather, its distribution cuts across genetic and regional boundaries.[18]

But there is also an areal patterning within Africa. While one can expect the Action Schema to have given rise to comparative constructions in any part of the continent, there is also a regional patterning. In Table 2 we find four instances of the Source Schema in Africa, and three of them involve Ethiopian languages: Amharic, an Ethio-Semitic language, and Beja and Bilin, both Cushitic languages.[19] A more detailed analysis suggests in fact that it is the Source Schema, rather than the Action Schema, which is the most common source for grammatical categories of a comparative of inequality in these language groups. This distribution might suggest that we are dealing with a genetic rather than an areal feature since Ethio-Semitic and Cushitic languages are both branches of Afroasiatic. But the Source Schema is also found in non-Afroasiatic languages of Ethiopia, e.g. in Kunama (where the ablative postposition or suffix *-kin* encodes the standard of comparison):

(6) Kunama (Nilo-Saharan; Böhm 1984: 94)
Marda- kin Kunama maida
Nera[20]- from Kunama be.noble
'A Kunama is more noble than a Nera'

These observations suggest that grammaticalizing metatypy provides yet another feature defining the Ethiopian highland region as a linguistic area: instead of the otherwise prevailing pattern of forming comparatives of inequality in Africa by means of the Action Schema, it is the Source Schema which is favoured in this region.[21]

3.2. REFLEXIVES

That the use of the Action Schema for expressing a comparative of inequality belongs to those features that characterize Africa as a linguistic area has already been mentioned by Greenberg (1959). Greenberg's examples also include that of reflexive marking: he observes that 'he himself' translates in African languages as 'he with his head' (1959: 23). This example relates to expressions for what tend to be referred to as emphatic reflexives, but one can generalize by saying that, in fact,

[18] As we noted above, many languages have grammaticalized more than one schema. With reference to example (5a), for example, one should mention that Swahili uses not only the Action Schema but also the Location Schema (having grammaticalized the form *kuliko* 'where there is' to a standard marker: see Heine 2000).

[19] The only African language outside the Ethiopian area is Nama, a Central Khoisan language (Stassen 1985: 40).

[20] Böhm uses the term 'Barea' instead of Nera; the former term is no longer considered by the speakers of this language to be appropriate, hence we have replaced it by 'Nera'.

[21] This observation might suggest that the Ethiopian area is an extension of the Asian 'macro-area' where the Source Schema is the clearly predominant one (see Table 2).

TABLE 3. Nominal sources for reflexive/reciprocal markers in African languages (Sample: 62 languages; for 25 of these, no nominal source could be found; see Heine 2000).

Nominal meaning	Number of occurrences	Percentages
body	20	51.2
head	9	23.1
owner	3	7.7
comrade	2	5.1
life	2	5.1
relative	1	2.6
soul	1	2.6
person	1	2.6
TOTAL	39	100.0

a number of African languages have grammaticalized reflexive pronouns (including markers for emphatic reflexives) which are etymologically derived from terms for the body-part 'head', as illustrated in the following example:

(7) Hausa (Chadic; Afro-Asiatic; Kraft and Kirk-Greene 1973: 231)
Sun kashè kân-sù
they kill head-their
'They have killed themselves.' (i.e. 'they have committed suicide')

In examples such as (7) we are dealing with a propositional schema where an object noun phrase which is co-referential with the subject noun phrase is grammaticalized to a reflexive marker. Such a schema can in fact be said to constitute an areal feature of Africa. However, Greenberg's example is in one respect not entirely satisfactory: in the majority of cases it is not a noun for 'head' that is employed in African languages as the head of the object noun phrase but rather the noun for 'body', as in the following example:

(8) Yoruba (Kwa, Niger-Congo; Awolaye 1986: 4)
Nwosu rí ara rɛ
Nwosu saw body his
'Nwosu saw himself.'

That this example illustrates the clearly predominant type found in Africa is suggested by the figures in Table 3, based on a sample of sixty-two African languages from all major genetic groupings and regions of the continent.[22]

Table 3 suggests that by far the most common nominal meaning to encode reflexive (and reciprocal[23]) concepts in African languages is to use the noun 'body'

[22] The reader is referred to Heine (2000) for more details.

[23] Reciprocal categories are not considered here; they differ in some ways from reflexive ones; see Heine (2000).

as the object noun in a propositional schema of the kind *X sees/hits/kills X's body*, and this schema has been grammaticalized to a reflexive category.[24] More than half of the sample languages for which a nominal source could be established use this schema, and we can in fact say that it may well constitute an areal property of the African continent: given any unknown African language, one may predict with more than chance probability that that language will have a grammaticalized form of the above schema, let us call it the 'body'-schema, to express reflexivity.

Compared to this, the number of African languages using a schema of the form *X sees/hits/kills X's head*, that is, which have grammaticalized the noun 'head' to a reflexive pronoun, is fairly small: less than one fourth of our sample languages appear to have done so. Furthermore, these languages are spoken in one specific area, the sub-Saharan belt of West Africa, roughly between Senegal and Cameroon, and they belong to two different language families: Niger-Congo (West Atlantic) and Afroasiatic (Chadic).[25] There is only one exception in our sample, which is Kemantney (Kimant), a Central Cushitic language of Ethiopia, which also has the 'head'-schema.

This suggests that, in addition to the pan-African distribution of the 'body'-schema, there is also an areal pattern based on the 'head'-schema, also cutting across genetic boundaries. Note that two languages of the sub-Saharan belt area, Margi and Mina (both belonging to the Chadic branch of Afroasiatic), have two reflexive categories, derived respectively from the 'body'-schema and the 'head'-schema.

3.3. SUMMARY

These observations suggest first, that among those linguistic features that are indicative of an areal rather than a genetically defined distribution there are patterns involving neither phonetic, nor phonological nor morphosyntactic forms; rather they involve meaning and the way meaning is encoded. Our concern was with meaning relating not to lexical semantics but to grammatical categories. Thus we considered not merely event schemas, that is, meaningful propositions, but rather the way these schemas are employed for the expression of grammatical functions. In other words, we were dealing with metatypy of a specific kind, a process involving a two-stage strategy, whereby speakers not only adopt[26] a certain semantic configuration or schema, but also the idea that this configuration be used for encoding grammatical meaning.

Second, both examples presented involve a pan-African areal patterning on the one hand, and a more restricted regional one on the other. Both patternings cut

[24] Lionel Bender (p.c.) observes that 'neck' and 'foot' are additional sources for reflexive markers in some African languages.

[25] These languages are Fulani, Diola, Hausa, Margi, Mina, Pero, Kwami, and Lele. The first two are West Atlantic languages, while the rest are Chadic. Note that the languages spoken further south, along the West African coast, are excluded from this belt.

[26] Whether this is done consciously or unconsciously is an issue we cannot dwell on here.

across boundaries of genetic units. The most plausible explanation, therefore, would seem to be one in terms of language contact. What we now need most urgently is, first, sociolinguistic micro-studies of speech communities in contact that would allow us to describe in more detail why and how exactly people adopt propositional schemas of the kind discussed in this section from other speech communities.

4. Conclusions

Our observations of some problems of historical linguistics in Africa may be summarized in the following way. First, we observed that quite a bit of progress has been made in the genetic classification of African languages; still, our knowledge of more remote relationship patterns is severely limited. Second, while the family-tree model, based on the one-parent assumption, has turned out to be the only one to describe genetic relationship appropriately, there are cases such as Nile Nubian which may be viewed as an additional challenge for this model. We should be aware, however, that models used in comparative linguistics, in the same way as models used elsewhere in the humanities, are based on probabilities rather than on exceptionless laws. Examples like Nile Nubian may suggest that searching for a family tree no longer makes much sense; still, looking at the overall situation of language history in Africa, such cases are statistically hardly significant. Thus, it seems advisable not to select, or to develop, models on the basis of such spectacular and unusual cases but rather on what is the expected case, that is, on what is most likely to have happened in the history of the language concerned, or of the people speaking that language.

Third, previous work has been overly concerned with a search for the origin of Africa's present linguistic diversity, and to this end new genetic classifications were proposed time and again to reduce present-day variety to earlier unity. Such work could rely on a set of readily applicable methods, and on an attractive, logically coherent model for describing linguistic relationship: the family-tree model. Compared to this work, research on contact-induced linguistic relationship is still in its infancy. What makes areal language classification particularly difficult are problems such as the following: (a) there are no reasonable findings to guide the student of areal linguistics as to how many features would be required to define an areal group, or how to determine its boundaries; from the little we know, boundaries of areal groups are notoriously fuzzy; (b) there are also no ready-made methods and models to classify languages according to contact-induced relationship.

Fourth, genetic linguistics rests primarily on the comparison of form–meaning units, that is, on observations on correspondences between morphemes and words of different languages. Similar approaches have been used to reconstruct areal diffusion processes, perhaps most successfully for the reconstruction of lexical borrowing. One goal of the present chapter is to suggest that language contact manifests

itself in the same way in semantic transfers, more specifically in grammaticalizing metatypy and, most likely, also in other kinds of metatypy. It may happen that people borrow a comparative or a reflexive morpheme from another language but, as we argue in this chapter, they are more likely to borrow conceptual templates, like event schemas, to develop a new comparative or reflexive category.

References

AWOLAYE, Y. 1986. 'Reflexivization in Kwa languages', pp. 1–14 of Dimmendaal (1986).

BECHHAUS-GERST, M. 1984. 'Sprachliche und historische Rekonstruktionen im Bereich des Nubischen unter besonderer Berücksichtigung des Nilnubischen', *SUGIA* (*Sprache und Geschichte in Afrika*, Cologne) 6.7–134.

—— 1996. *Sprachwandel durch Sprachkontakt am Beispiel des Nubischen im Niltal: Möglichkeiten und Grenzen einer diachronen Soziolinguistik*. Language Contact in Africa 3. Cologne: Köppe.

BLENCH, R., and SPRIGGS, M. 1997. Editors of *Archaeology and Language*, vol. 1: *Theoretical and Methodological Orientations*. London: Routlege.

BÖHM, G. 1984. *Grammatik der Kunama-Sprache*. Beiträge zur Afrikanistik 22. Vienna: Afro-Pub.

CLAUDI, U. 1993. *Die Stellung von Verb und Objekt in Niger-Kongo-Sprachen. Ein Beitrag zur Rekonstruktion historischer Syntax*. Afrikanistische Monographien 1. Institut für Afrikanistik, Universität zu Köln.

CREISSELS, D. 2000. 'Typology', in Heine and Nurse, forthcoming.

DIMMENDAAL, G. J. 1986. Editor of *Current Approaches to African Linguistics*, vol. 3. Dordrecht, Cinnaminson: Foris Publications.

DIXON, R. M. W. 1997. *The Rise and Fall of Languages*. Cambridge: Cambridge University Press.

DURIE, M., and ROSS, M. D. 1996. Editors of *The Comparative Method Reviewed: Regularity and Irregularity in Language Change*. New York: Oxford University Press.

EHRET, C., and POSNANSKY, M. 1982. Editors of *The Archeological and Linguistic Reconstruction of African History*. Berkeley and Los Angeles: University of California Press.

FERGUSON, C. A. 1976. 'The Ethiopian language area,' pp. 63–76 of *Language in Ethiopia*, edited by M. L. Bender, J. D. Bowen, R. L. Cooper, and C.A. Ferguson, London: Oxford University Press.

GILMAN, C. 1972. 'The Comparative Structure in French, English, and Cameroonian: Pidgin English: An Exercise in Linguistic Comparison'. Dissertation, Northwestern University, Evanston, IL.

—— 1986. 'African areal characteristics: sprachbund, not substrate?', *Journal of Pidgin and Creole Languages* 1.33–50.

GREENBERG, J. H. 1949. 'Studies in African linguistic classification: I. The Niger Congo family', *Southwestern Journal of Anthropology* 5.79–100.

—— 1955. *Studies in African Linguistic Classification*. New Haven: Compass Publishing Company.

—— 1959. 'Africa as a linguistic area', pp. 15–27 of *Continuity and Change in African Cultures*, edited by W. R Bascom, and M. J. Herskovits. Chicago: The University of Chicago Press.

—— 1963. *The Languages of Africa*. The Hague: Mouton.
—— and BERRY, J. 1971. Editors of *Linguistics in Sub-Saharan Africa*. Current Trends in Linguistics 7. The Hague, Paris: Mouton.
GÜLDEMANN, T. 1997. 'The Kalahari Basin as an object of areal typology: a first approach', *Khoisan Forum* (Cologne), Working Paper 3.
HEINE, B. 1976. *A Typology of African Languages Based on the Order of Meaningful Elements*. Kölner Beiträge zur Afrikanistik 3. Berlin: Dietrich Reimer.
—— 1981. 'African noun class systems', pp. 189–216 of Seiler and Lehmann (1981).
—— 1993. *Auxiliaries. Cognitive Forces and Grammaticalization*. New York: Oxford University Press.
—— 1997a. *Possession: Cognitive Sources, Forces, and Grammaticalization*. Cambridge: Cambridge University Press.
—— 1997b. *Cognitive Foundations of Grammar*. Oxford: Oxford University Press.
—— 2000. 'Polysemy involving reflexive and reciprocal markers in African languages', pp. 1–29 of *Reciprocals: Form and Function*, edited by Z. Frajzyugier and T. S. Curl. Amsterdam: John Benjamins.
—— and NURSE, D. 2000. Editors of *African Languages: An Introduction*. Cambridge: Cambridge University Press.
KRAFT, C. H., and KIRK-GREENE, A. H. M. 1973. *Hausa*. Teach Yourself Books. London: St Paul's House.
LEWIS-WILLIAMS, J. D. 1984. 'Ideological continuities in prehistoric Southern Africa: the evidence of rock art', pp. 225–52 of Schrire (1984).
NICHOLS, J. 1992. *Linguistic Diversity in Time and Space*. Chicago: University of Chicago Press.
ROSS, M. D. 1996. 'Contact-induced change and the comparative method: cases from Papua New Guinea', pp. 180–217 of Durie and Ross (1996).
—— 1997. 'Social networks and kinds of speech-community event', pp. 209–61 of Blench and Spriggs (1997).
SANDFELD, K. 1938. 'Problèmes d'interférences linguistiques', pp. 59–61 of *Actes du Quatrième Congrès Internationale de Linguistes*. Copenhagen: Einar Munksgaard.
SCHRIRE, C. 1984. Editor of *Past and Present in Hunter Gatherer Studies*. New York: Academic Press.
SEILER, H., and LEHMANN, C. 1981. Editors of *Apprehension: das sprachliche Erfassen von Gegenständen*, part I: *Bereich und Ordnung der Phänomene*. Tübingen: Gunter Narr.
STASSEN, L. 1985. *Comparison and Universal Grammar: An Essay in Universal Grammar*. Oxford: Blackwell.
THELWALL, R. 1982. 'Linguistic aspects of greater Nubian history', pp. 39–56 of Ehret and Posnansky (1982).
THOMASON, S. G., and KAUFMAN, T. S. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.
WERNER, R. 1987. *Grammatik des Nobiin (Nilnubisch): Phonologie, Tonologie und Morphologie*. Nilo-Saharan 1. Hamburg: Rüdiger Köppe.
WESTERMANN, D. 1927. *Die westlichen Sudansprachen und ihre Beziehungen zum Bantu*. Beiheft zu den Mitteilungen des Seminars für Orientalische Sprachen, Jahrgang 29. Berlin: Walter de Gruyter.
WOLFF, E. 1993: *Referenzgrammatik des Hausa*. Hamburger Beiträge zur Afrikanistik 2. Münster, Hamburg: LIT.