

RARE AND ENDANGERED – LANGUAGES OR FEATURES?
AN AFRICAN PERSPECTIVE

Friederike Lüpke
Endangered Languages Academic Programme
School of Oriental and African Studies London
friederike.luepke@soas.ac.uk

Motivated by the hypothesis that rare and endangered languages exhibit a higher ratio of unusual features than less uncommon ones, this paper uses four case studies to demonstrate that the hypothesis is not born out for African languages. There is no correlation between the rarity of a language and the rarity of areas of its structure. The reasons for the independence of the presence of unusual features from the rarity of the language in which they occur are the following: first, we lack sufficient data and unanimity on cross-linguistic categories and features for a systematic appraisal of their rarity; second, we lack sufficient linguistic information on the majority of the world's languages, and especially on African languages; and third, features can be rare or endangered and change regardless of the status of the language in which they occur.

Motivé par l'hypothèse que les langues rares et menacées d'extinction présentent un taux supérieur des caractéristiques inhabituelles que celles moins rares, cet article utilise quatre études de cas pour démontrer que cette hypothèse ne tient pas pour les langues africaines. Il n'y a pas de corrélation entre la rareté d'une langue et la rareté des domaines de sa structure. Les raisons pour cette indépendance entre caractéristiques inhabituelles et rareté de la langue dans laquelle ils se trouvent sont les suivants : tout d'abord, il nous manque des données suffisantes et de l'unanimité sur les catégories et traits à travers les langues pour une évaluation systématique de leur rareté ; en second lieu, il nous manque des informations linguistiques suffisantes sur la majorité des langues du monde, et en particulier sur les langues africaines ; et troisièmement, les traits peuvent être rares ou menacés et sujet à changement indépendamment du statut de la langue dans laquelle ils se produisent.

“[F]rom the evidence we have to date, it would appear that the most grammatically complex and unusual languages are [...] often spoken by small tribes whose traditional way of life is under threat.” (Nettle & Romaine 2000: 11)

0. INTRODUCTION

Endangered languages represent more than half of the world's languages, and the disappearance of every single one of them undoubtedly creates an enormous loss – in terms of linguistic diversity, of a unique way of expressing culture through language and of irreplaceable facets of identity and ancestry. Since the first alarm bells were rung at the LSA in 1991 (see Language Issue XXX 1992), considerable and growing linguistic efforts have been made and continue to be made to create lasting records of endangered languages while the languages themselves are fading away. Currently, the description of endangered languages is fusing with the emerging field of documentary linguistics, so that language documentation reads first and foremost as documentation of endangered languages (see also Austin and Grenoble, 2007).

This paper investigates endangered languages exclusively from the perspective of their contribution to our linguistic knowledge of the world's languages. In doing so, the article takes a particular theoretical slant: it tests the idea that small or endangered languages have more unusual, rare, or exceptional traits than other languages. This

idea has been advanced by a number of scholars, but generally lacks empirical underpinnings. Since it would add even more impetus to the urgency to document endangered languages, if it turned out there was a correlation between rarity and endangered status, it seems worthwhile to give this hypothesis some deeper thoughts. The chosen heuristic to investigate whether there is a link between rarity of a feature or characteristic and rarity of the language exhibiting it is to look at rare and/or endangered features from different linguistic domains in a number of African languages. I explore whether the rarity or endangerment of these features is connected in some way with the endangerment status of the language in which they occur. An African perspective is motivated by the observation that Africa tends to be underrepresented in linguistic description and documentation. Since Africa hosts ca. 2,600 languages, a better inclusion of African languages in the theory and practice of language documentation is highly desirable. Four different situations from across the African continent have been selected:

- an endangered and rare cultural practice with linguistic consequences in an endangered language;
- a rare, but seemingly not endangered, linguistic feature of an endangered language not attested in closely related languages, endangered or not;
- a rare, but not endangered, linguistic feature of an endangered language that also occurs in close and unthreatened neighbouring languages;
- and finally, an endangered, but not rare, subsystem of endangered and thriving African languages;

The features inspected in the case studies are the following:

A first case study explores the custom of inserting lip-plates among Chai women of Ethiopia (Yigezu, 1998, Yigezu, 2001) and its linguistic consequences. The lip-plates have important effects on the phonetics of women's speech and result in a reorganisation of the phonological system of the language. Other African cultures in which this custom was reportedly present have abandoned it, independently of the status of the language(s) in which they were used. From there, the discussion moves to morphology and its interface with semantics by discussing the different and typologically highly unusual deictic classifiers in the demonstrative word in the Nigerian language Goemai (Hellwig, 2003; and in this volume). The compulsory presence of postural classifiers in the demonstrative forces speakers to always pay attention to the position of a Figure with respect to a Ground when referring to it with a demonstrative. Although other Chadic languages possess the necessary linguistic structures – complex demonstratives and an inventory of posture verbs – to develop such a classification system, they have not done so. The domain of syntax is represented with the passive alternation of the endangered Guinean language Jalonke (Lüpke, 2005, Lüpke, 2007). Cross-linguistically extremely rare but also attested in other Mande and African languages (Cobbinah and Lüpke, submitted), this language has a zero-coded passive. Unthreatened languages of the Mande group also exhibit this construction, among them the major West African lingua franca Manding. Not associated with a single language is the section on written language as a linguistic subsystem, which presents the ongoing use of Arabic-based scripts for the writing of African language throughout the entire Sahel (Lüpke, 2004) in informal and marginalised contexts. While this exographic writing tradition is by no means rare, it has become endangered due to the insistence on the Roman script in formal contexts.

After a discussion of rarity, the case studies and their relevance for endangerment theory are treated in turn, before a final section draws the conclusions to be gained from the case studies for the theory and practice of language documentation: that the position that rarity is linked to endangerment or minority status of a language does not hold, and what the consequences of this finding are.

1. RARITY: LANGUAGES VS. FEATURES

1.1 A LINK BETWEEN RARE OR ENDANGERED FEATURES AND ENDANGERED LANGUAGES?

As mentioned above, it is often assumed, but almost never empirically investigated that endangered languages or languages spoken by small speech communities show disproportionately more rare traits than languages spoken by large speech communities. For instance, Wunderlich (2002) ascribes the presence of a large number of lexical irregularities and more complex morphology in smaller or endangered languages to the size of the community:

“In small populations (of, say, around 300 people) it is an advantage to use stored items that are commonly known because they are faster processible the more frequent they are, while in a larger population or a population with many contacts there is need for marking topic and focus, and the use of stored items is less felicitous, since people do not share the same set of items.” (2002: 14)

In the same vein, Nettle (1999) presents the results of a computer simulation that shows a longer survival rate for unusual or unstable features in small-scale communities.

A recent special issue of the journal *Linguistische Berichte* (Austin and Simpson, 2007), dedicated to the linguistics of endangered languages, comes up with an impressive number of unusual or unique characteristics in the endangered languages covered. These are the extraordinary relevance of kinship for morphosyntactic distinctions in Australian languages (Evans, 2007), basic OVS word order in Uranina (Olawsky, 2007), length harmony in Leggbó (Hyman and Udoh, 2007), quirky subjects in Garifuna, a language without case marking (Munro, 2007), person as a feature determining agreement in Archi (Chumakina et al., 2007), a zero-marked passive in Jalonke (Lüpke, 2007), an unusual distribution and use of the ergative marker in Kinnauri (Saxena, 2007), phoneme inventories of extreme sizes in the Asia-Pacific region (Hajek, 2007), complex spatial deixis in Northern Wakashan languages (Bach, 2007) and morphologically nonmarked direct and inverse systems in Mesoamerican languages (Zavala, 2007).

1.2 NO LINK BETWEEN RARE OR ENDANGERED FEATURES AND ENDANGERED LANGUAGES?

Undoubtedly, then, endangered languages do have unusual and rare features or characteristics like the ones listed above - but is their existence accidental or systematic? Authors like Nettle (1999), Nettle and Romaine (2000) or Wunderlich (2002) emphasise the regular association of rare features with rare/endangered host language, without statistically proving such an association in spoken languages. In

contrast, a number of scholars deny this association on theoretical or empirical grounds (Bickel and Nichols, 2003, Cysouw, forthcoming).

Haspelmath (2007: 123) points to the fact that “almost every newly described language presents us with some "crazy" new category that hardly fits existing taxonomies”, and gives examples from major languages such as Japanese along with more ‘exotic’ languages such as the Chadic language Mina.¹ Lüpke (2007) argues that often, categories such as the zero-marked passive of Jalonke, see also 5 below, are assumed to be rare because they are excluded from too restrictive cross-linguistic definitions and hence, if described at all, are not described in terms of the category in question. The result is that these categories go unnoticed and therefore cannot feed cross-linguistically sounder definitions. Closely linked to this argument is the ongoing debate in typology on to what extent linguistic description and theory need to rely on pre-established categories, and what their nature should be, and to what extent categories need to be language-specific (see Haspelmath, 2007 vs. Newmeyer, 2007, for an example). In view of the many unresolved issues, I largely neglect the nontrivial question of how assumptions about cross-linguistic categories influence judgment on what is rare and what is not and base this paper on the assumption that the categories classified as rare in, e.g., the *World Atlas of Language Structures* (henceforth WALS, Haspelmath *et al.*, 2005), are rare indeed and flag it as an issue that needs further discussion.² However, it turns out that even if the available typological information on features is taken at face value, it cannot serve to establish a correlation between rarity/ endangerment of a language and rarity of a feature.

1.3 INCONCLUSIVE QUANTITATIVE STUDIES

Using the WALS, which shows the distribution of 142 linguistic features in between 120 and 1,370³ languages, Cysouw (forthcoming) investigates the notion of rarity empirically by computing the probability of occurrence for values of the covered features. Table 1 shows the 6 languages that have the highest overall rarity index according to his statistical analysis with the number of speakers as given in Gordon (2005).

Table 1: Languages with the highest overall rarity index and their speaker numbers

Language	Phylum	Number of speakers
Wari'	Chapacura-Wanhan	1,833
Dinka	Nilotic	2,000,000 or more
Jamul	Tiipay Yuman	220
Nuer	Nilotic	804,907
Karó (Arára)	Tupi-Guarani	150

¹ Dryer (1997) makes a similar point in vehemently arguing against the assumption of absolute universals, since they are so easily invalidated as linguistic descriptive knowledge advances.

² Bickel & Nichols (2003) present a typology of rarity which is based on the dichotomy between absolute and relative rare. The former are universally rare, while the latter are rare at the universal level but common in a specific area. This distinction is not adopted here because, as will be argued below, it cannot be meaningfully applied to every linguistic category, see discussion in section 2.3..

³ In total, up to 2,600 languages are covered. Since far from all languages are represented for all features investigated, only 55,000 out of the possible 369,000 data points are coded in WALS, corresponding to 15% of the data matrix.

Winnebago

Siouan

230

Languages with small speaker numbers prevail, yet large languages are represented among the top 6 as well. However, the data suffer from the same weak point as all typological investigation of rarity, possible terminological and theoretical biases and limits of cross-linguistic comparability left aside: the scarcity of data. Both the structural information on the world's languages and their speaker numbers is far from offering a balanced overview of possible language structures on the one hand and on the sociolinguistic situation of languages on the other hand. It should be noted that 'rare language' in the terminology of Cysow means languages that scored as rare according to a maximum of 142 features investigated in WALS. Not only are languages certainly composed of considerably more features, but more importantly linguistic knowledge regarding a possible universal inventory of features and the values attached to them is in its infancy (Evans and Levinson, 2009). In addition, the majority of features investigated in WALS pertain to the formal side of linguistic structure, since typology has traditionally focussed on phonology, morphology and syntax. Only occasionally does WALS touch on semantic features, such as the existence of a have-perfect. This exclusion of semantics is not surprising, given the relative recency of the field of semantic typology ; for a seminal study see Talmy (1985); for recent examples see Pederson *et al.* (1998), and Levinson and Meira (2003). But even the largely neglected domain of semantics (and pragmatics) notwithstanding, clearly languages are made of many more features than 142, that yet remain to be determined before being compared! Until the inventory of features relevant for all areas of languages has been extended and systematised⁴, the assertion that a language is 'overall rare' only has meaning in relation to features covered by WALS, not in an absolute sense.⁵

Finally, since one of the major findings of WALS is the importance of geographical proximity for languages to share features and/or their values, it is worthwhile to mention Cysow's (forthcoming) finding of north-western Europe as the most unusual linguistic area worldwide, with a high number of quirks not attested anywhere else. While Frisian, one of the north-western European languages in question, is endangered, Dutch, English, French and German clearly are not. Cysow argues that it is the socioeconomic weight of these languages and the dominant position of European researchers that most plausibly led to the inclusion of features in WALS that turn out to be irrelevant for the majority of the world's languages, such as the have-perfect or word order change in polar questions.

While surveys like that in WALS which compile and systematise data of a quantity unheard of, they clearly cannot be used (yet?) to answer questions like the one of a possible correlation between rare status of a feature and endangered status of its host language, and hence an informal approach like the one taken here must suffice until a much larger data base is available. Let us now start this informal survey of possible correlations between rare features and minority or endangered languages.

⁴ The project "Grammatical features" of the Surrey Morphology Group is an attempt at doing exactly that. See <http://www.surrey.ac.uk/LIS/SMG/featuresdescription.html>.

⁵ In addition, it remains to be investigated whether languages tend to be overall rare in the first place, i.e. whether there tends to be a clustering of exceptional traits in a language, or whether rare characteristics can occur in isolation.

2. PHONETICS/PHONOLOGY: THE EFFECT OF LIP PLATES ON FEMALE SPEECH AMONG CHAI WOMEN OF ETHIOPIA (YIGEZU 1998, 2001)

2.1 THE CHAI PEOPLE

The Chai people of southwestern Ethiopia and their closest neighbours, the Tirma, and Baale, share the ethnonym Suri. The number of Suri people can only be estimated – Yigezu & Dimmendaal (1998), approximate them to 60,000 people, whereas Abbink (2000) gives their number as 28,000.⁶ Closely related culturally as well as linguistically are the Mursi, estimated by Turton (2004) to count less than 10,000 people. Chai is spoken by approximately 13,000 people according to Yigezu and belongs together with Tirma, Baale and Mursi to the Surmic branch of Nilo-Saharan.⁷ The Suri are agropastoralists and inhabit a very remote area of southwestern Ethiopia. They have barely any access to modern infrastructure and are under strong pressure from invading cattle-raiding neighbouring groups and rebels operating in the area bordering Sudan with who they are involved in violent conflicts (Abbink, 1993). It is a safe assumption that the closely related and almost unknown languages spoken by the Suri can be classified as endangered.

2.2 LIP-PLATE INSERTION

Suri and Mursi women practice the custom of lip-plate insertion. Girls attaining puberty wear a lip-plate inserted into the lower lip, which has previously been pierced and stretched. The size of the plate is gradually increased until the chosen size is reached. In order to accommodate the lip-plate, which can be up to six inches in diameter and weigh up to 200 grams, the lower incisors are removed as well. The plates are taken out for eating. Although folk theories explaining the custom are widespread among travellers and journalists – it is said to originate as an attempt to keep the women safe from slave raids, and a large size of the lip-plate is said to correlate with a high bridewealth - these are unconfirmed by anthropologists (see, e.g. Turton, 2004).

2.3 LINGUISTIC CONSEQUENCES

Due to the deformation of the lower lip and the removal of the lower incisors, Chai women are unable to articulate bilabial and dental consonants, several of which have phonemic status.

Yigezu (2001: 208) identifies three main strategies to compensate for this handicap:

1. The place of articulation of the consonant is maintained, but its manner changed.

⁶ The differing component groups of the ethnonym Suri as well as the diverging numbers illustrate the problematic nature of a monolithic concept of ethnicity and the impossibility to determine the number of members of an ethnic group (and even more so of speakers of (a) given language(s)). Both have been noted in general for Africa and for the Suri in particular.

⁷ Despite strong ethnic bonds between the four groups, their languages belong to two different branches of Surmic and do not pattern with the common ethnonym: Tirma, Chai and Mursi are Southeastern Surmic languages, whereas Baale is classified as Southwestern Surmic (Yigezu, 1998).

2. The manner of articulation of the consonant is maintained, but its place of articulation changed.
3. Both place and manner of articulation of the consonant are changed.

Table 2 below exemplifies bilabial, dental and alveolar consonants in normal speech and their counterparts in lip-plate speech according to the different strategies.

Table 2: Sounds in normal and lip-plate speech

	Normal (male) speech	Lip-plate speech	Relation between normal and lip-plate speech
Strategy 1	[b]	[β]	= place
	[ɸ]	[w]	≠ manner
	[β]	[ɸ]	
	[b]	[d]	
Strategy 2	[ɸ]	[d]	= manner
	[s]	[ɸ]	≠ place
	[p]	[t]	
	[m]	[ŋ]	
Strategy 3	[b]	[d]	≠ manner ≠ place

Cross-cutting the strategies are the mechanisms underlying them: Yigezu describes some of them, for instance the change from plosive to approximant of the same place of articulation, as passive or “mechanical and low-level changes, directly resulting from the change in the geometry of the front part of the mouth” (2001: 209). Other changes, such as those involving the replacement of a bilabial with a velar nasal, he claims, are active or “cognitive changes made in order to maintain some kind of acoustic-perceptual quality rather than articulatory quality” (2001: 209).

Vowels are equally affected by acoustic changes resulting from changes in the articulators in lip-plate speech. The back rounded vowels become unrounded, as evidenced acoustically through a higher second formant. For a list of affected vowels, see table 3.

Table 3: Chai vowels in normal speech and their lip-plate speech counterparts

Normal (male) speech	Lip-plate speech
[o]	[ɤ]
[ɔ]	[ʌ]
[ʊ]	[u]
[u]	[ə] ⁸

While Yigezu’s study shows how speakers systematically compensate for articulatory deficiencies by reorganising existing speech sounds of the language, and thus create a parallel phonological system coexisting with the ‘normal’ one, many questions remain unanswered so far. We do not know, for instance, how lip-plate and normal speech are acquired by children – prepubescent girls have the ‘normal’ sound inventory’, but switch to lip-plate speech once the disks are inserted. Since not all of

⁸ Yigezu (2001: 214) uses the schwa to symbolize the unrounded counterpart of the high back vowel u, not the corresponding IPA symbol.

the changes are mechanical, a learning process is required. We have no knowledge either if phonetic differences between the genders survive once the practice of wearing lip-plates is abandoned.

2.4 ENDANGERMENT RELEVANCE

Lip-plate insertion is an exceedingly rare practice at the beginning of the 21st century. In addition to the Suri, lip-plates are or were worn by male members of the Suya and Kayapo groups in the Amazon region of Brazil. Colonial testimonies (Meek, 1925, Tremearne, 1912), however, describe the practice as widespread in Africa, citing several ethnic groups of Nigeria among the tribes inserting lip-plates, for instance the Angas and Borok. Lip-plates were worn until several decades ago by women of the Makonde and Sara groups as well.

Several languages whose speakers practice or allegedly used to practice lip-plate insertion can be qualified as endangered. However, there is no principled reason to correlate the presence of the custom, resulting in the extremely rare case of two phonetic and phonological systems for one language, with the endangered status of that language. Turton describes pressures on the Mursi to abandon the practice from outside and within the community. The Ethiopian government and missionaries operating in the area regard lip-plates as uncivilised; and the Mursi themselves have come to see them as backwards under these influences. It is plausible to assume similar motivations for abandoning lip-plates in other cultures. Ironically, and counter to claims made about unstable and unusual features thriving in isolated communities and not surviving in contact situations, contact seems to contribute preserving them, at least in the short term, among the Chai and Mursi. Abbink (2000) and Turton (2004) report how western tourists invade the Suris and Mursis in growing numbers in order to photograph women wearing lip-plates, seemingly a quintessential symbol of 'tribal culture'. The women charge money for having their photos taken, thus creating an important revenue and motivation to perpetuate the custom. Rather than remoteness protecting an unusual, particular feature, exposure to and contact with outsiders contribute to preserve it, totally against the logic of Nettle & Romaine (2000).

3. MORPHOLOGY/SEMANTICS: DEICTIC CLASSIFIERS IN THE DEMONSTRATIVE WORD IN GOEMAI (HELLWIG, 2003)

3.1 THE GOEMAI PEOPLE

Goemai, a language of the West Chadic branch of Niger-Congo, is spoken by about 200,000 speakers⁹ on the Jos plateau in central Nigeria. Despite the seemingly high number of speakers, Goemai can be regarded as endangered, due to a longstanding pressure of the contact language Hausa, resulting in massive language shift among the younger generation of speakers, and exacerbated by a recent series of interethnic riots in the area.

⁹ As is the case for not only for Ethiopia and Nigeria, but for Africa in general, numbers of speakers have to be treated with extreme caution. Either they result from census data using too simplistic or misleading questionnaires only investigating ethnicity or one language spoken, or they originate from colonial census data with an added estimated population growth.

3.2 THE PHENOMENON

Goemai has a typologically unusual system of nominal classification. Deictic classifiers occur in the demonstrative word and categorise the referent of the head noun **or** the nominal concept (but not the noun) according to “its shape, extendedness, position and/or animacy” (Hellwig 2003: 239; this volume). There are five different deictic classifiers¹⁰ in the language, four postural and an existential one. They are based on locative verbs, three of which are derived from postures (‘sit’, ‘stand’, and ‘lie’); the fourth locative verb is ‘hang/move’. Locative verbs in Goemai have two main uses, an assertional and a classificatory one. In assertional use, the verbs encode the actual position of the referent of the verb’s argument. In classificatory use, it is not the current position, but the canonical (i.e. typical or default) position of the verb’s argument or Figure (Talmy, 1985) that is encoded by the locative verb. Examples (1) and (2) illustrate the two different functions of locative verbs.

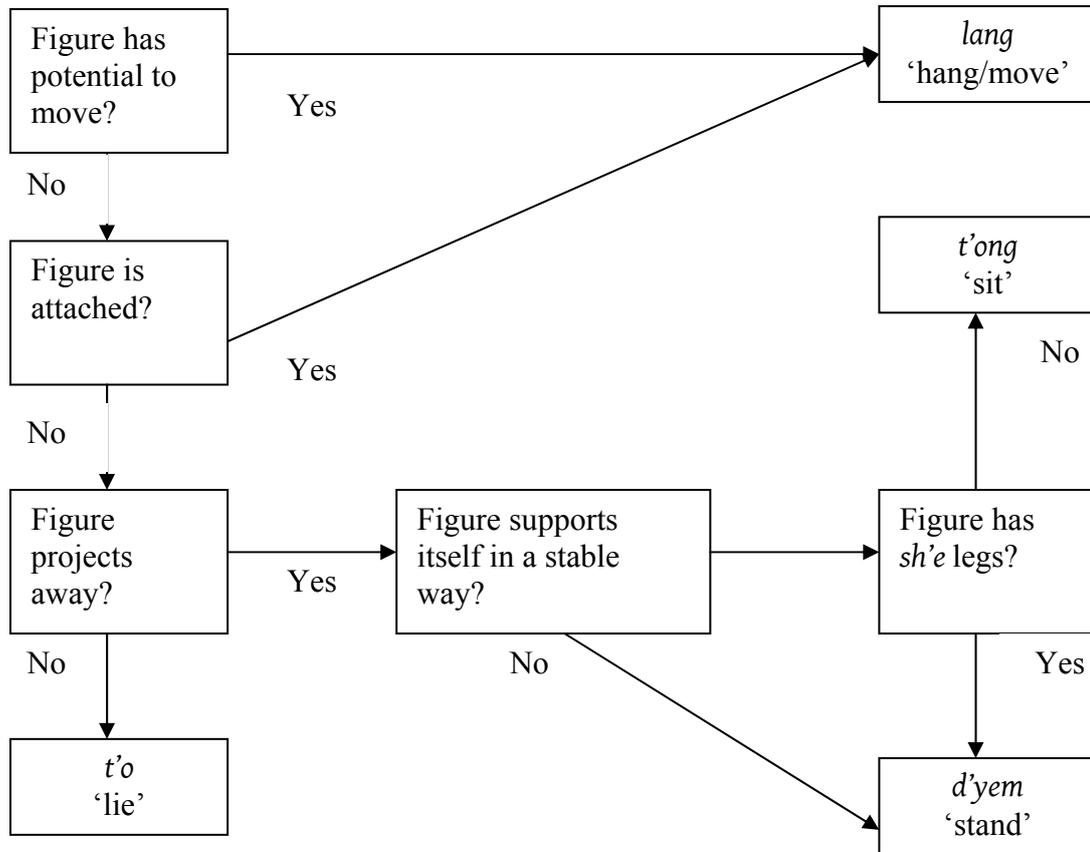
- (1) *T’eng Yuut t’o sek p’ang*
 tree in_mass lie(sg) BODY Stone
 ‘The trees lie in a mass near the hill.’ (= lying trees)
 (Hellwig 2003: 136)

- (2) *Ni t’ong d’i k’a*
 3SG Sit(sg) LOC.ANAPH HEAD(sg)
 ‘It sits there on top.’ (= upside down)
 (Hellwig 2003: 178)

In their basic sense, locative verbs encode the position of a Figure with respect to a Ground. The criteria for the selection of a posture verb in these unmarked contexts are summarised in the flowchart in Figure 3 below (Hellwig, 2003: 144).

¹⁰ Following Aikhenvald (2000), Hellwig uses the term ‘deictic classifier’ for classifiers that occur in deictic expressions, not for morphemes that classify deictic expressions.

Figure 3: selection criteria for Goemai posture verbs



Figures in Goemai are associated with a typical or canonical position, and this position can be described by a posture verb – animate entities and natural forces, for instance, by default ‘hang/move’, whereas objects construed in the language as having a base ‘sit’. These entities comprise most bottles, baskets, and fruits sitting on their growth point, if they have one. Trees, buildings, and entities being perceived as having legs ‘stand’, while masses, flexible objects, and unfeatured objects, such as balls, ‘lie’. The correlation of a default posture with a Figure gives rise to the second use of posture verbs in Goemai, the classificatory function. Even when the current position of the Figure deviates from the default position, the default postural can be used to classify it as belonging to the class of ‘sitting’, ‘standing’, lying’ or ‘hanging/moving’ objects. Classificatory uses occur in the following situations:

- The Figure is in a position different from the default position, but the current position is not in the focus of attention (cf. (3))
- Several Figures are in different positions (cf. (4))
- The Figure is not located at all (cf. (5)).

(3) *Wang k'oon t'ong k'a kuk sh'ep.*
 pot face_down(sg) sit(sg) HEAD(sg) stump wood
 ‘The pot sits face down on the tree stump’ (= upside-down pot)

- (4) *Cup* *Hok* *t'wot* *k'ek* *tebul*
 cup DEF sit(pl) HEAD(pl) table
 'The cups sit on the table.' (= some upright, some upside down)
 (Hellwig 2003: 172)
 (Hellwig 2003: 173)
- (5) *Hen* (...) *shyang* *nd'ûûn* *toeb'al* (...)
 1Sg hunt/watch INSIDE calabash
 'I (...) looked into the calabash (...).
Kwalba *t'ong* *d'i* *ba.*
 bottle sit(sg) LOC.ANAPH NEG
 The bottle didn't sit there.' (= non-existing bottle)
 (Hellwig 2003: 173)

In some syntactic environments, for instance in verbal clauses, speakers can choose between the two uses, and they do so based on a number of pragmatic criteria. If the location of the referent is in focus, the classificatory use prevails. If the position of the referent with respect to a Ground or its internal posture is in Focus, the assertional use occurs. In other environments, only one interpretation is possible. This is the case of demonstratives: as the term 'deictic classifier' for the postural verbs occurring in them suggests, the referent of the head noun is classified according to its canonical position. In these contexts, if the position of the Figure cannot be described by a posture verb, or if it is unknown, noncanonical or unimportant, a classifier based on the existential verb is chosen.

The structure of the demonstrative word in Goemai is summarised in Figure 4 (after Hellwig 2003: 242; and in this volume).

Figure 4: The structure of the demonstrative word in Goemai.

Nominaliser		Ad- verbialiser	Deictic classifiers				Deictic root
Sg	Pl			Singular	Plural	Gloss	
<i>goe-</i>	<i>moe-</i>	<i>n-</i>	Postural classifiers	<i>lang-</i>	<i>leng-</i>	'hang/move'	<i>-nang</i>
				<i>t'ong-</i>	<i>t'wot-</i>	'sit'	
				<i>d'yem-</i>	<i>d'yam-</i>	'stand'	
				<i>t'o-</i>	<i>t'oerop-</i>	'lie'	
			Existential classifier	<i>d'e-</i>		'exist'	

Examples (6) and (7) show the complementary distribution of postural and existential classifiers: In (6), the Figure, a calabash, is in its canonical position, hence the postural classifier is used in the demonstrative. In (7), a calabash is lying upside down, and instead of a postural classifier, the existential classifier appears.

coast of Guinea by migrants from this area. Just like the other varieties of Yalunka, hundreds of kilometres separate the Jalonke from them.

4.2 THE CONSTRUCTION

Jalonke and other Central Mande languages exhibit a construction in which the Theme object of a transitive verb is linked to subject position. In Jalonke, the Agent of the clause is not syntactically expressed at all. Consider the sentence pair in (8) and (9):

(8) *O sarax -aa baa.*
 2PL sacrifice -DEF extract
 ‘You do a sacrifice.’
 Ataya 241

(9) *Ka saraxa mun baa?*
 DISC sacrifice NEG extract
 ‘Or isn’t a sacrifice done?’
 Ataya 242

This construction is shared with the majority of Central Mande languages. One non-trivial difference across these languages concerns the admissibility of an adjunct phrase containing the Agent of the event denoted by the verb; some languages, for instance Bambara, allow the equivalent of an English *by*-phrase, whereas others, like Jalonke, do not. In the absence of an overt Agent, the question arises of whether this construction is best analysed as a passive or whether verbs appearing in syntactic contexts like (8) and (9) are labile in valence and orientation.

I have argued in detail elsewhere (Lüpke 2005, 2007), that the construction in (9) fulfils all but one criterion characterising ‘basic passives’ cross-linguistically (Keenan and Dryer, 2007, Siewierska, 2005) – it serves the function of foregrounding the Theme of the verb action; the corresponding active verb form is transitive and has an Agent subject and a Theme object; and the passive is agentless. Sentences like (9) cannot be interpreted as instances of argument ellipsis, resulting in a clause without a subject. The position of the negation marker, also recognised by Creissels (1991) as a diagnostic for grammatical relations in Mande, reveals that the erstwhile object has been linked to subject- the negation marker always follows subject but precedes the object and the verb.¹¹ The construction cannot be understood as a topicalisation, either. Topicalised elements in Jalonke are always resumed *in situ* by an anaphoric pronoun, as in the following example:

¹¹ The negation marker cannot be analysed as a clitic in Wackernagel position in order to suggest that the single argument in (9) might not be in subject position. The negation marker does not always occur in second position in the clause; and regularly occurs in third position in cases of topicalisation, as in (10) above, and when question, discourse and focus markers occur initially.

- (10) *Jombentεε fan, E m' aa koogu i*
 Jombente Too 3PL NEG 3SG marry at
 'Jomebente too, they haven't married her off.' Kooguna 681

Truthconditionally, active-passive pairs are equivalent: although syntactically, no Agent is expressed, its presence is entailed, as evidenced by semantic tests:¹²

- (11) *Bande -εε jin. (!Muxi Nda a jin.*
 Food -DEF cook Person Some 3SG cook
 'The food was cooked. (Somebody cooked it.
 A *mun jin. A kan tagi i)*
 3SG NEG cook 3SG owner middle at
 It didn't cook by itself.'
- (12) *Bande -εε jin. (*Muxi oo m' aa a jin.*
 Food -DEF cook person whatever NEG 3SG 3SG cook
 'The food was cooked. (*Nobody cooked it.
 A *jin. A kan tagi i)*
 3SG cook 3SG owner middle At
 It cooked by itself).'

The only criterion according to which the Jalonke passive falls out of the scope of a basic passive is the presence of morphological or periphrastic marking on the verb. As evident from the examples above, the two verb forms are identical. Yet, just as middles are formally marked in some languages (Greek, Fula), middles without formal marking have been recognised as well (English), so why turn the presence of formal marking into a *sine qua non* condition for the passive?

4.3 ENDANGERMENT RELEVANCE

A zero-marked passive is non-existent according to definitions that insist on the presence of some periphrastic or morphological marking on the verb or in the verb phrase. It is therefore not surprising that only very few claims about zero-marked passives analogous to the one of Jalonke have been made, among them for the Austronesian language Mangarrai (Arka and Kosmas, 2005). The rarity of passives without formal marking might justify their exclusion from cross-linguistically inspired definitions of the constructions such as Keenan & Dryer (2007). But I would like to argue that it is just as likely that definitions demanding the presence of a formal difference between active and passive verb forms have led to zero-marked passives being described in other terms. To appreciate this point, consider examples from Carlson (1994), also used by Keenan & Dryer (2007):

- (13) *nàŋa à sikàŋi bò*
 man.DEF PERF Goat kill
 'The man killed the goat.' (Carlson 1994: 251)

¹² This passive interpretation does not follow from pragmatic inferencing, but is semantically entailed for all transitive verbs of the language, with five exceptions. See Lüpke (2007) for a discussion of these verbs.

an Arabic script not for the writing of Arabic but for the writing of African languages. A wealth of religious and profane literature in Hausa, Fula, and Wolof (cf. Caron, 2000, Dalby, 1986, Philips, 2000, Seydou, 2000) and other languages, is visible proof of this literary tradition. The modified Arabic script employed for the writing of African languages is often referred to as Ajami¹⁴ (the term used for the script in Hausa and Fula) or Wolofal (the name in Wolof). Ajami and other related scripts are based on a modified Arabic alphabet introduced in the sphere of influence of Islam throughout Africa and disseminated by Islamic scholars. Historically, they enjoyed an enormous prestige and were often the only form of literacy attested in the relevant cultures. In contemporary societies however, Ajami writing has become marginalised and exists in an exclusively informal and hidden niche, that of Q'uranic education. All formal domains of literacy have been taken over, at least in theory, by the official languages of the countries in question or by Latinised orthographies for African languages, qualifying Ajami writing systems as endangered subsystems irrespective of the status of the languages in which they are used. In the following, a number of West African Ajamis serve to illustrate this point.

5.2 THE ONGOING USE OF AJAMI WRITING

For the Chadic language Hausa spoken in Nigeria and Niger, the use of Ajami is attested as early as in the 17th century (Philips, 2000: 19). Over 20,000 manuscripts in Ajami held in the Nigerian National Archives (Philips 2000: 27) are proof of this long and flourishing culture of writing in Hausa. The decline of this Ajami began when it was officially replaced with Romanised Hausa by the British colonial administrators, although “[t]he informal use of Ajami in manuscripts by scholars, merchants and others continues today wherever there are Hausa speakers” (Philips, 2000: 27).

Similar observations as for Hausa hold for the Ajami used for the Atlantic language Fula, spoken throughout the entire Sahel. Accounts of the historical importance of writing in Arabic letters are not available for all countries and dialect areas. Nevertheless, it can be safely stated that pre-colonial Fula literature in Ajami covered religious, political, administrative, poetic and personal texts and was most prolific wherever Fula states existed, as in Senegal, Guinea, and North Cameroon (Seydou, 2000: 64-65). For some areas, such as the Futa Jalon in Guinea, a brief history, a catalog of texts ranging from the 18th to the 20th century, and a partial evaluation of the contemporary role of the script are available. In contrast to Hausa, for which the use of the Arabic script was actively discouraged and replaced by Romanised Hausa by the British, the French colonisers of Guinea ignored indigenous Fula writing traditions, since their goal was to create a population literate in French. The continuing use of Fula Ajami cannot be systematically measured. For the Futa Jalon region of Guinea, the facts point to a continuing popularity of Ajami: although a standardised Roman orthography was created for the Fulfulde variety of Fula in Guinea and used in adult literacy campaigns, the Ajami tradition persists until today, seeing the birth of new genres, and resulting in a flourishing written environment. Salvaing & Hunwick remark:

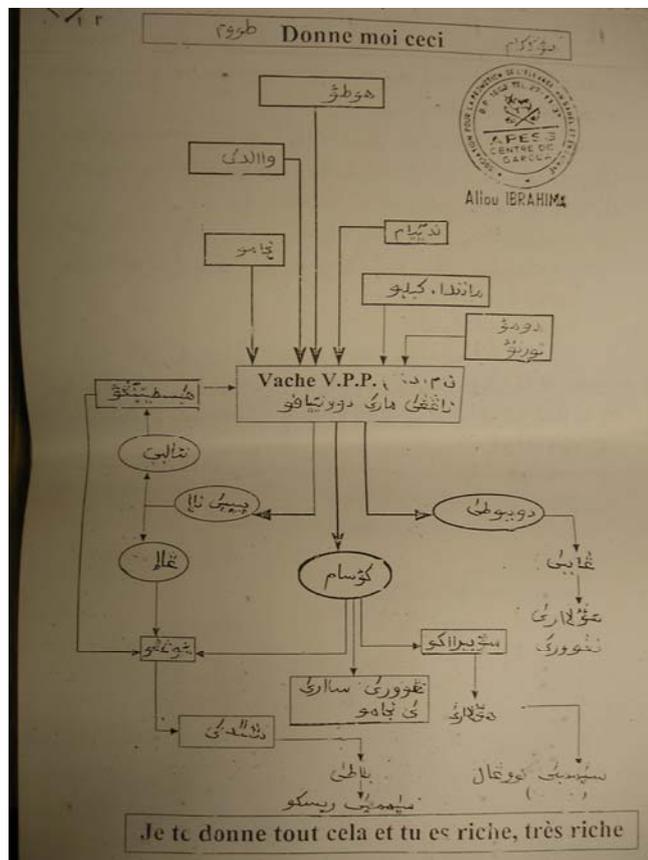
“Today, even slightly educated folk are capable of reading and writing Fulfulde in ajami script, at least for matters of everyday life and private correspondence. The great

¹⁴ Ajami is originally a cover term for the use of the Arabic script to write anuse Ajami only for codified versions of writing in Arabic script, as developed for Hausa and Fula, in contrast to ad hoc uses of Arabic script for the writing of African languages.

spread of written Fulfulde does not seem to have been hindered by the abandonment of teaching Fulfulde in public schools fifteen years ago, when the government, based on the work of the Military Committee for National Recovery, gave preference to French.” (Salvaing and Hunwick, 2003: 503-4)

The thriving tradition of writing in Fula Ajami is certainly an exception in the African landscape – the existence of a codified orthography, an important body of religious and profane literature as well as its ongoing and uninterrupted use make (Guinea) Fula Ajami a unique case. Yet, this literacy is nowhere officially documented or instrumentalised in formal education or adult literacy in Guinea – almost unbelievable since it seems to be the form of literacy that is the closest to the UNESCO’s definition of literacy linking reading and writing to ‘everyday life’. Figure 3 shows a sample of writing from a workshop on cattle herding in Ngaoundéré in Cameroon.

Figure 3: A diagram illustrating the importance of cattle in Fula Ajami, from Ngaoundéré, Cameroon, December 2004



Another documented case of the use of Arabic letters for a language other than Arabic is the Atlantic language Wolof, spoken in Senegal. The creation and use of Wolofal, according to Camara (1997), dates back to the 17th century, for this language is tightly linked to the Islamic brotherhood of *mourides*. The *mourides* have been very influential in the Senegalese religious landscape, and the use of Wolofal for religious and poetic writings in their realm has resulted in an important body of literature, for the most part preserved in private libraries and copied by hand.

Apart from the visible and formalised uses of Ajami and Wolofal, Arabic-based scripts are formally and informally used for letter writing in these and other languages. My own field research in Cameroon revealed that graduates of Q’uranic schools are able to read (and to a lesser extent to write) the Ajami used in their region – Hausa Ajami, Fula Ajami or even both. Even for languages lacking a formalised Ajami tradition, informal

and even ad hoc writing in Arabic characters is attested, for instance for the Mande languages Soso (Guinea), Mogofin (Guinea) and several varieties of Manding spoken in Mali, Gambia, Guinea, Guinea Bissau, Ivory Coast, and Senegal .

5.3 ENDANGERMENT RELEVANCE

It is expected that the semiformal to spontaneous use of the Arabic alphabet for the writing of African languages is much more widespread than reported so far, for two reasons:

1. the important role of Islamic education, leading to literacy in Arabic, throughout the concerned West African countries
2. the marginalisation of the role of Islamic education and the failure to take the resulting non-Latin-based and grass root literacy into account.

However, these same reasons are also responsible for the endangered status of Ajami writing: the formal education system in all the concerned countries relies almost exclusively on the official languages. Where one finds experiments regarding the use of national languages in primary or adult education, a Roman script is used without exception. This is deplorable, given its potential to raise literacy rates in zones officially regarded as having the highest illiteracy rates in the world, and with stagnating numbers of (officially) literates. But due to a total lack of institutional support, Ajami writing is doomed to a very limited and threatened status. Just like other endangered subsystems of language associated with schooling and education, for instance numeral systems (Comrie, 2005), the decline of Ajami is entirely independent of the status of the languages in which it is used – Hausa with 24,000,000 speakers and Fula with 13,000,000 speakers according to Gordon (2005) are huge *linguae francae*, while Wolof with 3,600,000 is dominant in Senegal and Gambia. Again, nothing permits us to deduct the status of this feature or subsystem from the status of the host language.

6. CONCLUSION

This paper has not demonstrated that there are more or less rare features in African endangered languages than in thriving ones. Rather, the results of my informal tour across African languages already show that the debate on rarity needs to be taken one step back. As tempting as it may be to advance the rarity of their linguistic structures as an argument in favour of the documentation and description of endangered languages, no systematic evidence supports this argument. It means close to nothing at present to state that there is no correlation between the status of a language as endangered and an exceptionality of traits of that language. We lack the prerequisites to start this debate in a meaningful way: in order to define rarity, much more information on all of the world's languages is needed; and in order to identify endangered languages, we need instruments more refined than (anyways unreliable) numbers of speakers. In addition, it may turn out that it is impossible to develop a universally applicable ontology of parameters to compare these languages.

These findings do not reduce in any way, however, the urgency of a linguistic description and documentation of endangered languages. Rather, it follows that the descriptive and documentary agenda needs to be extended to encompass all un(der)described and underdocumented languages – these languages, the majority of

the world's languages, if they are not threatened (yet), can become endangered any time. Their features, whatever this is meant to signify, should always be regarded as endangered, since language continuously changes and renews itself. Since African languages are marginalised and underresearched, independently of their endangerment status (Lüpke, 2009), they deserve special linguistic attention.

REFERENCES

- Abbink, Jon. 1993. Ethnic conflict in the "tribal zone": the Dizi and Suri in southern Ethiopia. *The Journal of Modern African Studies* 31:675-682.
- Abbink, Jon. 2000. Tourism and its discontents. Suri-tourist encounters in southern Ethiopia. *Social Anthropology* 8:1-17.
- Aikhenvald, Alexandra Y. 2000. *Classifiers. A typology of noun categorization devices*: Oxford studies in typology and linguistic theory. Oxford: Oxford University Press.
- Arka, I. Wayan, and Kosmas, Jeladu. 2005. Passive without passive morphology? Evidenced from Mangarrai. In *Voice in Western Austronesian*, eds. I. Wayan Arka and Malcolm Ross, 87-117. Canberra: Pacific Linguistics.
- Austin, Peter K., and Grenoble, Lenore A. 2007. Current trends in language documentation. In *Language Documentation and Description, vol. 4*, ed. Peter K. Austin. London: SOAS.
- Austin, Peter K., and Simpson, Andrew. 2007. Introduction to the special Issue "Endangered Languages". *Linguistische Berichte Sonderheft* 14:5-18.
- Bach, Emmon. 2007. Deixis in Northern Wakashan: recovering lost forms. *Linguistische Berichte Sonderheft* 14:253-266.
- Bickel, Balthasar, and Nichols, Johanna. 2003. Typological enclaves. *Paper presented at the 5th Biannual Conference of the Association for Linguistic Typology, Cagliari*.
- Camara, Sana. 1997. 'Ajami' literature in Senegal: the example of Serin Muusaa Ka, poet and biographer. In *Source: Research in African literatures; vol. 28, no. 3, p. 163-182: ill, 1997*.
- Carlson, Robert. 1994. *A grammar of Supyire*. Berlin/New York: Mouton de Gruyter.
- Caron, Bernard. 2000. La littérature haoussa. In *Panorama des littératures africaines. Etat des lieux et perspectives*, eds. Ursula Baumgardt and Abdellah Bounfour, 93-107. Paris: L'Harmattan/INALCO.
- Chumakina, Marina, Kibort, Anna, and Corbett, Greville. 2007. Determining a language's feature inventory: person in Archi. *Linguistische Berichte Sonderheft* 14:143-172.
- Cobbinah, Alexander, and Lüpke, Friederike. submitted. Not cut to fit - zero-coded passives in African languages. In *Proceedings of WOCAL 6*, eds. Matthias Brenzinger and Anna-Maria Fehn. Amsterdam/Philadelphia: John Benjamins.
- Comrie, Bernard. 2005. Endangered numeral systems. In *Bedrohte Vielfalt: Aspekte des Sprach(en)tods*, eds. Jan Wohlgemuth and Tyko Dirksmeyer, 203-230. Berlin: Weissensee.
- Creissels, Denis. 1991. *Description des langues négro-africaines et théorie syntaxique*. Grenoble: ELLUG.
- Cysouw, Michael. forthcoming. Quantitative explorations of the world-wide distribution of rare characteristics, or: the exceptionality of north-western European languages. In *Expecting the unexpected: Exceptions in grammar*, eds. Horst Simon and Heike Wiese. Berlin: Mouton de Gruyter.
- Dalby, David. 1986. *L'Afrique et la lettre/Africa and the written word*. Oshodi: Afprint.
- Dryer, Matthew. 1997. Why statistical universals are better than absolute universals. *CLS* 33:123-145.
- Evans, Nicholas. 2007. Warramurrungunji undone: Australian languages in the 51st millennium. *Linguistische Berichte Sonderheft* 14:19-44.

- Evans, Nicholas, and Levinson, Stephen C. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32:429-492.
- Gordon, Raymond G. Jr ed. 2005. *Ethnologue: Languages of the World* Dallas, Tex.: SIL International.
- Hajek, John. 2007. Sound systems of the Asia-Pacific: some basic typological observations. *Linguistische Berichte Sonderheft* 14:207-222.
- Haspelmath, Marin. 1990. The grammaticization of passive morphology. *Studies in Language* 14:25-72.
- Haspelmath, Martin, Dryer, Matthew S., Gil, David, and Comrie, Bernard eds. 2005. *The World Atlas of Language Structures*. Oxford: Oxford University Press.
- Haspelmath, Martin. 2007. Pre-established categories don't exist: consequences for language description and typology. *Linguistic Typology* 11:119-132
- Hellwig, Birgit. 2003. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria), Faculteit der Letteren, Katholieke Universiteit Nijmegen and Max Planck Institut für Psycholinguistik. (MPI Series in Psycholinguistics, 22.): PhD Thesis.
- Hyman, Larry M., and Udoh, Imelda. 2007. Length harmony in Leggbo: a counter-universal? *Linguistische Berichte Sonderheft* 14:73-92.
- Keenan, Edward L., and Dryer, Matthew. 2007. Passive in the world's languages. In *Language typology and syntactic description*, ed. Timothy Shopen, 325-361. Cambridge: Cambridge University Press.
- Levinson, Stephen, and Meira, Sergio. 2003. "Natural concepts" in the spatial topological domain - adpositional meanings in cross-linguistic perspective: an exercise in semantic typology. *Language* 79:458-516.
- Lüpke, Friederike. 2004. Language planning in West Africa - who writes the script? *Language Documentation and Description* 2:90-107.
- Lüpke, Friederike. 2005. A grammar of Jalonke argument structure, Radboud Universiteit Nijmegen [Max Planck Series in Psycholinguistics 30]: PhD thesis.
- Lüpke, Friederike. 2007. Vanishing voice - the morphologically zero-coded passive of Jalonke. *Linguistische Berichte Sonderheft* 14:173-190.
- Lüpke, Friederike. 2009. At the margin - African endangered languages in the context of global endangerment discourses. *African Research and Documentation* 109:15-41.
- Meek, Charles Kingsley. 1925. *The northern tribes of Nigeria*. vol. 1. London: Oxford University Press.
- Munro, Pamela. 2007. Oblique subjects in Garifuna. *Linguistische Berichte Sonderheft* 14:113-142.
- Nettle, Daniel. 1999. *Linguistic diversity*. Oxford: Oxford University Press.
- Nettle, Daniel, and Romaine, Suzanne. 2000. *Vanishing voices: the extinction of the world's languages*. Oxford: Oxford University Press.
- Newmeyer, Frederick J. 2007. Linguistic Typology Requires Crosslinguistic Formal Categories. *Linguistic Typology* 11:133-157.
- Olawsky, Knut J. 2007. ObVious OVS in Urarina syntax. *Linguistische Berichte Sonderheft* 14:45-72.
- Olson, David R., and Torrance, Nancy eds. 2001. *The making of literate societies*. Malden, Ma./Oxford: Blackwell.
- Pederson, Eric, Danziger, Eve, Wilkins, David P., Levinson, Stephen C., Kita, Sotaro, and Senft, Gunter. 1998. Semantic typology and spatial conceptualization. *Language* 74:557-589.
- Philips, John Edward. 2000. *Spurious Arabic. Hausa and colonial Nigeria*. Madison WI: University of Wisconsin-Madison.
- Salvaing, Bernard, and Hunwick, John. 2003. Writers of Guina. In *Arabic literature of Africa*, eds. John Hunwick and R.S. O'Fahey, 491-529. Leiden/Boston: Brill.

- Saxena, Anju. 2007. Contrastive focus in Kinnauri narratives. *Linguistische Berichte Sonderheft* 14:191-206.
- Seydou, Christiane. 2000. Litterature peule. In *Panorama des litteratures africaines. Etat des lieux et perspectives*, eds. Ursula Baumgardt and Abdellah Bounfour, 63-75. Paris: L'Harmattan/INALCO.
- Siewierska, Anna. 2005. Passive constructions. In *World atlas of language structures*, eds. Marin Haspelmath, Matthew Dryer, David Gil and Bernard Comrie, 434-437. Oxford: Oxford University Press.
- Talmy, Leonard. 1985. Lexicalization patterns: Semantic structure in lexical forms. In *Language typology and syntactic description, vo. 3: Grammatical categories and the lexicon*, ed. T. Shopen, 57-149. Cambridge: Cambridge University Press.
- Tremearne, Arthur J. N. 1912. *The tailed head-hunters of Nigeria*. London: Seeley, Service & Co.
- Turton, David. 2004. Lip-plates and "the people who take photographs". Uneasy encounters between Mursi and tourists in southern Ethiopia. *Anthropology today* 20:3-8.
- Wunderlich, Dieter. 2002. Major steps in the evolution of language. *Paper presented at the Linguistic Summerschool Düsseldorf. Symposium on Evolution*.
- Yigezu, Moges. 1998. Women in society and female speech among the Suri of Southwestern Ethiopia. In *Surmic languages and cultures*, eds. Gerrit Dimmendaal and Marco Last, 83-101. Köln: Rüdiger Köppe.
- Yigezu, Moges, and Dimmendaal, Gerrit. 1998. Notes on Baale. In *Surmic languages and cultures*, eds. Gerrit Dimmendaal and Marco Last, 233-280. Köln: Rüdiger Köppe.
- Yigezu, Moges. 2001. Articulatory and acoustic effects of lip-plate speech in Chai and its implications for phonological theory. *Journal of the International Phonetic Association* 31:203-221.
- Zavala, Roberto. 2007. Inversion and obviation in Mesoamerica. *Linguistische Berichte Sonderheft* 14:267-306.