

A Geographical Typology of African Languages: Unity and Diversity

Paper presented at the Tokyo University of Foreign Studies, ILCAA, 16 October 2008
Bernd Heine

The question raised in the title of this paper has been posed by a number of students of African languages, it has figured in the title of a seminal paper by Greenberg (1959). In the present paper it is argued that it is possible, on the basis of a quantitative survey on African languages of all major genetic groupings and geographical regions, to define a catalogue of phonological, morphosyntactic, and semantic properties that can be of help in defining African languages vis-a-vis languages in other parts of the world.

1 On linguistic areas

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

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

Our interest in this paper is exclusively with question (1). A variety of different terms have been proposed to refer to sprachbunds, such as linguistic area, convergence area, diffusion area, *union linguistique*, *Sprachbund*, etc. (see Campbell et al. 1986: 530). Perhaps the most frequently discussed sprachbunds are the Balkans (for convenient summaries, see e.g. Joseph 1992; Feuillet 2001), Meso-America (Campbell et al. 1986), Ethiopia (Ferguson 1976), South Asia (Masica 1976; Emeneau 1980), the East Arnhem Land (Heath 1978), the Amerindian Pacific Northwest (Sherzer 1973; Beck 2000), the Vaupés basin of northwest Amazonia (Aikhenvald 1996; 2002), Standard Average European (Haspelmath 1998; 2001), and the Daly River area of Australia (Dixon 2002: 674-9). Furthermore, there are quite a number of less widely recognized sprachbunds, such as the Circum-Baltic (Nau 1996; Koptjevskaja-Tamm & Wälchli 2001), the Middle Volga region (Johanson 2000), or the Circum-Mediterranean area (Stolz 2002).

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

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

On the other hand, sprachbunds are defined exclusively in terms of linguistic parameters without reference to the historical forces that gave rise to them. Emeneau's classic definition¹ is a paradigm case of such definitions; a more recent version is the following (see also Aikhenvald 2002: 7-8):

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

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

(4) Characterization of linguistic areas

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

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

2 Earlier work

Pre-Greenbergian comparative African linguistics suffered from the fact that no systematic distinction between different kinds of historical relationship was made, that is, it remained for the most part unclear whether the linguistic classifications proposed were intended to be

¹ "This term "linguistic area" may be defined as meaning an area which includes languages belonging to more than one family but showing traits in common which are found not to belong to the other members of (at least) one of the families." (Emeneau 1956: 16, n. 28)

genetically, areally or typologically defined or, more commonly, were an amalgamation of all three kinds of relationship. Accordingly, most of the works published prior to 1959 do not offer unambiguous evidence on areal patternings within Africa or between Africa and other parts of the world.

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

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

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

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

Greenberg (1983) went on to reconstruct the history of these properties by studying their genetic distribution. He hypothesized that (i), (iii) and (iv) are ultimately of Niger-Kordofanian origin even though they are widely found in other African language phyla, in

particular in Nilo-Saharan languages. For (ii) however he did not find conclusive evidence for reconstruction, suggesting that it may not have had a single origin but rather that it arose in the area of the Central Sudanic languages of Nilo-Saharan and the Adamawa-Ubangi languages of Niger-Congo.

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

3 "Africanisms"

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

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

A number of properties that are clearly more widespread in Africa than elsewhere are not considered here, for the following reasons. First, because they appear to be genetically determined. The presence of gender or noun class systems is a case in point. Most instances of such systems to be found in Africa are presumably genetically inherited. This can be assumed to apply on the one hand to the nature-based noun class systems found in Niger-Congo and Khoisan languages, and on the other hand to the sex-based gender systems of Afroasiatic and Central Khoisan languages.²

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

Furthermore, the fact that Khoisan languages are among the phonologically most complex languages in the world, some of them distinguishing more than 110 distinct phonemes, is

² It is possible that the presence of gender systems in the Eastern Nilotic languages (Maa, Teso-Turkana, Lotuxo, Bari) is the result of language contact with Cushitic languages, but the evidence on this issue is far from conclusive.

ignored here since it does not appear to be characteristic of Africa as a linguistic area, being restricted to a few North and South Khoisan languages.

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

3.1 Grammar

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

Ignoring click consonants, there are a number of consonant types that are widespread in Africa but uncommon elsewhere. This applies among others to coarticulated labial-velar (or labiovelar) stops (Meeussen 1975: 2; Greenberg 1983: 4; Gilman 1986: 41). Labial-velars may be voiceless (*kp*) or voiced (*gb*). There are also corresponding nasals and/or fricatives, but they do not show the wide distribution of stops, and their occurrence is largely predictable on the basis of stops (Greenberg 1983: 4). The distribution of this property is clearly areally constrained: Labial-velar stops occur in a broad geographical belt from the western Atlantic to the Nile-Congo divide, and they are also occasionally found outside this belt (see Welmers 1974: 47-8), e.g. in Katla and Giryama. Still, they are found in three of the four African phyla; only Khoisan languages have no labial-velar stops. Also, in the Afroasiatic and Nilo-Saharan phyla, their occurrence is restricted essentially to one branch each, namely Chadic and Central Sudanic, respectively (Greenberg 1983: 7). Outside Africa, coarticulated labial-velar stops are found only sporadically, especially in northeastern Papua New Guinea in the Kâte-Ono group of the Indo-Pacific languages³, in some languages of Melanesia, and in the Austronesian language Iai (Greenberg 1983: 5; Maddieson 1984: 215-6); see section 3.2.4 for more details.

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

³ According to Mike Cahill (p.c.), "perhaps 20 languages of Papua New Guinea have labial-velar stops, including Kate, Dedua, Kube, Ono, Fulumu, Amele, as well as Yeletnye, which uniquely not only has /kp, gb/, but also phonemic post-alveolar /tp, db/ as well." Furthermore, he adds that Santa Ana of the Solomon Islands has /gb/.

A third type of consonants that is widespread in Africa can be seen in implosives, which -- following Clements and Rialland -- we define as non-obstruent stops. To be sure, there are non-African languages, such as the Indonesian language Auye (Mike Cahill, p.c.), but such languages are rare. Furthermore, word-initial prenasalized consonants, for the most part voiced stops, are widely found in Africa (Meeussen 1975: 2; Gilman 1986: 41), although they occur most of all in Niger-Congo languages.

An outstanding property relating to the vowel system can be seen in the presence of cross-height vowel harmony based on distinctions of the tongue root position, commonly known as ATR (advanced tongue root) vowel harmony. It is widespread in Niger-Congo and Nilo-Saharan languages across the continent but appears to be rare outside Africa.

Morphological properties that have been mentioned include reduplication of nouns and adjectives, used to express a distributive function (e.g. Swahili *tano tano* 'five each, in fives'; Gilman 1986: 40). Within the verbal word, many African languages are characterized by a wide range of verbal derivational suffixes expressing functions such as reflexive, reciprocal, causative, passive, stative, andative (itive), and venitive (ventive), and these suffixes can be combined in sequence (Meeussen 1975: 2; Gilman 1986: 43). However, both these properties can also be observed widely in non-African languages.

A conspicuous feature of nominal morphology is the paucity of languages having case inflections, and ergative structures are fairly uncommon, but northeastern Africa is a noteworthy exception: There are a number of languages across genetic boundaries that have case inflections, and the only languages exhibiting an ergative organization, Shilluk, Pàri, Anywa, and Jur-Luwo, are found there. Northeastern Africa is also typologically remarkable in that there are quite a number of languages having a marked-nominative system, where it is the accusative rather than the nominative case that is unmarked – note that marked nominative languages are crosslinguistically exceptional. A perhaps unique property of case systems is the presence of case marked exclusively by tonal inflection, which so far has been found only in African marked-nominative languages but nowhere else in the world (König 2006).

With regard to word classes, African languages have been said to be characterized by a paucity of adjectives and, in a number of languages adjectives are claimed to be absent altogether; what tends to be expressed in non-African languages by adjectives is likely to appear as verbs of state in Africa (cf. Gilman 1986: 40). On the other hand there is a word class of ideophones that appears to be remarkably salient in many African languages (Meeussen 1975: 3). While languages in other parts of the world have ideophones as well, African languages have been found to have them in distinctly larger numbers. Furthermore, ideophones expressing color distinctions have so far only been found Africa (Kilian-Hatz 2001; Voeltz & Kilian-Hatz 2001).

In their arrangement of words, African languages of all four phyla exhibit a number of general characteristics such as the following: While on a worldwide level languages having a verb-final syntax (SOV) appear to be the most numerous, in Africa there is a preponderance of languages having subject – verb – object (SVO) as their basic order: Roughly 71% of all African languages exhibit this order (Heine 1976: 23; see also Gilman 1986: 37). Furthermore, the placement of nominal modifiers after the head noun appears to be more widespread in Africa than in most other parts of the world. Thus, in Heine's (1976: 23)

sample of 300 African languages, demonstrative attributes are placed after the noun in 85%, adjectives in 88%, and numerals in 91% of all languages. Another characteristic in the arrangement of meaningful elements relates to verbal structure: In most African languages, pronominal subject clitics or affixes precede the tense markers (93%), which again precede the verb (83%), while adverbs follow the verb 93% (Heine 1976: 24).

An arrangement of basic word order that occurs in a number of languages across the continent but is fairly uncommon outside Africa concerns what nowadays tends to be referred to as SOVX order. In languages having this order, the direct object precedes the verb but the indirect object and adjuncts follow the verb. SOVX languages are likely to have postpositions and to place the genitival modifier before its head while other nominal modifiers follow the head noun (cf. the type B of Heine 1976).

Serial verb constructions have been claimed to be more common in Africa than elsewhere (Gilman 1986: 41). Recent studies suggest in fact that they are not confined to Niger-Congo but exist also in Khoisan languages (Kilian-Hatz 2003; König 2003); still, the majority of African languages do not qualify as serial verb languages, and such languages are not uncommon in some other parts of the world.

With reference to information structure, front-focusing of nouns by means of some kind of cleft-construction has been mentioned, frequently used obligatorily in word questions, where *who went?* is expressed by *who is it who went?* (Gregersen 1977: 50-1; Gilman 1986: 39). In addition to noun phrase focusing there is also front-focusing by means of verb-copying, where the verb appears first in the focus position and is repeated in the main clause (Gilman 1986:39); the exact distribution of this phenomenon across Africa, however, is unknown. Note that focus marking by means of verbal inflections has so far only been found in African languages.

In addition there are construction types that are said to be found in a number of African languages but to be rare outside Africa. One of them is called anastasis by Meeussen (1975: 4), consisting in the swapping of subject and complement participants within the clause, e.g., the possibility to express 'Worms enter the corpse' by 'The corpse enters worms'. It is unknown how widespread anastasis is in Africa, and it would seem that it is not all that uncommon in other parts of the world (Felix Ameka, p.c.).

Logophoric marking constitutes another construction type that has been claimed to be specifically African. Logophoric pronouns indicate coreference of a nominal in the non-direct quote to the speaker encoded in the accompanying quotative construction, as opposed to its non-coreference indicated by an unmarked pronominal device (Hagège 1974; Güldemann 2003a; see also Güldemann & von Roncador 2002). Thus, whereas (5a) illustrates a logophoric structure, (5b) is a plain, non-logophoric structure.

(5) Ewe (Kwa, Niger-Congo)

a é gblɔ bé ye- dzó.
 3.SG say that LOG- leave
 'She_i said that she_i left.'

b é- gblɔ bé é- dzó.
 3.SG- say that 3.SG- leave
 'She_i said that she_j left.'

Logophoric structures are with very few exceptions concentrated in a large belt extending from the southeastern corner of Ethiopia to the east up to the Niger River in the west and are found in three of the four language phyla (Güldemann 2003a; von Roncador 1992: 173).

Finally, there are a number of conceptualization strategies that might qualify as Africanisms. This applies in particular to what is called the goose-file model of spatial orientation (Heine 1997: 12-14), to be found in at least three of the four African language phyla, described by Meeussen in the following way:

Imagine a place from which a house can be seen, and further away a small hill. In such a situation the hill will be referred to in African terms as being 'in front of the house', and the house as being 'behind the hill', whereas in European languages the reverse expressions will be used. (Meeussen 1975: 3)

The following example from the Kuliak language So may illustrate the goose-file model, where an item to be located is conceptualized not as facing the speaker but rather as facing the same direction as the speaker.

(6) So (Kuliak, Nilo-Saharan)
 nɛ́kɛ yóG sú- o sóG.
 be.at people behind- ABL hill
 'There are people in front of the hill.'

There is another conceptualization strategy that has been proposed as an Africanism (Meeussen 1975), being one manifestation of what is usually called the inclusive or inclusory construction, which is used in reference to a plural that refers to a set of individuals and includes two explicit constituents. The form the construction typically takes in African languages is illustrated in (7).

(7) Swahili
 sisi na wewe
 we and you
 'I and you'

It is unknown how widespread this construction type is; it is by no means restricted to Africa, being found in various other parts of the world (Blake 1987; Singer 1999; Moravcsik 2003: 479).

Another strategy that is not restricted to Africa but is perhaps more widespread in Africa than elsewhere, to be found in all four language phyla, consists in the fact that in affirmative answers to negative questions the speaker wants to know if the propositional content of the question is correct or not, e.g., 'Didn't you sleep?' – 'Yes, I didn't' or 'No, I did' (Meeussen 1975:4; Gregersen 1977:44; Felix Ameka, p.c.).

3.2 Polysemy and grammaticalization

Perhaps the most conspicuous area where one might expect to find Africanisms can be seen in lexical and grammatical polysemies. The following are a few examples that have been pointed out by students of African languages.

Within the domain of nominal polysemy, a paradigm case can be seen in the fact that the same noun is used for 'meat' and 'animal' or, alternatively, that there are different but etymologically related nouns for 'meat' and 'animal' (Greenberg 1959, 1983: 4) – a case described by Lichtenberk (1991) more appropriately as heterosemy. Perhaps remarkably, if one of the two meanings is derived from the other then it goes from 'meat' to 'animal' rather than vice versa⁴. This is suggested at least by the fact that whenever the two are distinguished by means of some derivational, compounding or other mechanism then it is the item for 'meat' that is likely to be unmarked and 'animal' to be marked; cf. the following examples (for an example from the Bantu language Tonga, see Greenberg 1983: 16):

Language	'meat'	'animal'
Hausa (Chadic, Afroasiatic)	nāmà	nāmàn dājī 'wild animal' ('meat of the bush')
!Xun (North Khoisan)	hā	hā-mà ('animal-DIM')

To be sure, such a polysemy can also be observed in other parts of the world, but it appears to be much more frequent in Africa than elsewhere (see section 2.4).

Another nominal polysemy that has been claimed to be pan-African is that of nouns denoting both 'hand' and 'arm', or nouns denoting both 'foot' and 'leg' (and 'wheel') (Gilman 1986: 43). Note, however, that these polysemies are also widespread outside Africa. Thus, in the worldwide survey by Witkowski and Brown (1985: 203), 50 out of 109 languages have a 'hand'/'arm' polysemy and 42 out of 109 languages a 'foot'/'leg' polysemy (see Heine 1997: 136).

Examples of polysemies involving verbs include verbs for 'eat', which are said to also denote 'conquer', 'capture a piece in a game', and 'have sexual intercourse' (Greenberg 1959), verbs for 'die', which tend to have many non-literal meanings in African languages such as 'be in a painful condition', 'break down' (cf. Meeussen 1975: 4), verbs for 'lie (down)' also meaning 'sleep', or verbs for 'hear' (to a lesser extent also 'see') also denoting other kinds of perception, such as 'smell', 'feel', 'taste', 'understand' (Meeussen 1975: 4-5). Meeussen (1975: 4) furthermore notes that the use of words for 'good' also tend to express 'nice', 'beautiful', and 'fine' in African languages. The status of some of these polysemies as Africanisms, however, is far from clear. For example, meaning ranges expressed by verbs for 'die' in African languages may also be found in Australia or the Americas (Felix Ameka, p.c.), and much the

⁴ Greenberg (1983:16) says however that this is not always so: "The most conspicuous exception is the Grasslands languages where the form *bep* or the like is found in many languages with the meaning 'meat' while the *nama* root survives as 'animal'. It would seem that this fact does not invalidate the hypothesis of a directionality 'meat' > 'animal'; rather, it might suggest that – for whatever reasons - an earlier meaning 'meat' received a new form of expression.

same applies to polysemy involving 'hear' (see e.g. Evans and Wilkins 1998 for evidence on Australian languages).

Another area where Africa provides a wide range of common properties concerns grammaticalization processes, whereby the same conceptual schemas and constructions are employed to develop grammatical categories. Perhaps the most widely discussed example concerns comparative constructions based on what in Heine (1997) is called the Action Schema, taking either of the forms [X is big defeats/passes Y] or [X defeats/passes Y in size], i.e., the use of a verb meaning either 'defeat', 'surpass' or 'pass' to express comparison (Meeussen 1975: 4; Greenberg 1983: 4; Gilman 1986: 39). To be sure, this contact-induced grammaticalization occurs also in other parts of the world, for example in Sinitic languages, Thai, Vietnamese, Hmong and Khmer, where a verb for 'to cross' has given rise to a standard marker of comparison (Ansaldò 2004: 490ff.), but outside Africa it is extremely rare, while roughly 80 % of the African languages have it (see table 1); we will return to this issue in section 2.4.

Furthermore, there is a grammaticalization process involving verbs for 'say' which are widely grammaticalized to quotatives, complementizers, purpose clause markers, etc. (Larochette 1959; Meeussen 1975: 3; Gilman 1986: 44; Güldemann 2001). However, this grammaticalization appears to be also fairly common outside Africa (see Ebert 1991; Heine & Kuteva 2002).

Body part terms used metaphorically for deictic spatial distinctions are found throughout the world; for example, nouns for the body part 'back' are the conceptual source for spatial terms for 'behind' in most languages. But this general grammaticalization process appears to be more common in Africa than elsewhere, and there are some developments that are likely to happen in Africa but unlikely to happen elsewhere (Meeussen 1975: 3; Gilman 1986: 42). Such developments include, but are not confined to, the grammaticalization of body parts for 'stomach/belly' to spatial concepts for 'in(side)', or of 'buttocks/anus' to 'below' and/or 'behind' (Heine 1997: 37ff.). Furthermore, sex distinctions used for the grammaticalization of the spatial concepts 'right' (< 'male, strong hand') and 'left' (< 'female, weak hand') have been proposed as pan-African features (Gilman 1986: 42), but such metaphorical transfers are by no means confined to Africa.

Further grammaticalization processes widespread in Africa involve the use of nouns for 'man' and 'woman' as attributive or derivational markers for sex distinctions (cf. Gilman 1986: 42), whereby e.g. the noun for 'girl' is historically a 'woman child' and 'bitch' a 'woman dog'. Finally, the grammaticalization of nouns for 'body' to reflexive markers has also been proposed as characterizing common African conceptualization processes (Gilman 1986: 42; Heine 2000) but, once again, this is a process that is by no means restricted to Africa.

3.3 Conclusion

The properties that have been discussed in this section may have given an impression of the kind of structural characteristics to be expected in African languages. It would seem that they can be classified into the following categories:

(8) Properties that seem to be essentially restricted to Africa:

- a clicks,
- b labial flaps,

- c several types of vowel harmony,
- d ideophones expressing color distinctions,
- e case inflections expressed exclusively by tone (so far only found in African languages, all of the marked nominative type),
- f lack of obligatory agreement of transitive verbs with their object.

(9) Properties that are distinctly more common in Africa than elsewhere. These are properties that are typologically remarkable, but many of them are in their occurrence either genetically or areally restricted:

- a labial-velar stops,
- b implosives, which Clements and Rialland define as non-obstruent stops,
- c ATR-based vowel harmony,
- d word-initial prenasalized stops,
- e noun class systems,
- f marked-nominative case systems,
- g marking negation at the end of the clause,
- h logophoric pronouns (indicate coreference of a nominal in the non-direct quote to the speaker encoded in the accompanying quotative construction, as opposed to its non-coreference indicated by an unmarked pronominal device),
- i focus marking by means of verbal inflections,
- j SOVX as a basic word order (where the direct object precedes while the indirect object and adjuncts follow the verb).

On the other hand there are also linguistic features that occur in other parts of the world but are hard to find in Africa. Thus, so far only four African languages (Shilluk, Pări, Anywa, and Jur-Luwo) have been found to show an ergative organization. Furthermore, while noun class systems are more common in Africa than elsewhere in the world, languages with noun classifiers are comparatively rare, systems such as the genitival classifiers of the Ubangian language Dongo-ko and the numeral classifiers of the Cross-River language Kana being exceptions. And finally, no clear cases of polysynthetic or noun incorporating languages have so far been found in Africa.

Still, in spite of all the work that has been done on Africa as a linguistic area, there is no entirely convincing evidence to answer the question raised in the title of this paper, for the following reasons. First, although there is some fairly comprehensive information on the areal distribution of some of the properties dealt with above (see e.g. Greenberg 1983; Heine 1976), we lack corresponding information on languages in other parts of the world in order to determine whether, or to what extent, we are really dealing with Africa-specific structures. Second, there is a genuine problem that any project aimed at defining Africa as a linguistic area is confronted with – one that has been described appropriately by Greenberg in the following way:

Ideally, if what is meant by an African areal characteristic is one which is found everywhere in Africa but nowhere else, then clearly none exists [...]. (Greenberg 1983: 3)

What this means is that it does not seem to be possible to define Africa as a linguistic area in the same way as, e.g., Meso-America has been defined (Campbell et al. 1986), that is, in accordance with the characterization proposed in (4), more specifically in terms of a set of

linguistic features that are not found outside that area (see (4b) and (4c)). In the remainder of this paper we will argue, however, that there nevertheless is a way of approaching this general issue.

4 A survey

Being aware that it does not seem possible to find a set of properties that clearly separate Africa from the rest of the world in accordance with (4), we decided to use an alternative approach. Following Greenberg (1983), we selected a set of eleven properties or characteristics and we asked colleagues working on African languages to provide information on the presence vs. absence of these properties in the language or languages studied by them. In this way we received information on 99 African languages. This sample is neither genetically nor areally entirely balanced but represents all major genetic groupings of Africa: Of the 99 languages, 55 belong to the Niger-Congo, 23 to the Afroasiatic, 15 to the Nilo-Saharan, and 6 to the Khoisan phylum. It also includes all major regions with the exception of north-central Africa, which is clearly underrepresented.

Choice of properties was determined by the following considerations. We were aiming at finding phenomena that are likely to set Africa off from other parts of the world. Accordingly, we chose properties that previous authors had claimed to be widespread in Africa but to be less so elsewhere in the world, that is, a range of properties discussed in section 2.3. But a number of these properties turned out to be unsuitable for our survey, either because there is lack of appropriate information on them in many of the languages concerned or because we suspected that their distribution might be genetically motivated. In the end we were left with eleven properties that could be expected to be relevant for an areal analysis.

4.1 The data

Table 1 lists the eleven properties used in the survey as well as the overall results of the survey, namely the relative frequency of occurrence of these properties. What it suggests is that the properties are roughly of three kinds: First, there are some properties (3, 5, 6, 7, 8, and 10) that occur in at least two thirds of the African languages of our survey; second, there are properties that are found in a minority of African languages (1, 2, 4, and 9); and finally, there is one property (11) that is found in roughly every second African language.

Table 1. Relative frequency of occurrence of 11 typological properties in African languages (Sample: 99 languages. Parameters 3, 7, and 8 have two options, A and B; if one of the options applies this is taken as positive evidence that the relevant property is present).

Property used as criteria	Number of languages having that property	Percentage of all languages
1 Labial-velar stops	39	39.4 %
2 Implosive stops	36	36.4 %
3 Lexical (A) and/or grammatical tones (B)	80	80.8 %
4 ATR-based vowel harmony	39	39.4 %

5 Verbal derivational suffixes (passive, causative, benefactive, etc.)	76	76.7 %
6 Nominal modifiers follow the noun	89	89.9 %
7 Semantic polysemy 'drink (A)/pull (B), smoke'	74	74.7 %
8 Semantic polysemy 'hear (A)/see (B), understand'	72	72.7 %
9 Semantic polysemy 'animal, meat'	40	40.4 %
10 Comparative constructions based on the schema [X is big defeats/surpasses/passes Y]	82	82.8 %
11 Noun 'child' used productively to express diminutive meaning	50	50.5 %

To test whether these properties are in fact characteristic of African languages, we asked experts of non-African languages and received information on an additional 50 languages. The results of the survey are summarized in table 2. What this table suggests is the following:

- (a) Africa stands out against other regions of the world in having on average 6.8 of the eleven properties, while in other regions clearly lower figures are found.
- (b) Outside Africa, no language has been found to have as many as five properties, while African languages have between five and ten properties. There are a few exceptions, to be discussed below.
- (c) While the African area can be set off from the rest of the world, it seems that there is also a worldwide north/south division: Languages of the southern hemisphere have clearly more of the properties than languages of the northern hemisphere.

Table 2. Distribution of 11 typological properties according to major world regions (Sample: 99 African and 50 non-African languages).

Region	Total of languages	Total of properties	Average number of properties per language
Europe	10	11	1.1
Asia	8	21	2.6
Australia/Oceania	12	37	3.0
The Americas	14	48	3.4
Africa	99	669	6.8
Pidgins and creoles ⁵	6	14	2.3
<i>All regions</i>	<i>149</i>		

⁵ Three of the six pidgin and creole languages are spoken in Africa and the rest in the Americas and in New Guinea.

A slightly different picture emerges if one draws a line within Africa, separating sub-Saharan from northern Africa. With northern Africa we refer to Afroasiatic languages with the exception of the Chadic branch, that is, it includes Ethio-Semitic, Cushitic, Omotic, and Berber languages. Accordingly, sub-Saharan Africa includes Chadic as well as all languages of the other three language phyla. As table 3 suggests, this distinction is justified on account of the distribution of properties: Whereas northern Africa does not behave much different from other parts of the world, exhibiting similar figures as e.g. the languages of Australia and Oceania, it is sub-Saharan Africa that stands out typologically, with an average figure of 7.2 properties.

Table 3. Distribution of 11 typological properties: Sub-Saharan Africa vs. rest of the world (Sample: 99 African and 50 non-African languages. Northern Africa = Afroasiatic minus Chadic; sub-Saharan Africa = Africa minus northern Africa).

Linguistically defined region	Total of languages	Total of properties	Average number of properties per language
World minus Africa	47	119	2.6
North-eastern Africa	13	46	3.7
World minus sub-Saharan Africa	60	165	2.8
Sub-Saharan Africa	86	635	7.2

4.2 Isopleth mapping

Isopleth mapping is a technique that has been employed in linguistic areas whose status is fairly uncontroversial, such as South Asia (Masica 1976), the Balkans (van der Auwera 1998), and Meso-America (van der Auwera 1998). Isopleth maps are designed on the basis of the relative number of features that languages of a linguistic area share: Languages having the same number of properties, irrespective of which these properties are, are assigned to the same isopleth and, depending on how many properties are found in a given language, the relative position of that language within the linguistic area can be determined.

What isopleth maps achieve is that they show the geographical distribution of the relative number of features making up a sprachbund. For example, on the basis of ten features characteristics of the Balkan languages, van der Auwera (1998: 261-3) finds that Bulgarian is the most central Balkanic language, being "included in all isoglosses" i.e., showing all ten Balkanic features (for a discussion of isopleths maps, see Heine & Kuteva 2006).

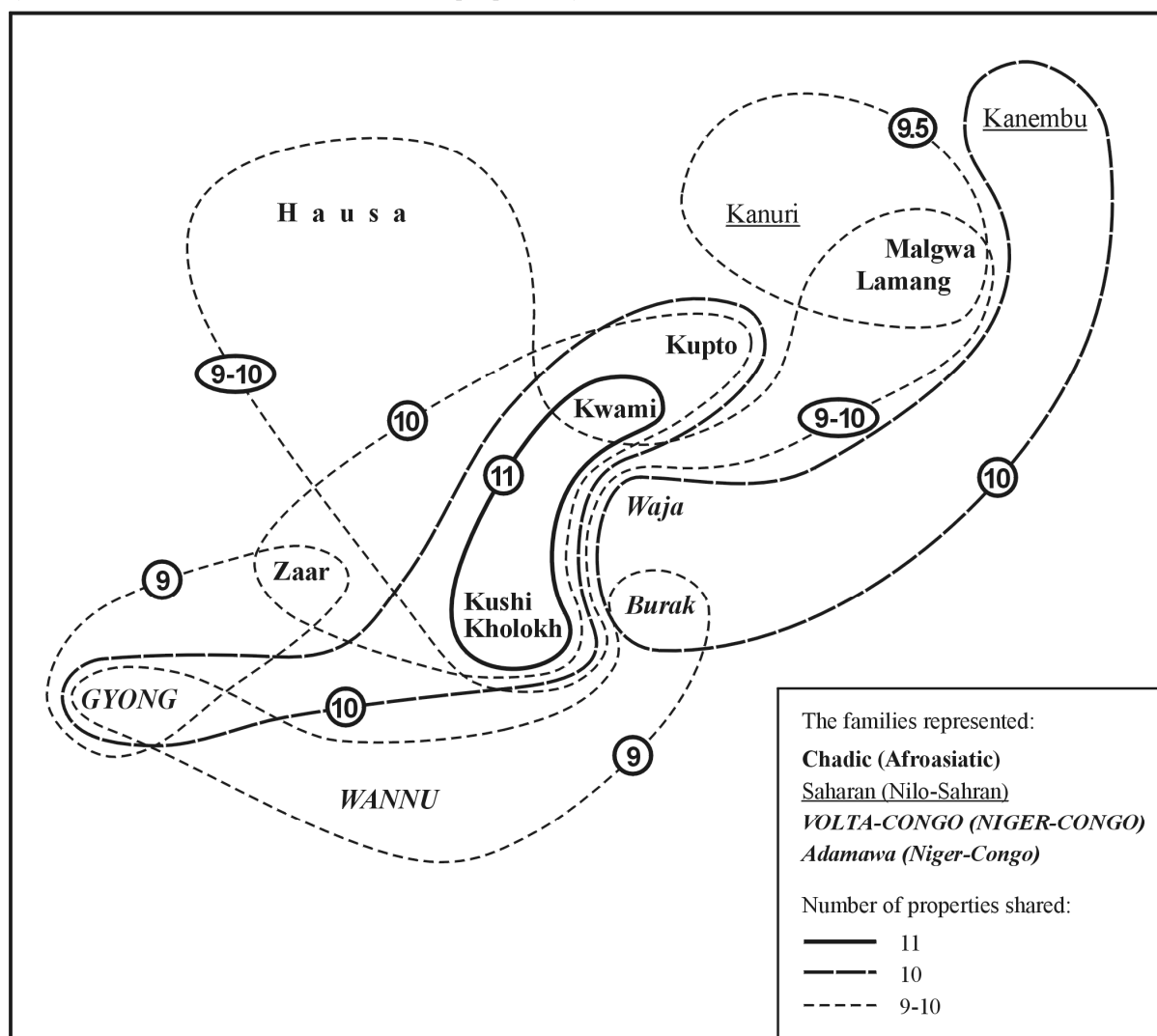
Applying isopleth mapping to Africa yields the following results: The most inclusive languages are Western Chadic, Gur (Voltaic), some Plateau and Guang languages, having 9 to 10 of the 11 properties considered. A secondary isopleth center is found in the Cameroon-Central Africa area, where up to 9 properties are found. Clearly less central are languages further to the west and south, that is, Atlantic and Mande languages on the one hand, and Bantu languages on the other, where around 6 properties are found. Peripheral Africa consists of the Ethiopian Highlands on the one hand, and northern (Berber) Africa, where less than 5 properties are found.

Isopleth research in general and in Africa in particular is far from encouraging, for the following reasons: First, what it achieves is roughly what one would expect without drawing on a quantitative technique: Languages spoken in the center of the area are likely to show the largest number of isopleths, and thus to be most central to the linguistic area concerned, and

3 Hausa	Chadic	7	8.5											
4 Kholokh	Chadic	7	7.5	9										
5 Kupto	Chadic	7	7.5	9	11									
6 Kushi	Chadic	6.5	7	9.5	10.5	10.5								
7 Kwami	Chadic	7	7.5	9.5	11	11	10.5							
8 Lamang	Chadic	8	9.5	9	8	8	7.5	8						
9 Malgwa	Chadic	7	9.5	10	9	9	8.5	9	10					
10 Zaar	Chadic	6	6.5	8	10	10	9.5	10	7	8				
11 Gyong	Plateau	7.5	8	8	9.5	9.5	9	9.5	6.5	7.5	9			
12 Burak	Adamawa	9/ 10	6.5/ 10	5/ 10	7/ 10	7/ 10	6.5/ 10	7/ 10	5/ 10	5/ 10	7/ 10	8/ 10		
13 Waja	Adamawa	10.5	8.5	7.5	7.5	7.5	7	7.5	7.5	7.5	6.5	8.5	8.5/10	
14 Wannu	Jukunoid	8.5	7	5.5	7.5	7.5	7.5	7.5	5.5	5.5	8.5	8.5	9.5	8

Map 1. A sketch map of northern Nigerian languages: Isoglosses of the number of shared typological properties (Encircled numbers = numbers of shared properties).

A sketch map of northern Nigerian languages: Isoglosses of the number of shared typological properties. (Encircled numbers = number of shared properties.)



The results of table 4 are presented in the form of an isopleth structure in map 1. Considered are only shared figures of nine or more properties on the basis of data presented in table 4. The overall picture that arises from this map yields one important finding: It suggests that the distribution of typological properties is not determined primarily by genetic relationship. While there are some genetic clusterings, combining e.g. the Chadic languages Kwami, Kushi, and Kholokh (11 properties), or the Niger-Congo languages Gyong, Wannu, and Burak (9 properties), more commonly the isopleth lines cut across genetic boundaries. This is suggested by the following observations:

- (a) The Saharan language Kanuri shares more properties (9.5) with the Chadic languages Malgwa and Lamang than with the fellow Saharan language Kanembu.
- (b) The Saharan language Kanembu shares more properties (10) with the Adamawan languages Waja and Burak than with the fellow Saharan language Kanuri.
- (c) The Adamawan languages Burak and Waja share more properties with the Saharan language Kanembu than with their fellow Niger-Congo languages Gyong and Wannu.

(d) At the same time, Waja shares more properties with Kanembu (10.5) than with any other fellow Niger-Congo language.

While we do not wish to propose any generalizations beyond the data examined in this section, what these data suggest is that, on the basis of eleven properties used, areal clustering provides a parameter of language classification that is hardly less significant than genetic relationship.

5 Conclusions

The analysis of our survey data suggests that there is evidence to define Africa as a linguistic area: African languages exhibit significantly more of the eleven properties listed in table 1 than non-African languages do, and it is possible to predict with a high degree of probability that if there is some language that possesses more than five of these eleven properties then this must be an African language. The data also allow for a number of additional generalizations based on combinations of individual properties. For example, if there is a language that has any two of the properties 1 (labial-velar stops), 2 (implosive stops), and 4 (ATR-vowel harmony), then this must be an African language.

Not all of the properties, however, are characteristic of Africa only; in fact, some are more common in other regions of the world. Property 5 (Verbal derivational suffixes) appears to be more common in the Americas than in Africa, property 11 (Noun 'child' used productively to express diminutive meaning) is as common in South America as it is in Africa, and property 6 (Nominal modifiers follow the noun) is equally common in the Americas and Australia/Oceania. What is relevant to our discussion is not the distribution of individual properties but rather the combination of these properties, where the African continent clearly stands out against the rest of the world on the basis of the eleven properties examined.

What this means with reference to (4) is that our characterization of linguistic areas needs to be revised to take care of the quantitative generalizations proposed in section 2.4, by rephrasing (4c) in the following way: This set of properties is not found at a comparable quantitative magnitude in languages outside the area.

The survey data presented are also of interest with reference to an issue concerning the genesis and explanation of creole languages. One of the main hypotheses advocated by students of these languages has it that the structure of creole languages, in particular of Atlantic and Indian Ocean creoles, can be explained at least in part with reference to substrate influence from African languages, more specifically from languages spoken along the West African coast (see e.g. Boretzky 1983; Holm 1988), and in its strongest form this hypothesis maintains that creole languages such as Haitian Creole have the structure of African languages, especially of Fon (Fongbe), with a European superstrate grafted on (see especially Lefebvre 1998). While we are not able to assess this hypothesis here, our data do not lend any support to such a hypothesis: With the exception of the Portuguese-based creole Angolar (Maurer 1995), creole languages do not exhibit any noticeable typological affinity with African languages on the basis of our survey data (see table 2).

To offer a diachronic interpretation of the results presented would be beyond the scope of this paper. An attempt in this direction has been made by Greenberg (1983), whose main goal was to identify sources for the spread of four areal properties in Africa (see section 2.2). In that

paper he argued that labial-velar stops (our property 1) originated in Niger-Congo and then diffused into Chadic and Central Sudanic languages, and he suggested that comparatives based on the Action Schema (our property 10) are of Niger-Congo origin (1983: 15). In a similar fashion, he found evidence for a Niger-Congo origin for the 'meat'/'animal' polysemy (1983: 18; our property 11). On the basis of our data there is nothing that would contradict these reconstructions. But there is also an alternative perspective to this situation.