

# **"Sprachraum" and geography: Linguistic macro-areas in Africa<sup>i</sup>**

Tom Güldemann  
Universität Zürich and MPI-EVA Leipzig

## **1 Introduction**

Continental surveying of linguistic features in Africa is a still little explored research topic. The prominent works in this domain are studies which are by now more than 20 years old, namely Greenberg (1959, 1983) and Heine (1976). Since then no new attempts have been made in which larger zones on the continent which share features and whose origin cannot be explained in terms of inheritance from a single proto-language - a.k.a. linguistic areas or "Sprachräume" - are identified and empirically substantiated. At the same time, the knowledge about African languages in terms of both genealogical and geographical coverage has increased considerably in the past decades. Recent work on macro-areal linguistics in Africa undertaken by the author and independently by Clements and Riailand (2008) on phonological features suggests that these new data provide a good basis for continuing and refining the results of previous work. The present study deals with a special topic in this wider research domain, namely the idea that the distributions of certain linguistic properties can pattern according to different kinds of geographical factors on a continental scale. In particular, it is argued here that the similarity of subsistence conditions along a latitudinal orientation as well as major topographic features can contribute to shape large-scale aggregations of linguistic features. It remains to be seen whether the conclusions arrived at here hold against more data and can be transferred to other continents.

## **2 Previous research and present methodology**

Greenberg is not only the author of the widely accepted genealogical classification of African languages (1963) but also provided the first studies (1959, 1983) concerned with the continental distribution of certain linguistic features, indicating the existence of large linguistic areas that supposedly emerged through contact-induced convergence. He starts out with an attempt to establish "special features of African languages" which are claimed to be recurrent in Africa, but almost absent outside it. These range from phonology (clicks, labial velars, prenasal clusters, tone) over morphology (noun classification, verb derivation) to the lexicon (polysemy involving 'meat' and 'animal'; polysemy involving 'eat', 'win' and 'have sex'; compounds based on the noun for 'child', adpositions derived from nouns of locative relations). Some of these features are said to establish a large "core area" of African languages, while other properties help to identify additional linguistic areas surrounding this core.

The overall results of his survey are summarized in Table 1. This table gives a short label for the proposed areas in the 1st column and the defining linguistic features in the 2nd column as far as Greenberg presents them.

<b>Area</b>	<b>Features</b>	<b>Partaking lineages</b>	<b>Lineage bias</b>
Core area	"special features"	<i>Niger-Congo, Central Sudanic, Songhay, Chadic</i>	<i>Niger-Kordofanian</i>
Khoisan	none	<i>South African Khoisan</i>	<i>Khoisan</i>
Horn of Africa	closed syllables, many consonants (ejectives), no tone, head-final	<i>Cushitic, Ethiosemitic, Kunama, Nara</i>	<i>Afroasiatic</i>
East Africa Nile valley	none	<i>Nilotic, Surmic, Moru-Madi</i>	<i>Nilo-Saharan</i>
Eastern Sahel	medial converbs, no gender, case system	<i>Saharan, Maban, Furan, Daju, Tama, Nubian</i>	<i>Nilo-Saharan</i>

**Table 1: Greenberg's (1959) proposed macro-areas**

The approximate geographical extent of a given area arises first of all from the data in the 3rd column of the table which lists the genealogical language groups (henceforth just "lineages", cf. Nichols 1992 for this term and its concept) involved. The rationale behind this is Greenberg's observation that certain features which are relatively rare cross-linguistically can occur frequently in an individual lineage or even in more than one lineage. To the extent that these are distributed in a more or less compact geographical region, the distribution of the linguistic feature can be inferred from the distribution of the lineage(s). This procedure is clearly a methodological shortcut with a considerable range of error. Nevertheless, it is also employed in the present article, because it is believed to yield results which are quite robust as soon as these are contrasted with the distribution pattern of such rare features on a global scale.

As can be discerned from the table, one shortcoming of Greenberg's survey is that in two of five cases - "Khoisan" and "East Africa Nile valley" - he leaves it to the reader to determine which features actually define the area. Another problem concerns the considerable overlap between the hypothesized areal language groups and the four super-groups established by his genealogical classification. This can be seen by comparing the 3rd with the 4th column: all but the first area have a strong genealogical bias in that they are predominantly or even exclusively constituted by families of just one such super-group (indicated in both columns by italics). In one case, an area is even fully coextensive with an alleged family: the "Khoisan" AREA is identical with the "South African Khoisan" FAMILY. This can make both types of classification almost vacuous in that invoking language contact on top of inheritance, which is given primary status in the first step of genealogical classification, allows one to take one's pick when deciding for a particular scenario for explaining modern distributions of linguistic isoglosses. Whatever defects Greenberg's survey of Africa may have, most of his linguistic areas show up again in one or another way in the macro-areal profile to be proposed in §3 and §4 below.

A second major areal survey of African languages was undertaken by Heine (1975, 1976). He followed a very different approach: on the one hand, he focused on a single feature complex which does not necessarily involve cross-linguistic rarity, viz. word order (evidently inspired by Greenberg's study of 1966); on the other hand, he carried out his survey on the basis of a large sample of individual languages aiming at some degree of genealogical and geographical exhaustiveness. As to be shown below, this approach

turns out to be a welcome supplementary tool for establishing sub-continental linguistic areas.

Heine's major results for the present discussion are twofold. First, he shows that Africa largely corroborates the existence of three cross-linguistically frequent clause word orders with their partly correlating features in noun phrase structure, affix position, etc., namely his types A (head-initial S-V-O), C (head-initial V-S-O), and D (head-final S-O-V), respectively. At the same time he identifies an additional type B which is far less frequent from a typological perspective; it can be roughly characterized as a "mixed" type which oscillates between head-initial features (e.g., the order verb-adverb) and head-final features (e.g., the order genitive-noun or postpositions), and in which the object position in the basic clause is not a defining criterion (it can be either O-V, V-O, or both).

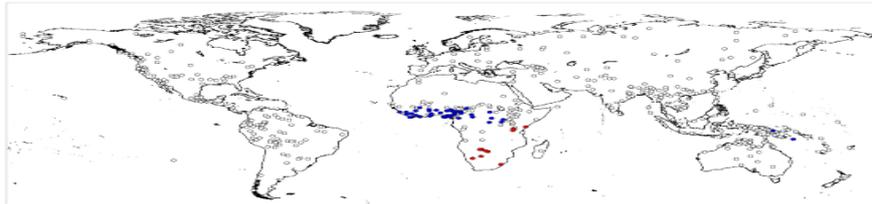
A second important outcome of Heine's study is that his four established language types partly pattern areally. This is shown in Map 1, reproduced here from the original study. Again, several areas established by this map will play a role in the following areal classification of Africa.

#### MAP 1 TO BE INSERTED ABOUT HERE

The present study combines the two approaches described above in that it takes different classes of linguistic features into account. The primary focus lies, parallel to Greenberg's procedure, on typologically highly marked properties - so-called "cross-linguistic quirks" (see Gensler 2003 regarding the concept and its general potential for historical linguistics). Such features have at least two advantages for the present purpose. First, quirks are rarely missed even in the sketchiest description of a language so that surveys of such features are already possible in areas or families where language documentation is still incomplete.

Secondly, their rarity makes it less likely that a multiple occurrence in genealogically and/or geographically close languages involves historically independent innovations. In other words, to the extent that their quiriness decreases a recurrent chance occurrence, any kind of "compact" distribution is likely to be due to inheritance or contact. Moreover, even if a relevant genealogical or geographical subset of languages does not show the feature throughout, the historical explanation remains viable.

This will be illustrated by the global distribution of phonemic clicks as shown by the red dots in Map 2. It can be seen that click phonemes are highly unusual in the languages of the world: there are only three wider locations where languages with these sounds are found: (1) one entirely isolated case in northern Australia, namely Damin, which is an initiation language formerly spoken in the Lardil Aboriginal community (not recorded on Map 2); (2) a loose cluster of three unrelated languages in eastern Africa, namely the Cushitic language Dahalo as well as Hadza and Sandawe, both commonly subsumed under the spurious label "Khoisan", but so far thought to be unrelated by specialists; and (3) a truly compact area in southern Africa involving languages of four unrelated language families, three so-called "Khoisan" families, Tuu, Ju-#Hoan, and Khoe-Kwadi, as well as some languages of Bantu – a sub-branch of Benue-Congo (Niger-Congo).



The extreme rarity of phonemic clicks as well as their biased geographical distribution strongly suggests that each of the two clusters of unrelated click languages in eastern and southern Africa reflects a certain amount of contact-induced proliferation of clicks within a geographical area. For the three languages in eastern Africa, which themselves are not and possibly never were in contact with each other, this would hold under the assumption that click languages were widely obliterated by such non-click lineages as Cushitic, Bantu, and Nilotic, which are the latest arrivals in the region. In fact, it is not an implausible idea that the two geographically separated African areas with clicks are also historically related (see §4.2 below). In more general terms I venture that even without synchronic evidence for language contact an explanation in these terms is the most likely hypothesis if a compact feature distribution along non-genealogical lines is paired with the feature's quirkiness.

Besides the preferred consideration of linguistic quirks, the following survey also deals with features which are typologically far less remarkable, but happen to show a clear areal pattern in Africa. For convenience, these are referred to here as "continental diagnostics". Examples to be presented below are syntactic properties like word order or phonological features like nasal vowels.

There is finally a third type of feature whose difference to the two previous types is that its distribution is not nearly coextensive with a proposed area, but is at least confined to it. These so-called "sub-areal features" may or may not be significant from a wider cross-linguistic and/or more local African perspective. They are taken to potentially define sufficiently large centers of innovation within a macro-area and, in not crossing its boundaries, help to corroborate its overall geographical profile.

The criteria for features which are potentially diagnostic for a linguistic area are the usual ones. First, there should, of course, be a sufficient amount of clustered geographical distribution of a feature in some area. As mentioned above, the rarer a feature the less it would be required to cover an area exhaustively. Secondly, features that must be suspected to be partly interdependent structurally are not counted separately, for example, word order correlations or correlations between word order and other grammatical categories (e.g., the distinction of medial vs. final verbs in clause-chaining verb-final languages). Another less important, but still welcome criterion is the existence of empirical evidence that the feature can be transferred by contact from one language to another. However, this is not necessary for accepting a feature, in particular because research on language contact has shown that - given appropriate sociolinguistic

conditions - the range of properties potentially subject to contact-induced proliferation is far wider than assumed previously (see, e.g., Curnow 2001).

A far more important criterion is that a feature must not correlate with genealogical boundaries. This in two ways: first, it should crosscut/ extend over boundaries of language families and thus define a "Sprachbund"; secondly, it should single out languages or language groups of the same family according to its presence or absence in line with a proposed areal pattern. It is clear then that particularly in continental surveys the identification of an areally significant isogloss is in the first place a function of the underlying genealogical language classification.

It is important in this respect that, in spite of the fact that until today Greenberg (1963) is the unchallenged basis of classifying African languages in and outside the discipline (see, e.g., Heine and Nurse 2000), his hypothesis about just four genealogical super-groups is not followed here. Against common belief, but in line with non-Africanist views like those of Thomason (1994) and Dixon (1997) it is ventured that his later lumping classification approach to African languages is in many respects as defective as his attempts in other areas of the world like the Pacific, the Americas, and Eurasia.

The three groups Khoisan, Nilo-Saharan, and Niger-Kordofanian, in the form he presents them, are not viewed here as reasonably proven genealogical entities (for easy reference they will nevertheless be indicated by the respective abbreviations KS, NS, and NK). Khoisan is treated here as 5 independent lineages. Similarly, Nilo-Saharan is, pace Bender (1996) and Ehret (2001), broken down to sub-groups which have been or obviously can be established on the basis of the available data as secure families. Central Sudanic and a core group of Eastern Sudanic (defined here as the intersection of Bender 2005 and Ehret 2001) are tentatively accepted, although even these far smaller units have not yet been proven by the historical-comparative method. In Niger-Kordofanian, too, one is only on relatively safe ground with Niger-Congo in a narrow sense, comparable to Stewart's (1976) "Volta-Congo" and Bennett and Sterk's (1977) "Central Niger-Congo" (pace Williamson and Blench 2000). That is, the term Niger-Congo is used here for a stock which for the time being excludes the families Mande, Atlantic, Dogon, Ijoid, and Kordofanian.

When treating a certain language (group) as not belonging to another lineage, this does not imply a claim about its status as an isolate lineage; it only means that its external genealogical relation is not yet established beyond reasonable doubt so that any kind of linguistic affinity to another group could have been mediated alternatively by language contact. In the context of the present topic, this means that a large linguistic area is always an alternative explanation to a hitherto unproven genealogical "super-group". In any case, the target of this cross-African typological survey is low-level family units in the sense of Nichols (1992), irrespective of their classification into higher-order groups.

The individual continental feature surveys are mostly based on published sources. When these lack information on certain families to be taken into account, additional literature was consulted. Nevertheless, a major problem is the lack of data on individual isolated languages and families, which is particularly serious in the geographical domain of Nilo-Saharan in eastern and east-central Africa; this area is genealogically highly diverse in hosting many small families and four larger fragmentation areas (see §4.1 below).

If a feature is found in a family, the kind of occurrence is classified into frequent, occasionally present, and absent. This step in the analysis of empirical data is still the most serious source of potential error, because the subjective assignment of the value "frequent" is reflected in the tables and maps by claiming a feature for the entire family. This is often a gross oversimplification of the facts and can only be justified with reference to the feature's global and/or continental rarity.

Before turning to the linguistic macro-areas in Africa themselves, a few words about the concept of a linguistic area a.k.a. "Sprachraum" and its identification are in order. For the present purpose a distinction between two conceptualizations of this term need to be made. In a very general and abstract sense, a linguistic area should be viewed as a purely observational phenomenon, namely a distribution of a linguistic feature over more or less compact geographical space. As mentioned above, what "compact" and thus geographical entity mean depends decisively on the larger reference area in which the distribution pattern of the feature is observed, for example, the globe, a continent, a sub-continent, etc. A linguistic area in this sense is an important empirical finding in its own right, still independent of a historical (or any other) explanation.

What linguists usually understand under the term has a much more specific sense. It is a clustered distribution of different linguistic features over some kind of compact geographical entity which is not well explained by a historical scenario of genealogical inheritance. In other words, the observed areal pattern of the features emerged to a considerable extent by way of contact between unrelated languages. This notion is overtly implied in the German term "Sprachbund", literally "union of languages", which contrasts with the more neutral "Sprachraum". While the English "linguistic area" literally translates the second German term, it has actually become the semantic counterpart of the first. Since it involves language contact only by convention and thus remains ambiguous, the explicit term "(linguistic) contact area" will henceforth be used.

A feature survey across Africa (and other continents for that matter) does not only yield linguistic contact areas. That is, a compact distribution of features can go hand in hand with the same genealogical affiliation of the languages concerned. In other words, there can be large geographical zones which are homogeneous not only in typological-structural, but also in genealogical terms (potentially reflected by isoglosses numbering in their hundreds). These areas largely correspond to what Nichols (1992) calls "spread zones", where homogeneity in linguistic (and other) terms is correlated with a considerable amount of demic population diffusion.

While genealogically homogeneous areas are relatively easy to recognize, it is the large contact areas for which criteria are more difficult to establish. As mentioned already, the identification of a contact-induced clustering of isoglosses on a larger geographical scale focuses here on cross-linguistic quirks supplemented by continental diagnostics. There are additional principles taken into account when establishing and delineating such a macro-area.

First, regarding differences in cross-linguistic markedness of linguistic features, both privative and multi-valued ones, the area possessing a less frequent value is overall better defined by the given property. To take the simple example of clicks, it is the presence of these consonants which is most diagnostic for southern Africa, rather than their absence for other parts of the continent. This principle holds irrespective of the size of the area established by the different feature values. For example, Mainland Southeast Asia is

characterized by the more common order noun-adjective, while adjacent northern Eurasia and South Asia have the overall rarer order adjective-noun (see Dryer 2005a). Hence, Eurasia and South Asia are better defined by their feature value than Southeast Asia, in spite of the fact that it is more difficult to argue for language contact across the huge territory of the former.

Secondly, the distribution of a feature, in order to be diagnostic, does not have to be coextensive with a single area. If, for example, a feature holds for two adjacent areas, it obviously does not delineate these two from each other; however, the feature is still useful to support their borders with respect to other neighboring areas which lack it; relevant examples will be pointed out in §3 below.

### 3 The contact-induced macro-areas of Africa

#### 3.1 Macro-Sudan belt

In the following, three macro-areas on the African continent will be proposed which are geographically defined by diagnostic linguistic features and at the same time are each diverse in genealogical terms. This suggests that their historical emergence is partly due to language contact.

The first area is the "Macro-Sudan belt" which has been argued for in more detail in Güldemann (2008). The basic results there are only repeated here briefly, without giving an extensive discussion of the features and their distribution. Recall from §2 that the rough outline of a macro-area will be defined in terms of partaking low-level families, which themselves have a certain distribution in geographical space. The genealogical composition of the Macro-Sudan belt is given in Table 2.

Family	Stock	Greenberg's supergroup
<b>CORE</b>		
Atlantic	-	NK
Mande	-	NK
Kru	Niger-Congo	NK
Gur	Niger-Congo	NK
Kwa	Niger-Congo	NK
Benue-Congo (except Narrow Bantu)	Niger-Congo	NK
Adamawa-Ubangi	Niger-Congo	NK
Bongo-Bagirmi	Central Sudanic	NS
Moru-Mangbetu	Central Sudanic	NS
<b>PERIPHERY</b>		
Dogon	-	NK
Songhai	-	NS
Chadic	Afroasiatic	
Ijoid	-	NK
Narrow Bantu (Benue-Congo)	Niger-Congo	NK
Nilotic	Eastern Sudanic	NS

**Table 2: Families partaking in the Macro-Sudan belt**

A comparison with Table 1 shows that the Macro-Sudan belt is largely identical with Greenberg's (1959, 1983) African "core area". As a difference to Greenberg, a more concrete distinction is made between families which can be conceived of as forming the core of the area and families which are peripheral to this core.

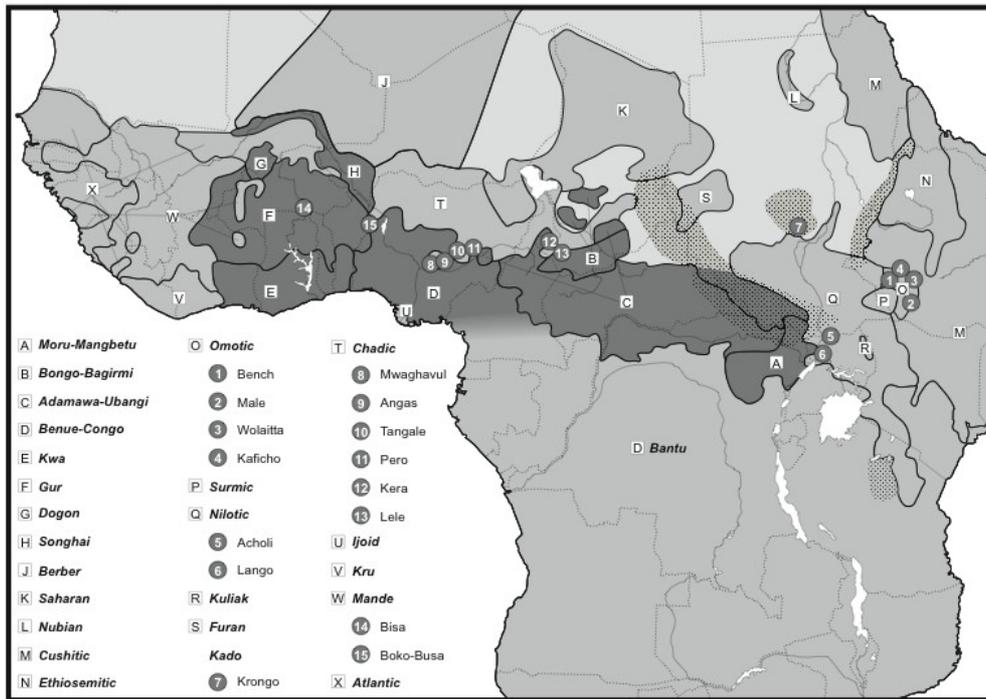
The way in which this particular area was actually developed conceptually corresponds pretty well with the overall procedure proposed above. At the beginning, a compact geographical distribution of a quirky linguistic feature was identified, which cannot be historically explained just by a scenario of inheritance from some proto-language. Güldemann (2003) surveys the occurrence of obligatory logophoric marking, which is a grammatical device indicating in non-direct reported discourse the coreference of a quote-internal nominal to its source, the speaker, as exemplified in (1)a. which contrasts with (1)b. uttered in a non-coreference context.

(1) Kera (Chadic, Afroasiatic)

a. *wə mintí wə kóoré*  
 3M.S<sub>x</sub> QUOT 3M.S<sub>y</sub> go.away  
 Er sagte, daß er weggehe [he said he would go]

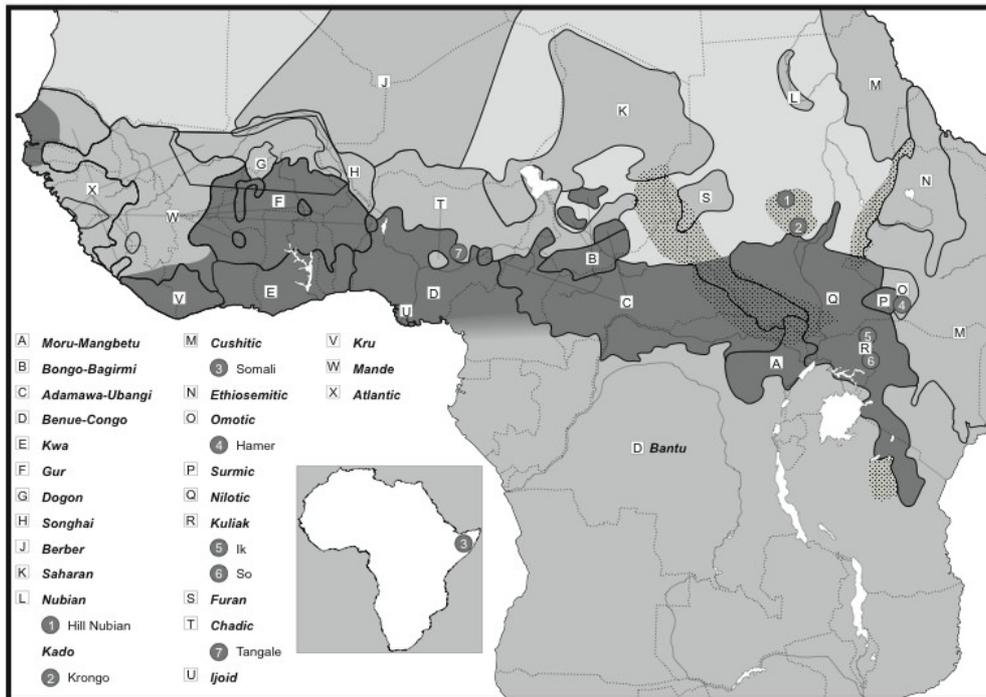
b. *wə mintí tó kóoré* vs.  
 3M.S<sub>x</sub> QUOT 3M.S.LOG<sub>x</sub> go.away  
 Er sagte, daß er weggehe [he said he himself would go] (Ebert 1979: 260)

The distribution of this feature is shown in Map 3. Note that this and similar maps to follow do not show Africa as a whole precisely because a given feature is absent from other parts of the continent.

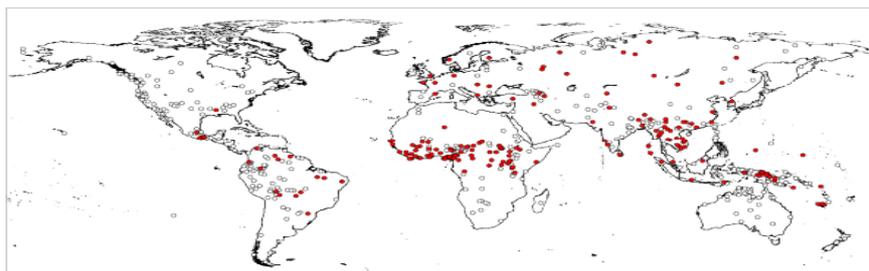


The area of obligatory logophoric marking happens to coincide with that of several other rare features in straddling a similar range of lineages and thus displaying a similar abstract geographical pattern, namely a large belt-like distribution south of the Sahara-Sahel and north of the Congo Basin.

For example, since the earliest comparative research on African languages it is well known that labial-velar consonants occur in a similar geographical region, which is highly biased from both a continental and global perspective (see Güldemann (2008: 156-8) and Map 2 above). What is different, though, is the geographical extent of the two features in the eastern and western direction.



A similar observation can be made with respect to the distribution of another feature, namely a type of vowel harmony called "ATR (= advanced tongue root)". Map 4 shows the geographical picture from the African perspective. Since languages with ATR vowel harmony potentially double their vocalic inventory, the feature can also be discerned indirectly from Maddieson (2005g) which records different degrees of complexity in vowel systems on the global level - the Macro-Sudan belt shows up in Map 5 as a compact block of red dots representing the highest range of seven and more vowels.



Additional candidate features for the Macro-Sudan belt as a whole are complex tone systems of three and more levels (Clements and Rialland 2008: 70-4, Maddieson 2005d) and vowel nasalization (Hajek 2005). All these features show roughly the same geographical pattern: a relatively robust northern and southern demarcation as outlined above, but different extensions towards the east and west.

Another feature which is relevant for the area is the typologically rare word order pattern S-(AUX)-O-V-X in which the non-subject constituents are separated from each other by the verb, as in (2).

(2) Koranko (Mande)

*ù sí wò lá-bùì yí rə*  
 1S PROSPECTIVE that.one CAUS-fall water in

I'm going to throw her (= that one) into the water (Kastenholz 1987: 117)

Güldemann (2008: 159-63) shows that this feature is not homogeneously distributed over the Macro-Sudan belt in that there is a large gap due to its virtually complete absence in the Ubangi and Bongo-Bagirmi families; this geographical pattern is partly anticipated in Heine's Map 1 (see his type B), although he does not record its presence in languages of Nigeria, Cameroon, and even Chad. Its global rarity (see Gensler and Güldemann 2003, Dryer and Gensler 2005) as well as its occurrence in the easternmost core family Moru-Mangbetu still justify viewing it as a typical property of the area as a whole.

A pattern that is of more general interest for the identification of macro-areas is provided by the case of 'exceed'-comparatives – a structure exemplified in (3).

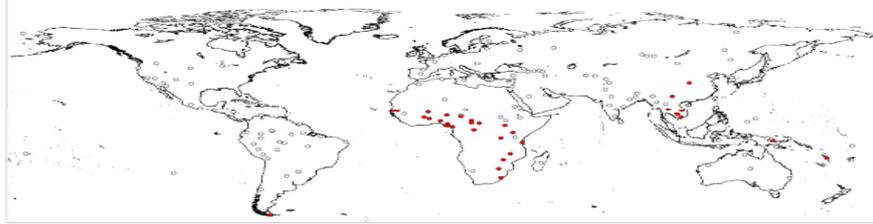
(3) Duala (Bantu, Benue-Congo, Niger-Congo)

*nín ndábò e koló búkà nín*

this house it big exceed that

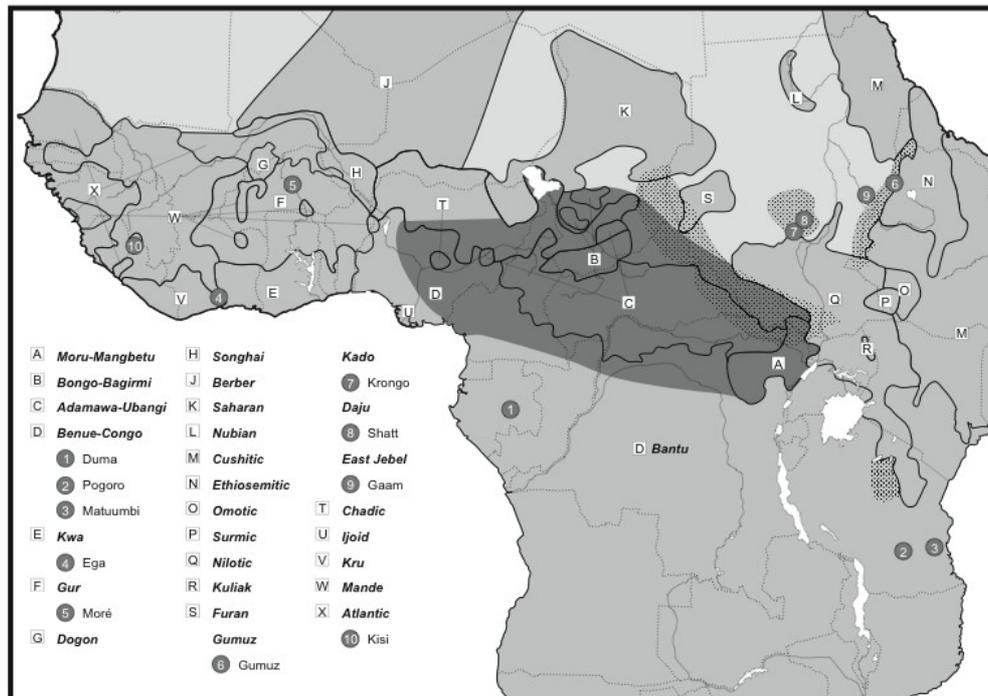
This house is bigger than that (Stassen 2005: 490)

This construction type is not widely distributed on the global level, but frequent in Africa, as already discussed by Greenberg (1983) and demonstrated again in the recent world-wide survey by Stassen (2005), shown by red dots in Map 6. More specifically, it is typical for the Macro-Sudan belt but, as opposed to previous features, is also widely attested in the Bantu family further south. Thus, while this property does not lend support to the southern boundary of the macro-area at issue, it still helps to establish its limits in the north.



A similar pattern seems to be found with implosives, except that these sounds are widely lacking in the western Bantu sphere and at the same time extend further east than most Macro-Sudan belt features (see Maddieson 2003: 28, 2005a).

Finally, two examples for sub-areal properties are given. In the Macro-Sudan belt, these indicate among other things that the four families Benue-Congo, Adamawa-Ubangi, Bongo-Bagirmi, and Moru-Mangbetu form a compact area that can be conceived of as an even more central innovation hotbed in the east. Relevant linguistic features are the globally rare word order pattern V-O-NEG (see Dryer forthcoming, Güldemann 2008: 163-5) and labial flap consonants (see Olson and Hajek 2003, Güldemann 2008: 165-6), shown in the Maps 7 and 8, respectively.



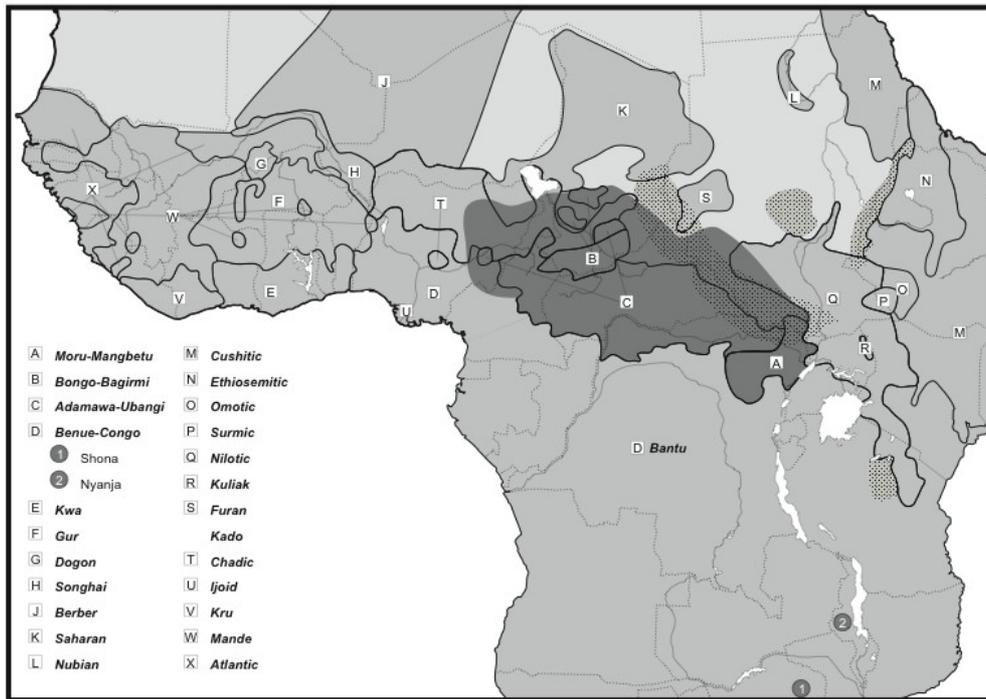


Table 3 demonstrates a different way the clustering of several geographical distributions on a continent can be visualized. It surveys major African lineages on the family level and some isolated languages to the extent that relevant data are available. These are ordered in the table to the effect that those geographically adjacent to each other and within the Macro-Sudan belt are put together, while those at its periphery or outside it are grouped around the former. The table takes six cross-linguistic quirks of the area into account (two of them sub-areal), whereby the cells are marked by shading according to a feature's frequency in a family. Due to the ordering explained, the middle of the table displays a cluster of many grey cells reflecting a cluster in real geographical terms.

<i>Family</i>	<i>Stock</i>	<i>Logopho-ricity</i>	<i>Labial-Velars</i>	<i>ATR Harmony</i>	<i>S-(AUX)-O-V-X</i>	<i>V-O-NEG</i>	<i>Labial flaps</i>
Berber	Afroasiatic						
Saharan							
Maban							
Furan							
Kordofanian							
Nubian	Eastern Sudanic						
Kunama (isolate)							
Nara (isolate)	Eastern Sudanic						
Semitic	Afroasiatic						
Cushitic	Afroasiatic						
Omotic	Afroasiatic						
Kado							
<i>Chadic</i>	Afroasiatic						
<i>Dogon</i>							
<i>Songhai</i>							
<b>Atlantic</b>							
<b>Mande</b>							
<b>Kru</b>	Niger-Congo						
<b>Gur</b>	Niger-Congo						
<b>Kwa</b>	Niger-Congo						
<b>Benue-Congo</b>	Niger-Congo						
<b>Adamawa-Ubangi</b>	Niger-Congo						
<b>Moru-Mangbetu</b>	Central Sudanic						
<b>Bongo-Bagirmi</b>	Central Sudanic						
<i>Ijoid</i>							
<i>Bantu</i> (Benue-Co.)	Niger-Congo						
<i>Nilotic</i>	Eastern Sudanic						
Surmic	Eastern Sudanic						
Kuliak							
Hadza (isolate)							
Sandawe (isolate)							
Khoe-Kwadi							
Ju							
Tuu							

Notes: Family column: grey cell = family of the wider area; **bold** = family of the core area; italic = family of the periphery

Feature columns: dark grey = frequent; medium grey = present; light grey = rare; blank cell = absent or unknown

**Table 3: Distribution of Macro-Sudan belt features across African lineages**

### **3.2 Kalahari Basin**

The second macro-area proposed here also has a precedent in Greenberg (1959) and Heine (1975) who seem to view commonalities across Khoisan southern Africa to be the result of both feature inheritance from an alleged proto-language as well as feature diffusion through contact. Since language specialists no longer follow the genealogical hypothesis (see, e.g., Güldemann and Vossen 2000), southern Africa must be viewed as a contact area to the extent it actually shares diagnostic features. A first such areal approach is presented by Güldemann (1998) where the term "Kalahari Basin" is applied.

The genealogical composition of the area is given in Table 4: its core comprises three independent families previously subsumed under "South African Khoisan"; its periphery is formed by certain Bantu languages which, after their relatively recent expansion into the area (not much earlier than 2000 years ago), have been influenced in different domains and to different degrees by contact with the core lineages. It should be taken into account that the geographical outline is a reconstruction of a situation which no longer holds synchronically, because Bantu as well as languages associated with European colonization have widely replaced languages of the core families.

Family	Stock	Greenberg's supergroup
<b>CORE</b>		
Khoe-Kwadi (includes "Central KS")	-	KS
Ju-#Hoan (includes "Northern KS")	-	KS
Tuu (= "Southern KS")	-	KS
<b>PERIPHERY</b>		
Narrow Bantu (Benue-Congo)	Niger-Congo	NK

**Table 4: Families partaking in the Kalahari Basin**

The features to be mentioned in the following are all characterized by the same distribution pattern: they are virtually universal in the three core families while being absent in most of Bantu in the north and east, with the exception of certain languages which have encroached onto southern Africa furthest.

The most striking quirk of the Kalahari Basin has already been discussed briefly in §2 above, namely clicks. In fact, the globally uneven distribution of clicks (Maddieson 2005c) coupled with the fact that these sounds are the backbone of the consonant systems in southern African languages would alone allow one to argue for a contact area. Other typical consonants of the area are ejectives and aspirated stops (Güldemann 2001, Maddieson 2005a) which have, like clicks, spilled over into some adjacent Bantu languages (Herbert 1987).

The high functional load of clicks is accompanied in all non-Bantu languages of the Kalahari Basin by another important feature – a strongly preferred phonotactic pattern of lexical roots,  $C_1V_1C_2V_2$ , whereby clicks and all other strong consonants are restricted to the first C-position (see, e.g., Güldemann 2001). While this feature is unlikely to be a quirk (a very similar pattern is, e.g., found in the Mande family in West Africa), it is a continental diagnostic because Bantu languages do not possess it. Due to the common development of a secondary phonotactic variant pattern,  $C_1V_1V_2$ , Kalahari Basin languages also show universally vowel nasalization (this compact area of nasal vowels has been missed entirely by Hajek 2005); Bantu like many other families in the world normally lacks this feature.

A continental diagnostic, contrasting in particular with the neighboring Bantu languages, can also be identified in the domain of nominal syntax. Heine (1976: 56) and Güldemann (1999) show that head-final genitives are universally present in the core languages of the Kalahari Basin, and this irrespective of other word order types in the clause and the noun phrase. As argued in the second work, this syntactic property also

has repercussions in the nominal morphology: since noun compounding according to the head-final order is a productive strategy for conveying functions like nominal derivation and locative relations, grammaticalization in this domain yields predominantly suffixes; and indeed noun morphology in general is mostly host-final. This feature has been transferred partly to some Bantu languages, which are otherwise strongly prefixing.

Another morphological contrast of the Kalahari Basin languages to Bantu also looks merely like a continental diagnostic, namely the lack of subject cross-reference on the verb. This feature involves, however, a more remarkable aspect in that several languages in the area index the object on the verb – a feature which cross-linguistically tends to imply subject marking.

### 3.3 Chad-Ethiopia

The third macro-area in Africa has been identified first by Heine (1975, 1976) who called it "Chad-Ethiopia"; it is, however, anticipated partly by Greenberg (1959) in that two of his areas taken together, Horn of Africa and Eastern Sahel, have roughly the same geographical extension. The families partaking to different degrees in this area are given in Table 5.

Family	Stock	Greenberg's supergroup
<b>CORE</b>		
Saharan	-	NS
Furan	-	NS
Maban	-	NS
Nubian	Eastern Sudanic	NS
Taman	Eastern Sudanic	NS
Nara	Eastern Sudanic	NS
Kunama	-	NS
Cushitic	Afroasiatic	
Semitic	Afroasiatic	
Omotic	Afroasiatic	
Nyimang	Eastern Sudanic	NS
<b>PERIPHERY</b>		
Kado	-	NS
Nilotic	Eastern Sudanic	NS
Surmic	Eastern Sudanic	NS

**Table 5: Families partaking in the Chad-Ethiopia area**

The feature complex which allowed Heine (1976) to establish this macro-area in the first place is verb-final syntax accompanied by several other correlating word order features beyond the clause (see, e.g., Azeb and Dimmendaal 2006 regarding clause chaining by means of a distinction between final and medial verb, the latter is called there "converb"). Although head-final syntax is very common on the globe, it is a continental diagnostic for Africa: it is overall rare on this continent and the only compact zone where such languages are found is the area at issue (see Heine's type D in Map 1 above); there

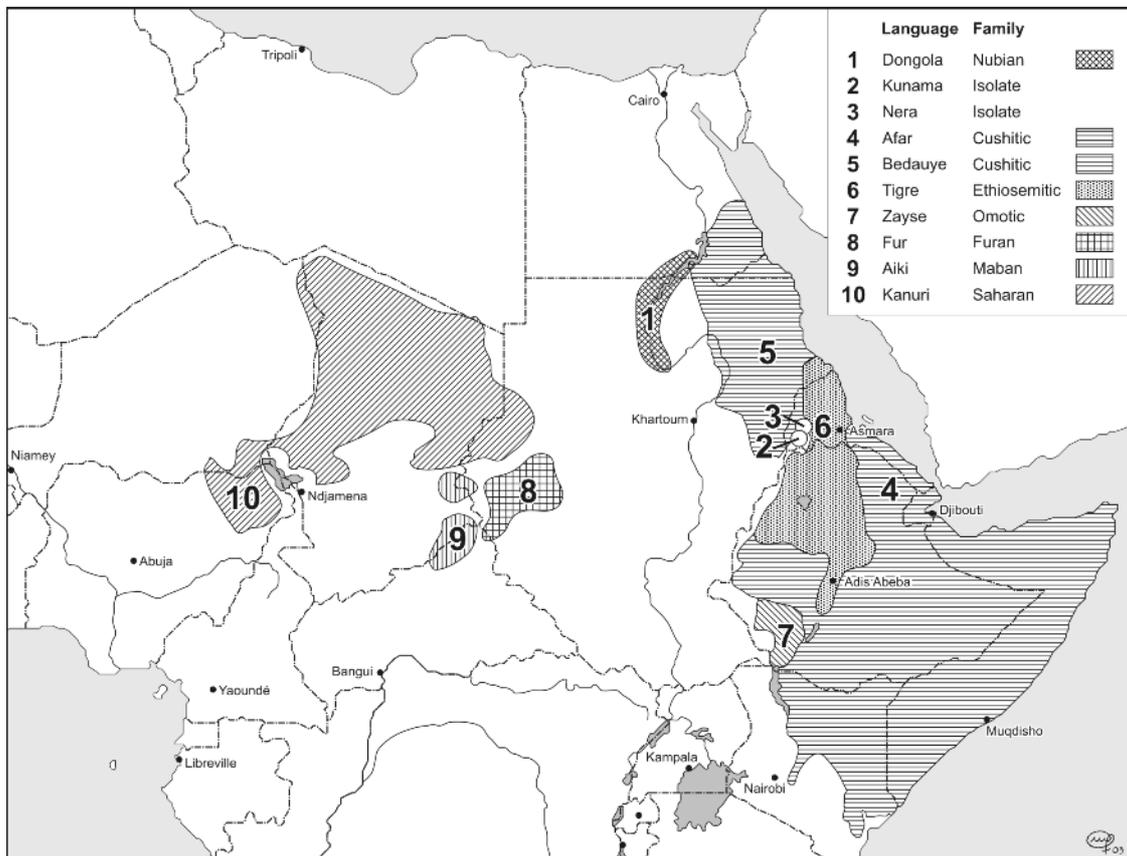
are only four additional and geographically isolated African lineages of this type: Dogon, Ijoid, Sandawe, and Khoe-Kwadi.

A second feature which is geographically largely coextensive with the head-final area is typologically less frequent. Güldemann (2005) shows that the same range of families possess a special type of complex predicate which comprises on the one hand a final semantically very generic auxiliary as the inflectional base (typically derived from such verbs as 'do', 'be (like)', 'become') and on the other hand different types of content words (like expressive mimetic signs, short quotes, loans, nouns, adjectives, and adverbs) which are converted in the construction to predicative expressions, as with the noun *tàssa* and the onomatopoeic word *kowkàw* in (4).

(4) Afar (Cushitic, Afroasiatic)  
*tàssa-iyy-*  
 happiness-AUX-  
 become happy

*kowkàw-iyy-*  
 ON:chatter-AUX-  
 chatter

(Hayward 1996: 541)



Possible concomitant features in many languages concerned are first that the two constituents tend to merge into one phonological word, as in (4), and second that the auxiliary is also used to index reported discourse (which explains its frequent translation as 'say'). The areal distribution of this type of complex predicate is shown in Map 9. The large gap between a western and an eastern block of affected languages, which also emerges in Heine's word-order map, reflects extremely low population density as well as

the relatively recent spread of Arabic which caused the obliteration of languages which formerly joined the two separate zones (see Dimmendaal (2005: 73) for a similar argument).

A third commonality of the Chad-Ethiopia area is that its languages typically display peripheral case in the sense of Blake (1994) in possessing three or more cases. This can be discerned from Dimmendaal (2005: 72-3) as well as Iggesen's (2005) worldwide survey; Map 10 displays the relevant languages by red dots. This feature seems to be typologically unremarkable but is again significant in Africa, because this continent is even poor in languages which have a system for just the two nominal core cases.



Additional candidate features for the Chad-Ethiopia area are the encoding of polar questions by means of verb affixes (Dryer 2005b) and the lack of the otherwise common consonant /p/ (Maddieson 2005f); the last feature includes the Berber languages to the northwest and is thus another instance of a feature that only supports one boundary of an area, in this case the delimitation towards the Macro-Sudan belt in the southwest.

In general, the Chad-Ethiopia area is the least secure of the three contact areas proposed here, in particular because it is still difficult to survey the wider region exhaustively due to the lack of data on isolated languages and small families in the west and southwest.

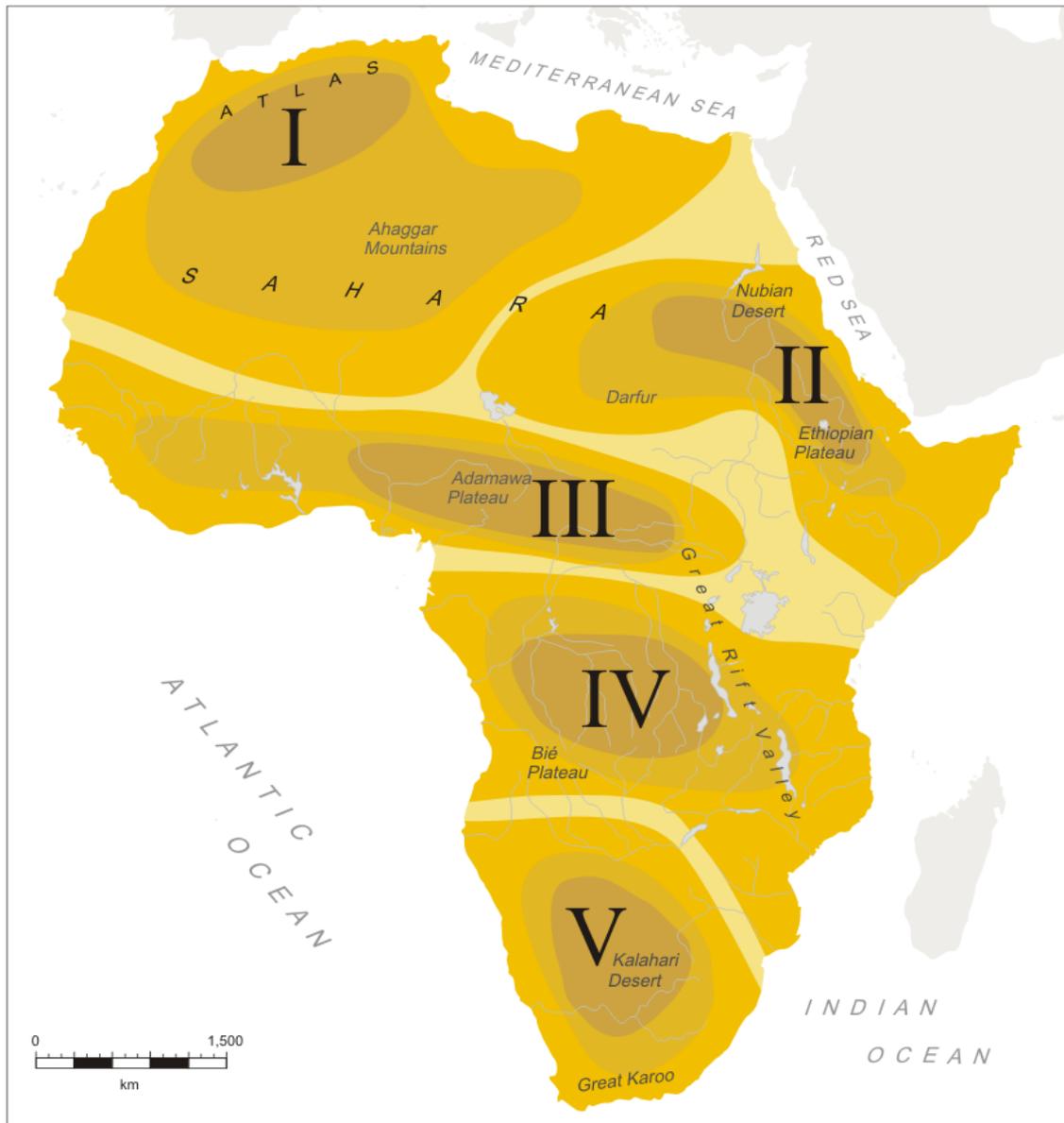
## 4 Macro-areas, history, and geography

### 4.1 Towards a macro-areal profile of Africa

Based on the previous information, a tentative macro-areal profile of Africa can be established as shown in Map 11.

It involves three large areas which are genealogically and structurally diverse, but can be defined in rough geographical terms on account of several linguistic features, often rare from a cross-linguistic perspective. From north to south these large contact areas are:

- Chad-Ethiopia (II)
- Macro-Sudan belt (III)
- Kalahari Basin (V)



The relevant feature distributions cannot be explained by a single historical process, for example, the more or less recent expansion of one or a few linguistic lineages. Given the size of these areas, their relatively abstract definition, and the involvement of many otherwise diverse language families, it should be clear that processes of convergence between languages must have played a considerable role, too. However, the effect of contact should not only be conceptualized as a gradual horizontal diffusion of a feature through space. Alternatively, due to an ancient founder effect one or the other feature may have been present in a considerable portion of a macro-area from very early on and subsequently seeped up into each new colonizing layer from the then relevant substrate – "vertical diffusion through time", so-to-speak. In an abstract, metaphorical sense one could model this as a pattern in which "features sit while populations move."

In addition to the three contact areas, there exist two larger regions which are highly homogeneous in both structural and genealogical terms and can be characterized as spread zones in terms of Nichols (1992):

Sahara spread zone (I)

Bantu spread zone (IV)

The indigenous African families predominating in these two zones, Berber and Bantu, respectively, have their most likely origin in one of their neighboring contact areas. Berber is a genealogical offshoot of Afroasiatic in the Chad-Ethiopia area and Bantu is a genealogical offshoot of Niger-Congo in the Macro-Sudan belt. While both Berber and Bantu no longer show the defining profile, they can sometimes continue one or the other feature of their original area, as with 'exceed'-comparatives in Bantu and missing /p/ in Berber. In a very abstract sense one could then claim a tripartite structure of Africa: Chad-Ethiopia + Berber vs. Macro-Sudan belt + Bantu vs. Kalahari Basin.

Each of the five macro-areas should be conceived of as being formed by a core region in which the set of diagnostic features should be distributed most regularly; the further away one moves from this core, the more will the overall intensity decrease according to which languages and families share these features.

This implies that the peripheries of the areas, or in other words, the territories between their cores, are likely to host languages and lineages which behave ambiguously. In the extreme case they may show a multiple affiliation. Just to mention a few examples: some Songhai languages line up with the Macro Sudan belt regarding logophoricity, labial-velars, and S-(AUX)-O-V-X, but some also lack /p/ like their northern Berber neighbors; or further east, parts of Kado share with the Macro Sudan belt logophoricity, ATR vowel harmony, and V-O-NEG, but at the same time show affiliation with the Chad-Ethiopia area regarding peripheral case and verb affixing for polar questions.

Finally, there is also one larger area of the continent which remains entirely unaccounted for in that it cannot be related clearly to any of the established macro-areas and is substantial enough for not qualifying as a peripheral transition zone. This area stretches from southern Sudan, over Uganda and Kenya into northern Tanzania. It can be roughly characterized as the expansion zone of the Nilotic family, framed by the following four fragmentation areas: Western Sudan border belt, Nuba Mountains, Ethiopian escarpment, and Manyara-Eyasi Basin. Apart from the observation that the languages and lineages involved are diverse in genealogical and structural terms, linguists are still confronted with a serious lack of data. While some diagnostic features like verb-initial clause order (Heine 1975, 1976), rare features in consonant systems (Schadeberg 1987), and marked nominative case systems (König 2006) appear to be typical for this area, their general distribution on the continent extends considerably into different neighboring areas, thus giving overall inconsistent geographical signals.

A largely similar macro-areal picture of Africa arises from Clements and Rialland's (2008) African survey of phonological features. The major difference is that these authors don't recognize a Chad-Ethiopia area and thus enlarge considerably what is called here the Macro-Sudan belt towards the northwest.

## **4.2 Stability and change**

It needs to be reiterated that the above macro-areal profile of Africa is an idealized snapshot of a reconstructed situation that presumably applied before major migratory

processes, since medieval times emanating especially from outside the continent, started to change drastically the geolinguistic structure of the continent. This poses the question to what extent these macro-areas have been stable before these historical events and can be projected further back into the past.

If one looks, for example, at Nichols' (1992) model of a long-term interaction between her two idealized area types, spread and residual zone, one would expect that a certain amount of historical stability should also apply to the African macro-areas proposed here, considering the similar size involved. However, this assumption is problematic, because it turns out that this kind of large area can be subject to considerable changes regarding size, internal profile, and even its very integrity.

First, the Sahara spread zone has undergone a dramatic change with the increasing desertification of the region as well as the expansion of Arabic speaking populations. Berber languages, which were presumably more widespread in the past, are distributed today in scattered pockets which only achieve a greater density in the west of the area; but even there they have been on the retreat because of language shift to Arabic.

The Chad-Ethiopia area, too, is hard to argue for from a synchronic perspective, for the same reasons given for the Sahara spread zone. It displays a large wedge running south through the Sudan, separating a western from an eastern block of languages. Even the size given for this contact area in the map is difficult to ascertain regarding its historical time depth. Given that large tracks of territory at the area's margins apparently result from the expansion of just two families, namely Saharan in the west and Cushitic in the east, it is possible that the area was smaller before these two spreads.

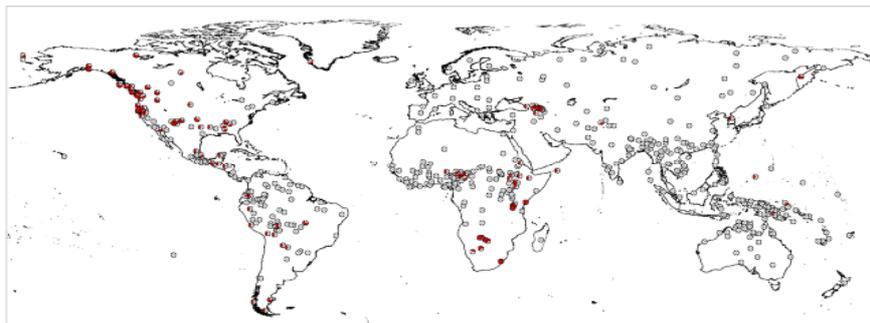
With respect to the Macro-Sudan belt, Güldemann (2008) hints at the possibility that it might have been larger originally in extending up to the Nuba Mountains and to southwest Ethiopia. These zones are separated today from the modern Macro-Sudan belt by the Nilotic and Surmic families due to their relatively recent expansion in this region. Under the plausible assumption that they obliterated there languages affiliated with the then larger macro-area this change could explain that some of its diagnostic linguistic features show up repeatedly further east, either in isolated pockets like Kado, Kordofanian, and Omotic or, due to substrate interference, in Nilotic and Surmic themselves.

The Bantu spread zone is possibly the only area which underwent expansion rather than contraction in the time window considered here. It is, however, difficult to model the profile of the huge territory at issue before the Bantu expansion, because the pre-Bantu languages were almost completely obliterated and systematic research about their possible substrate interference in Bantu has not yet begun. It can at least be suspected that the area was genealogically not as homogeneous as it is today.

The Bantu expansion from the north and the European colonization from the south are responsible for the fact that the Kalahari Basin, too, has entirely lost its territorial integrity. Today, the languages of the oldest southern African lineages, commonly subsumed under the spurious label "Khoisan", have become extinct in most parts of South Africa and some regions of Namibia. Moreover, modern technology has allowed groups subsisting on modes other than hunting and gathering to also encroach on virtually all arid zones of the interior of Namibia and Botswana so that the non-Bantu languages have almost everywhere become endangered minority languages.

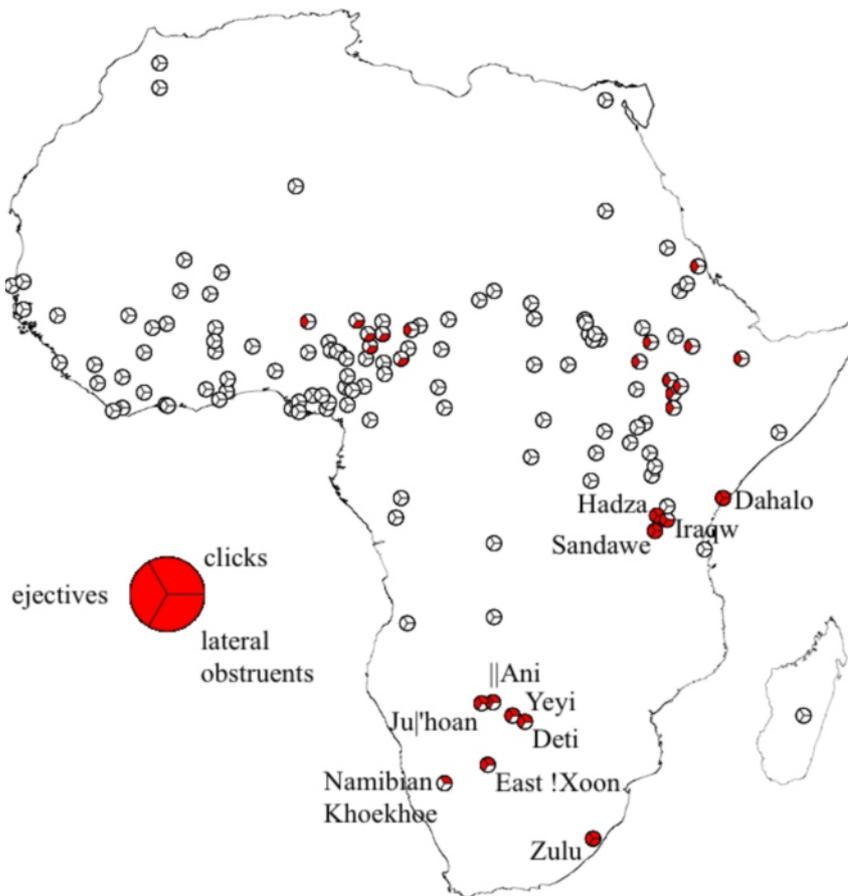
Some evidence suggests that one must even consider the possibility that the profile of the Kalahari Basin has changed in another, far more dramatic sense. Güldemann (1998) proposes to treat the area as a residual zone in terms of Nichols (1992). However, in the meantime it seems quite probable that not only Bantu, but the Khoe-Kwadi family, too, is not indigenous to southern Africa; it may have started out further north, possibly even in eastern Africa, and entered southern Africa just a few centuries before the relevant Bantu languages (Güldemann in preparation). Moreover, since serious historical-comparative research in non-Bantu languages of southern Africa has just begun, it is still open season to consider a genealogical relation between the two remaining families, Ju-#Hoan and Tuu. If both these possibilities turned out to be correct after more sound research has been accomplished, the Kalahari Basin might once have hosted just one lineage, which would amount to what Nichols characterizes as a spread zone. This would mean that it might have changed its very character, that is, from an area of relative homogeneity to an area which is far more diverse in structural and genealogical terms.

Finally, there is an even more significant phenomenon on a continental scale related to the previous remarks. Certain linguistic features that are diagnostic even on a global scale are recurrently shared between two areas of the continent at issue here, namely eastern Africa and southern Africa. The properties, partially mentioned already, are head-final genitives and related morphology (Güldemann 1999) and more importantly several types of rare obstruent consonants (see Map 12), such as ejectives (Maddieson 2005a), lateral obstruents (Maddieson 2005b), and clicks (Maddieson 2005c).



The African distribution of the consonant features is repeated in Map 13. Eastern and southern Africa are the only two areas in the world where all three sound types occur, although their overall density in the southern Kalahari Basin is higher. The geographical spread of the features outside the Kalahari Basin is different: clicks only involve Dahalo, Hadza, and Sandawe in Kenya and Tanzania - languages which can arguably be viewed

as remnants reflecting a typological profile of this region before the intrusion of the Bantu and Nilotic families; ejectives and lateral obstruents, however, are also found further north and west, because they concern in particular Afroasiatic as an entire lineage, involving Chadic, Ethiosemitic, Cushitic, and Omotic.



Important for the present discussion is that the two feature clusters in eastern and southern Africa are separated by the Bantu spread zone which, as mentioned, is not particularly old. Given the partially extreme rarity of the features concerned, it was therefore assumed already by Güldemann (1999: 80-1) that the two geographical distributions are historically related, more specifically that the Bantu expansion obliterated an earlier areal connection between the Kalahari Basin and the northeastern region. This would imply some kind of pre-Bantu macro-area stretching from eastern Africa down to the southern tip of the continent; as discussed below, this would be drastically different from the overall pattern shown in Map 11.

### **4.3 Macro-areas and geography**

The previous section tried to show that especially spreads of colonizing linguistic lineages can seriously affect the size and shape of a linguistic macro-area, be that of a genealogical nature or due to large-scale convergence processes, and possibly even lead to the extinction of its earlier profile and its shift from one type to another. Nevertheless,

it can be proposed that there are some more general influences which shape such macro-areas irrespective of the individual historical processes that decisively determine the concrete distribution of linguistic populations.

For such large geographical entities like continents, the influencing factors are, it is claimed here, of a geographical nature. This is opposed to certain claims in the recent discussion of linguistic contact areas which refer in particular to the problems encountered in identifying them in the first place and then explaining their historical emergence. Campbell (2006: 16), for example, writes that "... it is the diffusion that is of prime importance, and ... the geographical aspect of putative 'linguistic areas' is derivative." While the first part of this statement is in principle uncontroversial, the second part is in my opinion inappropriate, because it should be intuitively clear that geographical features are among the factors which determine the "diffusion" itself, in particular in terms of its trajectory and speed.

That geographical factors do have an effect on linguistic distributions has, of course, been demonstrated before. One clear example is the observation that linguistic diversity is significantly higher in the tropics than in more moderate zones (see, among others, Nettle 1999). For Africa in particular, the fact that the northern- and southernmost parts of the continent are overall less diverse in terms of number of languages and number of lineages has been known for a long time and is epitomized in Dalby's (1977) concept of the African "fragmentation zone".

The question therefore arises whether the macro-areas outlined above also correspond in some way with geographical factors. At first glance, no obvious regular correlations with topography, biogeographic zones, etc. arise: while one might be tempted to align the Sahara spread zone and the Kalahari Basin with the dry areas of northern and southern Africa, respectively, no clear criteria come to mind for establishing the internal coherence of the Chad-Ethiopia area, the Bantu spread zone, and the Macro-Sudan belt, and their differentiation from each other.

However, there is a pattern emerging on a more general continental scale. This can be best elucidated by considering the geographical isogloss patterning of the Macro-Sudan belt (see Güldemann 2008 for more discussion). While the features are fairly consistently delimited in the north by the Sahara-Sahel and in the south by the Atlantic Ocean and the Congo Basin, their extension in the west and east varies far more. This suggests that historical dynamics of the wider area are somehow "blocked" in a north-south direction, while large-scale feature spreads have a predominantly west-east trajectory. In other words, the area has a latitudinal rather than a longitudinal axis. It is this axis pattern which is largely corroborated by the regular layered structure emerging from Map 11, because it could in principle be expected that one or the other large area runs north-south.

Such a spatial configuration of macro-areas, due to their west-east, rather than north-south orientation, is essentially compatible with a central idea by Diamond (1999, chapter 7) regarding the different population histories of continental areas. He argues that their historical dynamics are decisively determined in the long term by the orientation of their geographical axis in that spreads are hampered along longitudinal axes, but facilitated along latitudinal axes, due to the fact that climatic-ecological factors and therefore conditions for human subsistence are overall similar in a west-east direction. It is proposed here that this phenomenon can also have its effect on large sub-continental areas, provided the size of a certain territory is sufficient for this factor to have a bearing

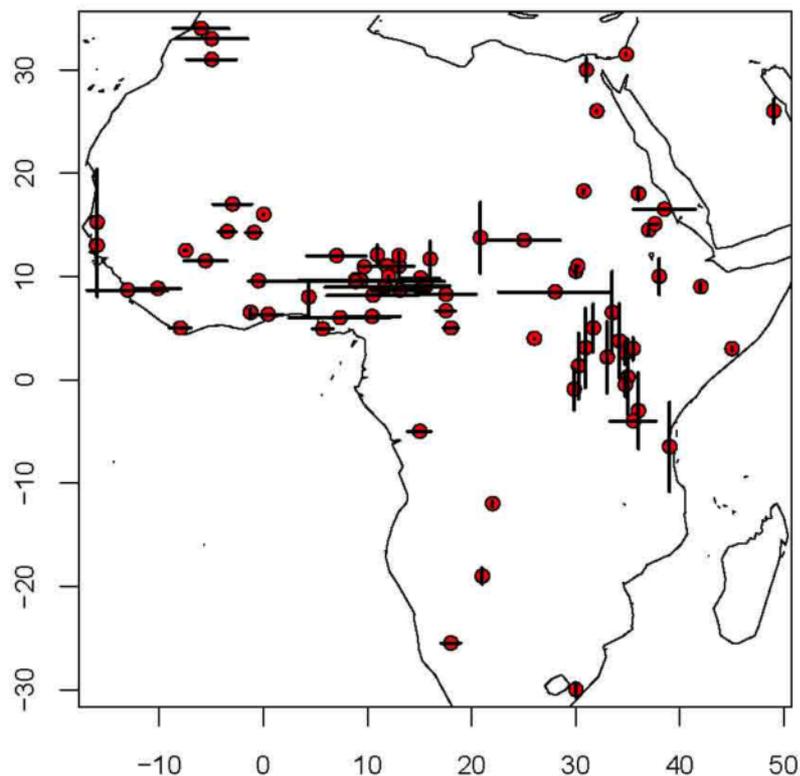
on it. That is, what can be called "latitudinal spread potential" is assumed here to be one of the major geographical factors that influenced the formation of macro-areal linguistic distributions on the African continent.

When speaking of a preference for latitudinal expansions, this should be conceived of as a statistical tendency which concerns long time spans and countless potential spreads. Clearly, north-south movements are possible and a number of such migrations into and out of a given area are securely attested historically. However, the relevance of the latitudinal spread potential can still be reflected in these cases - this in a quite different way. That is, if a linguistic population crosses such a boundary in a longitudinal direction this might be associated with a considerable change of its linguistic profile. If one looks at African families straddling more than one area, it can be observed that the typological and/or genealogical classification into sub-groups correlates with their respective areal alliance. One salient case emerges from the comparison of the non-Bantu Benue-Congo (and some northern Bantu) languages with the majority of narrow Bantu which has widely expanded into the south: while the former display a clear affinity with the Macro-Sudan belt, the latter has a quite distinct character and has even come to form an area of its own (cf. Maps 2-8 for this geographical divide within Benue-Congo). A second example is Cushitic: the majority of the family is structurally squarely within the Chad-Ethiopia area, while its southern branch in Tanzania no longer displays the diagnostic features (cf. Maps 1, 9, 10 for the lack of Chad-Ethiopia features in Southern Cushitic). In lineages predominantly distributed in a peripheral transition zone between two macro-areas, this applies in a more attenuated form in that individual languages show in line with their location a greater or smaller affinity with one or the other area. For example, only the southernmost Songhai language Dendi has labio-velars; in the same way, several Macro Sudan belt features which are untypical for Chadic as a whole show up repeatedly in languages in the southern fringe of this family.

The explanation in terms of the latitudinal spread potential seems to provide a reasonable account for the macro-areal profile of Africa in Map 11. However, §4.2 has entertained a historical scenario involving a large pre-Bantu area that stretched in the east of the continent from Kenya down to its southern tip. This idea implies a clear north-south axis and thus contradicts the very hypothesis just proposed. However, as soon as it is recognized that the latitudinal spread potential is not the only geographical factor that determines the trajectories of spreads, the two conflicting phenomena can be reconciled. That is, stable features of physical and biogeography like mountains, bodies of water, deserts, dense forests etc. should also steer the movement of human populations or the trajectory of feature spreads through diffusion in and across more or less "stable" populations. And indeed, the eastern flank of Africa is characterized by the most important topographic feature of the continent, the rift valley. This runs north-south and forms a chain of major landmarks, most importantly the Ethiopian escarpment, Lake Victoria, Lake Tanganyika, and Lake Malawi; this will have acted, parallel to the eastern coast line, as a geographical "barrier" channeling spreads of peoples and their features preferentially along a north-south axis. In general, "macro-topography" can be considered to be another factor which plays a role in how large-scale distributions of linguistic features assemble in geographical terms, and which at times can conflict with and win out over other factors, for example, the latitudinal spread potential.

The very conflict between two strong trajectories in eastern Africa, the impact of the west-east direction along the Macro-Sudan belt on the one hand and the north-south channeling of spreads along the rift valley on the other hand, might actually shed light on the existence of the large area referred to at the end of §4.1 which seems to defy any alignment with one or the other macro-area in Map 11. It could be conceived of as a zone of continuous "increased turbulence" which inhibits the sedimentation of certain features according to a more or less stable areal pattern.

Incidentally, Cysouw and Comrie (forthcoming) seem to corroborate this finding. With the background of my hypotheses laid out here, a sub-sample of the WALS (Haspelmath et al. 2005), composed of the 77 best-described languages from Africa, was subjected to a measurement of a language's typological similarity to its adjacent languages and the general geographical axis thereof. This is graphically shown in Map 14 where the length of the lines for a given language is representative for the strength of the preference in a north-south or a west-east direction (ultimately leading to a point, when there is no preference for either direction).



Taking into account that the density of sampled languages in the WALS is far more representative in the centre of the continent than in its north and south, the overall picture in Map 14 seems to give signals similar to the areal patterns discussed here, namely a stronger west-east preference in the Macro-Sudan belt against a north-south preference

further east, largely coinciding with the "turbulent" zone along the eastern African rift valley.

Needless to say, this and other hypotheses of this article must remain speculative and more research must show whether they hold against more data. Since some of the ideas entertained above seem to make fairly concrete predictions, it might, however, not be too long that they can be tested more systematically. For example, the potential effect of the latitudinal spread potential should also have some effect on macro-areas on other continents of comparable size. Another idea relates to intra-family diversity in that it might be expected to be higher along a north-south axis, while it tends to be lower in a west-east direction, provided, of course, the relevant territory is large enough. In fact, in so far as genealogical language groups can also be modeled in an abstract sense as clustered distributions of linguistic isoglosses, it would also be worth testing whether, all other things being equal, they themselves show a preference for greater west-east, rather than north-south extension. Or finally, important geographical landmarks like mountains, coast lines, interior water bodies etc. which follow a north-south axis and thus act against the latitudinal spread potential should also be associated on other continents with longitudinal macro-areas. This study can only be a first and very preliminary attempt to ascertain the impact of geographical factors on linguistic distributions.

## References

- Azeb Amha and Gerrit J. Dimmendaal 2006 Converbs in an African perspective. In Ameka, Felix K., Alan Dench and Nicholas D. Evans (eds.), *Catching language: the standing challenge of grammar writing*. Trends in Linguistics - Studies and Monographs 167. Berlin/ New York: Mouton de Gruyter, 393-440.
- Bender, M. Lionel 1996 *The Nilo-Saharan languages: a comparative essay*. Handbook in Linguistics 6. München/ Newcastle: LINCOM Europa.
- Bender, M. Lionel 2005 *The East Sudanic languages: lexicon and phonology*. Carbondale: Southern Illinois University.
- Bennett, Patrick R. and Jan P. Sterk 1977 South Central Niger-Congo: a reclassification. *Studies in African Linguistics* 8: 241-273.
- Blake, Barry 1994 *Case*. Cambridge: Cambridge University Press.
- Campbell, Lyle 2006 Areal linguistics: a closer scrutiny. In Matras, Yaron, April McMahon and Nigel Vincent (eds.), *Linguistic areas: convergence in historical and typological perspective*. Hampshire: Palgrave Macmillan, 1-31.
- Clements, Nick and Annie Rialland 2008 Africa as a phonological area. In Heine and Nurse (eds.), 36-87.
- Curnow, Timothy J. 2001 What language features can be 'borrowed'?. In Aikhenvald, Alexandra Y. and Robert M. W. Dixon (eds.), *Areal diffusion and genetic inheritance: problems in comparative linguistics*. Oxford: Oxford University Press, 412-436.
- Cysouw, Michael and Bernard Comrie forthcoming How varied typologically are the languages of Africa?. In Botha, Rudie and Chris Knight (eds.), *The cradle of language, vol. 2: African perspectives*. Oxford: Oxford University Press.
- Dalby, David 1977 *Language map of Africa and the adjacent islands*. London: International African Institute.
- Diamond, Jared M. 1999 [1997] *Guns, germs, and steel: the fates of human societies*. New York/ London: W. W. Norton.
- Dimmendaal, Gerrit J. 2005 Head marking, dependent marking and constituent order in the Nilotic area. In Voeltz (ed.), 71-92.
- Dixon, Robert M. W. 1997 *The rise and fall of languages*. Cambridge: Cambridge University Press.
- Dryer, Matthew S. 2005a Order of adjective and noun. In Haspelmath et al. (eds.), 354-357.
- Dryer, Matthew S. 2005b Polar questions. In Haspelmath et al. (eds.), 470-473.
- Dryer, Matthew S. forthcoming Verb-object-negative order in central Africa. In Cyffer, Norbert (ed.), *Negation in West Africa*. Amsterdam: John Benjamins.
- Dryer, Matthew S. and Orin D. Gensler 2005 Order of object, oblique, and verb. In Haspelmath et al. (eds.), 342-345.
- Ehret, Christopher 2001 A historical-comparative reconstruction of Nilo-Saharan. *Sprache und Geschichte in Afrika, Beiheft 12*. Köln: Rüdiger Köppe.
- Gensler, Orin D. 2003 Shared quirks: a methodology for "non-orthodox" historical linguistics. Paper presented at the 17th International Conference of Linguists, University of Prague, July 29, 2003.
- Gensler, Orin D. and Tom Güldemann 2003 S-Aux-O-V-Other in Africa: typological and areal perspective. Paper presented at the 4th World Congress of African Linguistics,

- Workshop "Distributed predicative syntax (S P O V X)", Rutgers University, June 21, 2003.
- Greenberg, Joseph H. 1959 Africa as a linguistic area. In William R. Bascom and Melville J. Herskovitz (eds.), *Continuity and change in African cultures*. Chicago: Chicago University Press, 15-27.
- Greenberg, Joseph H. 1963 *The languages of Africa*. Research Center in Anthropology, Folklore, and Linguistics Publications 25. Bloomington: Indiana University.
- Greenberg, Joseph H. 1966 Some universals of grammar with particular reference to the order of meaningful elements. In Greenberg, Joseph H. et al. (eds.), *Universals of language*. Cambridge: MIT Press, 73-113.
- Greenberg, Joseph H. 1983 Some areal characteristics of African languages. In Dihoff, Ivan R. (ed.), *Current approaches to African linguistics 1*. Publications in African Languages and Linguistics 1. Dordrecht/ Cinnaminson: Foris, 3-22.
- Güldemann, Tom 1998 The Kalahari Basin as an object of areal typology - a first approach. In Schladt, Mathias (ed.), *Language, identity, and conceptualization among the Khoisan*. Quellen zur Khoisan-Forschung 15. Köln: Rüdiger Köppe, 137-169.
- Güldemann, Tom 1999 Head-initial meets head-final: nominal suffixes in eastern and southern Bantu from a historical perspective. *Studies in African Linguistics* 28,1: 49-91.
- Güldemann, Tom 2001 Phonological regularities of consonant systems across Khoisan lineages. *University of Leipzig Papers on Africa, Languages and Literatures* 16. Leipzig: Institut für Afrikanistik, Universität Leipzig.
- Güldemann, Tom 2003 Logophoricity in Africa: an attempt to explain and evaluate the significance of its modern distribution. *Sprachtypologie und Universalienforschung* 56,4: 366-387.
- Güldemann, Tom 2005 Complex predicates based on generic auxiliaries as an areal feature in Northeast Africa. In Voeltz (ed.), 131-154.
- Güldemann, Tom 2008 The Macro-Sudan belt: towards identifying a linguistic area in northern sub-Saharan Africa. In Heine and Nurse (eds.), 151-185.
- Güldemann, Tom in preparation Changing profile when encroaching on hunter-gatherer territory?: towards a history of the Khoe-Kwadi family in southern Africa. In Güldemann, Tom, Patrick McConvell and Richard Rhodes (eds.), *Hunter-gatherers and linguistic history: a global perspective*. submitted to Cambridge University Press.
- Güldemann, Tom and Rainer Vossen 2000 Khoisan. In Heine and Nurse (eds.), 99-122.
- Hajek, John 2005 Vowel nasalization. In Haspelmath et al. (eds.), 46-49.
- Haspelmath, Martin, Matthew Dryer, David Gil and Bernard Comrie (eds.) 2005 *The world atlas of language structures*. Oxford: Oxford University Press.
- Hayward, Richard J. 1996 Compounding in Qafar. *Bulletin of the School of Oriental and African Studies* 59,3: 525-545.
- Heine, Bernd 1975 Language typology and convergence areas in Africa. *Linguistics* 144: 27-47.
- Heine, Bernd 1976 A typology of African languages based on the order of meaningful elements. *Kölner Beiträge zur Afrikanistik* 4. Berlin: Dietrich Reiner.
- Heine, Bernd and Derek Nurse (eds.) 2000 *African languages: an introduction*. Cambridge: Cambridge University Press.
- Heine, Bernd and Derek Nurse (eds.) 2008 *A linguistic geography of Africa*. Cambridge: Cambridge University Press.

- Herbert, Robert K. 1987 Articulatory modes and typological universals: the puzzle of Bantu ejectives and aspirates. In Shockey, L. and R. Channon (eds.), *Festschrift for Ilse Lehiste*. Dordrecht: Foris, 401-413.
- Iggesen, Oliver A. 2005 Number of cases. In Haspelmath et al. (eds.), 202-205.
- König, Christa 2006 Marked nominative in Africa. *Studies in Language* 30,4: 655-732.
- Longacre, Robert E. 1990 Storyline concerns and word order typology in East and West Africa. *Studies in African Linguistics, Supplement 10*. Los Angeles: Department of Linguistics, University of California.
- Maddieson, Ian 2003 The sounds of the Bantu languages. In Nurse, Derek and Gérard Philippson (eds.), *The Bantu languages*. Routledge Language Family Series 4. London: Routledge, 15-41.
- Maddieson, Ian 2005a Glottalized consonants. In Haspelmath et al. (eds.), 34-37.
- Maddieson, Ian 2005b Lateral consonants. In Haspelmath et al. (eds.), 38-41.
- Maddieson, Ian 2005c Presence of uncommon consonants. In Haspelmath et al. (eds.), 82-85.
- Maddieson, Ian 2005d Tone. In Haspelmath et al. (eds.), 58-61.
- Maddieson, Ian 2005e Uvular consonants. In Haspelmath et al. (eds.), 30-33.
- Maddieson, Ian 2005f Voicing and gaps in plosive systems. In Haspelmath et al. (eds.), 26-29.
- Maddieson, Ian 2005g Vowel quality inventories. In Haspelmath et al. (eds.), 14-17.
- Nettle, Daniel 1999 *Linguistic diversity*. Oxford: Oxford University Press.
- Nichols, Johanna 1992 *Linguistic diversity in space and time*. Chicago: University of Chicago Press.
- Olson, Kenneth S. and John Hajek 2003 Crosslinguistic insights on the labial flap. *Linguistic Typology* 7,2: 157-186.
- Schadeberg, Thilo C. 1987 Zwei areale Sprachmerkmale im Ostsudan unter besonderer Berücksichtigung des Nera-bena. In Mukarovsky, Hans G. (ed.), *Leo Reinisch: Werk und Erbe. Sitzungberichte, Philosophisch-Historische Klasse 492*. Wien: Österreichische Akademie der Wissenschaften, 213-229.
- Stassen, Leon 2005 Comparative constructions. In Haspelmath et al. (eds.), 490-93.
- Stewart, John 1976 *Towards Volta-Congo reconstruction (inaugural lecture)*. Leiden: Leiden University Press.
- Thomason, Sarah G. 1994 Hypothesis generation vs. hypothesis testing: a comparison between Greenberg's classifications in Africa and in the Americas. Paper presented at the Conference on Language and Prehistory in the Americas, Boulder, Colorado, March 24, 1990. University of Pittsburgh: unpublished manuscript.
- Voeltz, F. K. Erhard (ed.) 2005 *Studies in African linguistic typology*. *Typological Studies in Language* 64. Amsterdam: John Benjamins.
- Williamson, Kay and Roger Blench 2000 Niger-Congo. In Heine and Nurse (eds.), 11-42.

## List of maps and map titles

- Map 1: Basic word order types across Africa (Heine 1976)
- Map 2: Clicks and labial-velars across the globe (after Maddieson 2005c: 84-5)
- Map 3: Logophoricity across Africa (Güldemann 2008: 156)
- Map 4: ATR vowel harmony across Africa (Güldemann 2008: 160)

- Map 5: Complex vowel inventories across the globe (after Maddieson 2005g: 16-7)  
Map 6: 'Exceed'-comparatives across the globe (after Stassen 2005: 492-3)  
Map 7: Word order V-O-NEG across Africa (Güldemann 2008: 165)  
Map 8: Labial flaps across Africa (Güldemann 2008: 166)  
Map 9: Complex predicates with generic AUX across Africa (Güldemann 2005: 147)  
Map 10: Three and more cases across the globe (after Iggesen 2005: 204-5)  
Map 11: The macro-areal profile of Africa  
Map 12: Rare obstruents across the globe (after Maddieson 2005: 32-3, 36-7, 40-1, 84-5)  
Map 13: Rare obstruents across Africa (after Maddieson 2005: 32-3, 36-7, 40-1, 84-5)  
Map 14: Axis orientation of typological similarity in African languages (Cysouw and Comrie forthcoming)

---

<sup>i</sup> Various versions of this contribution were presented at the following occasions: 28th Annual Meeting of the Deutsche Gesellschaft für Sprachwissenschaft, University of Bielefeld (23/02/06); Leipzig Spring School on Linguistic Diversity, University of Leipzig (28/03/06); Work in Progress Series, MPI-EVA Leipzig (26/09/06); Fakultät für Kunst, Geschichte und Orientwissenschaften, Universität Leipzig (02/11/06); Seminar für Allgemeine Sprachwissenschaft, Universität Zürich (15/12/06); Laboratoire Dynamique du Langage (DDL), Université Lumière Lyon 2 (18/12/06); Seminar für Afrikawissenschaften, Humboldt Universität Berlin (19/12/06); Institute for Asian and African Studies, University of Helsinki (13/03/07); International Symposium on Language Contact and the Dynamics of Language, MPI-EVA Leipzig (11/05/07); Workshop on the Typology of African Languages, 7th International Conference of the Association of Linguistic Typology (ALT), CNRS Paris (24/09/07). I am grateful to Hans-Jörg Bibiko (MPI-EVA Leipzig) and the editors of this volume for the preparation of the Maps 2, 5, 6, 10, 12, and 13, and of Map 11, respectively.