Studies in African Linguistics
Volume 45, Numbers 1&2, 2016

A PHYLOGENETIC ANALYSIS OF STABLE STRUCTURAL FEATURES IN WEST AFRICAN LANGUAGES*

Kristoffer Friis Bøegh, Aymeric Daval-Markussen & Peter Bakker
Aarhus University

Lexical comparison has long dominated the study of West African language history. Approaching the subject from a different perspective, this paper compares a sample of West African languages based on a selection of typological features proposed to be temporally stable and hence possible markers of historical connections between languages. We utilize phylogenetic networks to visualize and compare typological distances in the language sample, in order to assess the extent to which the distributional properties of the selected features reflect genealogy, areality, or no plausible historical signal. Languages tend to cluster in accordance with genealogical relationships identified in the literature, albeit with a number of inconsistencies argued to reflect contact influences and chance resemblances. Results support the contention that typology can provide information about historical links between West African languages.

Keywords: typology, historical linguistics, West African languages

1. Introduction

Northern sub-Saharan western Africa ("West Africa") is known for its great linguistic diversity, and also for its unclear linguistic past. Dating back to the 19th century, lexical evidence has predominated in the comparative study of the region's languages (e.g. Koelle 1854; Westermann 1927; Greenberg 1963), while areal relationships have played a minor role in the reconstruction of language history (see Heine & Kuteva 2001). West Africa is characterized by a wealth of widespread lexical and typological features, often shared within areas, which suggest that genealogical connections do not coincide with their distribution (see Heine & Nurse 2008).

This study presents a comparison of languages of West Africa from a typological perspective based on phylogenetic network analysis. In recent years, the application of computational classificatory techniques, i.e. phylogenetic methods, carried over from biology has increased and diversified greatly in historical linguistics, in studies of family relationships and areal phenomena (see e.g. Gray and colleagues 2003, 2009, 2011; McMahon & McMahon 2003; Kitchen et al. 2009; Walker & Ribeiro 2011, to mention a few). Most linguistic phylogenetic studies have used cognate judgments (e.g. of basic vocabulary items) as the basis for analysis, but in other cases, espe-

* We wish to thank the editor and the reviewers for their valuable comments and suggestions for the paper. Acknowledgements for assistance with feature score checking are due to Sergio González de la Higuera Rojo, Yonatan Ungermann Goldshtein, and especially Robert D. Borges, who also contributed additional data for Songhay, Mande, Berber, and Semitic. Parts of this research were sponsored through the "Cognitive Creolistics Project" (2013-2015), VELUX Project 31971.

cially where processes of internal and external language change have rendered lexical cognates difficult to assess, linguists have turned to cluster analysis of typological features (i.e. structural linguistic properties unconstrained by formal correspondences) for modeling linguistic macro-history. They assert that abstract structures, not only linguistic forms, are subject to processes of vertical and horizontal transfer which create a historical signal (see e.g. Nichols 1992; Wichmann & Saunders 2007; Dunn et al. 2005, 2008, 2011; Donohue & Musgrave 2007; Donohue et al. 2008; Greenhill et al. 2009, 2010; Reesink & Dunn 2012; Dediu & Cysouw 2013; Wichmann 2015).

We compare a sample of West African languages on the basis of phylogenetic network analysis of character data for a selection of 30 features of phonology, morphology, and syntax from the *World Atlas of Language Structures* (WALS; Dryer & Haspelmath 2013; earlier version Haspelmath et al. 2005). The features chosen for our sample were assessed by Wichmann & Holman (2009) to be relatively stable, or conservative, in intergenerational transmission; as such, their distributional properties may retain traces of historical linguistic connections of phylogeny and interaction that can be modeled with computational clustering techniques. We utilize the algorithm NeighborNet (Bryant & Moulton 2004) in SplitsTree v. 4.13.1 (Huson & Bryant 2006) to quantify typological distances between the languages under comparison and to draw splits graphs of resulting clusterings in order to explore the extent to which continuation of the selected features appears to reflect inheritance, patterns of language contact, or no apparent historical signal.

2. The languages of West Africa: classification and areal relationships

In our comparisons, we consider genealogical as well as areal relationships of West African languages when discussing positions of languages in network graphs. To set the scene, this section briefly reviews methods, traditions, and current perspectives of African comparative linguistics. For an extensive discussion on the methods and results of African historical-comparative linguistics, we refer to Campbell & Poser (2008: 120-145).

2.1. Language classification. While the Comparative Method of linguistic reconstruction has been applied at a micro-level to specific subgroups, it has not been comprehensively applied in the study of African linguistic macro-history. Nurse (1997: 262-3) attributes this fact to a combination of reasons: the vast number of African languages (estimated to be over 2,000 in Lewis et al. 2016), the state of documentation for many of these, and the time depth of their diversification and interaction. The upper limit of the Comparative Method is commonly held to be around 8,000 to 10,000 years, before evidence for relatedness fades out when the lexicon is investigated (e.g. Nichols 1992). African language families are of considerable age, however (compare, for instance, the hypothetical dispersal dates of African language families put forward in Blench 2013: 62). Africanists have made use of a range of alternative classificatory methods. Some early studies ordered languages into families on the basis of typological criteria, such as presence or absence of noun class systems in contiguous language groups (for an example of a critique of such typological classifications, see Greenberg 1963: 1). The current understanding of West African linguistic relationships has, however, come mainly from lexical evidence, i.e. from identification of (mostly lexical) shared innovations, (multilateral) form-meaning comparison of lexical items, noun class markers, and pronominals, as well as lexicostatistical quantification of cognacy in vocabulary items (e.g. Greenberg 1963; Bennett & Sterk 1977; Bendor-Samuel 1989; Heine & Nurse 2000).

The first attempts at placing languages of West Africa into related groupings took place in the 19th century. Koelle (1854) grouped languages together based on resemblances in vocabulary, starting a long tradition of comparative studies (such as Westermann 1911, 1927) based mostly on lexical evidence. Based on multilateral lexical comparison, the influential classification of Greenberg (1963) grouped Africa's languages into four macro-phyla: Afro-Asiatic, Niger-Kordofanian (=Niger-Congo), Nilo-Saharan, and Khoisan. Greenberg's major conclusions became widely accepted in African historical linguistics, and the classification continues to serve, directly or indirectly, as the starting point for current considerations of most African family relationships.[1]

However, Greenberg's work also made linguists call into question what constitutes a proper methodological basis for classifying languages, and the question whether the prominence of large, genealogically undiverse African families reflects reality or rather an intellectual tradition of 'lumping' classification remains a topic of debate. The view that Africa in fact is home to greater genealogical diversity than what has traditionally been recognized is currently gaining ground. Still, West African linguistic relationships are not delineated to a degree where an agreed classificatory scheme can be presented (for diverging classifications, see e.g. Heine & Nurse 2000; Blench 2006; Sands 2009; Dimmendaal 2011). The internal structures and/or external affiliations of language groups such as Atlantic, Mande, Dogon, Ijoid, Adamawa, Ubangi, Central Sudanic, and Songhay continue to be debated, as do the affiliations of a number of individual languages.

2.2. Areal relationships. For a long time, mapping of contact patterns in West Africa was not prioritized as a way to achieve progress in language classification (Heine & Kuteva 2001), but recent macro-areal typological studies have, building on the pioneering studies of Greenberg (1959, 1983) and Heine (1975, 1977), provided evidence in favor of long-term contact to account for several current linguistic areal patterns in West Africa (see Heine & Nurse 2008). This led to the idea that the West African region forms part of a linguistic contact area, variously referred to as the 'Sudanic belt' (Clements & Rialland 2008) or, more often, the 'Macro-Sudan belt' (Güldemann 2008, 2010). The Macro-Sudan belt can be defined as a broad zone stretching from south of the Sahara-Sahel to north of the Congo Basin and spanning the continent from the Atlantic Ocean in the west to the Ethiopian Plateau in the east.

The central feature of a contact area is the existence of linguistic similarities shared among unrelated languages of a geographical area. Features that are widely shared across West African languages, and across family boundaries, span the linguistic system and include, among other features, lexical and grammatical tone systems, marked phonological features (i.e. ATR vowel harmony, labiovelar stops, labiodentals flaps, and voiced implosive stops), as well as a number of lexical and grammatical polysemies, verbal derivational suffixes, logophoric pronouns, and word order patterns (for an in-depth discussion of shared linguistic features in West Africa, see Heine & Leyew 2008).

A question related to the observation that some features are widespread across the region is whether convergence has obscured phylogenetic signals; one can also ask whether this question has different answers for different participating groups of languages. Güldemann (2008: 168) posits that the Macro-Sudan belt is structured as concentric circles of varying degrees of shared

1 Research since the 1960s has led to substantial revisions in the classification's proposed family units, however. An example of a group which has undergone extensive changes is Kwa of Niger-Congo. In the (New) Kwa of recent classifications, Ijoid, Kru, and the branches now subsumed under West Benue-Congo, all of which were formerly included, have been excluded (e.g. Williamson & Blench 2000: 29).

linguistic features. In his view, languages spoken across West Africa belong to an overall 'core' or 'periphery' zone of historical contact influence (Güldemann 2010: 568). The core zone is characterized by the language groups Atlantic, Mande, Kru, Gur, Kwa, Benue-Congo (except Narrow Bantu), Adamawa, Ubangi, Bongo-Bagirmi, and Moru-Mangbetu, whereas the peripheral zone comprises Dogon, Songhay, Chadic, Ijoid, Narrow Bantu, and Nilotic.

3. Methods and sampling

This section presents our choice of a phylogenetic algorithm for visualizing language connections, the sampling of structural features, and the language sample and data sources.

3.1. Phylogenetic algorithms. Phylogenetic clustering algorithms were originally developed in order to make inferences about evolutionary relationships between biological species according to their observed characteristics (such as presence of feathers, or DNA sequences; see Felsenstein 2004 for an overview). Computational phylogenetic methods have since been carried over to other branches of science, including historical-comparative linguistics, inspired by a number of conceptual parallels between biological diversification and language development (see Atkinson & Gray 2005: 514).

Phylogenetic analysis allows both hypothesis testing and generation, and opens up for a renewed exploration of language history and diversity based on different kinds of encoded language data (for overviews of various applications of phylogenetic techniques in historical-comparative linguistics, see Nichols & Warnow 2008; Bowern & Evans 2015; see also Dunn 2015 for an introduction to the mathematical procedures underlying the algorithms most commonly used in linguistic studies).

In biology, network models of evolution which do not assume a bifurcating pattern of diversification are used to trace reticulations in the historical development of species (Bryant & Moulton 2004). Networks make it possible to visually capture the tangled webs of complex linguistic relationships, and, as such, allow assessment of the degree of horizontal transfer between a set of languages under comparison. Networks draw a tree-like output graph where data are divergent, while interrelationships among characters are represented as web formations, suggestive of parallel development or contact-induced transfer. For these reasons, network-oriented models are well-suited for language comparison, especially so of languages used in parts of the world where contact has played an important role in language history.

We utilize the distance-based agglomerative clustering algorithm NeighborNet (Bryant & Moulton 2004) in SplitsTree v. 4.13.1 (Huson & Bryant 2006) to quantify typological distances between taxa (representing languages) and visualize their clusterings in network graphs. Quantification of typological distances means that a set of languages under comparison can be treated as more or less structurally similar to each other based on shared and non-shared values in a matrix of encoded typological features. The algorithm produces unrooted networks with taxa distanced in relation to each other by splits between branches and nodes, with branch lengths being proportional to the amount of divergence between taxa. Unlike a compromised tree graph where nodes can have only one parent, NeighborNet computes and depicts the total possible trees contained in a dataset, with conflicts in the signal being represented as additional edges (webbing) in the output network.

It is important to note that when applied to a compatible character matrix, network-oriented phylogenetic methods will produce a graph of connected taxa, regardless of the nature of the provided data. This means that a network graph can give the impression of depicting, for instance, a family relationship, where it actually represents system similarity, the source of which requires interpretation. Typological similarity in language development can be accounted for by the following criteria: 1) shared inheritance; 2) contact influences; 3) chance convergence, including universal tendencies and homoplasies due to a limited typological design space (see Croft 2008: 230; Reesink & Dunn 2012).

3.2. Structural phylogenetics and sampling of features. Language comparison based on phylogenetic analysis of purely structural features began with Dunn et al. (2005, 2008) who showed that despite prolonged contact influence between the well-established Western Oceanic subgroup of the Austronesian family and the group of various Papuan languages of the same region of Melanesia, typological data could be used to obtain groupings corresponding roughly with families established by previous comparative research based on lexical cognates. Other families that have been studied with phylogenetic methods on the basis of typological features include Indo-European (Longobardi et al. 2013), Bantu (Petzell & Hammarström 2013), Arawak (Carling et al. 2013), among others. Structural phylogenetic analysis thus opens up possibilities of exploring whether comparisons based on different levels of the linguistic system (e.g. lexicon, abstract structural features) consistently point in the same direction concerning West African linguistic relationships.

Moreover, Dunn et al. (2005: 2075) suggest that typological features may retain a phylogenetic signal beyond the current ceiling on the reconstruction of language history, opening up the possibility of uncovering otherwise undetectable linguistic relationships. In a critical analysis of the methods and results of Dunn et al. (2005), Donohue & Musgrave (2007) agree that cluster analysis based on structural features opens up for investigation of linguistic macro-history, but they contend that structural features, like lexicon, can diffuse, making it difficult to assert whether the source of typological similarities is inheritance or diffusion (see also Donohue et al. 2008; Greenhill et al. 2009, 2010; Dunn et al. 2011; Dediu & Levinson 2012; Reesink & Dunn 2012). Advances in research on differential stability in typological features may extend our ability to look into language history, however. If some features can be defined as intrinsically stable to change, these can turn out, depending on the set of languages under comparison, to be well-suited indicators of family relationships or, alternatively, patterns of historical contact (Wichmann 2015: 221).

A number of proposals have been made as to which typological features are the most stable through time (e.g. Maslova 2004; Cysouw et al. 2008; Parkvall 2008; Wichmann & Holman 2009; Dediu 2011; Dediu & Cysouw 2013). While their exact results differ, the various studies do show overall agreement as to which features, or areas of grammar, belong to either the more or the less stable category (for an evaluation, see Dediu & Cysouw 2013).

We follow the list of stable WALS features devised by Wichmann & Holman (2009). They identified a total of 33 out of 134 non-redundant WALS features as being 'very stable' over time, with 'stability values' (SV) ranging from 50.6% to 80.8% according to their metric C. In theory, the SV range from 100% for the most stable to -100% for the least stable, but in practice from 80.8% (feature 31A 'Sex-based and Non-sex-based Gender Systems') to -24.9% (feature 58A 'Obligatory Possessive Inflection'). We retained the features identified as very stable, with the following modifications to the selection.

Interdependent features were filtered as much as possible from the data set, as to minimize potential bias towards a particular area of grammar, and replaced with other features identified as 'stable', i.e. features with SV from 34.1% to 48.3%, which were part of the same areas of grammar as the original features (for a full overview of the feature selection, see Table 1). Hence, WALS feature 32A ('Systems of Gender Assignment') was omitted as it relates to gender systems and therefore overlaps with 31A ('Sex-based and Non-sex-based Gender Systems'). Features 81A ('Order of Subject, Object and Verb') and 84A ('Order of Object, Oblique, and Verb') were redundant because of the presence of 83A ('Order of Object and Verb'), since all three features relate to OV-word order. Instead, 82A ('Order of Subject and Verb'), the only remaining stable feature relating to word order, was included. Feature 87A ('Order of Adjective and Noun') was removed and replaced with 88A ('Order of Demonstrative and Noun') due to its strong cross-linguistic correlation with 83A ('Order of Object and Verb'). Feature 137A ('N-M Pronouns') was omitted as it relates to concrete shapes of morphemes, not typology. Moreover, features for which insufficient data were available were also replaced (those for which less than 25% of the information was initially known). As a result, features 21A, 42A, 47A, 61A, 79A, 99A, and 121A ('Exponence of Selected Inflectional Formatives', 'Pronominal and Adnominal Demonstratives', 'Intensifiers and Reflexive Pronouns', 'Adjectives without Nouns', 'Suppletion According to Tense and Aspect', 'Alignment of Case Marking of Pronouns', and 'Comparative Constructions', respectively) were removed and replaced with features 27A, 33A, 48A, 65A, 69A, 94A, and 104A ('Reduplication', 'Coding of Nominal Plurality', 'Person Marking on Adpositions', 'Perfective/Imperfective Aspect', 'Position of Tense-Aspect Affixes', 'Order of Adverbial Subordinator and Clause', and 'Order of Person Markers on the Verb').

The final sample of 30 features from the 'stable' and 'very stable' categories in Wichmann & Holman (2009) covers all areas of grammar, though not all areas are equally represented, see Table 1. For detailed feature descriptions, and for multistate feature values, see wals.info.

Table 1 Overview of 42 selected WALS features (the twelve removed features are marked with strikethrough)

ID	Description	SV	Areas of grammar from WALS
9A	The Velar Nasal	54.6%	Phonology
10A	Vowel Nasalization	57.0%	
18A	Absence of Common Consonants	55.3%	
~~21A~~	~~Exponence of Selected Inflectional Formatives~~	~~55.1%~~	Morphology
27A	Reduplication	36.2%	
28A	Case Syncretism	67.4%	
29A	Syncretism in Verbal Person/Number Marking	70.8%	
30A	Number of Genders	72.9%	Nominal Categories
31A	Sex-based and Non-sex-based Gender Systems	80.8%	
~~32A~~	~~Systems of Gender Assignment~~	~~66.9%~~	
33A	Coding of Nominal Plurality	41.3%	
39A	Incl./Excl. Distinction in Independent Pronouns	64.6%	
40A	Incl./Excl. Distinction in Verbal Inflection	65.0%	
~~42A~~	~~Pronominal and Adnominal Demonstratives~~	~~51.7%~~	
44A	Gender Distinction in Independent Personal Pronouns	50.6%	

~~47A~~	~~Intensifiers and Reflexive Pronouns~~	~~56.8%~~	
48A	Person Marking on Adpositions	40.6%	
57A	Position of Pronominal Possessive Affixes	55.0%	
~~61A~~	~~Adjectives without Nouns~~	~~54.2%~~	Nominal Syntax
63A	Noun Phrase Conjunction	54.3%	
65A	Perfective/Imperfective Aspect	36.0%	Verbal Categories
66A	The Past Tense	52.4%	
69A	Position of Tense-Aspect Affixes	47.3%	
73A	The Optative	56.7%	
~~79A~~	Suppletion According to Tense and Aspect	52.4%	
~~81A~~	~~Order of Subject, Object and Verb~~	~~53.3%~~	Word Order
82A	Order of Subject and Verb	35.7%	
83A	Order of Object and Verb	66.8%	
~~84A~~	~~Order of Object, Oblique, and Verb~~	~~55.1%~~	
85A	Order of Adposition and Noun Phrase	70.8%	
86A	Order of Genitive and Noun	65.3%	
~~87A~~	~~Order of Adjective and Noun~~	~~50.6%~~	
88A	Order of Demonstrative and Noun	42.4%	
89A	Order of Numeral and Noun	54.9%	
90A	Order of Relative Clause and Noun	54.5%	
94A	Order of Adverbial Subordinator and Clause	44.5%	
~~99A~~	~~Alignment of Case Marking of Pronouns~~	~~51.1%~~	Simple Clauses
104A	Order of Person Markers on the Verb	37.2%	
118A	Predicative Adjectives	74.3%	
119A	Nominal and Locational Predication	70.9%	
~~121A~~	~~Comparative Constructions~~	~~56.0%~~	
~~137A~~	~~N-M Pronouns~~	~~53.9%~~	Lexicon

3.3. Language sample and data collection. We compare a convenience sample of 94 genealogically and geographically diverse languages associated with West Africa, subsumed under the Afro-Asiatic, Niger-Congo, and Nilo-Saharan stocks.[2] Their approximate locations can be found on Map 1 and Map 2. Within Afro-Asiatic, the West and East Chadic, Berber, Biu-Mandara, and Semitic branches are represented. Within Niger-Congo, North and South Atlantic, Western and Eastern Mande, Dogon, Kru, Gur, Adamawa, Ubangi, Kwa, Ijoid, and the Benue-Congo branches Edoid, Nupoid, Yoruboid, Igboid, Platoid, Kainji, Cross-River, as well as the geographically diverse Bantoid groups A, B, C, E, F, G, H, K, L, M, P, R, S (Guthrie 1948), and Wide Grassfields are represented. Within Nilo-Saharan, Songhay, Western Saharan, and Bongo-Bagirmi are represented. The isolate Bangime and some individual languages with uncertain affiliations are also included.

[2] Languages are henceforth represented in maps and networks by lower case WALS codes, with the following upper case suffixes marking affiliations following WALS: ADM: Adamawa; BAN: Bantoid; BB: Bongo-Bagirmi; BER: Berber; BM: Biu-Mandara; CR: Cross-River; DOG: Dogon; ECHA: East Chadic; EDO: Edoid; EMAN: Eastern Mande; GUR: Gur; IGB: Igboid; IJO: Ijoid; ISO: isolate; KAI: Kainji; KRU: Kru; KWA: Kwa; NATL: North Atlantic; NUP: Nupoid; PLA: Platoid; SATL: South Atlantic; SEM: Semitic; SON: Songhay; UB: Ubangi; WCHA: West Chadic; WMAN: Western Mande; WSAH: Western Saharan; YOR: Yoruboid. ISO 639-3 codes rather than WALS codes are used for languages not featured in the WALS. See Table 2 for all language IDs.

The language sample was constructed with the aim of providing a testing ground for the stable feature analysis, but does not cover the region's full linguistic diversity. We have chosen to focus the sampling on languages belonging to the Macro-Sudan belt, and particularly Greenberg's (1959: 24) African 'core' area groups. The scope of the study is narrowed in that a number of relevant language groups for which data were not available to us were excluded from the sample. Language gaps include e.g. Kordofanian, Mangbetu, Gbaya-Manza-Ngbaka, and the isolate Laal.

Map 1 Approximate location of sampled languages, West Africa

Map 2 Location of sampled languages, Central, Eastern, and Southern Africa

Data collection for the languages included in the sample was initially achieved by looking up feature values in the WALS. As the data coverage for many of the African languages in WALS is low, we consulted a number of secondary sources, mostly published reference grammars or grammar sketches, in order to fill gaps in the data. At least two persons checked the score of each covered feature value. Data coverage is 86% in the final matrix, i.e. 2,441 out of 2,850 possible characters are scored (see Appendix).

Table 2 lists the languages of the sample, with information on linguistic affiliations, and sources consulted. Table 2 follows the broad classification of the WALS into genera (Dryer 2005), but we will also consider more specialized views in our comparison of networks.

Table 2 Language sample

Language	ID	Genus	Source(s)
Adamawa Fulfulde	fuaNATL	N. Atlantic	Stennes (1967); WALS
Akan	aknKWA	Kwa	Osam (1994); WALS
Angas	ancWCHA	W. Chadic	Burquest (1973); WALS; Wolff (1959)
Babungo	babBAN	Bantoid	Schaub (1985); WALS
Bagirmi	bagBB	Bongo-Bagirmi	WALS
Bambara	bamWMAN	W. Mande	Brauner (1974); WALS
Bangime	bgmISO	Isolate	Hantgan (2013)
Baule	bleKWA	Kwa	Timyan (1977); WALS
Bena	bluBAN	Bantoid	Morrison (2011); WALS
Birom	birPLA	Platoid	Bouquiaux (1970); WALS; Wolff (1959)
Bozo	bozWMAN	W. Mande	Daget et al. (1953)
Bushoong	bshBAN	Bantoid	Vansina (1959); WALS
Cicipu	awcKAI	Kainji	McGill (2009)
Dagbani	dgbGUR	Gur	Hudu (2014); Hyman & Olawsky (2004); WALS
Dan	danEMAN	E. Mande	Doneux (1968); Welmers (1973)
Defaka	defIJO	Ijoid	Bennett et al. (2012); Jenewari (1983); WALS
Dendi	ddnSON	Songhay	Zima (1976, 1998)
Diola	dioNATL	N. Atlantic	Sapir (1965); WALS
Doyayo	doyADM	Adamawa	Hewson (2010a); WALS
Efik	efiCR	Cross-River	Una (1900); WALS; Welmers (1973)
Ega	egaKWA	Kwa	Bole-Richard (1983); Salffner (2004); WALS
Engenni	egnEDO	Edoid	Thomas (1978); WALS
Ewe	eweKWA	Kwa	Duthie (1996); WALS
Ewondo	ewoBAN	Bantoid	Essono (1994); Redden (1979), WALS
Fon	fonKWA	Kwa	Lefebvre & Brousseau (2002); WALS
Fyem	fyePLA	Platoid	Nettle (1998); WALS; Wolff (1959)
Gã	gaKWA	Kwa	Kotey (1969); Kropp-Dakubu (2008); WALS
Grebo	grbKRU	Kru	WALS
Gude	gudBM	Biu-Mandara	Hoskison (1983); WALS
Gurma	grmGUR	Gur	Chantoux et al. (1968); WALS
Hausa	hauWCHA	W. Chadic	WALS
Hdi	hdiBM	Biu-Mandara	Frajzyngier (2002); WALS
Igbo	igbIGB	Igboid	Emenanjo (1987); WALS
Ijo-Kolokuma	ijoIJO	Ijoid	WALS; Williamson (1965)
Ikaan	iknEDO	Edoid	Salffner (2009)
Izi	iziIGB	Igboid	Meier et al. (1975); WALS
Jamsay	jmsDOG	Dogon	Heath (2005a); WALS
Jukun	jukPLA	Platoid	Nurse (2010); WALS; Welmers (1973); Wolff (1959)
Kana	kanCR	Cross-River	Ikoro (1994, 1996); WALS; Wolff (1959)

Kanakuru	knkWCHA	W. Chadic	WALS
Kande	kbsBAN	Bantoid	Grollemund (2006); WALS
Kanuri	knrWSAH	W. Saharan	WALS
Kera	kerECHA	E. Chadic	WALS
Kigiryama	nyfBAN	Bantoid	Lax (1996)
Kikuyu	kikBAN	Bantoid	Bergvall (1987); Mugane (1997); WALS
Kisi	kisSATL	S. Atlantic	Childs (1988, 1995); WALS
Kongo	konBAN	Bantoid	Dereau (1955); WALS
Korandje	kcySON	Songhay	Souag (2010)
Koromfe	kfeGUR	Gur	Rennison (1997); WALS
Koyra Chiini	kchSON	Songhay	Heath (1999b); WALS
Koyraboro Senni	kseSON	Songhay	Heath (1999a); WALS
Kpelle	kpeWMAN	W. Mande	WALS; Welmers (1973)
Lele	lelECHA	E. Chadic	WALS
Linda	lndUB	Ubangi	Cloarec-Heiss (1986); WALS
Luvale	luvBAN	Bantoid	Horton (1949); WALS
Makhuwa	muaBAN	Bantoid	van der Wal (2009); WALS
Mbum	mbmADM	Adamawa	Santandrea (1964); WALS
Mina	hnaBM	Biu-Mandara	Frajzyngier & Johnston (2005); WALS
Miya	miyWCHA	W. Chadic	Schuh (1998); WALS
Mooré	mooGUR	Gur	Nikiema (2001); WALS
Mpongwe	mpoBAN	Bantoid	Ambouroue (2007); WALS
Mumuye	mumADM	Adamawa	Shimizu (1983); WALS; Wolff (1959)
Ngambay	ngmBB	Bongo-Bagirmi	Ndjerareou et al. (2010); WALS
Ngbandi	ndiUB	Ubangi	Toronzoni (1989)
Ngizim	ngzWCHA	W. Chadic	Schuh (1972); WALS
Ngoni	ngoBAN	Bantoid	Ngonyani (2003); WALS
Nupe	nupNUP	Nupoid	Kandybowicz (2008); Kawu (2002); WALS
Nyamwezi	nymBAN	Bantoid	Maganga & Schadeberg (1992); WALS
Obolo	oboCR	Cross-River	Rowland-Oke (2003); WALS
Sena	senBAN	Bantoid	Funnell (2004); WALS
Shona	shnBAN	Bantoid	Fortune (1955); WALS
Shuwa	shuSEM	Semitic	Carbou (1913)
Soninke	snnWMAN	W. Mande	Diagne (2006); Diagana (1994)
Supyire	supGUR	Gur	Carlson (1994); WALS
Swahili	swaBAN	Bantoid	WALS
Tadaksahak	dsqSON	Songhay	Christiansen-Bolli (2010)
Tagdal	tdaSON	Songhay	Benítez-Torrez (2009)
Tamasheq	taqBER	Berber	Heath (2005b)
Tasawaq	twqSON	Songhay	Wolff & Alidou (2001)
Temne	tneSATL	S. Atlantic	Kamarah (2007); WALS
Tera	terBM	Biu-Mandara	Newman (1970); WALS
Tommo So	tmsDOG	Dogon	McPherson (2010); WALS
Tondi Song-way Kiini	tstSON	Songhay	Heath (2005c)
Tonga	tozBAN	Bantoid	Carter (2002); WALS

Tubu	tbuWSAH	W. Saharan	Le Coeur & Le Coeur (1956); WALS
Umbundu	umbBAN	Bantoid	Schadeberg (1990); WALS
Vai	vaiWMAN	W. Mande	WALS; Welmers (1976)
Wolof	wlfNATL	N. Atlantic	Pichl (1957); WALS
Yao	yaoBAN	Bantoid	Sanderson (1922); WALS
Yoruba	yorYOR	Yoruboid	WALS
Zande	zanUB	Ubangi	Hewson (2010b); WALS
Zarma	zarSON	Songhay	Bornand (2006); Sibomana (2008); WALS
Zenaga	zenBER	Berber	Faidherbe (1877); WALS
Zulu	zulBAN	Bantoid	Doke (1927); WALS

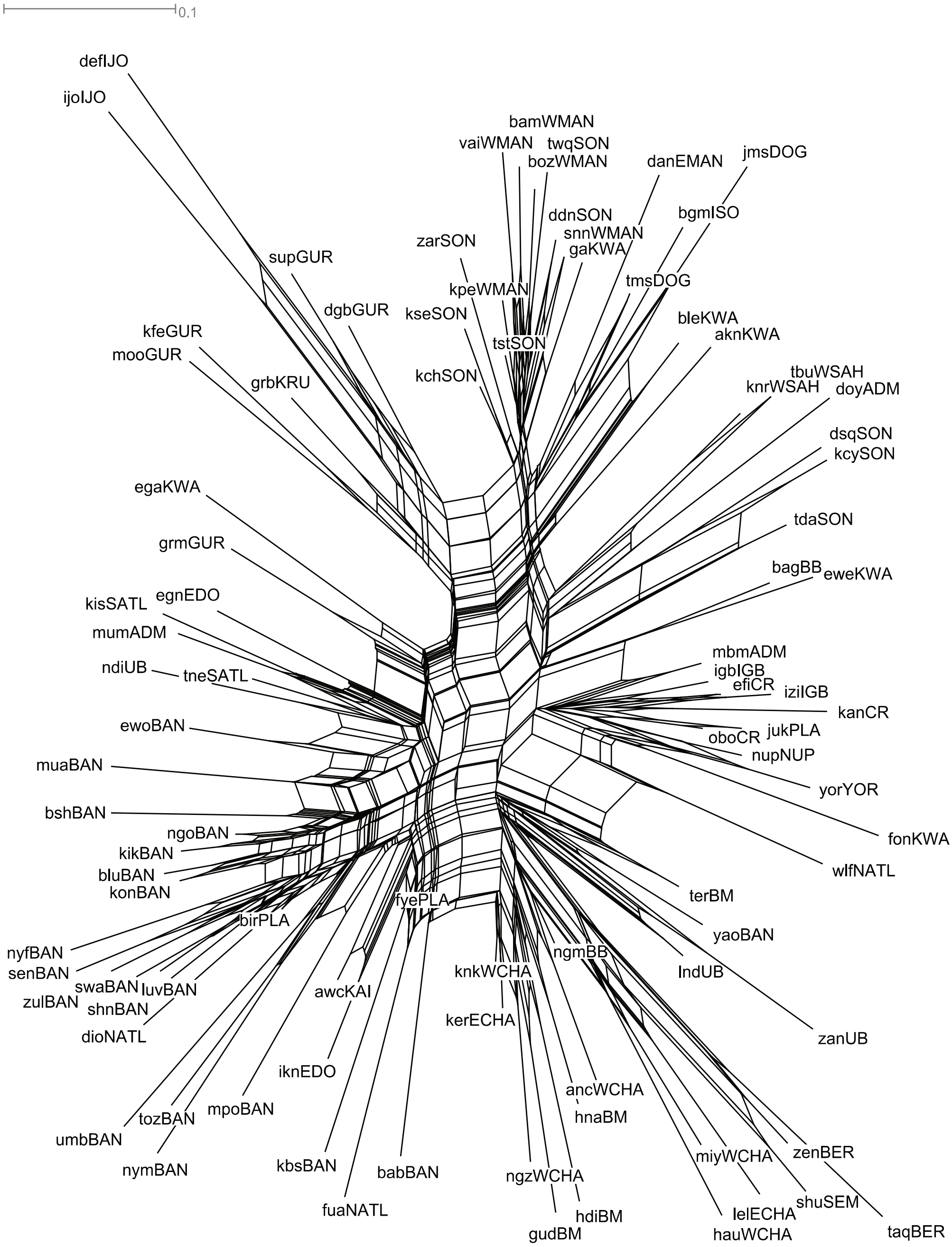
4. Results and discussion

The data were imported into SplitsTree, which produced graphic representations of the result. In the following, we present a series of network graphs generated from different subsets of the language sample, and discuss aspects of language clusterings. We first compare overall language connections, and we then place focus on Afro-Asiatic and 'outlier' language groups, i.e. groups aligned variously in the literature with Nilo-Saharan or Niger-Congo, or considered independent families. We then turn to comparing connections and interconnections in Volta-Congo (cf. Williamson & Blench 2000), i.e. Kru, Gur, Kwa, Benue-Congo, Adamawa, and Ubangi.[3] Finally, we compare networks depicting languages belonging to the core vs. periphery zones of the Macro-Sudan belt, respectively.

4.1. Overall language connections. Figure 1 depicts typological distances and connections of our maximal sample of 94 languages in a network graph. The Figure 1 network shows languages grouped into clusters of varying internal and external complexity in a web-like, interconnected structure, with a number of out-branchings.

3 Different labels for the unit include 'Volta-Congo' (Stewart 1976; Williamson & Blench 2000), 'Central Niger-Congo' (Bennett & Sterk 1977), and 'Narrow Niger-Congo' (Güldemann 2008: 176). Proposals vary with regard to the inclusion of Ubangi with these groupings.

Figure 1 NeighborNet of 94 West African languages



The distribution of languages in the network shows, on the one hand, that the linguistic landscape of West Africa is characterized by a high degree of variation and diversity. Niger-Congo languages, in particular, occupy a large part of the typological space. The reticulations in the network structure, on the other hand, indicate that some traits are widely shared between languages of different affiliations. Overall positions of languages in the network match well with Heine's typological classification based on word order patterns (see Heine 1975, 1977). Note, however, that only eight of our 30 features (27%) relate to word order patterns. The network displays a basic typological dichotomy between head-initial languages with noun classes, verbal extensions, and SVO word order, located in the lower part of the graph, and head-final languages without noun classes, a sparse distribution of verbal extensions, and SOV word order, placed in the top part of the graph.

The specific positions of a number of languages in Figure 1 are compatible with genealogical relationships identified in the literature to an extent that is unlikely to be attributable to chance convergence; see, for instance, the cluster of geographically diverse Bantoid (BAN) languages, located in the lower left part of the network. Several languages that are grouped together are, however, spoken in contiguous areas (e.g. Kru, Gur, Kwa, Yoruboid), and it is difficult to assess to what extent such clusterings are conditioned by geography or phylogeny. Areality appears to play a part in the distribution of some languages; for instance, the Bantoid languages in the sample that are geographically closest to the Macro-Sudan belt, i.e. Ewondo (ewoBAN), Babungo (babBAN), and Kande (kbsBAN), fall outside of the tight cluster of dispersed Bantoid languages. No evident overall areal pattern can be discerned, however; languages spoken in different areas are found in different parts of the network.

A number of languages do not pattern along areal lines, and also not in accordance with close family relationships, as the presence of e.g. North Atlantic Diola (dioNATL) of Senegambia among Bantoid languages reflects. The grouping of the diverse languages Engenni (egnEDO), Kisi (kisSATL), Mumuye (mumADM), Gurma (grmGUR), Ega (egaKWA), and Ngbandi (ndiUB) found in the left part of the network can only be interpreted as the result of chance similarities in typology, such as the presence of noun class systems, shared word order patterns, etc., rather than an explicit family relationship or areal diffusion. Homoplasies (chance resemblances) can thus be seen to affect the distribution of taxa in the network.

We will now compare some language group-specific positions and connections in order to further explore possible reasons for observed clusterings. To this end, linguistic relationships previously identified in the literature are discussed, where relevant.

4.2. Afro-Asiatic. The family connection between the Afro-Asiatic languages included in the present comparison, i.e. East Chadic (ECHA), West Chadic (WCHA), Biu-Mandara (BM), Berber (BER), and Semitic (SEM), is well-established in the literature (e.g. Childs 2003: 29).

The Figure 1 network indeed groups Afro-Asiatic languages together in a cluster, in the lower right part of the graph, displaying some internal structural diversity. One Afro-Asiatic language, Tera (terBM) of Nigeria, falls outside of this grouping. The distribution of all but one of the geographically diverse Afro-Asiatic languages in one cluster supports that the typological stability profiles of Afro-Asiatic languages tend to carry a phylogenetic signal. One non-Afro-Asiatic language, Ngambay (ngmBB) of Bongo-Bagirmi, spoken in Chad, is interspersed with the Afro-Asiatic languages, located toward the reticulate middle low part of the network. Given its geography, the position of Ngambay can indicate prolonged contact with contiguous Afro-Asiatic languages. Alternatively, the position of Ngambay is affected by indeterminacy due to lack of data (16 out of 30 characters are scored for Ngambay; see Appendix for data coverage in the language sample).

4.3. Bongo-Bagirmi, Western Saharan, Songhay. The other Bongo-Bagirmi language, Bagirmi (bagBB) in our sample is found close to a grouping of Songhay and Western Saharan languages, in the upper right part of the Figure 1 network. Bongo-Bagirmi, Western Saharan, and the non-contiguous cluster of Songhay languages are all associated with the undemonstrated Nilo-Saharan phylum (see e.g. Bender 2000).

On the one hand, finding these languages near each other in the network may reflect that the set of stable features retains a phylogenetic signal. On the other hand, one Adamawa language, Doyayo (doyADM), is located between the Songhay and Western Saharan languages in the graph, which speaks against such an interpretation. However, comparison of Doyayo in the different networks presented throughout the paper reveals that the taxon representing the language displays positional indeterminacy.

Songhay has, besides the proposed Nilo-Saharan affiliation, previously been linked with Mande (even as a Mande-based creole, see Nicolaï 1987), Gur, Chadic, and other groups of West Africa (Childs 2003: 46). Interestingly, the sampled Songhay languages are split into two clear groups in the Figure 1 network: 1) the northernmost Songhay languages Korandje (kcySON), Tadaksahak (dsqSON), and Tagdal (tdaSON), which group with Western Saharan and Doyayo; 2) the remaining Songhay languages of the sample, i.e. Dendi (ddnSON), Zarma (zarSON), Koyraboro Senni (kseSON), Tondi Songway Kiini (tstSON), and Koyra Chiini (kchSON), which are located in a tight cluster with Western Mande languages, in the top of the graph. Also Gã (gaKWA) is found at the periphery of this cluster, a result best explained by chance similarity. The location of southern Songhay languages with the Western Mande cluster can be interpreted as reflecting strong historical influence between these languages (cf. Creissels 1981; Nicolaï 1984, 1989; forthcoming work by Robert D. Borges). Songhay and Berber, although also known to have been in contact (Nicolaï 1990; Souag 2010, 2012, 2015), do not display clear interconnections in the network.

4.4. Mande, Dogon, Bangime. The Mande languages extend over the greater part of the western half of West Africa and are considered an established genealogical unit based on shared cognacy (Dwyer 1998). The external alignment of Mande is less clear; Mande is considered either a distant Niger-Congo branch or a remnant group from an earlier diversity from before the Niger-Congo expansion (e.g. Williamson 1989; Dimmendaal 2011).

Besides links with Songhay, Figure 1 shows the included Mande languages to share most features with Gur, Ijoid, and Dogon. Western Mande forms a tight cluster. The position of Western Mande may reflect historical contact with Gur and Kru (see Childs 2003: 201-2 for a discussion on contact between Mande, Gur, and Kru). The only representative of Eastern Mande included in this study, Dan (danEMAN) of the Ivory Coast and Liberia, is found in an adjacent cluster with Jamsay (jmsDOG) and Tommo So (tmsDOG) of Dogon, and the isolate Bangime (bgmISO), spoken in the Bandiagara Cliffs area and adjacent plains in Mali.

Like Mande, Dogon may be an independent family or a distant Niger-Congo branch (e.g. Williamson & Blench 2000; Dimmendaal 2011). Dogon is geographically and typologically close to Mande, Songhay, and Gur, and has previously been linked genealogically with both Mande and Gur (see Hochstetler et al. 2004). In addition to links between Dogon and (Eastern) Mande, the Figure 1 network shows that the included Dogon languages display affinity with the representative languages of the Tano branch of Kwa (aknKWA, bleKWA), when based on the selected features. Connections between Dogon and Gur are less evident.

Bangime was previously considered a Dogon language, but it has more recently been suggested to be an isolate (Blench 2007). The close connection between Bangime and Dogon in the Figure 1 network most plausibly reflects historical contact between these languages, see also Hantgan (2013: 13).

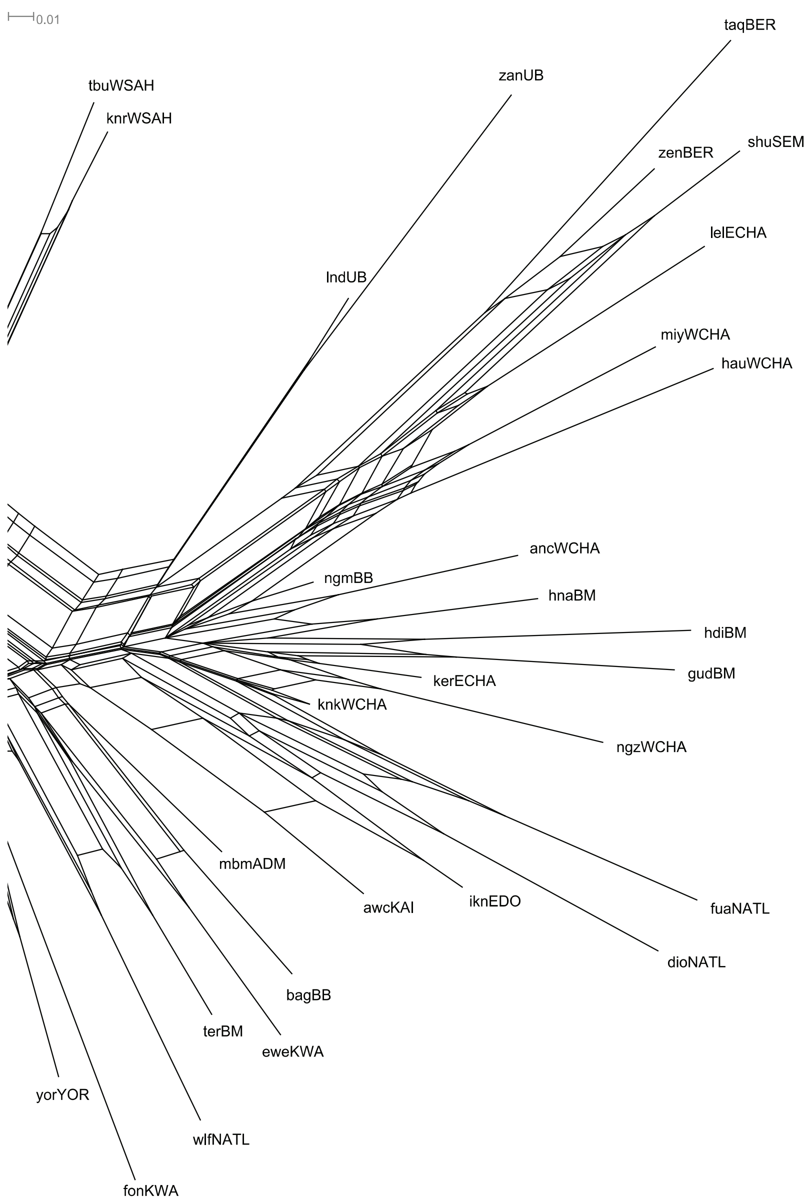
4.5. North and South Atlantic. North and South Atlantic languages are mainly spoken along the Atlantic coast from Senegal to Liberia, with dialects of Fulfulde spread out across West Africa. Greenberg (1963) posited that the Atlantic groupings formed a family unit, 'West Atlantic', a view not supported by later comparative work (see Childs 2003: 46-50).

The Figure 1 network shows substantial diversification between the languages sampled from the Atlantic groups. The North Atlantic languages, i.e. Adamawa Fulfulde (fuaNATL), spoken in the borderlands of Nigeria and Cameroon, and Wolof (wlfNATL) and Diola (dioNATL), both of Senegambia, all appear in different parts of the network. The distribution of South Atlantic languages testify further to the diversity of 'West Atlantic', cf. the positions of Temne (tneSATL) and Kisi (kisSATL) in the network.

Figure 2 displays a rotated, zoomed in, and modified version of the Figure 1 phylogeny. The network shows an attempt to uncover a possible North Atlantic phylogenetic signal by omitting Bantoid and Platoid languages from the phylogeny, as these, due to (presumed) non-inherited similarities in typology, were seen to cluster with Diola (dioNATL) in Figure 1. As a result, Adamawa Fulfulde (fuaNATL) and Diola can now be seen clustering together, indicating a family signal. They are, however, still displaced from Wolof (wlfNATL), which is found in the bottom of the network.

Bennett & Sterk's (1977) lexicostatistical study found that the diversification within the Atlantic groupings is nearly as great as in the remainder of Niger-Congo. Our results based on typology points in the same direction, affirming that the historical connections within, and between, the Atlantic groupings are uncertain.

Figure 2 The Figure 1 phylogeny minus Bantoid and Platoid (network detail shown)

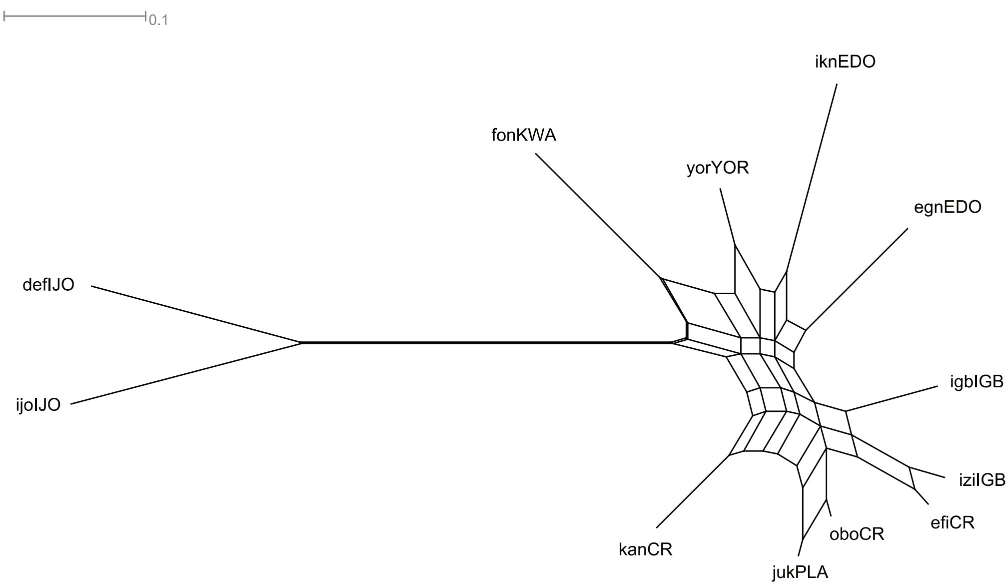


4.6. Ijoid. Ijoid is represented in the sample by Ijo-Kolokuma (ijoIJO) and Defaka (defIJO), both spoken in the Niger Delta of southwestern Nigeria. The family connection between Ijo and Defaka is supported by comparative research, but the external position of the family is unclear (see Connell et al. 2012). Greenberg (1963) classified Ijoid as Kwa, but today the cluster is viewed either as a Niger-Congo subgroup with an uncertain position within the phyla or as an independent family (Dimmendaal 2011). Geographically, the Ijoid languages are surrounded by Benue-Congo branches, e.g. Edoid, Igboid, Yoruboid, and Cross-River. While the Ijoid languages display some phonological and lexical similarities to their neighboring languages (Williamson 1971), they shows no sign of a noun-class system, have SOV word order, which is usually associated with more western branches of Niger Congo, and a verbal morphology that differs markedly from that of all of its immediate neighbors.

Figure 1 visualizes the structural unity of Defaka and Ijo-Kolokuma, in that both languages are found in the top left part of the graph, in a narrow cluster. Their long branches indicate divergence from the other languages in the sample. The network also visualizes how removed Ijoid is from Benue-Congo languages, which, otherwise, display structural commonalities. Ijoid is found closer to the West Volta-Congo groups Gur and Kru than to any East Volta-Congo branch (cf. Williamson & Blench 2000).

Figure 3 shows Ijoid and languages spoken in its vicinity in comparison. Ijoid branches far from the other languages, and draws no visible networks to them. This deviance of Ijoid suggests indeed an independent family, based on the stable feature analysis.

Figure 3 NeighborNet of Ijoid and its contiguous languages

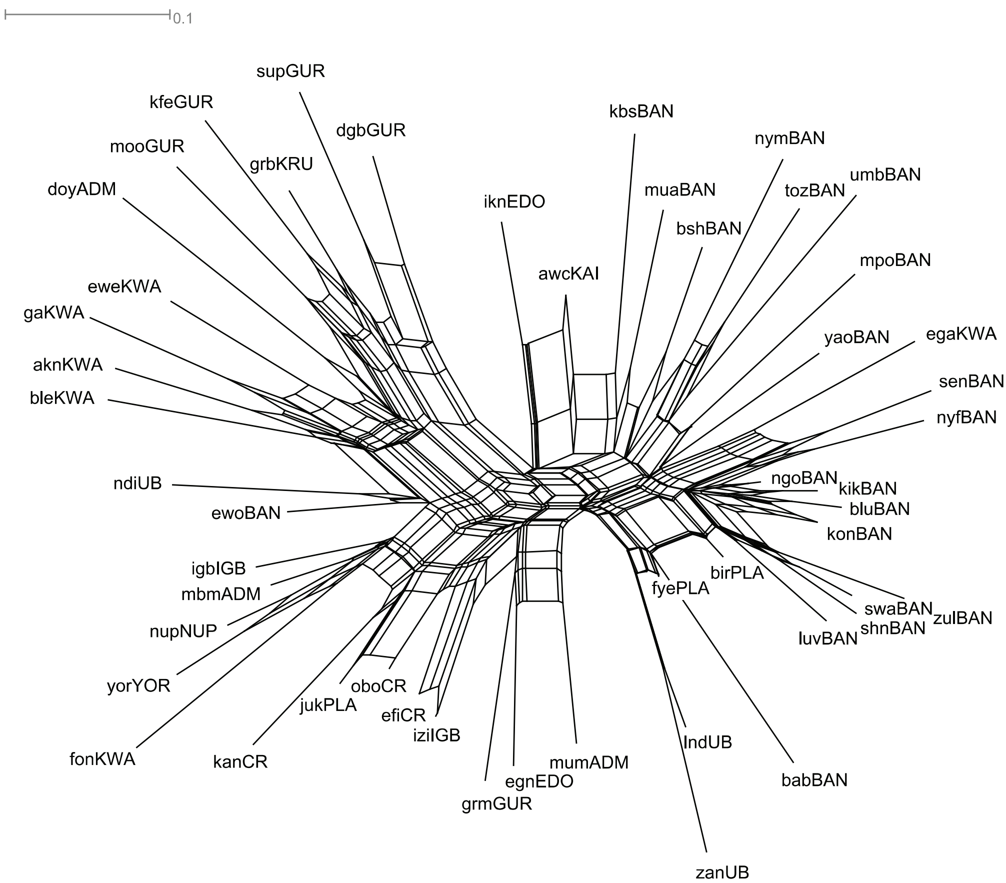


4.7. Volta Niger-Congo connections. Thus far, we have discussed branches of the Afro-Asiatic phylum, languages linked to the undemonstrated Nilo-Saharan phylum, and Niger-Congo outlier branches, which, possibly, constitute independent families. This section focuses on Niger-Congo languages for which substantial lexical evidence, including a wide range of cognate grammatical morphemes, supports a family relationship (Dimmendaal 2008: 841), i.e. Kru, Gur, Kwa, Benue-Congo, and Adamawa. Ubangi languages will also be considered. For a provisional tree of the internal structure of Volta-Congo, see Williamson & Blench (2000).

Figure 1 shows Volta-Congo branches grouped in different parts of the network. Gur and Kru of West Volta-Congo are grouped together, in the upper left part of the graph, and there is a continuum of East Volta-Congo languages, with languages of Kwa and West and East Benue-Congo spanning down the right side of the graph.

A NeighborNet of the sample of Volta-Congo languages is presented in Figure 4.

Figure 4 NeighborNet of Volta-Congo languages



4.7.1. Gur, Kru, Adamawa, Ubangi. The Gur languages extend through the central interior of West Africa, spanning from Mali and into Nigeria. Gur has been considered a genealogical unit since the time of Koelle (1854). Gur forms one of the main branches within West Volta-Congo in the widely used Williamson & Blench (2000) classification, together with Adamawa of Nigeria and Cameroon. Adamawa was grouped with the eastern Ubangi languages ('Eastern') in Greenberg (1963), but the inclusion of Ubangi in Niger-Congo has since been questioned (e.g. Moñino 2010). Formerly placed with Kwa, Kru languages are now considered to be part of a continuum with Gur and Adamawa. Kru is, however, understudied both with regards to its internal and external relationships (Sands 2009: 568).

Notwithstanding Gurma (grmGUR), the included Gur and Kru languages cluster together in the Figure 4 network's upper left part. The three Adamawa (ADM) languages that are included, are scattered around the network. This result supports the close connection between Gur and Kru. One Adamawa language included in the comparison, Doyayo (doyADM), is close to these branches (recall, however, that it clustered elsewhere in Figure 1). Considering the geographical proximity of the Adamawa languages, it is unexpected to find them removed from each other in the network. The position of Mbum (mbmADM) and that of Mumuye (mumADM) with Nigerian languages can indicate contact effects between these groups (compare also Figure 6, below). The two Ubangi languages, Zande (zanUB) and Banda-Linda (lndUB), cluster tightly together, in the lower right side of the graph, indicating either a family relationship or, as they are close to each other geographically, contact. Ngbandi (ndiUB) does not display similarities with Banda-Linda and Zande. Rather, it branches out close to the Cameroonian Bantoid language Ewondo (ewoBAN), in the left part of the graph. The results do not show a signal linking Adamawa with Ubangi.

4.7.2. Kwa and Benue-Congo. The boundaries of a Kwa family unit of Lower Guinea have been discussed since Westermann (1927). Greenberg's (1963) inclusive Kwa branch extended far, ranging from the Ivory Coast into the Benue and Cross River valleys of Nigeria, and comprised West Benue-Congo, Kru, and Ijoid. Greenberg's version of Kwa has since been substantially revised, but clear-cut divisions between the branches continue to be hard to define (Williamson & Blench 2000: 17).

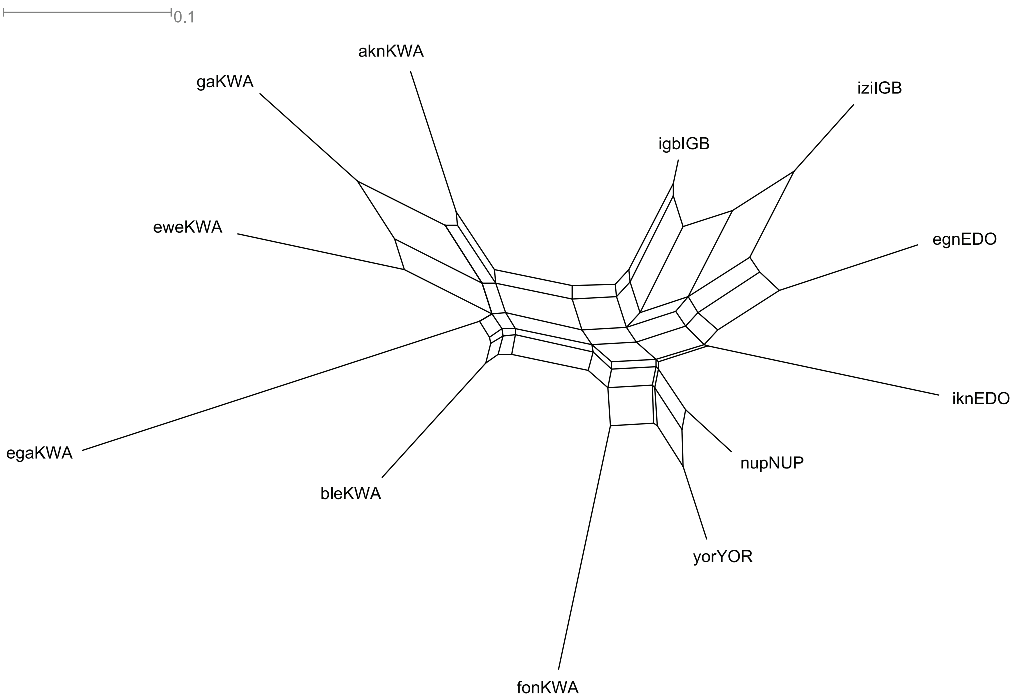
As many of these languages are spoken in contiguous areas, discerning between geographically and genealogically-conditioned clusters is difficult, as the distribution of the languages in the Figure 4 network testifies. In the bottom of Figure 4, a continuum of Benue-Congo languages spans from the northern Bantoid language Ewondo (ewoBAN) to the West Benue-Congo branches Nupoid, Yoruboid, Cross-River, Igboid, and the Platoid language Jukun (jukPLA). Bantoid and the Platoid languages Fyem (fyePLA) and Birom (birPLA) are located on the right side of the graph.

The Gbe language Fon (fonKWA) is found among West Benue-Congo languages. Ewe (eweKWA), the other included Gbe language, is located with the other Kwa languages, i.e. Gã (gaKWA), Akan (aknKWA), and Baule (bleKWA). In Figure 1, the Kwa languages are split between the Tano branch (aknKWA, bleKWA), Gã (gaKWA), Ega (egaKWA), and Gbe (eweKWA, fonKWA). In Figure 4, however, four of six Kwa languages cluster together, in the left part of the network.

Ega is an example of a language whose positions in networks echo its uncertain affiliation. Ega is the westernmost language associated with Kwa, based on lexical evidence (Bole-Richard 1983). This classification is problematic, however, as Ega has borrowed a substantial amount of its vocabulary from diverse branches of Niger-Congo (compare the wordlist in Blench 2004). Ega is located between Bantoid languages in Figure 4, a result perhaps conditioned by the fact that these languages all have noun class systems. Filtering out features relating to noun classes from the character matrix (features 30A, 31A) distances Ega from Bantoid, but does not cause it to align directly with the Kwa group (this result is not shown, as it does not differ markedly from the Figure 4 network).

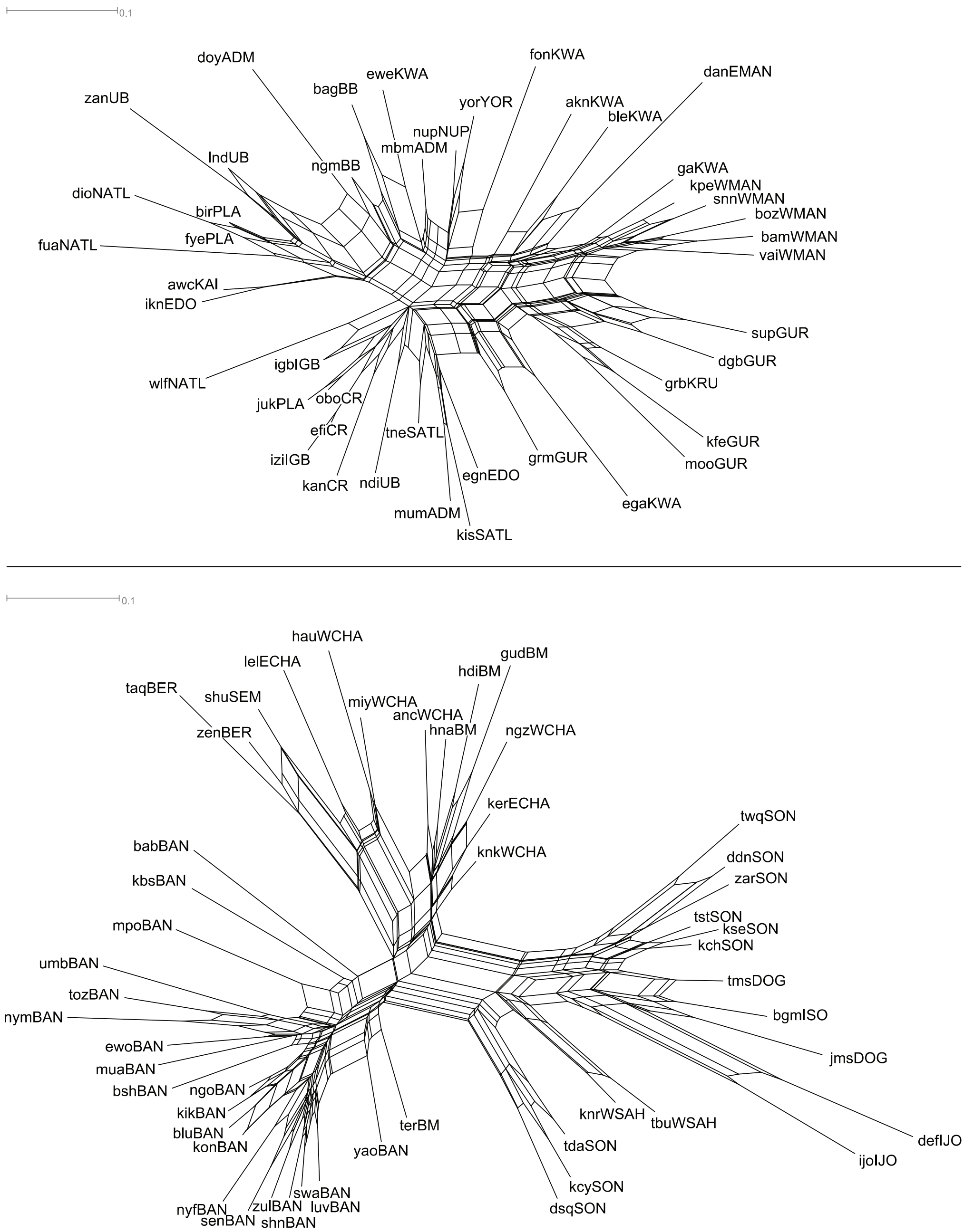
Figure 5 tests whether a clear split between West Benue-Congo and Kwa can be supported by the sampled features, if other Volta-Congo languages are omitted from the phylogeny. A division can indeed be identified: Kwa languages are found on the left side of the graph against West Benue-Congo languages on the right. Having removed Bantoid, Kru, and Gur from the phylogeny, Ega (egaKWA) now groups with Kwa. While Gbe aligns with West Benue-Congo in some recent classifications (e.g. Manfredi 2009; Blench 2006: 118), the two languages representative of Gbe are found in different parts of the Figure 5 network. Fon (fonKWA) is located next to Yoruba (yorYOR), and Ewe (eweKWA) is located nearest Gã (gaKWA) and Akan (aknKWA).

Figure 5 NeighborNet of Kwa and West Benue-Congo languages



4.8. Comparison of zones of the Macro-Sudan belt. Figure 6 shows the two final networks that will be compared in this analysis. The two networks depict the languages assigned respectively to the core and the periphery zones of the Macro-Sudan belt. The networks test whether the features chosen for our sample show the core vs. periphery distinction within the Macro-Sudan belt, or, in other words, whether typological feature analysis models linguistic geography to different degrees, and if such a pattern correlates with the supposed intensity of contact-induced diffusion.

Figure 6 NeighborNet of Macro-Sudan belt core (top) and periphery (bottom) languages



The core network is highly reticulate. Gur and Kru still group together, and the same is the case for Mande. Different Benue-Congo groups are closely interconnected, but Yoruba (yorYOR), Nupe (nupNUP), and Fon (fonKWA) now appear closer to Kwa languages. A number of branches are very short and densely linked, and, for instance, Fyem (fyePLA) almost appears drawn into the body of web formations.

The periphery network gives a different impression, with languages patterning more clearly according to family lines. As in previous networks, Bantoid languages are grouped together. One Afro-Asiatic language, Tera (terBM) of the Biu-Mandara branch, is found in the outer part of the Bantoid cluster, as are the northernmost Bantoid languages Babungo (babBAN), Kande (kbs-BAN), and Mpongwe (mpoBAN), as well as Yao (yaoBAN) of Malawi. Except for Tera, the Afro-Asiatic languages cluster as well. To the right in the graph, Songhay is, as previously, split into two distinct clusters. Bangime (bgmISO) falls between the Dogon languages, and Western Saharan and Ijoid repeat their previous clustering patterns.

The Figure 6 networks show that reticulations and inconsistencies with regards to the distribution of families are more pronounced in the core languages of the Macro-Sudan belt, when based on structural features assessed to be stable over time.

5. Conclusion

This paper has approached West African language comparison from a typological perspective. We have presented a series of splits graphs generated from phylogenetic network analysis applied to a set of typological features assessed to be temporally stable to change, encoded for a diverse sample of West African languages.

The results reveal that, despite the widespread nature of a number of marked features across the region, West African languages show variation and diversity in their distribution of the surveyed typological features. A number of genealogical groupings previously identified in the literature appeared as units in the networks, although with inconsistencies which we have attributed to a combination of gaps in the data and language sample, and to the absence of phylogenetic signal – either due to contact influences that have affected the selected features, homoplasies, or the absence of a signal to uncover in the first place. The positional indeterminacy of some languages echoed Africanists' ongoing debates concerning uncertain linguistic affiliations. Profiles of structural stability of West African languages hence appear to reflect a combination of lineage-specificity, historical interaction between languages, and chance similarities. If previous classifications can be trusted, our analysis reveals that also stable features change under areal influence. As none of the features display perfect stability (Wichmann & Holman 2009), this is to be expected.

Experimentation with different feature selections can be further explored for testing structural continuity in West African languages. To this end, the collected data for the WALS features compared in the paper are open for inspection, replication of results, and further analysis and interpretation (see Appendix). The results of this paper suggest interesting new avenues towards the disentangling of the relationships between African languages.

References

Ambouroue, Odette. 2007. *Éléments de description de l'Orungu: Langue bantu du Gabon* (B11b). PhD Thesis: Université libre de Bruxelles.

Atkinson, Quentin D. & Russell Gray. 2005. Curious parallels and curious connections – phylogenetic thinking in biology and historical linguistics. *Systematic Biology* 54/4: 513-526.

Bender, Lionel. 2000. Nilo-Saharan. In Bernd Heine & Derek Nurse (eds.) *African Languages: An Introduction*, 43-73. Cambridge: Cambridge University Press.

Bendor-Samuel, John (ed.). 1989. *The Niger-Congo Languages*. Lanham, MD: University Press of America and SIL.

Benítez-Torres, Carlos M. 2009. Inflectional vs. Derivational Morphology in Tagdal: A Mixed Language. In Masangu Matondo, Fiona McLaughlin, & Eric Potsdam (eds.) *Selected Proceedings of the 38th Annual Conference on African Linguistics: Linguistic Theory and African Language Documentation*, 69-83. Somerville, MA: Cascadilla Proceedings Project.

Bennett, Patrick R. & Jan P. Sterk. 1977. South Central Niger-Congo: A reclassification. *Studies in African Linguistics* 8: 241-273.

Bennett, Wm. G., Akinbiyi Akinlabi & Bruce Connell. 2012. Two Subject Asymmetries in Defaka Focus Constructions. In Jaehoon Choi et al. (eds.), *Proceedings of the 29th West Coast Conference on Formal Linguistics,* 294-302. Somerville, MA: Cascadilla Proceedings Project.

Bergvall, Victoria Lee. 1987. *Focus in Kikuyu and universal grammar*. PhD Thesis: Harvard University.

Blench, Roger. 2004. The Ega Language of Côte d'Ivoire. Etymologies and Implications for Classification. http://www.rogerblench.info/Language/Niger-Congo/Kwa/Ega data.pdf (accessed 31 August, 2015)

Blench, Roger. 2006. *Archaeology, language, and the African past*. Lanham, MD: AltaMira Press.

Blench, Roger. 2007. Baŋgi Me, a language of unknown affiliation in Northern Mali and its affinities. *Mother Tongue, Journal of the Association for the Study of Language in Prehistory* Issue XII: 147-178.

Blench, Roger. 2013. Why Is Africa So Linguistically Undiverse? Exploring Substrates and Isolates. *Mother Tongue, Journal of the Association for the Study of Language in Prehistory* Issue XVIII: 43-68.

Bole-Richard, Rémy. 1983. Ega. In G. Hérault (ed.) *Atlas des langues Kwa de Côte d'Ivoire,* Vol. 1, 359-401. Abidjan: ILA.

Bornand, Sandra. 2006. *Parlons zarma: Une langue du Niger*. Paris: L'Harmattan.

Bouquiaux, Luc. 1970. *La langue birom (Nigeria septentrional). Phonologie, Morphologie, Syntaxe. Paris*: Bibliothèque de la Faculté de Philosophie et lettres de l'Université de Liège.

Bowern, Claire & Bethwyn Evans (eds.). 2015. *The Routledge Handbook of Historical Linguistics.* London: Routledge.

Brauner, Siegmund. 1974. Lehrbuch des Bambara. Leipzig: Verlag Enzyklopädie.

Bryant, David & Vincent Moulton. 2004. NeighborNet: an agglomerative algorithm for the construction of phylogenetic networks. *Molecular Biology and Evolution* 21/2: 255-65.

Burquest, Donald Arden. 1973. *A Grammar of Angas*. PhD Thesis: University of California at Los Angeles.

Campbell, Lyle & William J. Poser. 2008. *Language classification: history and method*. Cambridge: Cambridge University Press.

Carbou, Henri. 1913. *Méthode pratique pour l'étude de l'arabe parlé au Ouaday et à l'est du Tchad.* Paris: Librairie orientaliste Guenther.

Carling, Gerd, Love Eriksen, Arthur Holmer & Joost van de Weijer. 2013. Contrasting linguistics and archaeology in the matrix model: GIS and cluster analysis of the Arawakan languages. In Lars Borin & Anju Saxena (eds.), *Approaches to Measuring Linguistic Differences*, pp. 29-56. Berlin: Mouton de Gruyter.

Carlson, Robert. 1994. *A Grammar of Supyire*. Berlin: Mouton de Gruyter.

Carter, Hazel. 2002. *An outline of Chitonga grammar*. Lusaka: Bookworld Publishers.

Chantoux, Alphonse, Alexandre Gontier & André Prost. 1968. Grammaire Gourmantché. (*Initiitions et Études Africaines*, XXIII.) Dakar: Institut Fondamental d'Afrique Noire, Université de Dakar.

Christiansen-Bolli, Regula. 2010. *A grammar of Tadaksahak: a northern Songhay language of Mali.* PhD Thesis: Rijksuniversiteit te Leiden. (Published in 2010 as *A Grammar of Tadaksahak - a Berberised Songhay Language (Mali)* (Berber Studies, vol.31). Köln: Rüdiger Koppe.

Childs, G. Tucker. 1988. *The phonology and morphology of Kisi*. PhD Thesis: University of California, Berkeley.

Childs, G. Tucker. 1995. *A Grammar of Kisi* (Mouton Grammar Library). Berlin: Mouton.

Childs, G. Tucker. 2003. *An Introduction to African Languages*. Amsterdam: John Benjamins Publishing Company.

Clements, Nick & Annie Rialland. 2008. Africa as a phonological area. In Bernd Heine & Derek Nurse (eds.), *A Linguistic Geography of Africa*. 36-86. Cambridge: Cambridge University Press.

Cloarec-Heiss, France. 1986. *Dynamique et équilibre d'une syntaxe: Le banda-linda de Centrafrique*. Paris: Société d'Études Linguistiques et Anthropologiques de France.

Connell, Bruce, William Bennett, Inoma Essien, Ebitare Obikudo, Akin Akinlabi & Ozomekuri Ndimele. 2012. Defaka and Ịjọ: A reassessment of the Ijoid relationship. In Matthias Brenzinger & Anne-Maria Fehn (eds.), *Proceedings of the 6th World Congress of African Linguistics*, *Cologne, 17-21 August 2009*, 145-154. Köln: Rüdiger Köppe Verlag.

Creissels, Denis 1981. De la possibilité de rapprochements entre le songhay et les langues Niger-Congo (en particulier Mandé). In Thilo Schadeberg & M. L. Bender(eds.), *Nilo-Saharan: Proceedings Of The First Nilo-Saharan Linguistics Colloquium, Leiden, September 8–10.* 185-199. Dordrecht: Foris Publications.

Croft, William. 2008. Evolutionary Linguistics. *Annual Review of Anthropology* 37: 219-34.

Cysouw, Michael, Albu, Mihai & Dress, Andreas. 2008. Analyzing feature consistency using dissimilarity matrices. STUF - *Language Typology and Universals* 61(3): 263-279.

Daget, Jacques, Mamadou Konipo & Mamadou Sanankoua. 1953. La Langue Bozo. (*Études Soudaniennes*, I.) Dakar: Institut Français d'Afrique Noire, Governement du Soudan.

Dediu, Dan. 2011. A Bayesian phylogenetic approach to estimating the stability of linguistic features and the genetic biasing of tone. *Proceedings of the Royal Society of London/*B 278(1704): 474-479.

Dediu, Dan & Michael Cysouw. 2013. Some structural aspects of language are more stable than others: A comparison of seven methods. *PLoS ONE* 8(1): e55009.

Dediu, Dan & Stephen C. Levinson. 2012. Abstract profiles of structural stability point to universal tendencies, family-specific factors, and ancient connections between languages. *PLoS ONE* 7(9), e45198.

Dereau, Léon. 1955. *Cours de Kikongo*. Namur: Wesmael-Charlier.

Diagana, Yacouba. 1994. Eléments de grammaire du soninké. (Les documents de linguistique africaine, 2.) Paris: Univ. Paris 7 Denis Diderot.

Diagne, Anna Marie. 2006. *Phonologie et morphologie du soninké: une analyse non linéaire*. PhD Thesis: Johannes Gutenberg Universität.

Dimmendaal, Gerrit J. 2008. Language Ecology and Linguistic Diversity on the African Continent. *Language and linguistics* compass 2/5: 840-858.

Dimmendaal, Gerrit J. 2011. *Historical Linguistics and the Comparative Study of African Languages*. Amsterdam and Philadelphia: John Benjamins Publishing Company.

Doke, Clement M. 1927. *Textbook of Zulu grammar*. 6th edn. Johannesburg: Maskew Miller Longman.

Doneux, Jean Léonce. 1968. *Esquisse grammaticale du Dan*. (Documents linguistiques 15). Dakar: Département de Linguistique Générale et Linguistique Africaine de la Faculté des Lettres et Sciences Humaines.

Donohue, Mark & Simon Musgrave. 2007. Typology and the linguistic macrohistory of Island Melanesia. *Oceanic Linguistics* 46: 348-387.

Donohue, Mark, Søren Wichmann, & Mihai Albu. 2008. Typology, areality and diffusion. *Oceanic Linguistics* 47(1): 223–232.

Dryer, Matthew S. & Martin Haspelmath (eds.) 2013. The World Atlas of Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info (accessed 31 August, 2015)

Dryer, Matthew S. 2005. Genealogical language list. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures*, 584-644. Oxford: Oxford University Press.

Dunn, Michael, Angela Terrill, Ger Reesink, Robert A Foley & Stephen C. Levinson. 2005. Structural phylogenetics and the reconstruction of ancient language history. *Science* 309: 2072-2075.

Dunn, Michael, Stephen C. Levinson, Eva Lindström, Ger Reesink & Angela Terrill. 2008. Structural phylogeny in historical linguistics: Methodological explorations applied in Island Melanesia. *Language* 84/4: 710-759.

Dunn, Michael, Simon J. Greenhill, Stephen C. Levinson & Russell D. Gray. 2011. Evolved structure of language shows lineage-specific trends in word-order universals. *Nature* 473: 79–82.

Dunn, Michael. 2015. Language phylogenies. In Claire Bowern & Bethwyn Evans (eds.), *The Routledge Handbook of Historical Linguistics*, 190-211. London/New York: Routledge.

Duthie, Alan S. 1996. *Introducing Ewe linguistic patterns*. Accra: Ghana Universities Press.

Dwyer, David J. 1998. The Place of Mande in Niger-Congo. In Ian Maddieson & Thomas J. Hinnebusch (eds.), *Language History and Linguistic Description in Africa*. 27-36. Eritrea: Africa World Press Inc.

Emenanjo, E. Nolue. 1987. *Elements of modern Igbo grammar: a descriptive approach*. Ibadan: Oxford University Press.

Essono, Jean-Jacques. 1994. *Description synchronique de l'ewondo, Bantu (A72a) du Cameroun: phonologie, morphologie, syntaxe*. PhD Thesis: University of Paris.

Faidherbe, Louis Léon César. 1877. *Le Zenaga des tribus sénégalaises: Contribution á l'étude de la langue Berbère*. Paris: Ernest Leroux.

Felsenstein, Joseph. 2004. *Inferring Phylogenies*. Sunderland, MA: Sinauer Associates, Inc.

Fortune, G. 1955. *An analytical grammar of Shona*. London: Longmans, Green & Co.

Frajzyngier, Zygmunt, with Erin Shay. 2002. *A Grammar of Hdi*. Berlin: Mouton de Gruyter.

Frajzyngier, Zygmunt & Eric Johnston, with Adrian Edwards. 2005. *A Grammar of Mina*. Berlin: Mouton de Gruyter.

Funnell, Barry J. 2004. *A contrastive analysis of two standardised varieties of Sena*. MA Thesis: University of South Africa, Pretoria.

Gray, Russell & Quentin D. Atkinson. 2003. Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* 426: 435-439.

Gray, Russell, Alexei J. Drummond, & Simon J. Greenhill. 2009. Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science* 323: 479-483.

Gray, Russell, Quentin D. Atkinson & Simon J. Greenhill. 2011. Language evolution and human history: what a difference a date makes. Philosophical Transactions of the Royal Society B: Biological Sciences 366(1567): 1090-1100.

Greenberg, Joseph H. 1959. Africa as a linguistic area. In William R. Bascom.& Melville J. Herskovits (eds.), *Continuity and change in African cultures*. 15-27. Chicago: Chicago University Press.

Greenberg, Joseph H. 1963. *The languages of Africa*. Bloomington, Indiana: Indiana University Press.

Greenberg, Joseph H. 1983. Some areal characteristics of African languages. In Ivan Dihoff R.(ed.), *Current Approaches to African linguistics* 1. 3-22. Dordrecht: Foris.

Greenhill, Simon J., Thomas E. Currie & Russell D. Gray. 2009. Does horizontal transmission invalidate cultural phylogenies? *Proceedings of the Royal Society B: Biological Sciences* 276: 2299-2306.

Greenhill, Simon J., Quentin D. Atkinson, Andrew Meade & Russell D. Gray. 2010. The shape and tempo of language evolution. *Proceedings of the Royal Society B: Biological Sciences* 277(1693): 2443-2450.

Grollemund, Rebecca. 2006. *Les Okandé du Gabon, locuteurs d'une langue en danger (langue bantoue du groupe B 30) - Langue et culture*. MA Thesis: Université Lumière Lyon 2.

Güldemann, Tom. 2008. The Macro-Sudan belt: towards identifying a linguistic area in northern sub-Saharan Africa. In Bernd Heine & Derek Nurse (eds.), *A linguistic geography of Africa*, 151-185. Cambridge: Cambridge University Press.

Güldemann, Tom. 2010. Sprachraum and geography. In Alfred Lameli Roland Kehrein & Stefan Rabanus (eds.), *Language and Space: An International Handbook of Linguistic Variation Volume 2: Language Mapping*, 561-585. Berlin, New York: Berlin: Mouton de Gruyter.

Guthrie, Malcolm. 1948. *The classification of the Bantu languages*. London: Oxford University Press for the International African Institute.

Hantgan, Abbie. 2013. *Aspects of Bangime Phonology, Morphology, and Morphosyntax*. PhD Thesis: Indiana University.

Haspelmath, Martin, Matthew Dryer, David Gil & Bernard Comrie (eds.). 2005. *The World Atlas of Language Structures*. Oxford: Oxford University Press.

Heath, Jeffrey. 1999a. *A Grammar of Koyraboro (Koroboro) Senni*. (Westafrikanische Studien, 19.) Köln: Rüdiger Köppe Verlag.
Heath, Jeffrey. 1999b. *A Grammar of Koyra Chiini: the Songhay of Timbuktu*. Berlin: Mouton de Gruyter.
Heath, Jeffrey. 2005a. *A Grammar of Jamsay*. Berlin: Mouton de Gruyter.
Heath, Jeffrey. 2005b. *Grammar of Tamashek (Tuareg of Mali).* (Mouton Grammar Series.) The Hague: Mouton de Gruyter.
Heath, Jeffrey. 2005c. *Tondi Songway Kiini (montane Songhay, Mali): Reference grammar and TSK-English-French Dictionary*. Chicago: The University of Chicago Press.
Heine, Bernd. 1975. Language typology and convergence areas in Africa. *Linguistics* 144: 27-47.
Heine, Bernd. 1977. *A Typology of African Languages based on the order of meaningful elements*. (Kölner Beiträge zur Afrikanistik 4). Berlin: Dietrich Reimer.
Heine, Bernd, and Zelealem Leyew. 2008. Is Africa a linguistic area? In Bernd Heine & Derek Nurse (eds.), *A Linguistic Geography of Africa,* 15-35. Cambridge, UK: Cambridge University Press.
Heine, Bernd & Tania Kuteva. 2001. Convergence and divergence in the development of African languages. In Alexandra Y. Aikhenvald. & R. M. W. Dixon (eds.), *Areal diffusion and genetic inheritance*, 393–411. Oxford: Oxford University Press.
Heine, Bernd & Derek Nurse (eds.). 2000. *African Languages: An Introduction*. Cambridge: Cambridge University Press.
Heine, Bernd & Derek Nurse. 2008. *A linguistic geography of Africa*. Cambridge: Cambridge University Press.
Hewson, John. 2010a. Doyayo. In Derek Nurse, Sarah Rose & John Hewson (eds.), *Verbal categories in Niger-Congo languages*, http://www.mun.ca/linguistics/nico/ (accessed 3 September 2015)
Hewson, John. 2010b. Doyayo. In Derek Nurse, Sarah Rose & John Hewson (eds.), *Verbal categories in Niger-Congo languages*. http://www.mun.ca/linguistics/nico/ (accessed 3 September 2015)
Hochstetler, J. Lee, J.A. Durieux & E.I.K. Durieux-Boon. 2004. Sociolinguistic Survey of the Dogon Language Area. SIL International. http://www-1.sil.org/silesr/2004/silesr2004-004.pdf (accessed 3 September, 2015)
Horton, Alonzo E. 1949. *A grammar of Luvale*. (Bantu Grammatical Archives 2). Johannesburg: University of Witwatersrand Press.
Hoskison, James T. 1983. *A Grammar and Dictionary of the Gude Language*. PhD Thesis: Ohio State University.
Hudu, Fusheini. 2014. What is the phonological word in Dagbani? A positional faithfulness account. *Ghana Journal of Linguistics* 3.1: 1-44.
Huson, Daniel H. & David Bryant. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23/2: 254-267.
Hyman, Larry M. & Knut Olawsky. 2004. Dagbani verb tonology. In Chege Githiora, Heather Littlefield & Victor Manfredi (eds.), *Trends in African Linguistics* 4, 97-108. Trenton, N.J.: Africa World Press, Inc.
Ikoro, Suanu M. 1994. Numeral Classifiers in Kana. *Journal of African Languages and Linguistics* 15: 7-28.
Ikoro, Suanu M. 1996. *The Kana Language*. Leiden: Research School CNWS, Leiden University.

Jenewari, Charles E. W. 1983. Defaka: Ijo's closest linguistic relative. In Ivan R. Dihoff (ed.), *Current Approaches to African Linguistics* 1, 85-111. Dordrecht: Foris.

Kamarah, Sheikh Umarr. 2007. *A Descriptive Grammar of KʌThemnɛ (Temne).* Munich: Lincom.

Kandybowicz, Jason. 2008. *The Grammar of Repetition: Nupe Grammar at the Syntax – Phonology Interface*. Amsterdam: John Benjamins.

Kawu, Ahmadu Ndanusa. 2002. *Variation in Nupe Phonology and Morphology*. PhD Thesis: Rutgers University, New Jersey.

Kitchen, Andrew, Christopher Ehret, Shiferaw Assefa & Connie J. Mulligan. 2009. Bayesian phylogenetic analysis of Semitic languages identifies an Early Bronze Age origin of Semitic in the Near East. *Proceedings of the Royal Society B: Biological Sciences* 276(1668): 2703–2710.

Koelle, Sigismund W. 1854. *Polyglotta Africana*. London: London Church Missionary House.

Kotey, Paul F. A. 1969. *Syntactic Aspects of the Gã Nominal Phrase*. PhD Thesis: University of Wisconsin.

Kropp-Dakubu, Mary Esther. 2006. *Parlons Gã: langue et culture d'Accra*. Paris: L'Harmattan.

Lax, Beverley. M. 1996. *The verbal system of Kigiryama*. PhD Thesis: University of Wisconsin-Madison.

Le Coeur, Ch. & Le Coeur, M. 1956. *Grammaire et textes teda-daza*. (Mémoires de l'institut d'Afrique Noire, 46.) Dakar: Institut Français d'Afrique Noire (IFAN).

Lefebvre, Claire & Anne-Marie Brousseau. 2002. *A grammar of Fongbe*. Berlin: Mouton de Gruyter.

Lewis, M. Paul, Gary F. Simons, & Charles D. Fennig (eds.). 2016. *Ethnologue: Languages of the World*, 19th edition. Dallas, Texas: SIL International. http://www.ethnologue.com (accessed 3 September, 2015)

Longobardi, Giuseppe, Cristina Guardiano, Giuseppina Silvestri, Alessio Boattini & Andrea Ceolin. 2013. Toward a syntactic phylogeny of modern Indo-European languages. *The Journal of Historical Linguistics* 3(1): 122-152.

Maganga, Clement & Thilo C. Schadeberg. 1992. Kinyamwezi: grammar, texts, vocabulary. (East African Languages and Dialects 1). Köln: Köppe.

Manfredi, Victor. 2009. Morphosyntactic parameters and the internal classification of Benue-Kwa (Niger-Congo). In Paola Crisma & Giuseppe Longobardi (eds.), *Historical Syntax and Linguistic Theory*, 329-343. Oxford: Oxford University Press.

Maslova, Elena S. 2004. Динамика типологических распределений и стабильность языковых типов. [Dynamics of typological distributions and stability of language types.] *Вопросы языкознания* 5: 3-16.

McGill, Stuart. 2009. *Gender and Person Agreement in Cicipu Discourse*. PhD Thesis, London: School of Oriental and African Studies.

McMahon, April & Robert McMahon. 2003. Finding Families: Quantitative Methods in Language Classification. *Transactions of the Philological Society* 101: 7-55.

McPherson, Laura. 2013. *A Grammar of Tommo So*. Berlin: Mouton De Gruyter.

Meier, Paul, Inge Meier, & Bendor-Samuel, John. 1975. *A Grammar of Izi: An Igbo Language*. Norman, Oklahoma: Summer Institute of Linguistics.

Moñino, Yves. 2010. The position of Gbaya-Manza-Ngbaka group among the Niger-Congo languages. Paper presented at the Workshop "Genealogical Classification of African Languages beyond Greenberg", 21-22 February, 2010, Humboldt-Universität zu Berlin. http://african-language-classification.hu-berlin.de/workshop-genealogicalclassification/presentations/ (accessed 3 September 2015)

Morrison, Michelle E. 2011. *A reference grammar of Bena*. PhD Thesis: University of Texas at Austin.

Mugane, John M. 1997. *A Paradigmatic Grammar of Gĩkũyũ*. (Stanford Monographs in African Languages). Stanford: CSLI Publications.

Ndjerareou, Mekoulnodji, Christy Melick & Sarah Moeller. 2010. A brief grammatical sketch of Ngambay. GIALens: Electronic Notes Series 4.2. http://www.gial.edu/documents/gialens/Vol4-2/MelickMoellerMekoulnodj-Ngambay.pdf (accessed 3 September 2015)

Nettle, Daniel. 1998. *The Fyem Language of Northern Nigeria*. (Languages of the World/Materials, 136.) Munich: Lincom.

Newman, Paul. 1970. *A Grammar of Tera*. (University of California Publications in Linguistics, 57.) Berkeley: University of California Press.

Ngonyani, Deo. 2003. *A grammar of Chingoni*. München: Lincom.

Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago: University of Chicago Press.

Nichols, Johanna & Tandy Warnow. 2008. Tutorial on computational linguistic phylogeny. *Language and Linguistics Compass* 2, 5:760-820. 10.1111/j.1749-818X.2008.00082.x

Nicolaï, Robert. 1984. *Préliminaires a une étude sur l'origine du songhay: matériaux, problématique et hypothèses*. (Marburger Studien Zur Afrika- Und Asienkunde 37). Berlin: Dietrich Reimer.

Nicolaï, Robert. 1987. Is Songhay a Creole language? In Glen Gilbert (ed.), Pidgin and Creole Languages, pp. 469-484. Honolulu: University of Hawaii Press.

Nicolaï, Robert. 1989. Songhay et Mande. *Mandenkan* 18: 69-80.

Nicolaï, Robert. 1990. *Parentés linguistiques (A propos du songhay).* Paris: CNRS.

Nikiema, Norbert. 2001. Moore. In Jane Garry & Carl Rubino (eds.), *Facts About the World's Languages, An Encyclopaedia of the World's Languages: Past and Present*, pp. 492-496. New York/Dublin: HW Wilson.

Nurse, Derek. 1997. The contribution of linguistics to the study of history in Africa. *Journal of African History* 38: 359-91.

Nurse, Derek. 2010. Jukun. In Derek Nurse, Sarah Rose & John Hewson (eds.), *Verbal categories in Niger-Congo languages*. http://www.mun.ca/linguistics/nico/ (accessed 3 September 2015)

Osam, Emmanuel Kweku Ahen. 1994. *Aspects of Akan Grammar – A Functional Perspective*. PhD Thesis: University of Oregon.

Parkvall, Mikael. 2008. Which parts of language are the most stable? *STUF - Language Typology and Universals* 61(3): 234-250.

Petzell, Malin & Harald Hammarström. 2013. Grammatical and lexical comparison of the Greater Ruvu Bantu languages. *Nordic Journal of African Studies* 22(3): 129-158.

Pichl, Walter. 1957. *Die Sprache der Wolof in Senegambien*. PhD Thesis: Universität Wien.

Redden, James E. 1979. *A Descriptive Grammar of Ewondo*. (Occasional Papers on Linguistics, 4.) Carbondale, Illinois: Southern Illinois University at Carbondale, Department of Linguistics.

Reesink, Ger & Michael Dunn. 2012. Systematic typological comparison as a tool for investigating language history. In Nicholas Evans & Marian Klamer (eds.), *Melanesian Languages on the Edge of Asia: Challenges for the 21st* Century (Language Documentation & Conservation Special Publication 5), pp. 34-71. Honolulu: University of Hawaii Press.

Rennison, John R. 1997. *Koromfe*. London & New York: Routledge.

Rowland-Oke, Mary. 2003. *Description systématique de la langue Obolo-Andoni (Nigéria*). Paris: L'Harmattan.

Salffner, Sophie. 2004. *Learning Ega: A linguistic Frame of Reference for Primers for Endangered Languages.* MA Thesis: Universitet Bielefeld.

Salffner, Sophie. 2009. *Tone in the Phonology, Lexicon and Grammar of Ikaan*. PhD Thesis: Department of Linguistics, School of Oriental and African Studies, University of London.

Sanderson, Meredith. 1922. *A Yao grammar*. 2nd edn. New York/Toronto: London.

Sands, Bonny. 2009. *Africa's Linguistic Diversity. Language and Linguistics Compass* 3/2 (2009): 559-580.

Santandrea, Stefano. 1964. *A Note on Kare Grammar*. Sudan Notes and Records 45: 103-112.

Sapir, David J. 1965. *A Grammar of Diola-Fogny: A Language Spoken in the Basse-Casamance Region of Senegal*. (West African language monographs 3). Cambridge: Cambridge University Press.

Schadeberg, Thilo C. 1990. *A sketch of Umbundu*. (Grammatische Analysen Afrikanischer Sprachen 1). Köln: Rüdiger Köppe.

Schaub, Willi. 1985. *Babungo*. London: Croom Helm.

Schuh, Russell G. 1972. *Aspects of Ngizim Syntax*. PhD Thesis: University of California, Los Angeles.

Schuh, Russell G. 1998. *A Grammar of Miya*. Berkeley and Los Angeles: University of California Press.

Shimizu, Kiyoshi. 1983. *The Zing dialect of Mumuye: a descriptive grammar*. Hamburg: Helmut Buske Verlag.

Sibomana, Leo. 2008. *Le Zarma parlé: Esquisse grammaticale, lexique, textes*. (Beiträge zur Afrikanistik, 18.) Berlin: Münster.

Souag, Mostafa Lameen. 2010. *Grammatical contact in the Sahara: Arabic, Berber, and Songhay in Tabelbala and Siwa*. PhD Thesis: School of Oriental and African Studies, University of London.

Souag, Mostafa Lameen. 2012. The subclassification of Songhay and its historical implications. *Journal of African Languages and Linguistics* 33(2): 181–213.

Souag, Mostafa Lameen. 2015. Non-Tuareg Berber and the genesis of nomadic Northern Songhay. *Journal of African Languages and Linguistics* 36(1): 121–143.

Stennes, Leslie H. 1967. *A Reference Grammar of Adamawa Fulani*. East Lansing: African Studies Center, Michigan State University.

Stewart, John. 1976. *Towards Volta-Congo Reconstruction* (Inaugural Lecture). Leiden: Leiden University Press.

Thomas, Elaine. 1978. *A Grammatical Description of the Engenni Language*. Dallas: SIL.

Timyan, Judith. 1977. *A discourse-based grammar of Baule: the Kode dialect*. PhD Thesis: University of New York.

Toronzoni, Ngama-Nzombio Tra Ndele. 1989. *Description du Ngbandi*. PhD Thesis: Université Libre de Bruxelles.

Una, F. X. 1900. *Efik language*. UYO: Modern Business Press.
van der Wal, G. Jenneke. 2009. *Word order and information structure in Makhuwa-Enahra*. PhD Thesis: Universiteit Leiden.
Vansina, Jan. 1959. *Esquisse de grammaire bushong*. (Annalen van het Koninglijk Museum van Belgisch-Congo: Reeks in 8 23). Tervuren: Commissie voor Afrikaanse Taalkunde / Commission de Linguistique Africaine.
Walker, Robert S. & Lincoln A. Ribeiro. 2011. Bayesian phylogeography of the Arawak expansion in lowland South America. *Proceedings of the Royal Society of London/B* 278, 2562–2567 doi:10.1098/rspb.2010.2579
Welmers, William Everett. 1973. *African Language Structures*. Los Angeles: University of California Press.
Welmers, William Everett. 1976. *A Grammar of Vai*. Berkeley: University of California Press.
Westermann, Diedrich H. 1911. *Die Sudansprachen: Eine sprachvergleichende Studie*. Hamburg: L. Friederichsen.
Westermann, Diedrich H. 1927. *Die Westlichen Sudansprachen und Ihre Beziehungen zum Bantu*. Berlin: Walter de Gruyter.
Wichmann, Søren. 2015. Diachronic stability and typology. In Claire Bowern & Bethwyn Evans (eds.), *Handbook of Historical Linguistics*, pp. 212-224. London and New York: Routledge.
Wichmann, Søren & Eric W. Holman. 2009. *Assessing temporal stability for linguistic typological features*. München: LINCOM Europa.
Wichmann, Søren & Saunders, Arpiar. 2007. How to use typological databases in historical linguistic research. *Diachronica* 24(2): 373-404.
Williamson, Kay. 1965. *A Grammar of the Kolokuma Dialect of Ijo*. Cambridge: Cambridge University Press.
Williamson, Kay. 1971. The Benue-Congo Languages and Ijo. In T. Sebeok (ed.), *Current Trends in Linguistics Volume 7: Linguistics in sub-Saharan Africa*, pp. 245-306. The Hague: Mouton.
Williamson, Kay. 1989. Niger-Congo overview. In John Bendor-Samuel (ed.), *The Niger-Congo languages: a classification and description of Africa's largest language family*, pp. 3–45. Lanham, MD: University Press of America.
Williamson, Kay & Roger Blench. 2000. Niger-Congo. In Bernd Heine & Derek Nurse (eds.), *African Languages–An Introduction*, pp. 11-42. Cambridge: Cambridge University Press.
Wolff, Hans. 1959. Subsystem typologies and area linguistics. *Anthropological Linguistics* 1(7): 1-88.
Wolff, Ekkehard H. & Ousseina Alidou. 2001. On the non-linear Ancestry of Tasawaq (Niger). Or: How "Mixed" can a language be?. In Derek Nurse (ed.), *Historical Language Contact in Africa*, pp. 523-574. Köln: Rüdiger Köppe.
Zima, Petr. 1976. Minority language enclaves. In Vladimir Klima (ed.), *African culture and integration*, pp. 123-152. Prague: Academia.
Zima, Petr. 1998. *Dialecte marginal ou pidgin vernacularisé? Le cas du dendi (songhay).* (Studies in African linguistics, 17). München: Lincom Europa.

Appendix Character matrix

Language	Characters (WALS features 9A, 10A, 18A, 27A, 28A, 29A, 30A, 31A, 33A, 39A, 40A, 44A, 48A, 57A, 63A, 65A, 66A, 69A, 73A, 82A, 83A, 85A, 86A, 88A, 89A, 90A, 94A, 104A, 118A, 119A)
aknKWA	1,1,1,1,1,1,1,1,6,3,1,6,2,1,2,1,4,1,1,1,2,4,1,2,2,1,5,1,3,1
ancWCHA	1,2,1,?,?,1,?,?,7,2,1,4,?,2,1,1,?,5,2,1,2,2,2,2,2,1,?,1,?,?
awcKAI	3,1,1,1,1,?,5,3,1,3,3,6,2,2,2,1,4,4,2,1,2,2,2,?,2,1,1,?,2,1
babBAN	1,2,1,1,1,?,5,3,2,5,5,2,2,1,2,1,3,2,1,1,2,2,2,2,2,1,1,1,3,2
bagBB	1,1,1,2,1,2,1,1,8,3,1,?,3,2,?,1,4,4,2,1,2,2,2,2,2,1,1,2,1,1
bamWMAN	1,1,1,2,1,1,1,1,2,3,1,6,2,4,1,1,4,2,2,1,1,1,1,6,2,4,1,1,3,1
bgmISO	1,1,1,1,1,1,2,2,8,3,3,6,2,3,1,1,4,4,2,1,1,1,1,2,2,?,?,1,2,?
birPLA	1,2,1,1,?,?,5,3,1,?,?,?,?,2,2,1,4,1,2,1,2,2,2,2,2,1,1,?,?,?
bleKWA	3,1,1,1,?,?,1,1,6,3,?,6,2,4,2,1,1,4,?,1,2,1,1,2,2,1,2,?,?,1
bluBAN	1,2,1,1,1,2,5,3,1,3,3,6,1,4,2,1,3,4,2,1,2,2,2,2,2,1,1,2,2,2
bozWMAN	1,1,1,2,1,1,1,1,2,3,1,6,2,4,1,1,4,5,2,1,1,1,1,2,2,3,1,1,2,2
bshBAN	1,3,1,3,1,?,5,3,1,3,3,6,?,1,2,?,3,1,?,1,2,2,2,2,2,1,1,?,1,1
danEMAN	3,1,1,1,1,?,1,1,7,5,1,6,2,4,1,1,1,2,2,1,1,1,1,2,2,1,5,1,2,2
ddnSON	3,2,1,?,1,1,1,1,2,3,1,6,2,4,2,1,4,5,2,1,1,1,1,2,2,?,?,1,1,?
defIJO	1,1,1,2,?,?,3,2,2,3,1,3,?,4,?,?,1,4,?,1,1,1,1,1,3,2,1,1,?,?
dgbGUR	1,2,1,3,1,1,5,3,2,3,?,?,?,4,2,1,2,4,2,1,2,1,1,2,2,3,?,?,3,1
dioNATL	1,2,1,1,1,2,5,3,1,5,5,2,2,2,2,2,4,4,2,1,2,2,2,2,2,1,1,2,1,2
doyADM	2,1,1,2,?,3,?,?,2,5,?,6,2,2,?,1,4,3,?,1,2,4,1,2,2,1,1,2,3,1
dsqSON	1,2,1,2,1,3,1,1,6,1,3,6,3,4,2,1,4,1,2,1,2,1,1,2,2,1,1,2,?,2
efiCR	1,2,1,1,1,2,1,1,9,3,3,6,2,4,1,1,2,4,2,1,2,2,2,2,2,1,1,1,?,1
egaKWA	1,2,1,?,?,?,5,3,1,3,?,?,?,1,?,1,1,5,2,1,2,1,1,2,2,1,2,?,?,?
egnEDO	3,2,1,1,1,1,3,3,9,3,1,6,2,4,2,1,4,3,2,1,2,2,2,2,2,1,1,1,1,2
eweKWA	1,1,1,1,1,2,1,1,8,3,1,6,3,2,2,2,4,4,2,1,2,1,1,2,2,1,1,2,3,1
ewoBAN	1,1,1,1,1,?,5,3,1,2,1,6,2,4,2,1,2,1,1,1,2,2,2,2,2,1,1,1,1,1
fonKWA	2,1,4,1,1,1,3,1,7,3,1,6,?,4,2,1,1,5,?,1,2,4,2,2,2,1,1,1,1,1
fuaNATL	2,2,1,1,?,?,5,3,2,5,1,2,2,2,2,2,1,2,1,1,2,2,2,2,2,1,1,3,1,1
fyePLA	1,2,1,1,?,?,?,?,1,?,?,?,?,2,?,?,?,2,?,1,2,2,2,2,2,1,1,?,?,?
gaKWA	1,1,1,2,1,?,1,1,2,3,?,6,?,1,?,1,4,4,?,1,2,1,1,1,2,1,1,1,3,1
grbKRU	1,1,1,?,1,3,3,3,2,3,1,2,2,4,?,1,2,2,2,1,2,1,1,2,2,1,?,1,1,1
grmGUR	1,2,1,1,1,1,5,3,2,3,1,6,2,?,?,1,4,5,?,1,2,4,?,2,1,1,?,1,?,?
gudBM	1,2,1,?,?,?,?,1,2,5,1,3,2,2,1,1,4,2,?,2,2,2,2,4,2,1,1,1,3,1
hauWCHA	3,2,1,1,4,1,2,2,2,3,1,1,3,2,2,2,4,5,2,1,2,2,2,6,2,1,3,1,2,1
hdiBM	1,2,1,1,?,?,1,1,2,5,5,6,2,2,1,1,1,5,?,2,2,2,2,5,2,1,1,?,2,1
hnaBM	2,2,1,2,?,?,?,1,8,5,?,6,2,2,1,1,2,5,?,1,2,2,2,2,2,1,1,1,?,1
igbIGB	1,2,1,1,1,1,1,1,6,3,1,6,2,4,1,?,?,2,2,1,2,2,2,2,2,1,1,1,3,1
ijoIJO	1,1,1,1,1,1,3,2,2,3,3,3,3,1,2,2,2,2,2,1,1,1,1,1,1,2,5,1,1,1
iknEDO	3,1,1,1,?,?,5,3,6,?,?,?,?,2,?,?,4,4,?,1,2,2,2,2,2,1,1,1,1,1
iziIGB	1,2,1,1,?,?,?,?,9,3,?,6,?,4,1,?,?,4,?,1,2,2,2,2,3,1,1,?,?,?
jmsDOG	2,1,1,1,1,1,3,?,6,3,3,6,2,4,?,1,1,4,2,1,1,1,1,6,2,7,2,1,3,1
jukPLA	1,1,1,1,?,?,1,1,9,?,?,?,?,4,?,1,4,1,2,1,2,2,2,2,2,1,?,?,?,?
kanCR	1,1,1,?,?,?,1,1,7,?,?,?,?,4,?,1,?,4,?,1,2,2,2,2,1,1,1,1,?,2
kbsBAN	2,2,1,1,1,3,5,3,1,3,3,6,2,2,2,2,1,2,?,1,2,2,2,4,2,?,?,1,2,1
kchSON	1,1,1,1,1,1,1,1,7,3,1,6,2,4,2,1,4,5,2,1,2,1,1,2,2,1,1,1,?,1
kcySON	3,2,1,2,1,3,1,1,2,3,3,6,4,4,2,1,1,1,2,1,2,1,1,2,2,1,1,2,1,2
kerECHA	?,?,1,?,1,1,?,?,1,5,1,1,2,2,?,?,?,2,2,1,2,2,2,2,2,1,1,1,?,?
kfeGUR	2,1,1,3,1,1,3,3,2,3,1,2,3,4,?,?,?,2,2,1,2,4,1,2,2,1,1,2,?,1
kikBAN	1,2,1,1,1,2,5,3,1,3,3,6,3,4,2,1,2,4,2,1,2,2,2,2,2,1,1,2,1,2
kisSATL	1,2,1,1,1,1,5,3,8,3,1,6,2,4,2,2,2,3,2,1,3,2,2,2,2,1,1,1,2,2
knkWCHA	?,?,1,?,?,?,?,?,2,?,?,?,?,4,?,1,1,2,?,1,2,2,2,2,2,1,1,?,2,1
knrWSAH	1,2,1,1,1,3,1,1,2,3,3,6,2,2,2,1,1,2,2,1,1,1,2,2,2,1,5,4,3,1

konBAN	1,2,1,1,1,2,5,3,1,3,3,2,2,4,2,1,3,4,2,1,2,2,2,2,2,1,1,2,?,2
kpeWMAN	1,2,1,2,?,?,1,1,2,3,?,6,?,4,2,1,?,4,?,1,1,1,1,2,2,1,1,?,3,1
kseSON	1,1,1,1,1,1,2,1,8,3,1,6,2,4,2,1,4,5,2,1,1,1,1,2,2,1,1,1,?,1
lelECHA	?,?,?,?,?,?,2,2,2,?,?,1,?,2,?,?,?,4,?,1,2,2,3,2,2,1,1,4,?,?
lndUB	1,?,1,1,?,?,4,2,1,5,?,?,?,2,2,1,?,3,2,1,2,2,2,2,2,1,1,?,3,1
luvBAN	1,2,1,1,1,2,5,3,1,3,3,2,2,2,1,2,2,1,2,1,2,2,2,2,2,1,1,2,1,1
mbmADM	1,1,1,1,?,?,1,1,7,3,?,6,?,2,?,?,?,2,?,1,2,2,2,2,2,1,1,1,?,1
miyWCHA	1,2,1,1,?,?,2,2,2,3,?,1,?,2,?,1,4,5,2,3,2,2,2,1,2,1,1,?,2,?
mooGUR	3,1,1,1,1,?,3,3,2,3,?,6,3,4,2,1,?,2,?,1,2,4,1,4,2,1,3,1,3,1
mpoBAN	1,2,1,1,1,3,5,3,1,3,5,6,2,2,1,1,2,1,2,3,2,2,2,2,2,1,2,1,2,2
muaBAN	1,1,1,2,1,2,5,3,1,3,3,6,2,4,2,1,1,1,1,1,2,2,2,6,2,1,1,2,2,1
mumADM	1,2,1,?,?,?,5,3,7,?,?,6,2,4,?,?,?,3,?,1,2,2,3,2,2,1,?,1,?,?
ndiUB	1,1,1,1,1,1,2,2,1,1,1,6,2,4,2,1,4,4,1,1,2,2,2,2,2,?,?,1,3,1
ngmBB	1,1,1,?,?,?,?,?,2,?,?,?,?,2,?,1,4,5,?,1,2,2,2,2,2,1,1,?,?,?
ngoBAN	1,2,1,1,1,2,5,3,1,3,3,6,2,4,2,1,2,4,2,1,2,2,2,2,2,1,1,2,1,1
ngzWCHA	3,2,1,?,?,?,?,1,2,5,?,1,?,2,1,1,?,2,2,1,2,2,2,2,2,1,5,?,2,1
nupNUP	3,1,1,1,?,?,1,1,8,3,?,6,2,4,?,1,4,5,2,1,2,2,2,2,2,1,1,1,3,1
nyfBAN	1,2,1,1,?,3,5,3,1,3,3,2,?,?,?,1,2,4,2,1,2,2,?,1,?,1,4,2,?,1
nymBAN	1,3,1,1,1,?,5,3,1,3,1,6,?,4,2,2,2,4,1,3,2,?,2,2,2,?,?,2,1,2
oboCR	1,2,1,1,1,1,1,1,9,3,3,6,2,4,2,1,4,1,2,1,2,2,2,2,2,1,1,1,1,1
senBAN	1,2,1,?,2,2,5,3,1,3,3,2,2,2,2,1,2,1,2,1,2,2,1,2,?,1,?,2,2,2
shnBAN	1,2,1,1,1,2,5,3,1,3,3,1,2,1,2,2,2,1,2,1,2,2,2,2,2,1,1,2,3,2
shuSEM	3,2,1,?,?,3,?,?,6,3,3,1,3,2,1,1,?,4,?,?,?,2,2,?,2,?,?,?,2,?
snnWMAN	1,2,1,2,?,1,1,1,2,3,1,6,?,4,2,1,?,5,2,1,1,1,1,1,2,1,1,1,?,1
supGUR	1,1,1,1,1,1,5,3,2,3,1,2,2,4,2,1,2,4,2,1,1,1,1,1,2,4,1,1,3,1
swaBAN	1,2,1,1,1,2,5,3,1,3,3,2,2,2,2,2,1,1,2,1,2,2,2,2,2,1,1,2,3,2
taqBER	2,2,1,3,2,3,2,2,6,3,3,1,3,2,2,1,1,?,2,2,2,2,2,2,1,7,1,1,1,1
tbuWSAH	2,1,1,1,1,3,1,1,2,3,3,6,?,2,1,1,?,2,2,1,1,1,2,2,2,1,5,?,3,1
tdaSON	3,2,1,?,1,3,1,1,2,3,3,6,2,?,?,1,4,1,2,1,2,1,2,2,?,?,1,2,1,?
terBM	1,2,1,1,?,?,?,?,8,3,?,6,?,2,2,2,2,5,?,1,2,2,2,2,2,1,1,?,2,1
tmsDOG	2,1,1,1,1,1,?,1,8,3,3,6,2,4,1,1,?,2,2,1,1,1,1,2,2,1,1,1,2,?
tneSATL	1,1,1,1,1,1,5,3,1,3,1,6,2,4,2,2,1,4,2,1,2,2,2,2,2,1,1,1,3,1
tozBAN	1,3,1,1,1,?,5,3,1,3,3,6,5,2,1,2,2,1,?,1,2,5,2,2,2,1,?,?,1,?
tstSON	1,1,1,1,1,1,1,1,2,3,1,6,2,4,2,1,1,?,?,1,1,1,1,2,2,1,1,1,?,1
twqSON	3,3,1,2,1,1,1,1,2,3,1,6,?,?,?,1,4,1,2,1,2,1,1,?,?,?,?,1,?,?
umbBAN	1,1,1,1,4,3,5,3,1,3,3,6,1,2,2,1,2,1,1,1,2,5,2,3,2,1,1,2,2,2
vaiWMAN	1,1,1,2,1,1,1,1,7,3,1,6,4,4,1,1,1,4,?,1,1,1,1,2,2,4,1,1,3,1
wlfNATL	1,2,1,2,1,1,1,1,6,3,1,6,2,2,2,2,1,5,1,1,2,2,2,2,1,1,1,1,1,1
yaoBAN	1,2,1,1,1,2,1,1,1,3,3,6,2,2,2,1,2,1,2,3,2,2,2,2,2,1,1,2,2,1
yorYOR	3,1,1,1,4,1,1,1,7,3,1,6,2,4,2,2,4,5,2,1,2,2,2,2,2,1,1,1,1,1
zanUB	1,2,1,1,?,?,4,2,1,?,?,3,2,2,2,1,2,1,?,1,2,2,3,5,2,1,1,2,3,1
zarSON	2,2,1,2,1,1,1,1,8,3,1,6,2,?,2,1,?,5,?,1,3,1,1,2,2,1,1,1,3,1
zenBER	3,2,1,?,1,3,2,2,6,3,3,1,3,2,2,1,1,4,?,1,2,2,2,2,1,1,?,2,?,?
zulBAN	1,2,1,2,1,2,5,3,1,3,3,2,2,2,2,1,2,1,2,1,2,2,2,6,2,1,1,2,3,2

Kristoffer Friis Bøegh
kristoffer.boeegh@gmail.com
Aymeric Daval-Markussen
linadm@cc.au.dk
Peter Bakker
linpb@cc.au.dk

Submitted: 10 Mar 14
Accepted: 12 Oct 15
Revisions received: 5 Sept 16