Studies in African Linguistics

Editor
David Odden

Associate Editors
G. Tucker Childs (Portland State University)
Bruce Connell (York University)
Katherine Demuth (Brown University)
Omar Ka (University of Maryland, Baltimore)
Ronald P. Schaefer (Southern Illinois University)
John V. Singler (New York University)

Publisher
The Department of Linguistics and the Center for African Studies
Ohio State University, Columbus, Ohio

Cover Design
Tina Messmer

Journal webpage: http://ling.osu.edu/sal

Studies in African Linguistics is published twice a year. Occasional supplements are published at irregular intervals and are available to current subscribers at reduced rates. It is printed and bound in the United States by Western Newspaper Publishing Co., Inc., Indianapolis, Indiana.

Contributions to the journal should be sent to: The Editor, *Studies in African Linguistics*, Department of Linguistics, 222 Oxley Hall, Ohio State University, Columbus, OH 43210. Contributors please refer to "Guidelines for Contributors" inside the back cover.

Correspondence concerning subscriptions or other business should be sent to: Managing Editor, *Studies in African Linguistics*, Department of Linguistics, 222 Oxley Hall, Ohio State University, Columbus, OH 43210. Subscription rates are the following:

	North America	*Elsewhere*
Individuals	$25.00 per year	$33.00 per year
Institutions:	$50.00 per year	$60.00 per year
Single issue:	one half the yearly rate	
Supplements:	variable price depending on supplement	

Make checks payable to **Studies in African Linguistics.**

Volume 34, Number 2 • 2005

ISSN 00390-3533

Studies in African Linguistics

Volume 34, Number 2 Fall 2005

TABLE OF CONTENTS

Studies in African Linguistics
Volume 34, Number 2, 2005

A DIACHRONIC ONOMASIOLOGICAL APPROACH TO EARLY BANTU OIL PALM VOCABULARY*

Koen Bostoen
Royal Museum for Central Africa Tervuren
Université Libre de Bruxelles

Despite its ancient and long-lasting importance to sub-Saharan African economies, there has been no systematic attempt to reconstruct Proto-Bantu vocabulary referring to the oil palm (*Elaeis guineensis* Jacq.). Scholars have identified four common noun stems for 'oil palm', i.e. *-*bídà* (cl. 5/6, 7/8), °-*bá* (cl. 5/6), *-*gàdí* (cl. 9/10), and *-*téndé* (cl. 3/4) but determined the stems' geographic distributions within the Bantu domain to be insufficiently widespread to reflect a Proto-Bantu origin. From the wider perspective of Niger-Congo, certain of these nouns undoubtedly reconstruct to a level higher than Narrow Bantu. This paper presents an onomasiological approach to the earliest Bantu 'oil palm' vocabulary, offering a diachronic semantic analysis of the main noun stems, and an evaluation of the historical implications of their current-day distribution, both with respect to each other and in the light of the available Niger-Congo data.

1. Introduction.

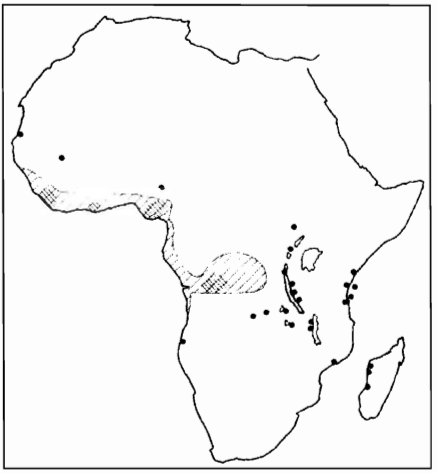
The oil palm (*Elaeis guineensis* Jacq.) is a tree crop, belonging to the Palmae or Arecaceae, a very ancient family of mainly tropical and subtropical plants (Pur-

* The research for this paper is funded by the Belgian Federal Science Policy Office and contributes to the "Words and Plants" project of the Linguistics service of the Royal Museum for Central Africa. It is based on lexical data collected for the "Sakasaka: Comparative Bantu Food Crop Vocabulary" database, developed by my colleague Jacky Maniacky and myself. My thanks go to Odette Ambouroué, Lee Bickmore, Roger Blench, Caroline Dedisa, Kathryn de Luna, Jacky Maniacky, Jeff Marck, Yves Moñino, André Motingea, Gérard Philippson and Martin Walsh for their comments and/or sharing of information.

seglove 1972: 416-7). This family includes other useful and well-known plants of tropical Africa, such as the betel-palm (*Areca catechu* Linn.), the coconut (*Cocos nucifera* Linn.), the borassus-palm (*Borassus aethiopum* Mart.), the raffia-palm (*Raphia* spp.), the date-palm (*Phoenix dactylifera* Linn.) and the Senegal date-palm (*Phoenix reclinata* Jacq.), also called wild, swamp or dwarf date-palm (Burkill 1997). The genus name *Elaeis* is derived from the Greek word *elaion*, 'oil', while the species name shows that Jacquin, who coined it, attributed the origin of the oil palm to the Guinea Coast (Hartley 1967: 37).

The geographic distribution of the oil palm is seen in Map 1, adapted from from Hartley (1967: 7), where dots indicate sporadic palm groves, and double-hatching indicates a high density of oil palm.[1]

Map 1: Distribution of the oil palm



On the western side of the continent, one can observe natural palm groves between 16°N and 10°S, but the real palm belt of Africa runs from the Fouta Jallon district of Guinea through Sierra Leone, Liberia, the Ivory Coast, Ghana, Togo, Benin, Nigeria, Cameroon and Gabon into the two Congos and even northern Angola. In eastern Africa, its distribution is much more scattered, since most of East and South-East Africa is too dry to support the species.

As one of the major indigenous oil-yielding plants of Africa, the oil palm's economic significance is both diverse and old. Palynological studies indicate that

[1] We thank Blackwell Scientific for permission to reproduce this map.

its presence in the natural vegetation of some parts of Africa is extremely ancient, with evidence for the oil palm in Guinea (Conakry) dating from 54.8 to 33.7 million years ago (Maley 1999; Maley & Chepstow-Lusty 2001; Sowunmi 1999). Moreover, archaeologists have uncovered its endocarp, often carbonised, in several sites of Western and Central Africa, which suggests that human exploitation of the oil palm dates as far back as 5,000 BP (see for instance Sowunmi (1999) for an overview of archaeological finds of palm-nut husks). The products of the oil palm have considerable commercial value, but also play an important role in African subsistence economies, providing palm oil, palm kernel oil, palm wine, palm harts, leaves for thatching, and petioles and rachises for fencing (Cornet 2001: 841-2; Purseglove 1972: 479-81; Sowunmi 1999: 201).

The economic relevance of the oil palm for a given society tends to be reflected in the amount of specialised vocabulary related to it. In Mongo (C61), to cite only one example, Hulstaert (1966) noted no less than 40 specific terms, referring to aspects as diverse as the different types of oil palms (chiefly according to age), the parts of the tree, the bunch and its parts, the different kinds of fruits (according to their appearance or their place on the bunch), and, finally, some derived products. However, the existence of a diverse vocabulary does not necessarily imply its documentation. Thus, the amount of terminology available to the researcher varies from one language to another. Frequently, one finds only the generic plant name (sometimes without notation of which kind of palm tree), the name of the fruit and, occasionally, the names of general oil palm products. Although a comparative lexical study of more precise vocabulary might be enlightening with respect to the history of oil palm exploitation, as a result of the uneven documentation of oil palm vocabulary, this paper will focus on the generic names of the tree itself and its fruits. Other types of palms will not be considered systematically either, although future research should deal in more detail with the lexical links that exist between the oil palm and other species.

2. Previous Reconstructions of Generic Bantu Oil Palm Names.

Although Kölle (1854) lists various Bantu nouns for palm-tree in his *Polyglotta Africana*, Meinhof (1910: 249) was the first to reconstruct the noun stem *-tende*, '*Palme*' on the basis of reflexes in Swahili (G42), Kongo (H16), Herero (R31) and Duala (A24). Meinhof and van Warmelo (1932: 246) adopted this reconstruction. Bourquin (1923: 230) mentions *-tende* together with the stems *-kindu* and *-lala*, but the latter two refer to other kinds of palm trees, the Senegal date palm (*Phoenix reclinata* Jacq.) (cf. Bastin *et al.* 2003) and certain species of fan

palm (*Hyphaene* spp.), e.g. Venda (S21) *mulala*, 'fan palm (*Hyphaene natalensis* Kunze)' (van Warmelo 1989: 221). Meeussen (1969) also reconstructed **-kindu* (cl. 3), **-téndé* (cl. 3; 9) and **-dada*. However, the most comprehensive account of reconstructed noun stems directly referring to the oil palm can be found in Guthrie (1967-1971), and was later adopted by Bastin *et al.* (2003): **-bá* (cl. 5/6; C.S. 1), **-bɨ́dà* (cl. 5/6, 7/8; C.S. 140), **-gàdɨ́* (cl. 9/10; C.S. 767 + C.S. 768 with the meaning 'nut of oil-palm') and **-téndé* (cl. 3/4; C.S. 1712).[2] In his comment to C.S. 140, Guthrie (1967-1971) affirms:

> This belongs to a group of three synonymous C.S., the others being C.S. 1 and 767. Apart from one or two entries from zone D, which could be due to intrusion, the whole group has a western distribution, and this no doubt reflects the fact that the oil-palm is a more important factor in the economy in the west than elsewhere in the Bantu field. From the relative spread of the three C.S. it seems probable that the PB-A (= *West Bantu*, *K.B.*) item was **-bɨ́dà*, which was replaced in the central and southern parts of the region by **-bá* on the western side and **-gàdɨ́* on the eastern. [vol 3, p. 49]

With respect to **-téndé*, Guthrie (1967-1971) doubts whether this C.S. consists of direct cognates, since the meanings of the five entries diverge. In an article on the contributions from comparative Bantu studies to the prehistory of Africa, Guthrie (1970) discusses other food crops, but does not consider the historical role of the oil palm. Furthermore, neither the historical interpretation by Dalby (1976) of Guthrie nor the historical linguistic research of Philippson & Bahuchet (1994-95) on cultivated crops in the Bantu domain investigate the oil palm. Beyond Narrow Bantu, on the contrary, linguistic evidence for the history of oil palm exploitation has gained more attention (see for example Blench (forthcoming), Connell (1998), and Williamson (1970, 1993) for lexical reconstructions in other Niger-Congo language groups). This paper will reassess the Guthrie (1967-

---

[2] In this paper, the 7-vowel system (*i ɪ e a o ʊ u*) of Bastin *et al.* (2003) is adopted for the representation of reconstructed forms, unless I explicitly refer to reconstructions of other authors, who may use other notation systems (e.g. *ɨ i e a o u ʉ*). Similarly, apart from the reconstructions by authors who always use the * sign, reconstructions preceded by * are hypothesized to go back to Proto-Bantu, while ° refers to uncertain and/or regional Bantu reconstructions. Abbreviations used in this text: BP = before present; cl. = class; $C_{1/2}$ = first or second consonant of a noun stem or verb root; C.S. = comparative series; H = high tone; L = low tone; NP = noun prefix; PB = Proto-Bantu; (P)NC = (Proto-)Niger-Congo; (P)DC = (Proto-) Delta-Cross; $V_{1/2}$ = first or second vowel of a noun stem or verb root.

1971) hypothesis, taking into consideration additional (narrow) Bantu data and recent non-Bantu Niger-Congo reconstructions.

3. Analysis and Distribution of the Bantu Comparative Series.

In this section, I will present a succinct analysis of each of the main 'oil palm' comparative series. Although phonological aspects will be treated if necessary, this analysis will focus on their (morpho-)semantic properties. I will also consider their geographic distribution as represented on the isogloss maps given throughout the article.

3.1 °*-bá*, 'oil palm', 'palm nut'. Table 1 presents a reflex of °*-bá*, 'oil palm' from each of the Guthrie zones in which it is attested.

Table 1: Reflexes of °*-bá*, 'oil palm', 'palm nut'

B86	Dzing	*di-ba*	cl. 5-6	'palmier élaïs'	(Mertens 1939: 17, 203)
C35b	Bolia	*ibá*	cl. 5-6	'*Elaeis guineensis* Jacq.'	(Ngila Bompeti 2000: 97)
H41	Mbala	*ba*	cl. 5-6	'nom générique des palmiers; palmier à l'huile'	(Mudindaambi 1977: 37)
R11	Umbundu	*ɛva*	cl. 5-6	nome comum das Palmácias	(Le Guennec & Valente 1972: 460)

As regards the phonological shape and the tone of °*-bá*, the reconstruction of the monosyllabic noun stem proposed by Guthrie (1967-1971) is confirmed by our data. The C_1 *b* has a direct reflex in most of the languages. With respect to the V_1, in all languages, the noun stems attest the vowel *a*. In Songye (L23), Stappers (1984: 36) mentions *èbwé*, 'palm tree', but since diphthongisation is not a common process in this language, one doubts whether *èbwé* is a reflex of °*-bá*. Finally, the reconstructed high tone is directly reflected in almost all languages in which the reflex is available with tonal notation.

With respect to noun class, the reflexes of °*-bá* can, roughly speaking, be subdivided in two groups, those which belong to cl. 5-6 and those which belong

to either cl. 11-10 or cl. 9-10.[3] This subdivision is morphosemantic because the variation of noun prefixes systematically corresponds to a change in meaning. While °*-bá* reflexes occurring in cl. 5-6 designate the tree itself, those preceded by $NP_{11\text{-}10}$ or $NP_{9\text{-}10}$ refer to its fruits.[4] Except in the C80 languages, where one only finds reflexes of the latter kind, nouns for the palm nut generally co-exist with nouns referring to the oil palm itself. Such is the case in Boma (B82), Tsong (B85d), Ntomba (C35a), Bolia (C35b) and Tetela (C71).[5] This is not always true the other way around. As will be shown below, several languages which have a cl. 5-6 reflex for the tree itself use another noun stem for its fruit. Based on the available Bantu data, the meaning 'palm nut' of the noun stem °*-bá* can be considered the result of a semantic evolution by means of simple noun class variation. This semantic shift seems to be secondary, since it is only attested in the B50, B60, and B80 languages and in the C30, C70, and C80 Inner Congo Basin languages.

Semantically, apart from the morphosemantic subdivision just treated, the present-day °*-bá* reflexes commonly refer to the oil palm, at least insofar as documented translations allow one to distinguish. For some languages, even though the kind of palm tree is not specified, e.g. Yoombe (H12b) *liba* (H-B), 'palmier' (Mabiala 1992), comparison with closely related languages suggests that a reflex probably refers to the oil palm. In other languages, on the contrary, it is clearly stated that the reflex serves as a generic noun for different species of palm trees, e.g. Ntandu (H16g) 'nom collectif des Palmiers à stipe' (Daeleman & Pauwels 1983: 207), Umbundu (R11) *eva*, 'nome comum das Palmácias' (Le

---

[3] One exception to this "rule" is the Mbere (B61) reflex *oba* 'E. var. nigrescens' (Raponda-Walker & Sillans 1961: 335), which probably belongs to cl. 3-4, as several other tree names do (cf. Adam 1954: 29). Guthrie (1967-1971; C.S. 1) gives the Mbere (B61) reflex *kọba* (cl. 7-8).

[4] From a comparative point of view, it is very ordinary in Bantu to observe the co-occurrence of the noun class pairings 11-10 and 9-10. Cl. 9 being much more common than cl. 11, nouns originally occurring in the latter are frequently reclassified in the former, since both generally take their plural in cl. 10 (Bastin 1985: 14; Grégoire 1976: 7). I described this class shift in some detail for the PB etymon **-bìgá*, 'pot' (Bostoen 2005: 199-200).

[5] In Bushong (C83), the name of the tree itself is *ʃáam* (Vansina 1959: 95), in Wongo (B85) *sa:mpa* (Burssens 1993: 472), and in Ndengese (C81) *toko* (Goemaere n.d.: 40). The latter noun stem also refers to the oil palm in Tetela (C71) and in certain C50 languages, such as Linga, Lyombo (C53), and Kele (C55). In Mongo (C61), it designates, more precisely, a young palm tree (Hulstaert 1966: 131). In Tetela (C71) and Leke (C14), it is also used to refer to the palm grove.

Guennec & Valente 1972: 460). In Ntandu (H16g), *bá* can be modified by three nouns to refer to the oil palm specifically: *bá dínsaamba*, *bá díkísáambu*, and *bá díngási* (Daeleman & Pauwels 1983: 207). The noun *nsáamba* designates the wine extracted from the male inflorescence, while *kisáambu* and *ngási* refer to the palm nut (ibid.: 208). As seen in the Mbala (H41) example *ba*, 'nom générique des palmiers; palmier à l'huile' (Mudindaambi 1977: 37), the reflex may also be a generic noun, refering simultaneously to one precise species as the most prototypical representative of the category, i.e. the oil palm.[6]

It should be noted, however, that the available data do not really allow one to assess the conceptual perimeters of the category covered by the generic noun. It is rather unlikely that in the mind of a native speaker the conceptual range of the °*-bá* reflex would exactly coincide with the botanically determined Palmae family or certain of its genera. Unpublished data from Bembe (H11) illustrate this point. The noun *bá* (cl. 5-6) seems to function as a generic noun for all palm trees bearing nuts, including the oil palm, the coconut (*bá dyá mindélé*, (litt.) 'the Whiteman's palm tree') and the raffia-palm ((*bá dyá*) *muhokó*) (Jacky Maniacky pers. comm.).[7]

Certain sources specify the variety of oil palm to which the °*-bá* reflex refers. According to Raponda-Walker & Sillans (1961: 335), the Duma (B51), Mbere (B61) and Ndumu (B66) reflexes refer particularly to the *nigrescens* variety. In botanical literature, oil palms are classified in two ways: according to the external appearance of the fruit (*nigrescens* or black-fruited, *viriscens* or green-fruited, *albescens* or white-fruited variety) and according to the thickness of the fruit's pericarp (*dura* or thick-shelled, *tenera* or thin-shelled, *pisifera* or shell-less variety), which may give the following combinations: *dura nigrescens, tenera nigrescens, pisifera nigrescens; dura virescens, tenera virescens, pisifera virescens; dura albescens* (Cornet 2001; Hartley 1967).

Although no such specifications are found in other languages, it is not surprising to note that the *nigrescens* type is referred to by a common Bantu noun, since it is the most common fruit type. In these languages, the less common *virescens* type has its own specific name, i.e. *-tɔnɔ* (cl. 3-4) (Raponda-Walker & Sillans 1961: 336). It must be a fairly local noun stem, Pove (B22c) being the only other language where it is also attested: *-tɔ̀nɔ́* (cl. 15-10), 'palmier *Elaeis* var. *vi-*

---

[6] In order to refer to other species of palm trees, the generic noun needs to be determined, e.g. *ba dipudu* 'cocotier' (Mudindaambi 1977: 37), cf. Ntandu (H16g) *bá dínkáándi*, *bá dímputu* 'cocotier' (Daeleman & Pauwels 1983: 207).

[7] Cf. Swartenbroeckx (1973: 310): *màyóko*, 'raphias géants pour *maléké*'.

rescens Jacq.' (Mickala Manfoumbi 2004: 638).[8] Compounds are used to refer to this type of oil palm in the Angolan Kongo (H16) variant, in which Gossweiler (1953: 512) noted *diba malau*, 'var. *virescens dura*' and *diba matundaba*, 'var. *virescens tenera*'.[9] The previously mentioned Ntandu (H16g) compounds, in which *bá* is modified by different names for the palm nut, possibly distinguish different kinds of oil palms according to their nut type. However, the available data do not permit us to corroborate this hypothesis.[10] Other *bá* compounds in Ntandu (H16g) differentiate oil palms with regard to age: *bá dínteende* (or *ntééndé ba*), 'jeune palmier'; *bá díyúmbu*, 'vieux palmier, très élancé' (Daeleman & Pauwels 1983: 208). As discussed in more detail further on, *ntééndé* is the reflex of the common Bantu noun stem **-téndé*.

As shown on Map 2, °*-bá* reflexes occur predominantly in the central and southern parts of Western Bantu. They are found in Guthrie's zones B (50-80), C (30+60-80), H (10+40), and R (R11). In historical terms, according to Vansina's classification (1995), based on the lexicostatistical data published later on in Bastin *et al.* (1999), the distribution of °*-bá* links the West Coastal languages (including B40-80 and H, except for H41) with the central Forest Bantu languages of the Inner Congo Basin (C33-35, C50-C80, B82).[11] The only languages that do not belong to one of these groups are Mbala (H41) and Umbundu (R11). Both are adjacent to the West Coastal group, but are thought to be part of the South-west group.

---

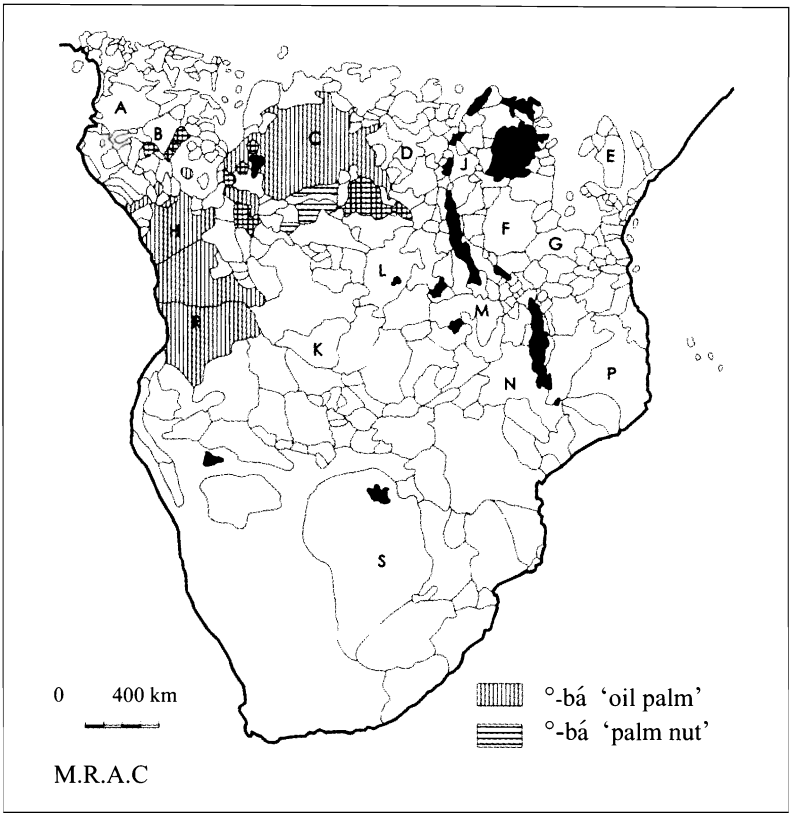
[8] Another specific noun stem for the *virescens* type recurs in the B10-B20 languages and in Fang (A75): *-kɔma* (Raponda-Walker & Sillans 1961: 335-336). However, it cannot be considered local because it also occurs in zones L and K as the name of the oil palm or of other species of palm-tree. Interestingly, in Tetela (C71), *dyeékómá* means 'noix de palme (verte à maturité)' (Hagendorens 1975: 54).

[9] Cf. Swartenbroeckx (1973: 295, 499): *malàfu*, *malàvu*, 'vin de palme, malafou; spiritueux'; *ntùndúbà*, *-dúwà*, 'noix palm. vert-rouge, refusée par certains'.

[10] Cf. Swartenbroeckx (1973: 179, 416): *kisámbu*, 'régime palm., noix de palme fraîche'; *ngási*, 'noix palm. entière'.

[11] As I will comment more extensively further on, the C50 are the only languages of this subgroup, attesting **-bídà*, rather than °*-bá* reflexes.

Map 2: The distribution of °-*bá* inside the Bantu domain



3.2 *-*bídà*, 'oil palm'. Table 2 lists reflexes of *-*bídà*, 'oil palm' from each Bantu zone in which it occurs.

Table 2: Reflexes of *-*bídà*, 'oil palm', 'palm nut'

A44	Nen	*ìbíl*	cl. 7-8	'régime mûr de noix de palme'	(Dugast 1967 : 77)
A85d	Tsong	*oyila*	cl. 7-8	'E. *var.* nigrescens'	(Raponda-Walker & Sillans 1961 : 335)
C101	Babole	*dibílà*	cl. 5-6	'palmier à l'huile'	(Leitch 1991 : 8)
D25	Lega	*kibíla*	cl. 7-8	'palm tree or nut'	(Botne 1994: 116)

The supplementary data which I collected do not contradict the phonological reconstruction of *-*bídà*, as proposed by Guthrie (1967-1971). The only problem

may consist in the identification of certain reflexes. Due to consonant attrition, a present-day **-bídà* reflex may be difficult to distinguish from a °*-bá* reflex. The Mbonge (A121) term *lí-já/mí-já*, 'oil palm' (Williamson 1973: 58), for example, could be interpreted as a reflex of °*-bá*. Nevertheless, data from closely related languages like Londo (A11) and Kundu (A11c), where scholars have documented *-íà (díà/míà)*, 'palm tree' (Kuperus 1985: 261) and *díà/míà*, 'ausgewachsene und gereinigte Ölpalme' (Ittmann 1971: 159) respectively, suggest that the Mbonge term is a **-bídà* reflex. Between two vowels, **d* becomes Ø in these languages, which may lead to the semivocalisation of *i*.

Morphologically, the reflexes of **-bídà* may vary in noun class. As is the case for °*-bá* reflexes, noun class alternation is a common device for semantic variation between the different parts of the tree. In Nen (A44), for instance, *nìbíl* (cl. 5) is the name of the tree, while *ùmíl* (cl. 3) signifies 'noix de palme' and *ìbíl* 'regime mûr de noix de palme' (Dugast 1967: 77, 141, 192). Nevertheless, this morphosemantic subdivision is slightly less systematic than for °*-bá*. In most languages, the name of the tree itself also occurs in cl. 5-6, e.g. Nen (A44) *nìbíl* (Dugast 1967: 141); Enyele (C10) *ebíya* (Caroline Dedisa pers. comm.); Ngombe (C41) *libía* (Rood 1958: 256). In the Myene (B10) languages, on the other hand, the noun for the tree belongs to cl. 14, taking its plural in cl. 6, e.g. Orungu (B11b) *oyílá/ambílá* (Odette Ambouroué pers. comm.). In other languages, a cl. 9-10 noun is attested, e.g. Kombe (A33b) *mbía* (Fernandez 1951: 412), Binza (C30) *mbíla* (Motingea 1996: 259), Bua (C44) *mbia* (De Cort *et al.* 1912: 23). Both unique, Lega (D25) uses a cl. 7-8 reflex to refer to both the tree and its nut (see table 2), while Bubi (A31) attests *öbiilá*, a cl. 3-4 noun (Bolekia Boleká 1991: 159).

With respect to the fruit of the oil palm, the most common class pairings are also cl. 9-10 and cl. 11-10, e.g. Baakpe (A22) *mbía/mbía* (cl. 9-10) (Kagaya 1992: 112), Tsong (B85b) *ńdzyá/ńdzyá* (cl. 9-10) (Iliku Mimpiya Dibata 1979: 111), Lingala (C36d) *lobíla/mbíla* (cl. 11-10) (Kawata 2003: 137), Lyombo (C53) *lo-íla/m̀-bíla* (cl. 11-10) (Stoop 1977: 32). Reflexes of **-bídà* belonging to noun classes other than 9-10 or 11-10 with the meaning 'palm nut' have only been noted in Nen (A44) and Enyele (C10), respectively *ùmíl* (3-4) (Dugast 1967: 192) and *mumbíya/mambíya* (3-6) (Caroline Dedisa pers. comm.).[12]

[12] Synchronically, the stem of the Enyele (C10) noun clearly is *-mbíya*. Diachronically, however, this stem should be analysed as *m-bíya*, the cl. 9(/10) noun prefix having been subsequently integrated into it.

In a restricted number of languages, such as Baakpe (A22), Nen (A44), Orungu (B11b), and Bobangi (C32), the **-bídà* reflex also refers to the bunch of oil palm nuts. Interestingly, in this case, the noun stem always occurs in cl. 7-8, e.g. Orungu (B11b) *ezílá/yílá* (Odette Ambouroué pers. comm.). Just as for °*-bá*, the reflexes referring to the nut (or the bunch of nuts) generally co-exist, language internally, with a reflex designating the tree itself, but this is not necessarily true the other way around. It is only in Duala (A24) and some C50 languages that the **-bídà* reflex is uniquely used for designating the palm nut, not the oil palm. In Duala (A24), a reflex of another common Bantu noun stem for the oil palm, i.e. **-téndé*, is attested. The C50 languages attest the local noun stem *-tòkò*, discussed in footnote 6. In Kele (C55), both nouns are combined into the following compound: *litoko lia mbila*, 'le palmier' (Anon n.d.). Although the uneven nature of documentation may somehow lead to an underestimation of the actual number of reflexes referring to the palm fruit, the tree seems to constitute the semantic focus of **-bídà*, while 'palm nut' and 'bunch of palm nuts' seem to be semantic derivations by noun class variation.

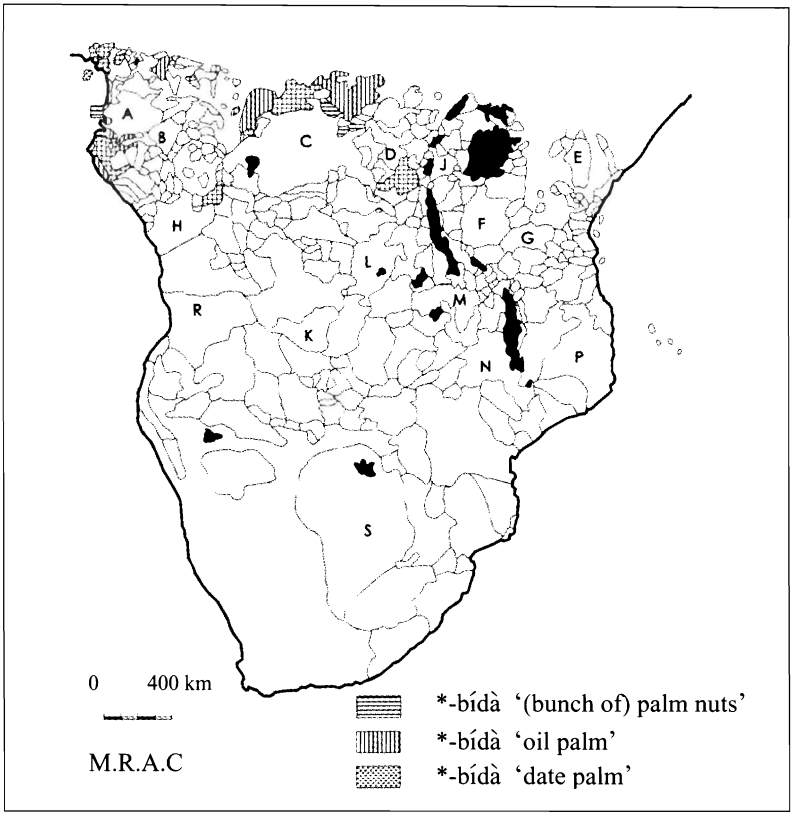
Concerning the semantics of **-bídà*, except for the morphosemantic derivations just treated, it is quite obvious, from the available comparative data, that this noun generally refers to the *Elaeis guineensis* Jacq. The only language in which it unmistakably concerns another type of palm tree is Bubi (A31): *öbiilá*, 'palmera datilera' (Bolekia Boleká 1991: 159). However, in some languages, the translation does not allow one to assess whether it specifically refers to the oil palm or serves as a generic name for palm trees, e.g. Mobenge (C43) *bira*, 'palmier' (Bareau & Reding 1912: 116). Botanical details with respect to the different varieties of oil palms are, once again, mainly found in the work of Raponda-Walker & Sillans (1961). In several of the Gabonese Bantu languages which they cover (A34, B11a-e, B21-22), the simple **-bídà* reflex always refers to the *nigrescens* variety, while compounds are used to refer to other varieties or sub-varieties, e.g. Benga (A34) *mbiya*, '*E.* var. *nigrescens*', *mbiya-nkóma*, '*E.* var. *virescens*', and *mbiya-ibóbu* '*E. sub.* var. *tenera*' (Raponda-Walker & Sillans 1961: 335-6).[13]

With respect to the geographic range, the reflexes of **-bídà* occur in Guthrie's zones A-(10-40), B (10-20), C (10+30-50) and D (D25) (cf. Map 3). In other words, their distribution is widespread but quite scattered within Forest Bantu.

[13] Note that the determinant *nkóma* of the first compound corresponds to the recurrent noun stem *-koma* referring to the *virescens* variety in the B10-B20 languages and in Fang (A75). The precise meaning of *ibóbu* could not be retraced.

Interestingly, they are found in perfect complementary distribution with the °*-bá* reflexes. These noun stems are never encountered together in the same language, although they may exist separately in neighbouring languages.

Map 3: The distribution of **-bídà* inside the Bantu domain



Consider, for example, Guthrie's C30 languages. As described above, Ntomba (C35a) and Bolia (C35b), two southern members of the C30 group, attest °*-bá*. In constrast, the more northerly C30 languages, such as the Ngiri languages spoken between the Ubangi and Congo rivers and Bobangi (C32) generally attest **-bídà* (Motingea 1996: 259; Whitehead 1899: 407).[14] Still more intriguing is the fact that this subgroup seems to correspond to a genetic split. In several classifi-

[14] Among the Ngiri languages, Bomboma is the only to have another term, i.e. *bo-langa* (Motingea 1996: 259).

cations (Bastin *et al.* 1983, 1999; Vansina 1995), Guthrie's C30 group falls apart in two distinct subgroups, splitting the C33-35 languages from the rest. The C33-35 languages are, in general, more similar lexically to the C60-80 languages, with whom they also share the use of °*-bá*. The other C30 languages, however, tend to be lexically closer to the rest of zone C, notably to C40 and certain C10-20 languages. As mentioned before, Vansina (1995) groups the C33-35 and C60-80 languages with the C50 languages and Boma (B82) in a unit which he designated "Inner Basin." The rest of C30 would be made up of Ngombe (C41) and Bwela (C42), and possibly with Ngondi (C11) and Pande (C12), a unit he labelled "Rivers." "Inner Basin" and "Rivers" would constitute co-ordinate branches of the node he called "North Zaire." Several of these languages attest **-bídà*.

The C10 languages attesting **-bídà* are Aka (C104), Leke (C14) and Babole (C101), whose positions in the classifications presented by Bastin *et al.* (1999) fluctuate according to the statistical procedure applied. The C40 languages (apart from Ngombe (C41)), in which **-bídà* was found include Mobenge (C43) and Bua (C44). These languages are often set apart with certain D20-30 languages in a unit termed "Boan," which is believed to be one of the primary Bantu offshoots (Vansina 1995). Even if °*-bá* and **-bídà* cannot be considered shared innovations diagnostic for genetic sub-grouping, their geographic distribution among Guthrie's zone C appears to be historically significant because their distribution tends to coincide with certain genealogical sub-divisions. However, the C50 languages, predominantly spoken along the upper Congo between Bumba and Kisangani and the very lowest end of the Lomami, deviate from this pattern. They attest **-bídà*, but are lexicostatistically closer to the C33-35 and C60-80 languages than the rest of the zone C languages.[15] Nevertheless, the historical background of this small language group is quite opaque. Based as they are on short lists of basic vocabulary, lexicostatistical classifications can only partially reflect linguistic history.

Quite relevant in this respect is the observation that the C50 languages also join the Congo River C30 and C40 languages, as opposed to the Inner Congo Ba-

[15] Although the C50 languages went with the C40 and Ngiri C30 languages in the Bastin *et al.* (1983) classification, they are set apart from them in the Bastin *et al.* (1999) classification, which was founded on far more extensive data and is, therefore, more reliable statistically. In most of the trees presented in this study, the C50 languages are lumped together as a co-ordinate branch of a cluster grouping C33-35, C60-80 and certain B80 languages, and as a primary offshoot of the node being at the same genealogical depth of the node uniting most of the other zone C languages.

sin C33-35 and C60-80 languages, for at least one other field of cultural vocabulary, i.e. pottery terminology. As I have recently demonstrated (Bostoen 2005, 2006), they share at least two lexical innovations designating two different types of clay pots with the Congo River languages, while the Inner Congo Basin languages are characterised by a distinct set of parallel lexical innovations. Evidence for close relatedness with one language group based on fundamental vocabulary, but with another language group with respect to certain domains of cultural terminology, suggests the superposition of different lexical layers as a result of language shift(s), leaving substrate vocabulary in the superimposed language(s). Motingea (2004) has described the present-day sociolinguistic situation of the Inner Congo Basin as a unity within diversity, stemming from the unachieved linguistic absorption of the oldest small-scale Bantu-speaking communities by larger and socio-politically dominant settler communities. Apart from the Bobangi (C32) slave traders, the Mongo (C61) and Ngombe (C41) can be considered as the major immigrant groups. It was mainly the latter two groups who settled along the Upper Congo River, which could explain certain particular lexical similarities between these Congo River languages and the more eastern C50 languages spoken along this same aquatic highway.

The rest of the distribution of **-bídà* reflexes provides additional historical information. Apart from Lega (D25), which is somewhat isolated, this noun stem occurs in the languages of zones A and B, which, lexicostatistically, form a primary Bantu branch labelled 'North-west' by Vansina (1995). In other words, **-bídà* is attested in two of the main Bantu subgroups, i.e. "North-west" and "West".

Interestingly, a reflex of **-bídà* has also been noted in two Ubangian languages: in Monzombo (spoken in the border region between the Central-African Republic, the Democratic Republic of Congo and the Congo Republic), and in Bangando (spoken in the extreme south of Cameroon). They attest *mbíā́*, 'palmier à huile' (Motte 1980: 128) and *mbìà*, 'Palmée *Elaeïs guineensis* Jacq.' (Moñino 1995: 643) respectively. Having emigrated from the Ubangi River bend area in the extreme north-western part of present-day RDC, the Bangando community is known to have spoken a Bantu language before they underwent their 'Gbayaisation'. This language was the forerunner or at least closely related to the forerunner of the present-day Ngombe (C41) language (Moñino 1995, pers. comm.). In this respect, their **-bídà* reflex could be seen as a retention from this period, rather than as a loan word from adjacent Bantu languages.

Other Ubangian languages of this region, however, attest a *-banga* like form for 'oil palm' (Moñino 1995: 643). This could be a Bantu loan, since a

similar form is not only attested in neighbouring A80-90 languages, e.g. Mpiemo (A86c) *baŋa*, '*Elaeis guineensis* Jacq. (*Arecacae*)' (Thornell 2004: 74), Kako (A93) *báŋá*, 'palmier à huile' (Ernst 1989: 79), but also occurs in more distant Bantu languages, referring to the palm kernel, e.g. Duala (A24) *m'bangá ma mbía*, 'le palmiste' (Helmlinger 1972: 276).[16] With respect to the other Ubangian reflex, the Monzombo are not known to be former Bantu speakers. Hence, *mbíã* probably is a loan word from the Bantu language spoken by the Aka pygmies, with whom they exploit the Mongoumba region's fauna and flora. Although Motte (1980: 128) does not mention a **-bídà* reflex as name for the oil palm in Aka (C104), *dì.bílà/mà.bílà*, 'régime de noix de palmes; palmier à huile', and *è.bílà/bè.bílà*,'régime de noix de palme' are mentioned by Thomas *et al.* (1993: 48). Maniacky (2005) recently identified other terms in the Monzombo food crop vocabulary that are probably Bantu in origin, namely words for three different yam species.

3.3 *-*gàdí*, 'oil palm'/'palm nut'/'palm oil'. Table 3 shows some reflexes of the noun stem *-*gàdí*. Each Bantu zone in which it figures with a meaning related to the oil palm is represented.[17] The reflexes which I collected confirm the phonological shape of the reconstruction *-*gàdí*, as proposed by Guthrie (1967-1971) and Bastin *et al.* (2003), on both the segmental and supra-segmental level. Guthrie (1967-1971) actually reconstructed *-*yàdi̧* (C.S. 1898) and *-*gàdi̧* (ps 202), both meaning 'oil'. In his comment on C.S. 1898, he raises the question whether it shares a common origin with *-*gàdi̧* 'oil palm' (C.S. 767), in spite of the divergence of shape and meaning. He concludes that the lack of evidence of associated nominals in any language with identical stems referring respectively to 'oil palm' and 'palm-oil' rules out treating them as osculant. In my view, there can be no doubt about the common origin of these nouns and they do co-occur in certain languages. Conversely, the necessity of reconstructing an alternative form having **y* as C_1 is questionable. For one thing, several reflexes that Guthrie lists under the starred form *-*yàdi̧*, such as in Lwena (K14), Luyana (K31), Nkoya (L62), Kwanyama (R21), and Herero (R31), may be derived from *-*gàdi̧*, since

[16] Crabb (1965: 83) lists several Ekoid languages attesting a *-banga*-like form for 'palm kernel', e.g. Abanyom *ḿ-bǎŋ*, which suggests that it could be considered for reconstruction to PB.

[17] In zones D, F, and M, the above-cited reflexes are the sole attestations of *-*gàdí*. These forms should not be taken as indications for the term's spread among zones D, F, and M languages. As I will explain further on, it is possible that certain of these easternmost reflexes spread through lexical diffusion.

*g regularly becomes Ø in these languages. However, it is true that such is not the case in the zone B, C, and H languages he cites. *g regularly becomes *k* in these languages. However, as I have previously argued for the osculant pair *-*bìgá* ~ °-*bìjá*, 'pot' (Bostoen 2004, 2005), this does not necessarily imply a distinct origin. The existence of forms such as *agali* in Mpongwe (B11a) or *omagadhi* in Ndonga (R22), favours *g as C_1.

Table 3: Reflexes of *-*gàdí*, 'oil palm'/'palm nut'/'(palm) oil

B302	Simba	*gékadi*	cl. 7-8	'E. *var.* nigrescens'	(Raponda-Walker & Sillans 1961: 335)
C81	Ndengese	*bongaji*	cl. 14	'palmpit'	(Goemaere nd.: 40)
D25	Lega	*kikasí*	cl. 7-8	'palm tree or nut used for oil'	(Botne 1994: 65)
F11	Tongwe	*sigasi*	cl. 7-8	'palm tree (oil palm)'	(Kakeya & Nishida 1976: 49)
H16g	Ntandu	*maási*	cl. 6	'huile de palme'	(Daeleman & Pauwels 1983: 208)
J66	Ha	*umugázi*	cl. 3-4	'oil palm'	(Nakagawa 1992: 25)
K14	Lwena	*ngáji*	cl. 9-10	'nut of the oil palm'	(Horton 1953: 214)
L23	Songye	*kyàjí*	cl. 7-8	'régime de noix de palme'	(Stappers 1984: 74)
M14	Rungu	*chazi*	cl. 7-8	'oil palm'	(Nurse & Philippson 1975)

With respect to phonological correspondences, I would also like to draw attention to a series of terms that resemble the regular *-*gàdí* reflexes, both formally and semantically. There are a series of nouns similar to the Swahili (G42) terms *mnazi* (cl. 3-4) 'coco-nut tree', and *nazi* (cl. 9-10), 'a coco-nut' (Johnson 1950: 292) that generally refer to the *Cocos nucifera* Linn., e.g. Nyakyusa (M31) *unnasí* (cl. 9-10), 'coconut palm' (Felberg 1996: 153); Yao (P21) *naási* (cl. 9-10), 'coco, palmeira' (Viana 1961: 153). In certain languages, however, they seem to refer to the oil palm. In the Tanzania Language Survey (Nurse & Philippson 1975), the Samialugwe (E34) noun *omunazi* (cl. 3-4), the Sango (G61) noun *mnasi* (cl. 3-4) and the Luhya (J31) noun *munazi* (cl. 3-4) are mentioned as equivalents of the Swahili (G42) noun *mchikichi* (cl. 3-4),'oil palm', while the

Tooro (J13) noun *omunazi* is given as the translation of the Swahili (G42) noun *mtende* (cl. 3-4), 'date palm'. All these terms share a noun stem beginning with /n/, which cannot be considered as the regular reflex of **g*, certainly not in post-nasal position, where it is normally conserved. Further research should determine whether these forms can be linked to regular reflexes of **-gàdí*.

Cross-linguistically, the reflexes of **-gàdí* figure in a number of different noun classes. Once again, noun class variation induces semantic change. Excluding certain rare exceptions, such as Nande (J42) *engasi* (cl. 9-10), 'oil palm tree, nut palm' (Fraas n.d.: 276, 279) or Ha (J66) *umugázi* (cl. 3-4), 'oil palm' (Nakagawa 1992: 25), nouns referring to the tree itself generally occur either in cl. 5-6 or in cl. 7-8, e.g. Kerewe (J24) *igazi* (cl. 5-6), 'oil palm' (Nurse & Philippson 1975), Tongwe (F11) *sigasi* (cl. 7-8), 'palm tree (oil palm)' (Kakeya & Nishida 1976: 49). In three L20-30 languages, however, cl. 7-8 reflexes designate the bunch of oil palm nuts, e.g. Hemba (L34) *kyàzi*, 'le régime du palmier élaïs' (Vandermeiren 1913: 829). In Ha (J66), the cl. 5-6 reflex *igázi* refers to the nut of the oil palm (Nakagawa 1992: 25).

The most widespread class pairings regarding the fruit of the oil palm are once more cl. 9-10 and cl. 11-10, e.g. Lwena (K14) *ngáji/jingáji* (cl. 9-10), 'nut of the oil palm' (Horton 1953: 214), Sanga (L35) *lwajì/ngajì* (cl. 11-10), 'noix du palmier à huile' (Coupez 1976: 19).[18] In this capacity, the reflex may be used as a modifier in connective constructions whose head noun is the name for '(palm) tree' or 'nut', e.g. Kimbundu (H21) *muxi ua ngaji*, 'palmeira' (Da Silva 1994: 460), Kiluba (L33) *manì a ngazì*, 'noix de palme' (Gillis 1981: 344), Shi (J53) *mavurhä gëengäzi*, 'huile de noix de palme' (Cuypers 1970: 48) or the aforementioned Ntandu (H16g) compound *bá díngási* (Daeleman & Pauwels 1983: 207).

Unlike °*-bá* and **-bídà*, **-gàdí* does not seem to refer primarily to the tree. It is only in a minority of languages, notably just to the east of the rainforest, that its reflexes exclusively designate the oil palm tree. In several western languages, this term only occurs with the meaning 'palm nut', while another noun is used for the tree, e.g. Pove (B22c) *ɔ́bɔ*, 'palmier *Elaeis* var. *nigrescens* Jacq.', *ngadí*, 'noix

[18] In Ciluba (L31a), the nasal prefix of the cl. 10 plural form has been integrated in the noun stem: *lungajì/ngajì*, 'une noix de palme' (De Clercq & Willems 1960: 149). A similar integration of the cl. 9-10 noun prefix is observed in other languages and may suggest the derived status of these nouns, e.g. in Lwena (K14) *mungaji/mingaji* (cl. 3-4), 'nut-palm (oil palm)' (Horton 1953: 214), in the Congolese Ngwana variant of Swahili (G42) *mngasi* (cl. 3-4), 'palmier à huile (Elais guineensis)' (Sacleux 1941: 534) or in Ndengese (C81) *bongaji* (cl. 14), 'palmpit' (Goemaere, n.d.: 40).

de palme' (Mickala Manfoumbi 2004: 638). This could suggest that 'palm nut' is a more primary meaning than 'oil palm' and that the latter is the result of a metonymically motivated shift or the syncope of periphrases, such as in Ntandu (H16g) and Kimbundu (H21), of which the first element, meaning '(palm) tree', has been dropped. Moreover, *-*gàdí* may also designate the oil extracted from the palm nut. When this is the case, the noun stem is always preceded by the prefix of cl. 6, e.g. Ntandu (H16g) *máasi* 'huile de palme' (Daeleman & Pauwels 1983: 207). In Mpongwe (B11a), *agali m'imbila* refers to the palm oil, while *agali m'anyanga* designates the palm kernel oil.[19] Unsurprisingly, the modifying nouns signify 'palm nut' and 'palm pit' respectively. In Kimbundu (H41), the cl. 6 reflex apparently needs a determinant meaning 'oil palm' in order to refer specifically to the palm oil: *maji ma ndende* (Da Silva 1994: 67).

These findings correspond to what we observe in certain Western Bantu languages of zones K and R, where the *-*gàdí* reflexes in cl. 6 seem to have undergone semantic broadening. These reflexes refer to any (vegetable) oil or fatty substance, e.g. Mbukushu (K333) *maghadhi*, 'oil, greasy liquid, able to burn' (Wynne n.d.: 364), Kwanyama (R21) *omaadi*, 'vegetable oil, animal fat, adipose tissue, mineral grease; ointment' (Turvey *et al.* 1977: 85). However, contrary to Vansina (2004: 44), there is no causal connection between this semantic shift and the fact that these languages are spoken south of the extreme limits where the tree could grow. This semantic generalisation is also observed in languages spoken inside the oil palm belt, e.g. Ndumu (B63) *mari*, 'toute sorte de matière grasse; huile; beurre; graisse' (Biton 1907: 62), Yans (B85) *meay*, 'huile, graisse' (Nguma 1986: 133), Bobangi (C32) *māli*, 'palm oil/fat, grease, oil' (Whitehead 1899: 407).[20]

---

[19] Remarkably, Connell (1998) reconstructed to Proto-Lower-Cross a similar form for 'palm kernel oil', i.e. **mì-mànyàŋà*. In spite of this striking resemblance, more data and research are needed to decide whether this form really is a cognate to the term *a-nyanga* for 'palm kernel' attested in Mpongwe (B11a) and some other Gabonese languages as Mbede (B61) and Ndumu (B63). The stem of these Bantu forms is *nyanga* or maybe *-anga*, *ny* being a cl. 9 noun prefix. These forms could be related to the *-banga* term observed in several northwestern Bantu language for 'palm kernel' (cf. infra).

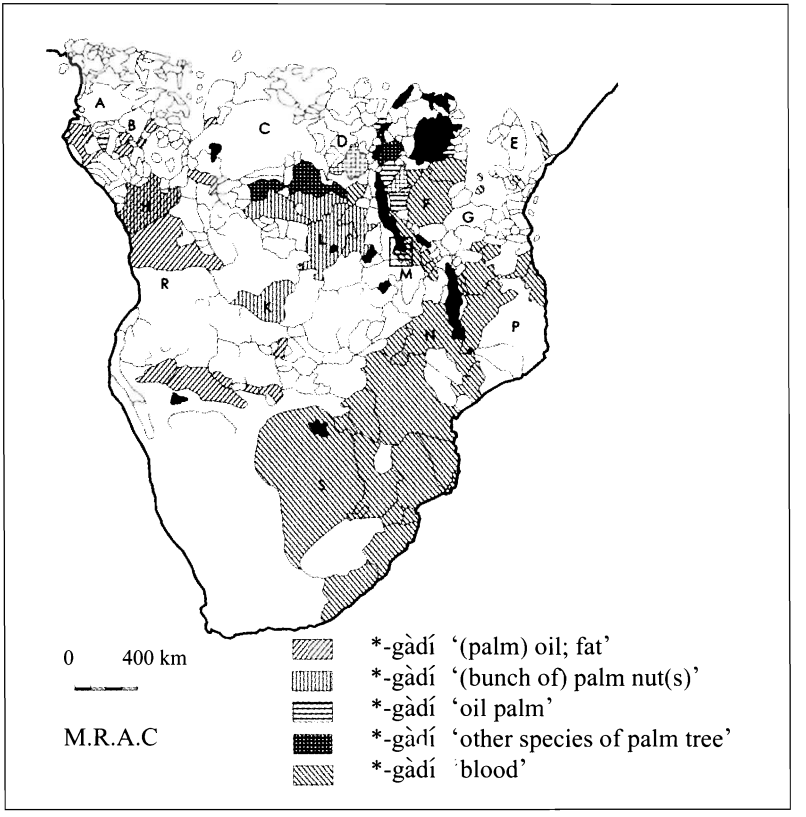
[20] In each of these languages, the cl.6 *-*gàdí* reflex is modified by a noun meaning 'palm nut' or 'palm kernel' in order to designate the palm oil: Ndumu (B63) *mari ma mbă*, 'huile de palme', *mari ma anânga*, 'huile faite avec l'amande des noix de palme'; Yans (B85) *meay a mba*, 'huile de palme'; Bobangi (C32) *māli ma ndīka* 'oil from kernels', *māli ma mibōkū* 'oil from nuts of oil palm'.

In the same respect, it is intriguing to observe that an identical noun stem, often belonging to cl. 6, occurs in the Eastern Bantu languages of zones E, F, G, N, M, P, and S with the meaning 'blood', e.g. Chewa (N31b) *magazi*, (Paas 2004: 243), Makonde (P23) *myadi*, (Guerreiro 1963: 114), Sotho-N (S32) *madí* (Ziervogel & Mokgokong 1975: 146). Despite their complementary geographical distribution, Guthrie (1967-1971) did not suggest a historical link between **-gàdí* 'blood' (C.S.766) and the oil palm related **-gàdí* forms. Rather, he recognized a common origin with **-gìdá*, 'blood' (C.S. 824) through metathesis of the vowels, but the historical link between these etymons is rather unlikely. The only common ground between both might be the fact that **-gàdí* replaced **-gìdá* for reasons of taboo (Bastin 1997: 20). Nevertheless, a metaphorically induced semantic shift from palm oil, usually reddish in colour, to 'blood' seems highly probable (Bastin 1997: 20). If this hypothesis bears out, it is all the more likely that the basic meaning of **-gàdí* was 'palm oil' and/or 'palm nut', rather than 'oil palm'. According to the available data, Rungu (M14) is the only language where two **-gàdí* reflexes uniting both semantic fields co-exist, i.e. *chazi* (cl.7), 'oil palm' (Nurse & Philippson 1975) and *úwazi* (cl. 14), 'blood' (Kagaya 1987: 59). Kagaya (1987) does not mention a word for 'oil palm' but Lee Bickmore (pers. comm.) confirms the coexistence of both meanings. He notes *uwaazi* for 'blood', *umuti waa ngazi* (in short, *umwaangazi*) for 'palm tree' and *amafuta yaa ngazi* or *amafuta yaa mwangazi* for 'palm oil'. Given that the meaning 'blood' only occurs in eastern Bantu languages, it is likely that this shift results from a semantic innovation going back to their latest common ancestor.

The precise tree variety referred to by reflexes signifying 'oil palm' is only known for the Gabonese languages of the Raponda-Walker & Sillans (1961: 335) survey. In Simba (A302), Apindji (A304) and Tsogo (B31), *gékadi* designates the most common *nigrescens* variety. In Kongo (H16) *mazi ma mona* refers to the oil extracted from the nut of the *albescens* variety (Gossweiler 1953: 513). No further botanical details for **-gàdí* are known. In some languages spoken at the southern border or just outside the equatorial rainforest, however, **-gàdí* reflexes refer to another kind of palm tree, e.g. Tetela (C71) *dikadí*, '*Raph. Laurentii de W.*, gros raphia des marais dont les rachis servent aux constructions' (Hagendorens 1975: 35), Bushong (C83) *ikady*, 'palmier de rivière' (Vansina 1959: 97), Ciluba (L31a) *dikadì*, 'le palmier bambou des marais' (De Clercq & Willems

1960: 54).[21] In Rwanda (J61), to the east of the rainforest, *ingazí* is said to refer to the Senegal date palm (*Phoenix reclinata* Jacq.) (Jacob 1984: 342).

Map 4: The distribution of **-gàdí* inside the Bantu domain



Reflexes of **-gàdí* whose meaning relates to oil palm occur in the following languages or language groups: B10 ('palm oil'), B30 ('oil palm'), Ndengese (C81) ('palm nut'), Lega (D25) ('oil palm' + 'palm nut'), Tongwe (F11) ('oil palm'), Kongo (H16) ('oil palm' + 'palm oil'), Kerewe (J24) ('oil palm'), J40-60

[21] Since **g* regularly gives Ø in intervocalic position in Ciluba (L31a), this form cannot be a regular reflex of **-gàdí*. It probably is a loan word copied from one of its northern neighbours, such as Tetela (C71) or Bushong (C83), where **g* regularly becomes *k*. Given that *di-kadì* attests the reversed tone pattern typical of Ciluba (L31a), it might be a relatively old loan.

('oil palm' + 'palm nut'), Lwena (K14) ('oil palm' + 'palm nut'), L20-30 ('(bunch of) palm nut(s)'), Rungu (M14) ('oil palm'). If one takes into account the reflexes having undergone a semantic generalisation to 'oil' or 'fat', the B50-80, K30 and R20 languages and Bobangi (C32) can be included as well. In other words, **-gàdí* manifests a quite scattered distribution, which is predominantly central and southern West Bantu. It has an extension, however, to some of the most western East Bantu languages spoken on both sides of Lake Kivu and Lake Tanganyika, where it predominantly functions as the name of the oil palm itself. It is possible that some of these easternmost languages obtained their reflex through lexical diffusion. As I noted above, the Lega (D25), Tongwe (F11), and Rungu (M14) reflexes are the sole documented attestations of **-gàdí*, 'oil palm' in their respective zones.

These languages being situated in the vicinity of Lakes Tanganyika and Kivu, precolonial regional trading networks interconnecting the Great Lakes and the Lake Corridor regions may have played a role in their diffusion. Very telling in this respect is the testimony of H.M. Stanley (1872) on the regional and long-distance trade at Ujiji market (my italics):

> There were the agricultural and pastoral Wajiji, with their flocks and herds; there were the fishermen from Ukaranga and Kaole, from beyond Bangwe, and even from Urundi, with their whitebait, which they called dogara, the silurus, the perch, and other fish; *there were the palm-oil merchants, principally from Ujiji and Urundi, with great five-gallon pots full of reddish oil, of the consistency of butter*; there were the salt merchants from the salt-plains of Uvinza and Uhha; there were the ivory merchants from Uvira and Usowa; there were the canoe-makers from Ugoma and Urundi; there were the cheap-Jack pedlers from Zanzibar, selling flimsy prints, and brokers exchanging blue mutunda beads for sami-sami, and sungomazzi, and sofi (...) Here were found Waguhha, Wamanyuema, Wagoma, Wavira, Wasige, Warundi, Wajiji, Waha, Wavinza, Wasowa, Wangwana, Wakawendi, Arabs, and Wasawahili, engaged in noisy chaffer and barter. [p. 473]

This could explain how the term trickled down from zone J for instance, where it is attested in J20/40/50/60 languages, into some more southerly languages, such as Tongwe (F11) and Rungu (M14). Yet, even if some of the easternmost **-gàdí* reflexes for 'oil palm' can be played down as lexical borrowings, the noun is attested in East Bantu. Moreover, if one also integrates the reflexes meaning 'blood', hypothesizing that they stem from the same etymon but have undergone a semantic shift, one definitely finds a general Bantu distribution. Consequently,

*-*gàdí* is a likely candidate for reconstruction in PB, not as name of the oil palm, but as a word for the fruit and its oil.[22]

3.4 *-*téndé*, 'oil palm'. Table 4 presents present-day reflexes of the noun stem -*téndé* meaning 'oil palm' or having a related signification.

Table 4: Reflexes of *-*téndé*, 'oil palm'/'palm oil'

A24	Duala	*lẹ́ndẹ́*	cl. 5-6	'Eléide de Guinée'	(Helmlinger 1972: 624)
B11b	Orungu	*ntɛ́ndɛ́*	cl. 9-10	'oil palm'	(Odette Ambouroué pers. comm.)
D14	Enya	*ceténdé*	cl. 7-8	'palmier'	(Kolni 1971: 89)
G42	Swahili	*mtende*	cl. 3-4	'dattier'	(Sacleux 1941: 610)
H16g	Ntandu	*ntééndé ba*	cl. 9-10	'jeune palmier'	(Daeleman & Pauwels 1983: 207)
K13	Lwena	*ndende*	cl. 9-10	'palm oil'	(Pearson 1973: 181)
R11	Umbundu	*ondende*	cl. 9-10	'azeite da palmeira; óleo de palma'	(Le Guennec & Valente 1972: 459-60)

The formal reconstruction of the noun *-*téndé*, as proposed by Guthrie (1967-1971), does not pose a problem in the light of the additional data which I have collected. As a result, it does not need to be reconsidered here.

Except for the Eastern Bantu *-*téndé* attestations belonging to cl. 3-4, which I will comment on later, the noun stem most often occurs in cl. 5-6 and cl. 9-10, and it generally refers to the tree itself, e.g. Noho (A32) *ilende* (cl. 5-6), 'Ölpalme' (Adams 1907: 72), Nkomi (B11e) *ntèndè* (cl. 9-10), '*E.* var. *nigrescens*' (Raponda-Walker & Sillans 1961: 335). Apart from the aforementioned Enya (D14) reflex in cl. 7-8, the Nen (A44) noun *hèlɛndɛ*, belonging to cl. 19, is one of the rare exceptions to this trend. As the last two examples of table 4 show,

[22] An anonymous reviewer claimed that *-*gàdí* is cognate with Proto-Delta-Cross *-*kɔ́d* 'raphia palm' (Connell 1998) and that therefore the possibility of *-*gàdí* as PB 'raphia palm' should be considered. Indeed, some Bantu reflexes refer to this palm tree or to a related species, e.g. 'bamboo palm' in Dzing (B86) and Ntomba (C35a), but these seem to be local innovations, rather than retentions. More evidence is also needed to demonstrate that the PDC form really is a cognate.

cl. 9-10 reflexes may also refer to the palm oil. However, the Lwena (K14) and Umbundu (R11) nouns are the only ones of their kind.[23] Otherwise, unlike the three common Bantu nouns treated so far, **-téndé* is never used to designate either palm oil or the palm nut. Its semantic scope seems to be limited to the tree as such.

In western Bantu, **-téndé* refers almost without exception to the oil palm.[24] It is only in Akoose (A15) that it is said to be the name of the *Raphia* spp.: *atên* (cl. 5-6), 'palm, raffia' (Hedinger & Hedinger 1982: 42). In Umbundu (R11), *ʊndente* (cl. 3-4) functions, just like the above-treated *°-bá* reflex *eva*, as the generic name for different kinds of palm trees (Le Guennec & Valente 1972: 460).[25] As regards botanical specifications, inasmuch they are known, unmodified **-téndé* reflexes seem to refer to the most common *nigrescens* variety. Such is the case in Fang (A75), Orungu (B11b) and Nkomi (B11e) (Raponda-Walker & Sillans 1961: 335). In the former language, the compound *alèn-bingom*, uniting the **-téndé* reflex and the *°-koma* reflex mentioned above, designates the *virescens* variety. In Kimbundu (H21), *(ndende ia) hôhô* refers to the 'var. *macrosperma & nigrescens forma dura*' and *ndende ia fumbe* to the 'var. *nigrescens forma pisifera*' (Gossweiler 1953: 513).[26] Interestingly, a recurrent connotation associated with **-téndé* is not botanical in nature, but concerns the age of the palm tree. In several widely dispersed languages, such as Nen (A44), Punu (B43), Ntandu (H16g), and Yaka (H31), the **-téndé* reflex is said to refer to a young oil palm. In the former language, it is a single noun, i.e. *hèlɛndɛ*, 'jeune palmier' (Dugast 1967: 62). In the other languages, it is part of a compound. In Punu (B43) and Yaka (H31), similar terms have been observed, respectively *diréndə dí mbari*, 'petit palmier ayant déjà produit' (Blanchon 1994) and *ntééndá mbáti*, 'jeune palmier' (Ruttenberg 1969: 49).[27] In Ntandu (H16g), it combines in different

---

[23] In Ngangela (K12b), a cl. 9 reflex modifies the word for 'tree' in order to refer to the oil-palm: *múci wa ndéende* (Maniacky 2003: 170). The word as such refers to the palm nut.

[24] Interestingly, in Brazil, the American Oil Palm (*Corozo oleifera* (HBK) Bailey) is called *Dende do Para* (Hartley 1967: 65).

[25] This noun is formally somewhat problematical, since **nd* in C_2 position is in general directly reflected in Umbundu (see for instance the last example of table 4). I do not have a straightforward explanation for this irregularity.

[26] In Kongo-C (H16b), *fùmbe* means 'larve de charançon du palmier (coléoptère)' (Laman 1936: 161).

[27] The nouns that modify the **-téndé* reflex have been translated as 'palmier à huile' (Blanchon 1994) and 'palmier' respectively (Ruttenberg 1969: 140). Both nouns seem to be part of a lo-

ways with the °-*bá* reflex: *ntééndé ba*, *bá díntеende*, 'jeune palmier' (Daeleman & Pauwels 1983: 207). As I have mentioned before, the noun *bá* also combines with *díyúmbu* to refer to an old oil palm. The scattered repetition of this connotation of youth could indicate that it is quite old.

Remarkably, in a large number of languages, belonging to Guthrie's zones E (40-70), F (10-20), G (20-60), J (20+30+60), M (10-30), N (13), and P (10-20), *-*téndé* has a reflex meaning 'date palm' (*Phoenix dactylifera* Poir.). Some rare exceptions notwithstanding, such as Gikuyu (E51) *ndende* (cl. 9-10), Kerewe (J24) *itende* (cl. 5-6), and Jita (J25) *gintende* (cl. 7n-8n) (Nurse & Philippson 1975), all these nouns occur in cl. 3-4, e.g. Nyakyusa (M31) *ɯnteende/imíteende* (Felberg 1996: 133). The dense distribution of the noun with this sense in Eastern Bantu languages is surprising, since the date palm does not really thrive in this part of the continent and its fruits are generally imported. In Eastern Africa, it is only planted as an ornamental, mainly on the coast and the offshore islands (Martin Walsh pers. comm.). The tree is subtropical in origin and cultivation. It has been introduced into some tropical countries, where it is usually completely sterile, as rain prevents pollination and the fruit will only ripen in a very dry climate. It mainly grows in a nearly rainless belt between 15 and 35 N Lat in the Sahara and the southern fringe of the Near East (Arabian Peninsula, southern Iraq, Jordan, etc., see Purseglove 1972).

The vast majority of this kind of *-*téndé* reflexes were found in the *Tanzanian Language Survey* database of Nurse & Philippson (1975). These records might be an artefact of the way in which they elicited their data (i.e. asking informants to write down local language equivalents of Swahili terms). In Swahili (G42), *mtende* (cl. 3-4) refers to the date palm, *tende* (cl. 9-10) to its fruits. Being bilingual, it is possible that the informants just took the Swahili (G42) word for want of a appropriate term in their own language. Of course, it might also be that lots of languages have a borrowed term for this palm because of its appearance in translations of the Bible, hymns etc. (Martin Walsh pers. comm.). Since the Nurse & Philippson (1975) data lack tonal notation, it is difficult to check whether these nouns have a tone pattern typical of recent loans, i.e. reflecting the Swahili (G42) penultimate accentuation instead of attesting regular tone correspondences characteristic of common Bantu inherited words. The few other pieces of data available, which stem from other sources, are rather uninformative with respect to reconstructing tone. In Jita (J25), for instance, where *i: n-té: nde*, 'date (fruit)' was

---

cal comparative series uniting the B40 group with H10 and H30. Raponda-Walker & Sillans (1961: 335) have reported *mbari*, '*E. var. nigrescens*' in several B40 languages.

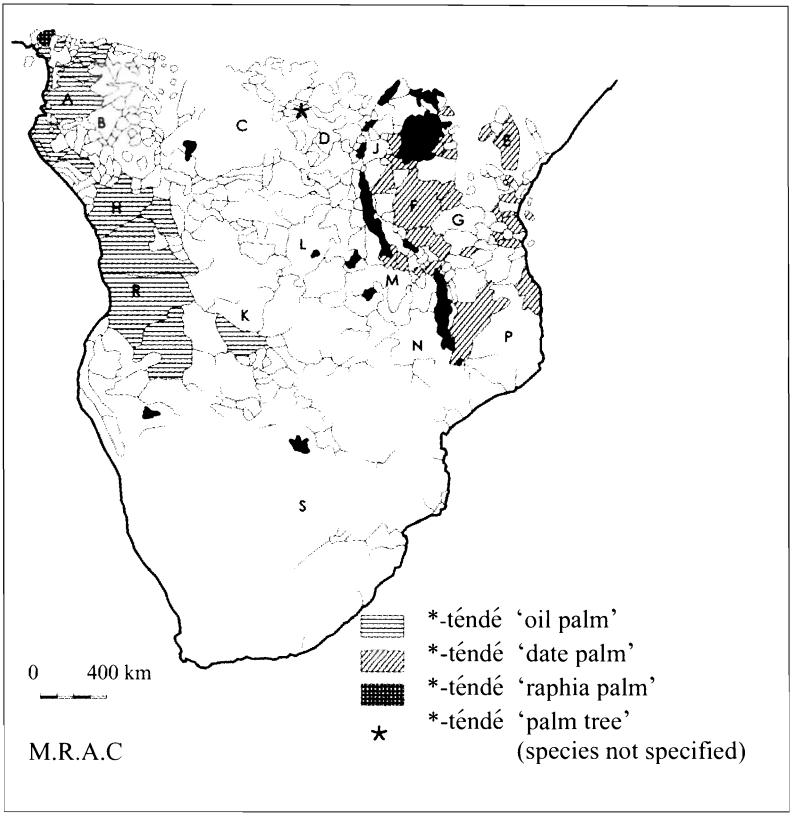
noted (Downing 1996: 233), the contrast *-*cvcv́* /*-*cvcv̀* is lost (Philippson 1998: 316). The resulting pattern is -*cvcv̀*, which is not distinctive from that of Swahili (G42) loans. Contrary to other Jita (J25) words, Downing (1996) does not consider it a Swahili (G42) loan. The Nyakyusa (M31) stress system is historically likewise unrevealing.

Whatever the case may be, even if the bulk of the *-*téndé* forms signifying 'date palm' are the result of a relatively recent diffusion from Swahili (G42), this still leaves open the question of how the date palm got to be called by this name. It looks most likely that *-*téndé* lost its original meaning 'oil palm' in favour of 'date palm', but where and when did this happen? One can assume that this semantic shift must have taken place along the East coast due to contact with the Arab world. This might have been in Swahili (G42), the language of Arab-Swahili merchants, who laid the foundations of the *lingua franca* it is today. Its origin might also be pre-Swahili, in Proto-North-East-Coast Bantu for instance. However, due to the lack of relevant tonal data, the evidence for such a pre-Swahili origin is fairly weak. In Shambaa (G23), for instance, Yukawa (1984: 89) recorded *m̀téndé* (cl. 3-4), 'date palm'. The tonal pattern of this noun could regularly correspond to *H-H, e.g. *-*címbá*,'lion' > *shímbá*, °-*pémbé* 'horn' > *m̀pémbé*, °-*kóndó* 'anus' > *ùkúndú*, and it does not attest the tone pattern that is most common (but not exclusive) among Swahili (G42) loanwords, i.e. H-L. However, its tone pattern could also reflect *H-L, e.g. *-*kónì*, 'firewood' > *ùkúní*, *-*kóndè*, 'bean' > *ǹkúndé*, which means that the tonal evidence is not conclusive on the possible pre-Swahili origin of the term. Whatever the origin of the term's semantic shift may be, it is remarkable that the date palm took the ancestral Bantu name for 'oil palm' upon its introduction to eastern Africa, and not that of a more closely related wild African relative, such as the *Phoenix reclinata* Jacq.[28]

Guthrie (1967-1971), having collected only five attestations, in both eastern and western Bantu languages, doubted whether C.S. 1712 consisted of direct cognates, since their meanings diverged. The supplementary reflexes we have collected permit us to confirm the unity of the comparative series concerned. As can be seen on Map 5, the geographic distribution of *-*téndé* consists of two continuous blocs of reflexes which differ semantically.

[28] The noun °-*kìndó* is reported to be recurrent for the African wild date palm among the eastern Bantu languages of zones E, G, J, K, L, M, P, and S (Bastin *et al.* 2003, Guthrie 1967-71).

Map 5: The distribution of **-téndé* inside the Bantu domain



If one considers solely the distribution of the **-téndé* reflexes related to the oil palm, one perceives a continuous north-south belt, which is limited to western Bantu and stretches more or less along the Atlantic coast from zone A (10-40+70-80) through zones B (10+40) and H (10-30) into northern zone R (R11), with a slightly westward extension into the K10 group. Historically speaking, according to the Vansina (1995) classification, these reflexes thus cover the North-west subgroup of Bantu (most of zone A + B10-30) and two subgroups of the West subgroup, i.e. West Coastal (B40-80 and zone H, except H41) and South-west (zone R, most of zone K, H41, and some zone L languages). Since it occurs in two main subgroups of Bantu, it cannot be considered a regional innovation; it must be quite old. This is all the more true, if one looks upon the eastern **-téndé* forms as cognates. In spite of their divergent meaning 'date palm', they are in my

view semantically too close to be considered historical homonyms. Although the major part of these *-*téndé* reflexes might be the result of a recent diffusion, their origin certainly is East Bantu.

Important in this respect is the geographical gap that exists between the western and eastern attestations, which rules out the possibility of a western diffusion source for the eastern reflexes. Being attested in at least three of the main Bantu subgroups, *-*téndé* seems to be an appropriate candidate for reconstruction in PB. Accordingly, the early presence of *-*téndé* in East Bantu is in all likelihood a corollary of the first Bantu expansions in this area. It must have reached this part of the Bantu domain as a (formal) retention, which subsequently underwent a semantic innovation. This hypothesis poses a problem due to the lack of eastern *-*téndé* reflexes, which unmistakably manifest the assumed original meaning 'oil palm' and which could reflect a diachronic semantic stepping stone between the inherited and the new sense. The Enya (D14) reflex cited in table 4, whose translation does not permit us to identify the genus of palm tree, might constitute such a missing link. Apart from this attestation, no other (possibly) 'oil palm' related *-*téndé* traces have yet been discovered in East Bantu.

3.5 Other recurrent Bantu names for 'oil palm'. Throughout the text I have pointed out other oil palm related terms, which are recurrent in several Bantu languages. However, either their geographic distribution is quite limited, so that they can be seen as local innovations, or it is more scattered but at present too fragmentary, so that drawing historical conclusions would be somewhat premature. Concerning the nouns referring to the tree as such, the terms I do not treat in detail in this paper belong essentially to eastern comparative series.

The Swahili (G42) term *mchikichi* has lexical correspondences in several other Tanzanian languages of zones E, F, G, M, and P. The Nurse & Philippson (1975) database reports for instance Kimbu (F24) *muchikichi*, 'oil palm', or Ndamba (G52) *mchikichi*, 'oil palm'. Given the fact that Swahili (G42) /ch/ generally corresponds to /s/ in these languages, it can be safely assumed that these forms are loanwords from Swahili (G42). The origin of the Swahili (G42) term, which seems to be a partial reduplication, is not clear however. Anyhow, it may have replaced the ancestral term *-*téndé*, whose meaning shifted from 'oil palm' to 'date palm'.

Another set of recurrent terms, occuring in zones E, F, G, J, and M, seems to be related to another Swahili (G42) term for 'oil palm', i.e. *muwese* (Heine & Legère 1995: 332). The correspondence is straightforward for forms such as Kuria (E43) *ama-wese*, 'oil palm', Nyamwezi (F21) *m-βese*, 'oil palm', Ha (J66)

ama-bese, 'palm oil', and Pimbwe (M11), *um-bese*, 'oil palm' (Nurse & Philippson 1975), whose C_1 corresponds regularly to the C_1 of the Swahili (G42) form. This is less evident for several zone J forms, such as Kiga (J14) *omu-meshe*, 'oil palm', Nyambo (J21) *omu-mese*, 'oil palm', Haya (J22) *eki-meshe*, 'oil palm' (ibid.), and Shi (J53) *mamesa*, 'l'huile de palme' (Hostens & Hoste n.d.: 266), where the /m/ in stem-initial position could be the result of the integration of a NP in the noun stem. It should be noted however that the C_2 of the Swahili (G42) noun does not regularly correspond to the C_2 of all other forms. Since /ch/ is expected in Swahili (G42) when /s/ occurs in the other languages, one can suppose that the Swahili (G42) *muwese* is a borrowing from these languages, and not the other way around. Most of the other terms are attested in languages spoken near Lake Tanganyika. Although more data need to be considered, its origin can tentatively be situated in this region. Its presence in Swahili (G42) might then be an outcome of the caravan trade. While *mchikichi* spread from the coast to the interior, the diffusion of *muwese* probably went in the other direction.

4. Proto-Bantu Oil Palm Vocabulary.

This overview of the most common 'oil palm' related Bantu names leads to some interesting insights, both semantic-botanically and historical-linguistically.

With regard to botanical semantics, in-depth results are hampered by the lack of detailed data. Apart from the botanic inventories listing vernacular names, the translations are almost always limited either to the common plant name or to its scientific designation. Nevertheless, the available data show that the most common Bantu names for 'oil palm' generally refer to the most common variety of the tree, i.e. the *nigrescens*. In this capacity, it also serves as a generic name for all varieties of oil palm trees. In order to refer specifically to the other, less frequent varieties, this common name is modified or a distinct name is used. Some of these specific variety designations seem to be recurrent in several adjacent languages, but with exception of the *°-koma* form, they never have a supra-local distribution. The *Elaeis guineensis* Jacq. often being the economically most valuable palm tree, the semantic range of the most common Bantu oil palm names has been extended in some languages, to become the generic name for several kinds of palm trees. Moreover, all of them have reflexes that in certain languages exclusively refer to another kind of palm tree. However, in all cases, 'oil palm' seems to be the primary meaning from which the others ('date palm', 'raffia palm', etc.) were derived.

For most names, such shifts are limited to rare and isolated languages. With respect to **-téndé*, however, the meaning 'date palm', which is widespread in Eastern Bantu, is probably the result of historically significant semantic shift that was generalised subsequently, either through inheritance from a common ancestor or by diffusion. Finally, as the most common Bantu oil palm designations, the noun stems treated above are not only used to refer to the tree itself, but also to designate its fruits and the oil extracted from them. Noun class variation is the morphological device relied on for this type of semantic alteration. A particular noun stem may have the different meanings 'oil palm', '(bunch of) palm nuts', and/or 'oil palm' in one and the same language, but the different meanings of one particular noun may also manifest a (partially) complementary distribution. Hence, one language may have several of the most common Bantu nouns, but each of them with a different meaning, e.g. Ntandu (H16g) *bá*, 'nom collectif des Palmiers à stipe'; *ngási*, 'noix de palme', *máasi*,'huile de palme'; *ntééndé ba*, 'jeune palmier' (Daeleman & Pauwels 1983: 207-8).

From a historical-linguistic point of view, this study has yielded insights that differ somewhat from the preliminary conclusions advanced by Guthrie (1967-1971). The use of additional data has led to a more comprehensive idea of the geographic distribution of the noun stems considered, while the valorisation of diachronic semantic evolutions has generated a better understanding of their mutual historical relationships. Guthrie (1967-1971) observed that **-bịdà*, **-bá* and **-gàdị́* have an almost exclusively western distribution and linked this observation to the fact that the oil palm tree is a more important economic factor in this part of the Bantu area. He deemed the **-téndé* C.S. too weak to draw historical conclusions from. Consequently, none of them could be considered as a PB-X item.

To start with, the geographic range of **-gàdí* is not exclusively western, even if one only takes into account the oil palm-related **-gàdí* reflexes, since these occur in the Eastern Bantu zones D, F, J, L, and M, and certainly not if one establishes a diachronic link between the eastern reflexes meaning 'blood' and their western counterparts meaning 'palm oil'. It is more appropriate to see the former as semantic derivations of the latter, than to consider both sets of reflexes as originating from two historically distinct homonyms. Hence, in all likelihood this form dates back to PB (and can thus be noted with a *: **-gàdí*). Similar diachronic semantically-inspired reasoning allows the reconstruction of **-téndé* into PB. It is not only attested in two of the western main Bantu branches as an oil palm name, i.e. 'North-west' and 'West' according to the Vansina (1995) labels, but it also occurs in East Bantu as a designation of the date palm.

As regards **-bídà* and *°-bá*, their distribution is indeed exclusively western. Nevertheless, the former is attested in the same two western primary branches as **-téndé*, i.e. 'North-west' and 'West'. Even if it has almost no eastern reflexes, this distribution suffices, in my opinion, to postulate a PB reconstruction. While the reduced economic relevance of the oil palm may have facilitated the semantic shifts that **-gàdí* and **-téndé* underwent in this part of the Bantu area, it may have induced the total disappearance of **-bídà*. The presence of a **-bídà* reflex in Lega (D25), generally considered as an Eastern Bantu language, seems to corroborate its status as a PB retention. As a result, amongst the four Guthrie (1967-1971) oil palm C.S., *°-bá* is probably the only one that really is a post-PB innovation. The geographic range of its reflexes is bound to the 'West' branch. Within this main Bantu subgroup, it co-occurs with the three PB noun stems. In this part of the Bantu area, **-gàdí* is predominantly used to designate the palm oil or any other kind of oil or fat, in some languages to designate the palm nut, but almost never to refer to the tree itself. **-téndé* and *°-bá* reflexes co-exist in Kongo (H10) and Umbundu (R11). In the latter language, the available data present them as perfect synonyms. In the Kongo (H10) area, *°-bá* is the generic name for (oil) palm trees, while **-téndé* conveys the particular notion of a young palm tree. Accordingly, *°-bá* cannot be considered an innovation having supplanted **-téndé* or **-gàdí*. Conversely, as stated above, it is in complementary distribution with **-bídà*, almost perfectly along the lines of lexicostatistically based sub-grouping. Hence, in accordance with the claim of Guthrie (1967-1971), *°-bá* probably replaced **-bídà* at a certain stage of the internal fragmentation of this branch. The latter has only been retained in the C50 languages and in the 'Sangha' (C10) and 'Rivers' (the northern C30 languages + C41-42) subgroups of the Vansina (1995) classification.

In sum, on the basis of the available Bantu data, **-bídà*, **-téndé*, and **-gàdí* can be tentatively reconstructed in PB, while *°-bá* should be considered a subsequent innovation. The meaning of the three PB noun stems was certainly related to the oil palm, but it can be questioned whether they were as synonymous as Guthrie (1967-1971) claimed. The comparative semantic analysis of the respective nouns has shown that the focal signification of each of them is slightly divergent, even if they may have identical meanings in different present-day languages. This probably indicates that their initial meaning was not identical. As I have argued above, the primary reference of **-gàdí* seems to be the palm nut and the oil which one extracts from it. The meaning 'oil palm', scattered within the **-gàdí* distribution area but mainly attested in the most western East Bantu lan-

guages, approximately spoken on both sides of Lake Kivu and Lake Tanganyika, probably results from subsequent (independent) semantic shifts.

As regards **-bídà* and **-téndé*, I have claimed that both refer primarily to the oil palm itself. This is certainly the case for the latter, whose western reflexes have this meaning almost exclusively. Through noun class variation, the **-bídà* reflexes are also used to refer to the palm nut or the bunch of palm nuts. Both terms might have co-existed in PB as references to the tree itself. They still do, as perfect synonyms, in the Myene (B11) languages. In Nen (A44), however, **-bídà* reflexes are used for the tree, its fruits and the bunch of nuts, while the **-téndé* reflex refers to a young palm tree. As mentioned above, this connotation of youth turns up in other languages and dates possibly back to PB. Although both designating the tree, **-bídà* and **-téndé* probably were only near-synonyms in the proto-language. Nevertheless, the redundancy caused by the loss of this particular connotation may explain why the one has disappeared in favour of the other in most other present-day Bantu languages.[29] Within the 'West' subgroup of the Vansina (1995) classification, the complementary distribution of both stems seems to run along genetic lines. While **-bídà* is solely attested in the "Sangha" and "Rivers" subgroups, **-téndé* figures in the 'West Coastal' and 'South-west' subgroups. Within the 'North-west' subgroup, the repartition of both noun stems looks to be somewhat more arbitrary. As far as the available data permit us to claim, **-téndé* is the only noun stem attested in A70 and **-bídà* the only one in A60 and B20, both are attested in separate A10, A20, A30 and A80 languages.

5. Bantu vs. Non-Bantu Niger-Congo Oil Palm Vocabulary.

In order to put the Bantu reconstructions just treated in a broader historical perspective, it might be worthwhile comparing them with non-Bantu Niger-Congo data. A systematic comparative study of these data is beyond the scope of this article. Nevertheless, as mentioned in the beginning of this paper, several linguists have made comparative studies of the oil palm vocabulary in non-Bantu Niger-Congo languages. In this section, I will present a summary of their reconstructions relevant in the light of our data, and I will comment on some historical relevant resemblances.

Williamson (1970, 1993) studied food plant names in the languages of Southern Nigeria, which apart from the Ijoid languages, belong to different sub-

[29] They co-exist in Akoose (A15) too, but here the **-bídà* reflex refers to the oil palm, while the **-téndé* reflex is the name of the raffia palm (Hedinger & Hedinger 1982).

groups of the (New) Benue-Congo branch of the Niger-Congo family. With respect to the oil palm vocabulary, just like in Bantu, she observed names that were consistent throughout a language group and could be assumed to be reconstructed to the proto-language of that group. Such was the case for other food crops believed to be indigenous to West-Africa. Given the very poor state of knowledge of the Proto-Niger-Congo sound system and its subsequent evolutions, Williamson (1993: 143) does not advance full-fledged reconstructions for oil palm names, but only presents two comparative series of resembling forms suggesting the possible existence of a common ancestral form. The first one consists of non-Bantu Benue-Congo forms possibly related to the Proto-Ịjọ reconstruction *ḍẹkụ́* (HLH) (cf. *-leg* like forms cited from Cross River and Plateau languages in Williamson (1973: 258-261)), but having no corresponding attestations within Narrow Bantu. According to Williamson (1993: 143), if the Proto-Ịjọ form is genuinely cognate with the Benue-Congo forms, this root goes back to Proto-Atlantic-Congo. At that time, she tentatively placed Ịjọid as Atlantic and Volta-Congo, the latter consisting of Kru, Kwa, Benue-Congo, Gur, Adamawa-Ubangi and possibly Dogon. The node uniting these three secondary branches, i.e. Atlantic-Congo, was a primary branch of Niger-Congo, coordinate with Mande and Kordofanian (Williamson 1989: 21).

However, the second series of Benue-Congo forms, meaning 'oil palm; palm oil' manifests strong resemblances with the PB reconstruction **-bídà*. To cite only some examples, Crabb (1965: 83) has listed several possible Ekoid cognates, e.g. *ù-bîr* in Nkim, *ò-βî* in Bendeghe, *ù-bi̇̀* in Nta. In Tiv, *ivile*, 'oil-palm/ oilnut' is attested (Abraham 1940: 306). Akpes has *e-bir* (Ibrahim-Ararabiyi 1989) and Elugbe (1989) reconstructed A-*ʹbʹidi* 'oil (palm oil)' in Proto-Edoid. Shimizu (1980: 197) advanced the Proto-Central-Jukunoid reconstruction **bìT*, 'oil', having reflexes such as *bir* in the Jibu dialect of Jukun. The author did not specify however whether it concerned palm oil. While Williamson (1973) supposed that **-bídà* represented a Bantoid innovation, she more recently linked all these forms to the De Wolf (1971) Proto-Benue-Congo reconstruction **-pide*, 'palm tree' and to the Proto-Ịjọ reconstruction *pulo* (Williamson 1993).

The bringing together of these reconstructions should be seen as very tentative, certainly as regards Proto-Ịjọ, and needs to be corroborated by solid sound correspondences.[30] Nevertheless, Williamson (1993) concludes that this root is

[30] According to Williamson (1993: 143), the Proto-Ịjọ **e-bin* (L-H), 'palm fruit' and the Edo *i-vin* (H-L), 'palm tree; palm kernels' might also be cognate. If this is the case, then the forms in the same languages meaning 'oil' are presumably not cognate.

clearly Proto-Atlantic-Congo, thus rising far up in the Niger-Congo family tree. She also refers to the Mukarovsky (1976) Proto-Western-Nigritic reconstruction *-*bil(a)*, 'oil palm; frond or nut of oil palm'.[31] This reconstruction is built on the – in my opinion, shaky – association of a series of reflexes from Gur, Togo Remnant and Western Kwa languages with Guthrie's Common Bantu form *-*bíḍà*, e.g. Tenyer (Gur) *mpen-ti-gi*, 'oil-palm' (*ti-gi* = 'tree'), Adele (Togo Remnant) *di-béŋ*, 'young oil palm', Awutu (Western Kwa) *á-bɛ̂*, 'oil palm'.[32] Given the state of knowledge of the internal relationships within Niger-Congo, such associations should be considered as highly speculative, for lack of evidence of regular sound correspondences. Amongst the numerous oil palm related vernacular names from West-African languages listed by Burkill (1997: 354-370), several items may be related to *-*bídà*, e.g. Akan (Kwa) *a-bɛ́*, 'oil-palm'; Guere (Kru) *bli*, 'oil-palm'; Akpafu (Kwa) *kù-bɛ̀rɛ̀*, 'oil-palm'; Anyi (Kwa) *ayee*, 'oil-palm'; Adangme (Kwa) *wie*, 'oil-palm'; Konkomba (Gur) *ebébirè*, 'oil-palm'; Ejagham (Benue-Congo, Ekoid) *obi*, 'oil-palm'; Icheve (Benue-Congo, Tivoid) *ò-vílè*, 'oil-palm'; Igbo (Benue-Congo, Igboid) *obea okpọ́rọ́*, '(oil-palm with) thin-fleshed (nut): *f. leucocarpa*'; Edo (Benue-Congo, Edoid) *ivin*, 'a general term for nuts', *ivin ọ́rúnmila*, 'nuts of Oronmila; after an Yoruba high-priest skilled in divination'; Edo (Benue-Congo, Edoid) *èɓìlì*, 'palm oil'; Yoruba (Benue-Congo, Defoid, Yoruboid) *ọ̀pẹ̀ ifá*, 'palm of Ifa (the god of divination)'. In sum, the PB etymon *-*bídà*, 'oil-palm', inherited by several present-day Bantu languages, seems to be a retention from an earlier ancestor. According to the available data, it probably reaches back to the Proto-Benue-Congo level, and perhaps beyond.

Connell (1998) focused on the yam- and palm-related vocabulary in one particular Benue-Congo subgroup, i.e. the Delta Cross languages of Southeastern

---

[31] Mukarovsky's Western Nigritic languages incorporate Western Guinean (part of Greenberg's West Atlantic), Mel (part of Greenberg's West Atlantic), Gur (Greenberg's Voltaic), Togo Remnant (part of Greenberg's Kwa), Western Kwa (part of Greenberg's Kwa) and Benue-Congo (more or less the same as Greenberg's) languages. Apart from Mande and Adamawa-Ubangi, it roughly corresponds to Greenberg's Niger-Congo.

[32] Another Awutu form cited by Mukarovsky (1976) is *é-wînî*, 'corps', which would be cognate to Guthrie's Common Bantu reconstruction *-*bìdî*, 'body'. The Bantu intervocalic /l/ corresponds here to /n/ in Awutu. Moreover, Mukarovsky (1976) also cites the Awutu form *á-bɛ̂*, together with a series of similar forms from other Western Kwa languages, as a possible reflex of his Proto-Western-Nigritic reconstruction *-*bá*-, 'oil palm', which he associates to Guthrie's Common Bantu reconstruction *-*bá*. Distinguishing between the present-day reflexes of *-*bá* and *-*bídà* often being tough within Bantu, as I have commented above, it is a still more risky business between languages of different Niger-Congo subgroups.

Nigeria. Terms for the trees themselves, as well as for frond, palm fruit, palm oil, palm kernel, broom, wine and possibly pressing were reconstructible to Proto-Delta Cross (PDC); he concluded that both the oil palm and the raffia palm were known and exploited by the speakers of that proto-language. Connell's PDC reconstructions of terms for the oil palm tree itself are interesting with respect to the Bantu vocabulary treated in this paper. Apart from *-*jòb*, which seems to be a Delta Cross innovation, Connell (1998: 335) has reconstructed *-*dde* and *-*ten* beyond PDC. The former, attested in the Upper Cross and Central Delta subgroups, seems to be cognate with forms found in much of Benue-Congo and beyond. The *-*dde* reconstruction may be related to the first comparative series listed by Williamson (1993: 143), which, as mentioned above, has no cognates within Narrow Bantu. The latter reconstruction, however, is probably related to the PB reconstruction *-*téndé*.

It is significant that Connell (1998) has reconstructed *u-*tɛ́n* in Proto-Lower-Cross with the meaning 'young oil palm', while the Proto-Ogoni *-*té*, and the Central Delta *àtèn* refer to a mature oil palm. As the author indicates, this variation between young and mature oil palm in the different Delta Cross subgroups gives evidence of semantic shift, but does not permit one to reconstruct a precise meaning in their proto-languages. Evidence for such shifting is also attested within Lower Cross, where Ushaghade *útén* is a general term for oil palm (Connell 1998: 335). However, in the light of the Bantu data discussed in this paper, the specific signification 'young oil palm' seems more likely. Moreover, the Delta Cross data deepen the antiquity of the term's connotation of youth, already observed within Bantu, and buttress its reconstruction in PB. As Connell (1998) rightly remarks, this reconstruction, *-*ten* in PDC and *-*téndé* in PB, manifests a strong similarity with terms found throughout Niger-Congo. He cites the following forms: Bambara (Mande) *nten*, Gã (Kwa) *tɛ̀ŋ*, Mambila (Bantoid) *ter*. De Wolf (1971) advanced the Proto-Benue-Congo reconstruction *-*tende*, 'palm tree', which is also given by Williamson (1973: 260). The latter refers to the Proto-Western-Nigritic reconstruction *-*tandi* by Mukarovsky (1976: 355), who cites the Gola (Atlantic) example *má-tende*, 'leaves of (raphia) palm tree' as a possible reflex. The Burkill (1997: 354-370) lists also contain several nouns whose historical relatedness with PB *-*téndé* merits closer examination, e.g. Mandinka (Mande) *téẽ*, 'oil palm'; Maninka (Mande) *tin-tulu*, 'palm of oil'; Yoruba (Benue-Congo, Defoid, Yoruboid) *ùdẹn*, 'kernel oil'.

It is, therefore, very likely that *-*téndé* is older than *-*bídà* because it not only can be reconstructed in Proto-Benue-Congo, but, as suggested by Connell (1998), probably goes back to Proto-Mande-Congo (i.e. Niger-Congo excluding

Kordofanian). Connell even considers the possibility that both PDC *-*dde* and PDC *-*ten* are reflexes of this old root. He refrains, however, from explaining how PDC would have acquired a double reflex of the same proto-form. Blench (forthcoming) seems to concur, uniting in one comparative series Bantu reflexes of *-*téndé* with would-be cognates of PDC *-*dde*, numerously represented in the Burkill (1997) lists, but absent from Bantu. The fact that both *-*dde* and *-*ten* can be reconstructed in one and the same proto-language, i.e. PDC, is a strong indication for two distinct (Proto?-)Niger-Congo etymons, of which only one was retained in PB.

As for the other common Bantu oil palm-related noun stems outlined in this article, there are no clear NC traces of *-*gàdí* beyond Bantu in the available documentation. Thus, this form, referring both to the palm nut and the oil extracted from it, is a Bantu innovation. Only Mukarovsky (1976) perceived possible reflexes of °-*bá* beyond Bantu. However, as noted above, the nouns he cited could also be reflexes of *-*bídà*. Mukarovsky based his reconstruction on the presumption that °-*bá* is a PB form, but, as I have argued, it should be considered an innovation ulterior to PB.

6. Conclusions.

This survey of non-Bantu Niger-Congo data has allowed us to put the common Bantu oil palm-related vocabulary in wider historical perspective. At least two of the terms reconstructed to PB, i.e. *-*téndé*, '(young) oil palm', and *-*bídà*, 'oil palm', have turned out to be retentions of pre-Bantu ancestors. Insofar as the present state of knowledge permits one to suggest, *-*téndé* seems to be the oldest term, in all likelihood dating back to a pre-Benue-Congo ancestor, maybe even to PNC, while *-*bídà* is probably a Benue-Congo innovation. The reconstruction of these two retentions into PB not only indicates that the speakers of the latest common ancestor of the present-day Bantu languages knew the oil palm, but also suggests that their ancestors were familiar with this tree in the distant past. However, these reconstructions do not tell us whether the oil palm was exploited in the PB era. More telling in this respect is the reconstruction of *-*gàdí* with the meaning 'palm nut; palm oil'. This points out that PB speakers knew how to use palm oil, and that they exploited the oil palm for nutritive and culinary purposes. The fact that *-*gàdí* does not exist beyond Bantu suggests that this practice was quite new, and not a retention from ancestral times. Of course, more detailed comparative research on Bantu terms for the associated products and the culture of the oil palm is required to draw more comprehensive historical conclusions on

its exploitation. Nonetheless, it is pertinent to note that none of the numerous reconstructions related to the dietary aspects of the oil palm vocabulary advanced by Connell (1998) seem to reconstruct to higher than the PDC level, while the names of the tree itself are pre-PDC retentions. Just as in Bantu, this observation may indicate that the ancestors of the PDC speakers had long been familiar with the oil palm, but that only later generations began to rely on the tree in their subsistence economy.

REFERENCES

Abraham, Roy Clive. 1940. *A Dictionary of the Tiv Language*. Herford: Stephen Austin & Sons.

Adam, J. 1954. *Grammaire composée mbede, ndumu, duma*. Brazzaville: Mémoires de l'Institut d'Etudes Centrafricaines 6.

Adams, Gustaf. 1907. "Die Sprache der Banôho". *Mitteilungen des Seminars für Orientalische Sprachen* 10: 34-84.

Anonym. n.d. *Lokele*. Manuscript. Yakusu: Boston Mission Society.

Bareau, M. & Reding. 1912. *Vocabulaire français-mobenge et mobenge-français*. Bruxelles: Imprimerie Veuve Monnom.

Bastin, Yvonne. 1985. *Les relations sémantiques dans les langues bantoues*. Bruxelles: ARSOM, Classe des Sciences Morales et Politiques 43, 4.

Bastin, Yvonne. 1997. "Réflexions diachroniques à propos de quelques thèmes pour 'sang' en bantou." *Paroles et Cultures bantoues, Mélanges en hommage à F.M. Rodegem*. Tervuren: Musée royal de l'Afrique centrale, Annales Sciences humaines vol. 159. Pp. 19-25.

Bastin, Yvonne, André Coupez, Evariste Mumba & Thilo Schadeberg. 2003. *Reconstructions lexicales bantoues 3 / Bantu Lexical Reconstructions 3*. Online available database: http: //Linguistics.africamuseum.be/BLR3.html (last consultation: 15-01-2006).

Bastin, Yvonne, André Coupez & Michael Mann. 1999. *Continuity and Divergence in the Bantu Languages: Perspectives from a Lexicostatistic Study.* Tervuren: Musée royal de l'Afrique centrale, Annales, Série in-8°, Sciences humaines 162.

Biton, Alexandre. 1907. *Dictionnaire français-ndumu et ndumu-français, précédé d'éléments de grammaire.* Franceville.

Blanchon, Jean Alain. 1994. *Dictionnaire ipunu-français.* Online available: http://www.cbold.ddl.ishlyon.cnrs.fr/CBOLD_Lexicons/Punu.Blanchon1994/ (last consultation: 08-06-2005).

Blench, Roger. forthcoming. *Archaeology, Language and the African Past.* Lanham: Altamira Press.

Bolekia Boleka, Julio. 1991. *Curso de lengua bubi.* Malabo: Centro Cultural Hispano-Guineano ediciones.

Bostoen, Koen. 2004. "Linguistics for the use of African history and the comparative study of Bantu pottery vocabulary". *Language and Revolution // Language and Time*, ed. by Frank Brisard, Sigurd d'Hondt & Tanja Mortelmans. Antwerp: University of Antwerp, Antwerp Papers in Linguistics. Pp. 131-54.

Bostoen, Koen. 2005. *Des mots et des pots en bantou. Une approche linguistique de l'histoire de la céramique en Afrique.* Frankfurt: Peter Lang, Schriften zur Afrikanistik–Research in African Studies.

Bostoen, Koen. 2006. "What comparative Bantu pottery vocabulary may tell us about early human settlement in the Inner Congo Basin". *Afrique et Histoire* 5: 221-263.

Botne, Robert. 1994. *A Lega and English Dictionary with an Index to Proto-Bantu Roots.* Köln: R. Köppe, East African Languages and Dialects 5, 3.

Bourquin, Walther. 1923. *Neue Ur-Bantu-Wortstämme nebst einem Beitrag zur Erforschung der Bantuwurzeln.* Berlin: D. Reimer, Beihefte zur Zeitschrift für Eingeborenen-Sprachen, Heft 5.

Burkill, H.M. 1997. *The Useful Plants of West Tropical Africa. Vol. 4: Families M-R*. Kew: Royal Botanical Gardens, 2nd edition.

Burssens, Nico. 1993. "Lexique et texte wongo (Bandundu Z)". *Annales Aequatoria* 14: 457-81.

Connell, Bruce. 1998. "Linguistic evidence for the development of yam and palm culture among the Delta Cross peoples of Southeastern Nigeria". *Archaeology and Language II, Archeological Data and Linguistic Hypotheses*, ed. by Roger Blench & Matthew Spriggs. London, New York: Routledge. Pp. 324-65.

Cornet, D. 2001. "Palmier à huile (*Elaeis guineensis* Jacq.)". *Agriculture en Afrique Tropicale*, ed. by Romain Raemaekers. Bruxelles: Direction Général de la Coopération Internationale, Ministère des Affaires Étrangères, du Commerce Extérieur et de la Coopération Internationale. Pp. 812-42.

Crabb, David W. 1965. *Ekoid Bantu Languages of Ogoja, Eastern Nigeria - Part I: Introduction, Phonology and Comparative Vocabulary*. Cambridge: Cambridge University Press in association with the West African Languages Survey and the Institute of African Studies, Ibadan, West African language monographs 4.

Daeleman, Jan & Luc Pauwels. 1983. "Notes d'ethnobotanique ntandu (Kongo). Principales plantes de la région de Kisaantu: noms ntandu et noms scientifiques." *Africana Linguistica* 9: 149-256.

Dalby, David. 1976. "The prehistorical implications of Guthrie's Comparative Bantu. Part II: Interpretation of cultural vocabulary". *Journal of African History* 17, 1: 1-27.

Da Silva, Maia Antonio. 1994. *Dicionário Complementar Portugues-Kimbundu-Kikongo (Linguas nativas do centro e norte de Angola)*. Cucujães: Editorial Missões, Cooperação Portuguesa.

De Clercq, Auguste & Emile Willems. 1960. *Dictionnaire tshiluba-français*. Léopoldville: Imprimerie de la Société missionnaire de St. Paul.

De Cort, Liaudet & Van Goethem. 1912. *Vocabulaire français-ababua et ababua-français*. Bruxelles: Imprimerie Veuve Monnom.

De Wolf, Paul. 1971. *The Noun Class System of Proto-Benue-Congo*. The Hague: Mouton.

Downing, Laura J. 1996. *The Tonal Phonology of Jita*. München: Lincom Europa.

Dugast, Idelette. 1967. *Lexique de la langue tunen (parler des Banen du sud-ouest du Cameroun)*. Paris: Klincksieck, Langues et littératures de l'Afrique noire.

Elugbe, Ben Ohi. 1989. *Comparative Edoid: Phonology and Lexicon*. Port Harcout: University of Port Harcout Press, Delta Series N° 6.

Ernst, Urs. 1989. *Lexique kako-français et français-kako avec tableaux de conjugaison*. Yaoundé: Societe Internationale de Linguistique.

Felberg, Knut. 1996. *Nyakyusa-English-Swahili and English-Nyakyusa Dictionary*. Dar-es-Salaam: Mkuki na Nyota Publishers.

Fernandez, L. 1951. *Diccionario español-kômbe*. Madrid: Instituto de Estudios Africanos.

Goemaere, Alphonse. n.d. *Spraakleer van het Londengese*. Manuscript. Tervuren: Musée Royal d'Afrique Centrale.

Gossweiler, John. 1953. *Nomes indigenas de plantas de Angola*. Luanda: Imprensa Nacional.

Grégoire, Claire. 1976. "Le champ sémantique du thème bantou *-bánjá". *African Languages* 2: 1-13.

Guerreiro, Viegas. 1963. *Rudimentos de língua maconde*. Lourenço Marques: Instituto de Investigação Científica de Moçambique.

Guthrie, Malcolm. 1967-1971. *Comparative Bantu: An Introduction to the Comparative Linguistics and Prehistory of the Bantu Languages, 4 vol.* London: Gregg International Publishers Ltd.

Guthrie, Malcolm. 1970. Contributions from Comparative Bantu Studies to the Prehistory of Africa. In David Dalby, *Languages and History in Africa*, 20-49. New York: Africana Publishing Corp.

Hagendorens, Joseph. 1975. *Dictionnaire otetela-français.* Bandundu: CEEBA, série III, vol. 5.

Hartley, Charles W.S. 1967. *The Oil Palm*. London: Longman.

Hedinger, Robert & Sylvia Hedinger. 1982. *English-Akoose Wordlist*. Manuscript. SIL Cameroun.

Heine, Bernd & Karsten Legère. 1995. *Swahili Plants*. Köln: Rüdiger Köppe

Helmlinger, Paul. 1972. *Dictionnaire duala-français, suivi d'un lexique français-duala.* Paris: Klincksieck, Langues et Littératures de l'Afrique noire 9.

Horton, Albert E. 1953. *A Dictionary of Luvale*. El Monte: Rhan Brothers Printing & Lithographing Co.

Hostens & Hoste. n.d. *Dictionnaire français-shi*. Manuscript. Bukavu: Centre d'étude des langues africaines.

Hulstaert, Gustaaf. 1966. *Notes d'ethnobotanique mongo*. Bruxelles: Académie royale des sciences d'Outre-Mer.

Ibrahim-Arirabiyi, Femi. 1989. *A Comparative Reconstruction of Akpes Lects*. Port Harcourt: University of Port Harcourt, Akoko North Ondo State, Ph.D. dissertation.

Iliku Mimpiya Dibata. 1979. *Esquisse grammaticale de la langue tsong: phonologie et morphologie*. Mémoire. Lubumbashi: UNZ.

Jacob, Irenée. 1984. *Dictionnaire rwandais-français en 3 volumes, extrait du dictionnaire de l'Institut National de Recherche Scientifique Kigali.* Kigali: Imprimerie scolaire.

Kagaya, Ryohei. 1992. *A Classified Vocabulary of the Bakueri Language*. Tokyo: ILCAA, Bantu Vocabulary Series 8.

Kakeya, Makodo & Nishida, Toshisada. (1976) *A Glossary of Sitongwe.* Manuscript.

Kawata, Ashem Tem. 2003. *Bagó dictionnaire: Lingala / Falansé-Français / Lingala*. Saint-Pierre-Les-Nemours: Laboratoire des langues congolaises.

Kölle, Sigismund Wilhelm. 1854. *Polyglotta Africana, or a Comparative Vocabulary of Nearly Three Hundred Words and Phrases, in More than One Hundred Distinct African Languages*. London: Church Missionary House.

Koloni, Jean. 1971. *Eléments de morphologie et de vocabulaire de la langue enya.* Mémoire. Kinshasa: Université Lovanium.

Kuperus, Juliana. 1985. *The Londo Word; Its Phonological and Morphological Structure.* Tervuren: Musée royal de l'Afrique centrale, Annales, Série in-8°, Sciences humaines 119.

Laman, Karl Edvard. 1936. *Dictionnaire kikongo-français.* Bruxelles: Institut royal colonial belge, tome II.

Le Guennec, Grégoire & Jose Francisco Valente. 1972. *Dictionário Português-Umbundu*. Lunda: Instituto de Inverstigação Cientifica de Angola.

Leitch, Myles. 1991. *Lexique Babole-Français*. Brazzaville: Université Marien Ngouabi, Centre pour l'Etude des Langues Congolaises.

Mabiala, Jean-Noel. 1992. *Etudes du kiyoombi, langue kongo du Congo.* Mémoire. Lyon: Université Lumiere-Lyon 2.

Maley, Jean. 1999. "L'expansion du palmier à huile (*Elaeis guineensis*) en Afrique centrale au cours des 3 derniers millénaires: nouvelles données et

interpretations". *L'homme et la forêt tropicale*, ed. by Serge Bahuchet, Daniel Bley, Hélène Pagezy & Nicole Vernazza-Licht. Marseille: Editions de Bergier, Travaux de la Société d'Ecologie Humaine (SEH) / APFT. Pp. 237-54.

Maley, Jean & Alex Chepstow-Lusty. 2001. "Elais guineensis Jacq. (oil palm) fluctuations in central Africa during the late Holocene: climate or human driving forces for this pioneering species?". *Vegetation History and Archaeobotany* 10, 2: 117-20.

Maniacky, Jacky. 2003. *Tonologie du ngangela - variété du Menongue (Angola)*. München: Lincom Europa, Lincom Studies in African Linguistics 61.

Maniacky, Jacky. 2005. "Quelques thèmes pour « igname » en bantu". *Studies in African Comparative Linguistics with Special Focus on Bantu and Mande*, ed. by Koen Bostoen & Jacky Maniacky. Tervuren: Musée royal de l'Afrique centrale. Pp. 165-89.

Meeussen, Achille Emile. 1969. *Bantu Lexical Reconstructions*. Tervuren: Musée Royal d'Afrique Centrale.

Meinhof, Carl. 1910. *Grundriss einer Lautlehre der Bantusprachen*. Berlin: Reimer & Vohsen.

Meinhof, Carl & Nicolaas Jacobus van Warmelo. 1932. *Introduction to the Phonology of the Bantu Languages Being the English Version of 'Grundriss einer Lautlehre der Bantusprachen.'* Berlin: D. Reimer, E. Vohsen.

Mertens, Joseph. 1939. *Dictionnaire idzing-français suivi d'un aide-mémoire français-idzing*. Bruxelles: IRCB, mémoire 4, 3.

Mickala-Manfoumbi, Roger. 2004. *Lexique pové-français français-pové*. Libreville: Editions Raponda Walker.

Moñino, Yves. 1995. *Le proto-gbaya. Essai de linguistique comparative sur vingt et une langues d'Afrique Centrale*. Paris: SELAF.

Motingea, André Mangulu. 1996. *Etude comparative des langues ngiri de l'entre Ubangi-Zaïre*. Leiden: CNWS.

Motingea, André Mangulu. 2004. "The linguistic domain of the Central Congo Bassin. Uniqueness in diversity". *Nature and Culture in the Democratic Republic of Congo*, ed. by Els Cornelissen *et al.* Tervuren: Musée royal de l'Afrique centrale. Pp. 93-9.

Motte, Elisabeth. 1980. *Les plantes chez les pygmées aka et les Monzombo de la Lobaye (Centrafrique)*. Paris: SELAF.

Mudindaambi, Lumbwe. 1977. *Dictionnaire mbala-français, part. I A – fugula*. Bandundu: CEEBA publications, serie III, vol. 4a.

Mukarovsky, Hans G. 1976. *A Study of Western Nigritic, Volume I*. Wien: Institut für Ägyptologie und Afrikanistik, Beiträge zur Afrikanistik, Band 1-2.

Nakagawa, Hiroshi. 1992. *A Classified Vocabulary of the Ha Language*. Tokyo: ILCAA, Bantu Vocabulary series 9.

Ngila Bompeti, Pierre. 2000. *Expérience végétale bolia (République Démocratique du Congo). Catégorisation, utilisation et dénomination des plantes*. Köln: R. Köppe, Grammatische Analysen Afrikanischer Sprache Band 14.

Nguma, Wendo. 1986. *Dictionnaire français-yansi (Rép. du Zaïre)*. Bandundu: CEEBA.

Nurse, Derek & Philippson, Gérard. 1975. *The Tanzanian Language Survey*. Online database: http: //www.cbold.ddl.ish-lyon.cnrs.fr/Docs/TLS.html (last consultation: 08-06-2005).

Paas, Steven. 2004. *Chichewa/Chinanja-English*. Blantyre: Christian Literature Association in Malawi.

Pearson, Emil. 1973. *English-Ngangela Dictionary*. Morelos, Mexico: Cuernavaca.

Philippson, Gérard. 1998. "Tone reduction vs. metrical attraction in the evolution of eastern Bantu tone systems". *Theoretical Aspects of Bantu Tone*, ed. by Larry M. Hyman & Chrles Kisseberth. Stanford: CSLI. Pp. 315-29.

Philippson, Gérard & Serge Bahuchet. 1994-1995. "Cultivated crops and Bantu migrations in Central and Eastern Africa: a linguistic approach". *Azania* 29-30: 103-20.

Purseglove, John William. 1972. *Tropical Crops: Monocotyledons*. London: Longman.

Raponda-Walker, André & Roger Sillans. 1961. *Les plantes utiles du Gabon. Encyclopédie biologique.* Paris: Paul Lechevalier.

Rood, N. 1958. *Ngombe-Nederlands-Frans Woordenboek. Dictionnaire ngombe-néerlandais-français.* Tervuren: Musée Royal d'Afrique Centrale, Sciences de l'Homme, Linguistique vol. 21.

Ruttenberg, Piet. 1969. *Lexique yaka-français, français-yaka.* Kinshasa.

Sacleux, Charles. 1941. *Dictionnaire swahili-français*, vol. 2. Paris: Université de Paris, Travaux et Mémoires de l'Institut d'Ethnologie.

Shimizu, Kiyoshi. 1980. *Comparative Jukunoid: Reconstructions and Comparative Wordlists of the Jukonoid Languages, Volume II (part 2).* Wien: Institut für Ägyptologie und Afrikanistik, Beiträge zur Afrikanistik, Band 5- 7.

Sowunmi, Adabisi M. 1999. "The significance of the oil palm (*Elaeis guineensis* Jacq.) in the late Holocene environments of West and West Central Africa: a further consideration". *Vegetation History and Archaeobotany* 8, 3: 199-210.

Stahl, Ann Bower. 1993. "Intensification in the West African Late Stone Age: a view from central Ghana." *The Archeology of Africa. Food, Metals and Towns*, ed. by Thurston Shaw, Paul Sinclair, Bassey Andah & Alex Okpoko. London-New York: Routledge. Pp. 139-53.

Stanley, Henry Morton. 1872. *How I Found Livingstone. Travels, Adventures and Discoveries in Central Africa Including Four Months Residence with Dr. Livingstone*. London: S. Low, Marston, Low, and Searle

Stappers, Leo. 1984. *Lexique kisongye*. Manuscript. Tervuren: Musée royal de l'Afrique centrale.

Stoop, Henk. 1977. *Le lyombo, un dialecte des Topoke*. Manuscript. Tervuren: Musée royal de l'Afrique centrale.

Swartenbroeckx, Pierre. 1973. *Dictionnaire kikongo et kituba-français*. Bandundu: CEEBA III, 2.

Thomas, Jacqueline M.C., Serge Bahuchet & Alain Epelboin. 1993. *Encyclopédie des Pygmées Aka, 4 volumes*. Paris: Editions Peeters.

Thornell, Christina. 2004. "Wild plant names in the Mpiemo language". *Africa and Asia* 4: 57-89.

Turvey, Basil Henry Capes, Wolfgang Zimmerman & Gabriel B. Taapopi. 1977. *Kwanyama-English Dictionary*. Johannesburg: Witwatersrand University Press.

Vansina, Jan. 1959. *Esquisse de grammaire bushong.* Tervuren: Musée royal de l'Afrique centrale.

Vansina, Jan. 1995. "New linguistic evidence and the Bantu expansion". *Journal of African History* 36: 173-95.

Vansina, Jan. 2004. *How Societies Are Born: Governance in West Central Africa before 1600.* Charlottesville: University Press of Virginia.

Van Warmelo, Nicolaas Jacobus. 1989. *Venda Dictionary, Tshivenda-English.* Pretoria: J.L. van Schalk.

Whitehead, John B. 1899. *Grammar and Dictionary of the Bobangi Language, as Spoken over a Part of the Upper Congo West Central Africa*. London: Kegan P., Trench, Trübner CO.

Williamson, Kay. 1970. "Some food plant names in the Niger Delta". *International Journal of American Linguistics* 36: 156-67.

Williamson, Kay. 1973. *Benue-Congo-Comparative Wordlist, Volume II.* Ibadan: West African Linguistic Society.

Williamson, Kay. 1989. "Niger-Congo overview". *The Niger-Congo Languages. A Classification and Description of Africa's Largest Language Family*, ed. by John Bendor-Samuel & Hartell L Rhonda. Lanham-New-York-London: University Press of America. Pp. 3-45.

Williamson, Kay. 1993. "Linguistic evidence for the use of tree and tuber food plants". *The Archeology of Africa. Food, Metals and Towns*, ed. by Thurston Shaw, Paul Sinclair, Bassey Andah & Alex Okpoko. London. New York: Routledge. Pp. 139-53.

Wynne, Ronald C. 1980. *English-Mbukushu Dictionary.* Amersham: Avebury.

Yukawa, Yasutoshi. 1984. *A Classified Dictionary of the Sambaa Language*. Tokyo: ILCAA, Asian and African lexicon 15.

Ziervogel, Dirk & Pothinus C. Mokgokong. 1975. Groot Noord-Sotho Woordenboek, Comprehensive Northern Sotho Dictionary. Pretoria: Van Schaik Ltd/Unisa.

African Linguistics
Royal Museum for Central Africa
3080 Terveuren, Belgium
koen.bostoen@africamuseum.be

[received November 10, 2005
accepted January 11, 2006]

Studies in African Linguistics
Volume 34, Number 2, 2005

FOUR WORDLISTS OF EXTINCT CAPE KHOEKHOE FROM THE 18TH CENTURY*

François-Xavier Fauvelle-Aymar
Centre National de la Recherche Scientifique (France)
French Institute of South Africa (Johannesburg)

The Khoekhoe language, a member of the Khoisan family, was widely spoken by Southern Africa pastoralists and hunters-gatherers a few centuries ago. Apart from varieties still spoken in the 20th century (such as Nama in Namibia), very little is known of the dialect spoken by the Cape Khoekhoe. This paper is a critical edition of four manuscript wordlists collected at the end of the 18th century by Robert Jacob Gordon, then commandant of the Dutch garrison at the Cape. These lists refer to several local varieties of the Cape Khoekhoe language, collected just before it became extinct, and display two distinct systems of clicks rendering.

When European travelers started to visit the coast of Southern Africa at the end of the 15th century, they met people who, they said, were "clucking and whistling" like "angry turkeys" (Raven-Hart 1967: 19). It was obviously an allusion to a phonological characteristic of local languages, namely the presence of "clicks" (compound sounds consisting of influxes and effluxes). Such clicks are found in all the languages of the Khoisan family (Greenberg 1966: 66-84), in two linguis-

* My acknowledgment to the staff of the Brenthurst library, and especially to Diana Madden; to Werner Prinsloo, who helped me with the Dutch; to Karim Sadr and Laurent Chauvet, who kindly read and corrected a previous version of this article, and to the editor and the anonymous referees for helping to improve this paper.

tic isolates of Tanzania which may or may not belong to the same family (Köhler 1981: 465; Güldemann & Vossen 2000: 99-101), and in some Southern Bantu languages (Niger-Congo family), such as Xhosa, Zulu and Swazi (Herbert 2002), that have long been in contact with Khoisan languages.

There is no reason to believe that all the people whom the navigators met during the first two centuries of sporadic contact on the shores of South Africa and Namibia (Raven-Hart 1967) were speakers of one and the same language. However, there are some reasons to think that the numerous pastoralist communities stretching from Algoa Bay, on the Indian Ocean coast, to the Orange River (and possibly further north), on the Atlantic Ocean coast, which are known to have formed, in the 17th century, a long chain of genealogically- and politically-related groups of cattle and sheep herders, were speakers of very close varieties of the same language (Elphick 1977: 10, *passim*). These people used to call themselves Khoekhoe, an auto-ethnonym formed on the nominal stem *Khoe*, 'person, human being' preceded by the adjectiveally-used root in the sense of 'human', meaning 'true human being'. From the time the Dutch established a colony at the Cape of Good Hope in 1652, to the disintegration of their societies and their almost complete physical disappearance due to epidemics by the beginning of the following century (Elphick 1977), the Khoekhoe were in very close contact with the settlers and all the travelers calling at the Cape while on their way to the East Indies. This situation, though very short in time, explains the existence of a significant amount of written material on the Khoekhoe living in the vicinity of the Cape in the 17th century and at the beginning of the 18th (Raven-Hart 1971; Fauvelle-Aymar 2002). As for the last part of the 18th century, the process of acculturation of the remaining Khoekhoe groups or individuals is not very well documented, mainly due to the fact that this occurred on the frontier zone, far from the view of officials and foreign callers. Late 18th century travelers did, however, meet remnants of very distant Khoekhoe groups, but generally left but very scarce information.

Not all Khoekhoe societies lived along South African coastal plains within reach of 17th and 18th century Cape colony. By the beginning of the 19th century, some herding groups (such as the Einiqua, Namaqua, and the Korana or !Kora) were found living inland in the middle Orange River valley. Other Namaqua (known as Great Namaqua) groups were also found established in Southern Namibia in the 19th century. How all these Khoekhoe groups were historically related with the Cape Khoekhoe is subject to debate.

Namibia is the only place where a Khoekhoe dialect (namely Nama/Damara or Khoekhoegowab) is still spoken today. Works by specialists are

available in this regard: dictionaries (Krönlein & Rust 1969; Haacke & Eiseb 2002); studies on the phonology (Beach 1938), grammar (Hagman 1977) and tonology (Haacke 1999) of Namibia Khoekhoe. Apart from Korana, which was summarily described just before it became extinct (Wuras 1920; Meinhof 1930; Engelbrecht 1936) — and which some specialists consider a different language, no other Khoekhoe dialect has been studied to a significant extent. Particularly striking is the discrepancy between the very few studies of Cape Khoekhoe and the fair quantity of linguistic material left by travelers and settlers: a work such as that of Peter Kolbe (1719), for instance, is still considerably overlooked even though it contains a list of hundreds of Khoekhoe words translated into German and Latin, to say nothing of its first-range historical and ethnographical value. The very small number of copies of this book available in public libraries, the fact that no modern edition of the book exists, in addition to the fact that it was written in German and that the translations available in other languages (French, English and Dutch) are considerably biased, explain for a large part why this work is underused or misused. As for modern studies on the Cape Khoekhoe language, one must mention G.S. Nienaber's *Hottentots* (1963), which is an Afrikaans/Khoekhoe dictionary referring, under each entry, to all the relevant historical sources. Although very useful, this work suffers from several flaws. First, no critical analysis was made of the published evidence, leading the author to use unchecked versions (Kolbe's book, for instance, is used in its Dutch edition of 1727, displaying huge divergences with the original text, concerning the way to render the clicks in particular). Second, since there is a great divergence in the written forms found among the sources, Cape Khoekhoe words are generally reconstructed after the Nama or Korana languages, especially in the determining of clicks. Almost necessarily, such a method artificially increases the relatedness between various dialects and conceals possible regular sound shifts, thus seriously minimizing its value both on linguistic and on historical grounds. One can also suspect that Christopher Ehret's lexicostatistical and glotto-chronological studies of Khoekhoe dialects/languages (eg. Ehret 1982) are based on Nienaber's work, or possibly on a narrower sample of equally-biased sources, such as Kolbe's Dutch edition. Thus, apart from the question of the intrinsic validity of the method used, one can wonder whether Ehret's historical hypothesis shares the same flaws as the sources on which it is based.

These and other examples should be strong incentives not to go too far into comparative linguistics on Khoekhoe language, as long as the basic needs of properly edited linguistic materials are not fulfilled (Fodor 1975: 157-161 *et pas-*

sim). The following is a first attempt by the author, who is not a linguist but a historian, to provide linguists with such material.

The Brenthurst Library in Johannesburg (South Africa) holds the papers of Colonel Robert Jacob Gordon (1743-1795), the once famous commandant of the Dutch *Vereenigde Oost-Indische Compagnie* garrison of the colony of the Cape of Good Hope.[1] During his sojourn in the Cape (1777-1795), Gordon made several discovery trips inland and behind the frontier of the Colony.[2] He corresponded with philosophers and naturalists in Europe, sending them written descriptions and specimens of minerals, plants or animals, and was the main informant of many foreign travelers of the time. But, due to the fact that he committed suicide when the British took over the Cape in 1795, his extensive travel diaries and other papers, some of which were obviously being prepared for publication, remained unpublished and their importance remained virtually unknown. Brought back to England by his widow, Gordon's papers passed through various hands in the course of the 19th century, until the collection was finally divided in two parts. The drawings and maps, consisting of six volumes collectively known as the "Gordon Atlas", were bought in 1914 by publisher Martinus Nijhoff (The Hague, Netherlands) and were afterwards filed at the Rijksprentencabinet in Amsterdam (Raper & Boucher 1988: 11-13). Despite their being of considerable interest, they remain for the most part unpublished. An incomplete photocopy of the maps and drawings is said to have once been available at the Museum Africa (formerly Africana Museum) in Johannesburg (H.G.O. 1948) but has not been recovered; another partial photocopy of the maps and drawings is hosted in the Cape Archives.[3] The other papers of the original collection of manuscripts were thought to be lost until 1964 when they re-appeared in the Staffordshire County Record Office (England). They were purchased in 1979 by the Brenthurst Library in Johannesburg, where they are registered under the number MS 107/1 to MS 107/18. A photocopy of the travel diaries (known as "Gordon's Dagboek") is

---

[1] Robert Jacob Gordon's grandfather had come to the Netherlands from Scotland, hence Gordon's Scottish name. But Gordon was all through his life a loyal Dutch soldier attached to the family of the Prince of Orange. That he may have displayed favorable feelings towards the Orangists and the British at the time of the French Revolution and invasion of the Netherlands could have placed him in a difficult personal situation by the end of his career, and may have led to his end (see below). On Gordon's life, see Cullinan (1992).

[2] Gordon had previously spent ten months at the Cape in 1773-4. It is when back in his homeland that he met the French philosopher Denis Diderot (Cullinan 1989).

[3] National Archives of South Africa. Cape Town Archives Repository. M1/138 to M1/147, M1/1105 to M1/1132, M1/3201 to M1/3210, AG7146/1 to AG7146/223B.

held in the Cape Archives.[4] Almost all papers left by Gordon are written in Dutch, but some documents (most probably intended to be transmitted to European correspondents) are written in English or in French.

A number of researchers have been working on these papers for the last three decades or so, leading to the publication of a meticulously-edited English translation of the four travel diaries of Gordon in South Africa (Raper & Boucher 1988),[5] portions of his letter to Hendrik Fagel, a prominent state clerk in the house of Orange (Smith & Pheiffer 1994)[6] and a draft essay entitled in French "Particularités relatives à quelques hordes hottentottes" (*Particular information on some Hottentot hordes*) (Smith & Pheiffer 1992).[7] A considerable amount of work remains to be done on these archives, which are of prime interest for the history of the Cape colony, the history of natural sciences and the ethnography of Southern African peoples, of whom Gordon was a very acute and sympathetic observer.

The essay entitled "Particularités relatives à quelques hordes hottentottes" consists of field notes that were then intended for publication at a later stage, which came to nothing. But since some were apparently incorporated into the diary of Gordon's third or fourth journey (leading him along the Orange River that was then the border of the Cape colony — and is today the border between South Africa and Namibia) between June 1779 and January 1780, one can ascribe these observations to that time, that area (today Northern Cape) and the people who then inhabited that area, namely the "Little"-Namaqua[8] (Smith & Pheiffer 1992: 5-6). The very short wordlist contained among the same set of papers and referring to the name of the months can thus be attributed to the same people (Smith & Pheiffer 1992: 40-41).

---

4 National Archives of South Africa. Cape Town Archives Repository. VC 592 to VC 598.

5 Another translation (by Patrick Cullinan) of Gordon's travel diaries, accompanied by the Dutch transcription, has been made available on the Internet, at: http://web.uct.ac.za/depts/age/people/Gordon/.

6 The unabridged English translation by R.H. Pheiffer (without editor's notes) is to be found on the Internet, same website: http://web.ucu.ac.za/depts/age/people/Gordon/gordon_letter.htm (accessed in July 2005).

7 Edition of the Dutch with an English translation annotated with ethnographical comments. The titles given to the folders are said not to be from Gordon's hand, but rather from his wife's, who was born in Switzerland, and who probably tried to put the papers in order after his death.

8 As opposed to the "Great"-Namaqua of Great Namaqualand in Namibia.

There is plenty of other linguistic material in the Gordon papers. Most noticeable is a blue folder entitled (again in French) "Collection de mots des divers langages de l'Afrique méridionale" (*Collection of words from various languages of Southern Africa*).[9] It contains eleven documents that, for the most part, consist of lists of Bantu words from South African, Angolan and Mozambican languages, with their translation into Dutch.[10] Among these documents are also found four pieces containing material in Khoekhoe languages. They are (in archival order):

— MS 107/10/1 [hereafter list A]. 1 sheet, width 210 × height 330 mm. Recto. No title. From the general aspect of these notes, one may infer, without any definitive evidence, that they are original field notes. From the comparison between the names of the two Oliphants (Afrikaans: Olifants) Rivers (probably the one that flows into the Atlantic Ocean 250 km north of Cape Town and the one that runs in the Little Karoo) (see Raper 2004: 288), and from the mention of the Gouritz River (that flows into the Indian Ocean to the west of the modern town of Mossel Bay), it can be assumed that this short collection of words was gathered during a voyage when he visited both areas, possibly during the second voyage between August 1778 and January 1779 (Raper & Boucher 1988: 182). In any case, this list is likely to be ascribed to the period between October 1777 and January 1780, for the accounts of the three trips made during this short period display the same way of rendering clicks with diacritical signs upon consonants. Except for the name of the North Olifants River, the list A is composed of river names and other (apparently local) information concerning the Little Karoo.

— MS 107/10/5 [hereafter list B]. 1 sheet, w275 × h432. Recto. Title: "Enige hottentots woorden" (*Some Hottentot words*). Given the neat

[9] Brenthurst Library, Gordon papers. MS 107/10.

[10] MS 107/10/2 (A few scattered notes on "Caffers" language); MS 107/10/3 (Xhosa wordlist? Various notes and calculations); MS 107/10/4 (Xhosa wordlist); MS 107/10/6 (two lists of Angola/Congo languages: 1. "Naam van het land van een swarte slaaf die te Loando St. Paúlo door de Portugesen aan de Fransen verkogt is. Lakombo was syn eigen naam en syn vaders naam Kongo en syn moeders Dampi. Syn land hiete Quacombo...". 2. "Naam van het land van myn grote swarte Cúpido Gakinge door syn anders genaamt"); MS 107/10/8 (Information on the Maratjsa, "een sort van Mackúas", or "swarte Caffers" who live around the tropic of Capricorn. Wordlist of the "Maratjsa Caffer taal"); MS 107/10/9 ("Madagascar taal" apparently collected from several Malagasy slaves); MS 107/10/11 (Tswana wordlist: "Moetjsonana taal. Die de Hottentotten Briquás noemen").

aspect of the document and of the actual writing, it cannot be considered as field notes but as a carefully re-written document. Minor additions (including the title) and corrections, possibly not from the same hand, have been made with a different (green) ink that does not alter significantly the reading. No external element makes it possible to attribute the collected words to any area in particular or to any period of time spent by Gordon in the field. However, the use of a completely different system to render the clicks (by a superscript *t* before some consonants) is a strong argument toward a different period of time, rather later than the previous one. This new system is not used in Gordon's traveling accounts, but it can *ex silentio* be considered coeval with his fourth voyage between November 1785 and March 1786, for he does not make use of the previous click system in his traveling account. This fourth and last voyage to the border of the Colony took Gordon to Algoa Bay through the great Karoo and the Baviaanskloof, and then back to Cape Town along the coast (Raper & Boucher 1988: 386). One can also notice that the word *Toe* ('Dog') is only attested in Eastern Khoekhoe (Nienaber 1963: 308), but linguists are left to decide whether this and other possible clues are decisive or not.

— MS 107/10/7 [hereafter list C]. 1 sheet, w383 × h225. Recto. No title. As for list A, it appears to be field notes recorded between 1777 and 1780 (but not on the same paper nor with the same ink). From the name of a Seacow River, which could be the Seekoei River — tributary of the Orange river visited several times by Gordon — but which in fact is the Dutch translation of a Khoekhoe name corresponding to the Bushmans River (which flows into the Indian Ocean to the East of Algoa Bay), it can be proposed that the entire list was collected when Gordon was present in the lower valley of this river for the first time in January 1778 (Raper & Boucher 1982: 140 sq). On that occasion he met with a Khoekhoe chief named Ruiter, "who has Gounaquas and Bastard Hottentots under him" (Raper & Boucher 1982: 140). Gordon spent some time drinking sour milk with the chief, who provided him with a guide to take him to the sea. Since it is likely that it is from this guide that Gordon collected the Khoekhoe name of the river (Raper & Boucher 1982: 141), the same can be asserted for the rest of the list. If this is so, this piece of Khoekhoe vocabulary is certainly one of the most informative we have concerning the Gonaqua/Hoengeiqua of the area. The list is not a word list but rather a list of sentences.

— MS 107/10/10 [hereafter list D]. 1 sheet, w206 × h243. Recto. No title. Fieldnotes. There is very little internal or external evidence to date this document, except the fact that Gordon again uses the superscript *t* to note clicks, which could be an argument to attribute this list to the same area and time as list B.

In the following transcription, Gordon's spelling, punctuation, use of small or capital letters, underlining, and use of diacritical signs have been scrupulously respected. *Idem* for the order of Khoekhoe/Dutch entries. Strikeout represents Gordon crossing letters or words out. The sign / (slash) is used here to indicate that words presumably part of one sentence or one comprehensive set of words are not placed on the same line in the manuscript. My interventions are limited to the placing of words in two columns when it proved possible and useful. I have also used italic font to signify Khoekhoe words or sentences. Wherever it proved necessary to restitute part of a word or to mention alteration or uncertain reading, I did it between square brackets [...]. In the footnotes, I give the English translation of the Dutch words and various other comments, and I use the sign *** to replace any Khoekhoe word which does not need to be repeated.

List A
Little Karoo Khoekhoe? ca. 1777-80

Ćara Ćamma / noord / Oliphants / rivier[11]
Sneirwig, /sort van / bosjes[12]
Oost Oliphant *kwàcas*[13]
Ñari; bloem ook *ǵa.* / *ñuriqua*[14]
verkeerde valey[15]

[11] Previously written *Ćàra.* North Olifants River (today Olifantsrivier), as opposed to the East Olifants River (see below), is probably the one flowing northward from Tulbagh to Papendorp in the Western Cape. Raper (2004: 288) indicates that the Khoekhoe name of that river was *Tharakkama* or *Trakamma*, which is consistent with the name given by Gordon.

[12] Sneirwig, sort of bushes. Sneirwig: reading uncertain. Bosjes: *b* superimposed on a *g*.

[13] East Olifants. Obviously the modern Olifants River that becomes the Gourits River and flows in the Indian Ocean near Mossel Bay.

[14] *** flower, also ***.

[15] Wrong valley.

k̀eina ćamma of mooy water[16]
en *k*[...[17]] *ćamma* is Oliphan[ts] / riv[ier] by gourits.[18]
Ćobeeb hexerivier.[19]
het casteel hiet / *kúi keip* of klipkraal.[20]
Sonder klap[21]
<u>*Toa Togoú*</u> de naam van Captein kees vader. besokend[22]
Sÿn goed sal niet vergaan.[23]
h̀ei ónna sonder naams[24]
is de naam van Captein kees[25]
Càmteep was ook myn naam[26]
hiet pluimdrager[27]

List B
Eastern Cape Khoekhoe? ca. 1785-6?

Soré	Zon[28]
ᵗCaan	Maan[29]
ᵗCa caan	Sterren[30]
hoeri / O ᵗCamma	Zeewater[31]

---

[16] *** or nice water.

[17] Writing unclear: possibly *kwúni*.

[18] And *** is Olifants River near the Gourits.

[19] Hex River, in the region of Worcester (Western Cape). See Raper (2004: 136).

[20] The castle [of Good Hope?] is called *** or stone kraal.

[21] Without click. Refers probably to the pronunciation of the words below, which are underlined.

[22] *** the name of Captain Kees' father. Visited.

[23] His goods will not disappear.

[24] *** without name.

[25] Is the name of Captain Kees. Probably follows the name above.

[26] Was also my name.

[27] The feathers-bearer. Obvious translation of the above word; compare to nama *!ami*, *!ammi*, 'feather' (Nienaber 1963: 493 'veer'; Haacke & Eiseb 2002). This name is likely to refer to Gordon's military hat.

[28] Sun.

[29] Moon (= month?).

[30] Stars.

[31] Sea water. *Hoeri* is a later addition.

Goedì	Schaap[32]
ᵗCaune	Lam[33]
ᵗkũi	Een[34]
Tam	Twee[35]
ᵗNonà	Drie[36]
Hakka	Vier[37]
Gissi	~~Vÿff~~ Tien[38]
ᵗKwaab	Oliphant[39]
ᵗNabab	Rinoster[40]
Jees-Owa	God[41]
Héba ha mi com	Spreeken[42]
Sase qũoi qũoi / mienqua kanje	Hottentots spreeken[43]
Areti ᵗ'ja	Gÿ Zijt mÿn vriend[44]
Sase ᵗca ᵗcouha	Wilt gÿ drinken[45]
ᵗza am 'li he Sats	hoe vaart gÿ[46]
Tiri ᵗkai	Ik ben wel[47]
Goere	Boog[48]
ᵗQũoũqũe	Man[49]

---

[32] Sheep.

[33] Lamb. Possible dot on u.

[34] One. Gordon had previously written *ᵗgũi*.

[35] Two. Gordon had previously written *Cam*.

[36] Three.

[37] Four.

[38] Ten. Gordon had previously written *Gisi* (reading uncertain).

[39] Elephant.

[40] Rhinoceros.

[41] God.

[42] Speak. Gordon had previously written *Héwa*.

[43] Speak Hottentot.

[44] You are my friend. Gordon had previously written *ᵗ'da*.

[45] Will you drink.

[46] How are you.

[47] I am well.

[48] Bow.

[49] Man.

Tiri ᵗarun'a	Ik heb honger[50]
Sase	Gÿ[51]
Tiri ᵗcan ᵗcamme ᵗca	Ik heb dorst[52]
ᵗCaini awei ᵗcamma	Dat is lekker[53]
Ham ᵗcobebe ᵗcobe / ha haha, ᵗnoe casi	een liedje[54]
ᵗCobe Licanse	Mÿn land is vet[55]
Hebati robe a / hala, hammosfi	~~halen~~
ᵗCoebe ahaha	Waar is ũw land[56]
~~*(Ti|kúis)*~~	myn vroũw *keus* vrouw ook *tera* of *tra kuis*[57]
ᵗCain si ᵗca ᵗcali ᵗcamma	Ik drink ũ gesondheÿd[58]
Jesowa ᵗcam Siszeme/ ᵗconge	God zÿ met ũ[59]
Hese ᵗkein ᵗcoeba	T'is mooÿ weer van dag[60]
Hese ᵗkein ~~tite~~ tite ᵗCoeba	Geen goed weer[61]
Toekei	Regen[62]
Tera|kuis	Vroũw[63]
ᵗQuouque	Man[64]
Bi ᵗa	Het Hooft[65]
Mouqua	Oogen[66]

---

[50] I am hungry.

[51] You.

[52] I am thirsty.

[53] That is nice.

[54] A song.

[55] My land is fat.

[56] Where is your land.

[57] *** my wife, *** wife, also *** or ***. Gordon had previously written *Tikois*.

[58] I drink (to) your health.

[59] God is with you.

[60] It's fine weather today.

[61] Not fine weather.

[62] Rain (noun and verb).

[63] Woman, wife.

[64] Man.

[65] The head.

[66] Eyes.

ᵗNanqua	Ooren[67]
hoeri ᵗkei	Zee[68]
ᵗCam ᵗna	Mond[69]
ᵗCoong	Tanden[70]
ᵗa	Voeten[71]
ᵗoenqua	hand[72]
Tora	Schamelheyd der Mans[73]
ᵗNeub	Schamelheyd van een / vrouw[74]
Tamma	Borst[75]
Samma	Prammen[76]
Domma	Keel[77]
ᵗGouti	Singen[78]
ᵗNour ~~à~~	~~Knife~~ nes[79]
ᵗCamseũwè	T'is waar[80]
aaan ᵗa	Ja[81]
aaan te	Neen[82]
Toe	dog[83]
~~On~~ ᵗ~~quaw~~ [Sahs] ᵗ'onha maré	hoe is ũw naam[84]
Cour-Cour	hen hoen[85]

---

[67] Ears.

[68] Sea. Gordon had previously written *goeri*.

[69] Mouth.

[70] Teeth.

[71] Feet.

[72] Hand.

[73] Men genitals (literally: shamefulness).

[74] Genitals of a woman (literally: shamefulness).

[75] Breast, chest.

[76] (Woman) Breasts.

[77] Throat.

[78] Sing.

[79] Sic: *mes*. Knife. Note English word used at first.

[80] It's true.

[81] Yes. Possibly dots above second and third *a* (reading uncertain); possibly *aaan ta*.

[82] No.

[83] Dog. Note English word used instead of Dutch *Hond*. Compare with *tu*, *tuna* in Eastern Cape Khoekhoe (Nienaber 1963: 308, "Hond IV").

[84] What is your name?

List C
Eastern Cape Khoekhoe (Gonaqua/Hoengeiqua)? 1778?

Nantroo.	Landrost[86]
Ǹoandi Ǹobi	ganse valey[87]
Tauwn	fontein[88]
demaatse[89] *ho keúi*	waar leg je nú[90]
de éi saatse ha be	wanneer ga je weer weg[91]
de éi saatse co hema ba	wanneer ben je hier gekomen[92]
Ǹaamoe	een bloem[93]
Kauwgha	hiet zeekoeirivier is bosjemans rivier.[94]
Boo[[95]]*éss*	hiet de fontein aan bosjesmans riviers mond.[96]
dieti tse na ǵsouwa ǩei	hoe hiet die fontein[97]
dieǩa hema tse ńati mi	waarom seg zÿ so[98]
Goinka tse c̄aa ǵau ha	wil je een glas wÿn hebben[99]
Tabaca ha tse 'ǵaú ha	wil je een stúk tabak hebben.[100]

---

[85] Hen. Notice English word (initially spelled hem) used at first.

[86] Landdrost (magistrate).

[87] Geese valley.

[88] Fountain, spring.

[89] Reading uncertain.

[90] Where do you lay (stay) now?

[91] When are you going away again?

[92] When did you come here?

[93] A flower.

[94] ***, that is Seacow River, is the Bushmans River. Compare with *Ǹaugha*, "hippopotamus river" *in* Raper & Boucher (1982: 141).

[95] One or two letters added, unreadable.

[96] *** is the name of the spring at the Bushmans River mouth.

[97] How is the spring called?

[98] Why do they say so?

[99] Will you have a glass of wine?

[100] Will you have some tobacco?

ik heb ~~dat os gehad~~ een os	*tiri go ǩa ha.*[101]
ik sal een os krÿgen	*tiri ha goa ho*[102]
ik heb een os	~~tiri~~[103]
ik heb gisteren een os gehad	*ǩoaase*[104] *tiri| ko| go ǩa ha*[105]
ik denk dat het waar is	*Camma se ti tiri ei.*[106]
Het is waar	*Camma saawse*[107]
Ik heb van daag geen os	*he see tiri go ka hátse.*[108]
~~Het is geloogen~~ Jÿ liegt	*gǹ ǸŌoo saatje ma hie owi*[109]
Ik sal daar na toe gaan	*tiri ha ńaatse ǸŌon.*[110]
Sal jÿ daar na toe gaan	*Saatse ha ńaatse ǸŌon.*[111]
hoor jÿ mÿ niet	*Saatkena tiri ńauwte.*[112]
Waarom lag zÿ	*die*[...[113]] *éimatse ǩei*[114]
Waarom huil zÿ.	*die éimatje ǩyn*[115]
Sagjes wat.	*ǸCaboose.*[116]
Saats keintse[117] ~~*ma*~~ *keúi*	Je moet daar mooÿ vredig leggen[118]

[101] I have an ox.
[102] I shall receive an ox.
[103] I have an ox.
[104] Previously *goaase*.
[105] I had an ox yesterday.
[106] I think that it is true. Last Khoekhoe word uncertain.
[107] It is true. Last Khoekhoe word uncertain.
[108] I have no ox today.
[109] You sing.
[110] I shall go there.
[111] Will you go there?
[112] Do not you hear me?
[113] One possible letter unreadable.
[114] Why do they laugh?
[115] Why do they cry?
[116] Slowly! (interjection)
[117] Possibly accent on k.
[118] You must lie very peacefully there (I wish you a very peaceful rest there).

List D
Attribution uncertain, ca. 1785-6?

^{ts}Kormoutstì[119]	goeden avond[120]
^{t}Kwamoutsi[121]	goeden morgen[122]
bareb	brood[123]
^{t}eip	vuur[124]
tabete	goeden dag[125]
^{t}camma	water[126]
^{t}ca	drinken[127]
arikn[128]	hond[129]
issa	mooy[130]
ti ^{ts}com ^{t}ym	een man[131]
Kõiess	een vrouw[132]
Kõieb	een man[133]
^{t}oãm	hand[134]
^{ts}tam	mooy weer[135]
gamma	leew[136]
^{t}houkha	een wolf[137]

---

[119] Letters *or* uncertain.
[120] Good evening.
[121] Letter *w* uncertain.
[122] Good morning.
[123] Bread.
[124] Fire. Last letter of Khoekhoe word uncertain.
[125] Good day. Possibly *^{t}abete*.
[126] Water.
[127] Drink.
[128] Last letter of Khoekhoe word uncertain (but not u).
[129] Dog.
[130] Nice.
[131] A man.
[132] A woman.
[133] A man.
[134] Hand.
[135] Nice weather.
[136] Lion.
[137] A wolf (hyena?).

ha Xatsi ᵗga	kom hier vriend[138]
g[...]s beep	beeste melk[139]
ᵗei kwakou	maak vuur[140]
danna ka tse he [...]rogadaw	waar is het wagenpad[141]
ka tse moe he comma	wyst my die plaats[142]
ᵗgoukweis	je bent myn vriend[143]

The importance of these lists is obvious. Not only can they document Cape Khoekhoe lexicon and toponymy, but also increase our knowledge on some aspects of the grammar, which was hardly the case for any other Khoekhoe wordlist compiled prior to one century ago. They are also evidence that, contrary to what is sometimes assumed, the Cape Khoekhoe language was still frequently used by the end of the 18th century, not only on the remote border of the colony but also in areas, such as the Little Karoo, that had long been confiscated by white farmers. Possibly, these lists could also help document regional diversity in the Cape Khoekhoe language, and thus give credit or not to the generally-admitted distinction between "Cape Khoekhoe" proper and "East Cape Khoekhoe" (see Kaap vs Oos *in* Nienaber 1963), or Western, Central and Eastern Cape Khoekhoe (Elphick 1977: *passim*), which are admittedly purely contextual or arbitrary classifications (eg Elphick 1977: xvii).

Of interest here is the range of special and diacritical signs used by Gordon to render the various influxes and effluxes that make up clicks, and possibly other characteristics of the language. Thus, the tilde (~) is mainly used on the vowel *u*, possibly to represent the sound /u/ in Khoekhoe, in order to differentiate it from the sound /y/ common in Dutch (eg. vuur (fire), /vy:r/). But this use does not appear to be systematic (even in Gordon's Dutch, where it is erratically used), a feeling that is reinforced by the fact that the Khoekhoe word for 'man' appears twice in the list B, once with and once without a tilde on the *u*. The tilde also appears in list D on the vowels *i* and *a*, though the alteration introduced is unclear. Similarly, accents are used on certain vowels, but it also remains unclear whether these are used to modify their value or to mark length, nasalization or tone. In two

[138] Come here (my) friend.

[139] Animal milk.

[140] Fire making.

[141] Where is the wagon path?

[142] Show me the place.

[143] You are my friend.

cases Gordon seems to use a subscript vertical hyphen that could also mark the tone.

Many systems have been used among travelers and scholars until a more or less standardized system was eventually adopted (e.g. Nienaber 1963: 162-163; Köhler 1981: 472-5). Gordon apparently invented two completely different systems for his own use. The first one, as displayed in lists A and C (corresponding to the years 1778-9), makes use of a range of accents placed on the first (or in one case on the second) consonants at the beginning of the words. The combinations thus produced are as follows: Ć, ć, C̀, c̄, G, ǧ, ȟ, ḱ, k̀, Ñ, ñ, ń, ǹ, t́, ẁ. If one considers that small and upper capitals are possibly insignificant variations and that the accent on w may be aberrant, we are left with six consonants able to bear accents: C, G, H, K, N and T. The second system, as displayed in lists B and D (corresponding to the year ca. 1785), makes use of a superscript *t* (ᵗ), that, once again, always immediately precedes six possible consonants (C, H, K, N, Q, Z). Only C, H, K and N are found in both systems. The superscript *t* is employed twice before the vowel *e* (in both cases it refers to the word *ᵗei*(*p*), 'fire') and once before the vowel a. In one instance, the superscript *t* is followed by an apostrophe and, in another, the two letters *ts* are superscript.

After a mere comparison of the word 'water' written in the two systems (lists A and B), one can suggest that the sign Ć corresponds to ᵗC and that both of them are used to mark the click *//g* (Nama orthography). But only a systematic comparison between the two systems and a sound comparison between the data provided by Gordon's wordlists and other collections of Khoekhoe words could eventually allow reconstructing at least part of the Cape Khoekhoe dialect. Here stops the work of the historian, knowing that he will probably benefit later from the work of the linguist.

REFERENCES

Beach, Douglas M. 1938. *The Phonetics of the Hottentot Language*. Cambridge: W. Heffer & Sons.

Cullinan, Patrick. 1989. "Robert Jacob Gordon and Denis Diderot: The Hague, 1774". *Quarterly Bulletin of the South African Library* 43 (4): 146-152.

Cullinan, Patrick. 1992. *Robert Jacob Gordon, 1743-1795: The Man and his Travels at the Cape*. Cape Town: Struik.

Ehret, Christopher. 1982. "The First Spread of Food Production to Southern Africa", *The Archaeological and Linguistic Reconstruction of African History* ed. By Christopher Ehret and Merrick Posnansky. Berkeley: University of California Press. Pp. 158-187.

Elphick, Richard. 1977. *Kraal and Castle. Khoikhoi and the Founding of White South Africa*. New Haven: Yale University Press.

Engelbrecht, Jan A. 1936. *The Korana*. Cape Town : Maskew Miller.

Fauvelle-Aymar, François-Xavier. 2002. *L'invention du Hottentot. Histoire du regard occidental sur les Khoisan (XVe-XIXe siècle)*. Paris : Publications de la Sorbonne.

Fodor, István. 1975. *Pallas und Andere Afrikanische Vokabularien vor dem 19. Jahrhundert. Ein Beitrag zur Forschungsgeschichte*. Hamburg: Helmut Buske.

Greenberg, Joseph. 1966. *The Languages of Africa*. Bloomington: Indiana University.

Güldemann, Tom & Vossen, Rainer. 2000. "Khoisan". *African Languages: An introduction*, ed. by Bernd Heine and Derek Nurse. Cambridge: Cambridge University Press. Pp. 99-122.

Haacke, Wilfrid H.G. 1999. *The Tonology of Khoekhoe (Nama/Damara)*. Cologne: Rüdiger Köppe.

Haacke, Wilfrid H.G. & Eiseb Eliphas 2002. *A Khoekhoegowab Dictionary with an English-Khoekhoegowab Index*. Windhoek: Gamsberg Macmillan.

Hagman, Roy S. 1977. *Nama Hottentot Grammar*. Bloomington: Indiana University.

Herbert, Robert K. 2002. "The Sociohistory of Clicks in Southern Bantu". *Language in South Africa*, ed. by Rajend Mesthrie. Cambridge : Cambridge University Press.

Köhler, Oswin. 1981. "Les langues khoïsan", *Les langues dans le monde ancien et moderne*, ed. by Jean Perrot. Paris : Éditions du CNRS. Vol. 1, pp. 457-615.

Kolbe, Peter. 1719. *Caput Bonae Spei Hodiernum, das ist Vollständige Beschreibung des Africanischen Vorgebürges der Guten Hofnung*. Nuremberg: Conrad Monath.

Krönlein, Johann G. & Rust F. 1969. *Nama Wörterbuch*. Pietermaritzburg: University of Natal Press.

Meinhof, Carl. 1930. *Der Koranadialekt des Hottentottischen*. Berlin: Dietrich Reimer.

Nienaber, Gabriel S. 1963. *Hottentots*. Pretoria: J.L. Van Schaik Beperk.

O.[...], H.G. 1948. "The Gordon Drawings". *Africana Notes and Records* VI (1): 17-19.

Raper, Peter E. 2004. *New Dictionary of South African Place Names*. Johannesburg & Cape Town: Jonathan Ball Publishers.

Raper, Peter E. & Boucher Maurice. 1988. *Robert Jacob Gordon: Cape Travels, 1777 to 1786*. Johannesburg: Brenthurst Press, 2 vols.

Raven-Hart, Roger. 1967. *Before Van Riebeeck. Callers at South Africa from 1488 to 1652*. Cape Town: Struik.

Raven-Hart, Roger. 1971. *Cape Good Hope 1652-1702: The First Fifty Years of Dutch Colonisation as Seen by Callers*. Cape Town: A.A. Balkema, 2 vols.

Smith, Andrew B. & Pheiffer Roy H. 1992. "Col. Robert Jacob Gordon's Notes on the Khoikhoi 1779-80". *Annals of the South African Cultural History Museum* 5 (1): 1-56.

Smith, Andrew B. & Pheiffer Roy H. 1994. "Letter from Robert Jacob Gordon to Hendrik Fagel, 1779". *Brenthurst Archives* 1 (2): 29-46.

Wuras, C.F. (ed. by Walther Bourquin). 1920. *Vokabular der Korana-Sprache*. Berlin: Dietrich Reimer.

French Institute of South Africa
PO Box 542
Newtown 2113 Johannesburg
Republic of South Africa
francois-xavier@ifas.org.za

[received February 28, 2006
accepted April 6, 2006]

Studies in African Linguistics
Volume 34, Number 2, 2005

INFORMATION STRUCTURING IN AKAN QUESTION-WORD FRONTING AND FOCUS CONSTRUCTIONS*

Charles Marfo and Adams Bodomo
University of Hong Kong

Wh-question fronting and focus constructions in Akan have three structural characteristics in common: constituent fronting, introduction of a clitic morpheme after the fronted constituent, and pronoun resumption in a canonical clause position. In comparing these constructions to each other and to related canonical constructions, one is confronted with the question whether the same discourse-contextual information is consistently expressed in both constructions. Using the framework of Lexical-Functional Grammar, we show that both *wh*-question fronting and focus constructions share representations in the constituent and functional structure. Considering the individual discourse-contextual information expressed in *wh*-question fronting and focus constructions, as compared to the discourse-contextual information expressed in the respective *in situ* and canonical clause counterparts, however, we show that a variance is drawn between them in the information structure. In a further constraint-based analysis, Optimality-Theoretic LFG is used to clarify the proposals made.

1. Introduction.

In this paper, we discuss *wh*-question fronting and (contrastive) focus constructions (formally noted as marked sentence-types) and other facts that are related to

* This paper has benefited from comments and discussions with a number of people at different fora. We would like to thank participants at the LFG2004 conference held in Christchurch, New Zealand, especially Tracy Holloway King and Miriam Butt. We are also very grateful to two anonymous reviewers and the editor of this volume, David Odden, for raising many issues that have led to substantial revisions of parts of the paper.

them in Akan, a Kwa language spoken in Ghana and some parts of Côte d'Ivoire. Three features characterize *wh*-question fronting and focus constructions in Akan: fronting of a constituent, introduction of a clitic morpheme after the fronted constituent, and pronoun resumption in a canonical clausal position. In comparing the two constructions to each other and to related non-extracting constructions, the question that one is confronted with is whether the same discourse-contextual information is realized in both constructions. In other words, as compared to related non-extracting constructions, does the resulting phrase structure configuration bring about semantic contrast in both constructions or not? This has been an internal controversy in Akan; specifically, compare Saah (1988) to Boadi (1990).

Using the framework of Lexical-Functional Grammar (LFG: Kaplan & Bresnan 1982, Bresnan 2001, Dalrymple 2001), we explore the similarities and differences between *wh*-question fronting and focus constructions. In this paper we show that in the constituent (c-) structure and the functional (f-) structure, both *wh*-question fronting and focus constructions essentially share common representations. However, considering the individual discourse-contextual information that is expressed in *wh*-question fronting and focus constructions, as compared to the discourse-contextual information expressed in the respective *in situ* and canonical clause counterparts, we show that a variance is drawn between them in the information (i-) structure, which is accessible to the semantic (s-) structure (King 1997, Butt & King 1998). In LFG, c-structure, f-structure, and i-structure respectively model the categorical representation, the grammatical functions, and the discourse-contextual information aspects of the grammar (e.g., see (28)). The LFG account in this paper is novel and, with it, the separate semantic content of Q-word fronting and focus constructions can be explicitly presented through an (attribute value matrix (AVM)-based) i-structure. In a further constraint-based analysis, Optimality-Theoretic LFG (OT-LFG: Bresnan 2000, Kuhn 2001) is used to clarify and strengthen our suggestions.

The rest of the paper is organized as follows: in section 2, we give a descriptive account of *wh*-question constructions in Akan, including its constituent *in situ* and constituent fronting occurrences. The (contrastive) focus construction in Akan is then described in section 3. In sections 4 and 5, we explain how the two constructions are similar to, or different from, each other and throw light on the intricacies involved in their constructions within LFG. With insights from OT-LFG, section 6 illuminates the discussions in sections 4 and 5. Section 7 provides the conclusion to our observations and analyses.

2. *Wh*-Question Constructions.

A *wh*-question construction in Akan is primarily identified by any of the interrogative phrases or pronouns in (1). Following Boadi (1990), we refer to the interrogative pronouns in (1) as question words or question phrases (hereafter, Q-words/Q-phrases). As discussed in sections 2.1 and 2.2, each of the Q-words can remain *in situ* in a canonical clause or fronted in an extra-sentential clause.

(1)	hwáń / hwáánóḿ	'Who / which people'
	sɛ́n	'How much, how many or what'
	á'dɛ́ń / (sɛ́) dɛ́ɛ́ń /á'dɛ́ń (ńtí)	'Why / for what reason'
	ɛ̀hé(é'fá)	'Where'
	(ɛ̀)dɛ́ɛ́ń / (ɛ̀)dɛ́ɛ́bɛ́ń	'What'
	bŕɛ́bɛ́ń / dàbɛ́ń	'When'
	NP + bɛ́ń	'Which (of that item)'

2.1 Q-word *in situ*. The Q-words are substitutes for the various syntactic categories, particularly the argument functions. Therefore, as illustrated in (2b) and (2c) for the subject and the object respectively, these Q-words can remain *in situ* in a canonical clause; i.e., as substitutes for the constituents they question. When the verb is questioned in the *in situ* representation, as shown in (2c), it is replaced by another verb, *yɛ*, literally meaning 'do.' In addition, the Q-word occurs in the final position.

(2) a. Kùsí rè-sèré àbòfŕá nó 'Kusi is laughing at the child.'
Kusi PROG-laugh child DEF

b. *Hwáń* rè-sèré àbòfŕá nó 'Who is laughing at the child?'
Who PROG-laugh child DEF

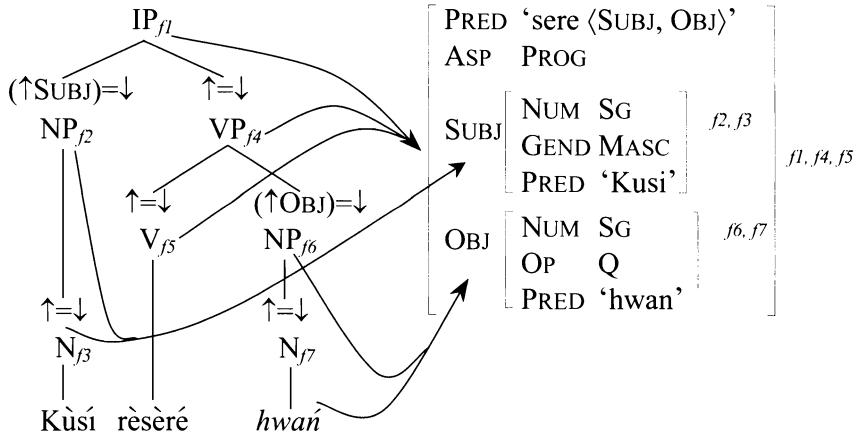
c. Kùsí rè-sèré *hwáń* 'Kusi is laughing at whom?'
Kusi PROG-laugh who

d. Kùsí rè-*yɛ́* àbòfŕá nó *dɛ́ɛ́ń* 'What is Kusi doing to the child?'
Kusi PROG-do child DEF what

The c- and f-structure instantiations of the Q-word *in situ* construction in (2c) are shown in (3) below. The illustration in (3) also shows how c-structure maps to f-

structure through the Structure-Function Mapping theory (Bresnan 2001, Dalrymple 2001, Falk 2001).

(3)



2.2 Q-word fronting. Besides the *in situ* representation of the *wh*-construction in Akan, with which the canonical phrase structure is maintained, there is another option of representation. This option involves the dislocation of the Q-word. Q-word dislocation in Akan refers to the fronting of the Q-word (hence, Q-word fronting) in an extra-sentential construction. A clitic morpheme, *nà*, referred to as a focus marker (FOC) (Boadi 1974, 1990, Saah 1988), is also introduced at the right edge of the fronted Q-word. In other words, as illustrated in (4), an obvious phrase structure variation is realized where the Q-word appears in some position that is above the canonical clause.

(4) a. [IP Kùsí rè-sèré *hwáń*] 'Kusi is laughing at whom?'
Kusi PROG-laugh who

Hwáń$_i$ nà [IP Kùsí ré-séré nó$_i$] 'Whom is Kusi laughing at?'
who FOC Kusi PROG-laugh 3SG

b. [IP Kòfí bé-ˈdúá *déén*] 'Kofi will sow what?'
Kofi FUT-sow what

Déén nà [IP Kòfí bé-ˈdúá] 'What will Kofi sow?'
what FOC Kofi FUT-sow

In a bid to enforce an argument that Q-word fronting in Akan encodes emphatic information, as against the information expressed in a related *in situ* construction, Saah (1988: 19-20) observes with two examples, slightly modified in (5),[1] that some Q-word *in situ* constructions related to greetings are canonically fixed in phrase structure (see (5a)). Thus, according to him, a corresponding Q-word fronting option (see (5b)) is ungrammatical.

(5) a. *Q-word* in situ

[$_{IP}$ Wò	hó	tè	*sέ́ń*]		'How are you?'
2SG	self	be.PRES	how		
[$_{IP}$ Wɔ̀-fr̀ɛ́		wò	*sέ́ń*]		'What is your name?'
3PL-call.HAB		2SG	how		(*lit.*: 'What do they call you?')

b. *Q-word fronting*

Sέ́ń nà	[$_{IP}$ wò	hó	té(ɛ́)]		'How are you?'
how FOC	2SG	self	be.PRES		
Sέ́ń nà	[$_{IP}$ wɔ̀-fr̀ɛ́		wó]		'What is your name?'
how FOC	3SG-call.HAB		2SG		(*lit.*: 'What do they call you?')

Perhaps Saah's observation is true in other dialects of Akan.[2] In Asante-Twi, however, fronting of greetings related Q-words is attested although it is a fact that it is not often done, as indicated in the grammatical constructions in (5b).

Saah (1988) also notes that where a Q-phrase is functioning as an adverbial of reason, it must be fronted obligatorily, as shown in (6a). According to him, the construction is ungrammatical where the Q-word remains *in situ*, as also shown in (6b). While being cautious about the supposed semantic difference between Q-word/phrase *in situ* and Q-word/phrase fronting, he further suggests that the Q-phrase needs to be at a stressed or emphatic position, hence the fronting; specifically, the specifier position of some projected pragmatic/discourse function. As

[1] They are a little modified in the sense that we have used a different Akan text – i.e., the use of *sέ́ń* 'how' (in Asante-Twi), instead of *dέ́ń* in Saah's example.

[2] Akan is composed of several dialects. The prominent ones are Asante-Twi, Fante, and Akuapim-Twi. It seems to us that Saah (1988) was referring to Fante, considering his selection of Akan texts (e.g., the use of *dέ́ń* in Fante instead of *sέ́ń* in Asante-Twi). However, according to our observations, even in Fante, fronting of Q-words is generally acceptable.

will be reiterated in section 5, we claim that a fronted Q-word does not invoke any further emphasis than what it inherently does at an *in situ* position in Akan (from Saah, 1988: 20).

(6) a. *Déɛ̀ǹ àdé ńtí* nà Kwàdwó bɔ̀-ɔ̀ Á!má
what thing because FOC Kwadwo hit-PST Ama
'For what reason/why did Kwadwo hit Ama?'

b. *Kwàdwó bɔ̀-ɔ̀ Á!má *déɛ̀ǹ àdé ńtí*
Kwadwo hit-PST Ama what thing because
'For what reason/why did Kwadwo hit Ama?'

Indeed, it is true that (6b) is ungrammatical, as Saah rightly notes. However, the ungrammaticality is only due to the fact that the whole interrogative phrase (Q-phrase), *déɛ̀ǹ àdé ńtí*, asking for the reason behind the agent's (Kwadwo) action, is incomplete in the present position. The complete Q-phrase should be introduced by the complementizer (COMPL), *sɛ́*, and read as *sɛ́ déɛ̀ǹ àdé ńtí*. The complementizer is optional when the Q-word/phrase is fronted and it is actually related to *ńtí* in the phrasal form, *sɛ́ ... ńtí* 'because'. Observe in (7a) below, the alternative to (6b), that the same Q-word *in situ* construction is grammatical with the complementizer as part of the Q-phrase. As has been noted earlier, Q-words/Q-phrases are only substitutes for canonical clause-internal constituents. So, the Q-phrase in (7a) actually replaces a phrase (of reason), like the one in (7b), which also has to be introduced by the COMPL. Otherwise, as also shown in (7c), the construction is ungrammatical.

(7) a. Kwàdwó bɔ̀-ɔ̀ Á!má *sɛ́* *déɛ̀ǹ àdé ńtí*
Kwadwo hit-PAST Ama COMPL what thing because
'For what reason/why did Kwadwo hit Ama?'

b. Kwàdwó$_j$ bɔ̀-ɔ̀ Á!má$_i$ *sɛ́* *ɔ̀$_i$-à-séré* (*nó$_j$*) *ńtí*
Kwadwo hit-PAST Ama COMPL 2SG-PRF-laugh 3SG because
'Kwadwo hit Ama because she has laughed (at him).'

c. *Kwàdwó$_j$ bɔ̀-ɔ̀ Á!má$_i$ *ɔ̀$_i$-à-séré* (*nó$_j$*) *ńtí*
Kwadwo hit-PST Ama 3SG -PRF-laugh 3SG because
'Kwadwo hit Ama because she has laughed (at him).'

In addition, as we have suggested and will be discussed in detail in section 5, the *in situ* construction in (7a) conveys the same discourse-contextual information that is expressed in the case of Q-phrase fronting construction in (6a). In other words, no semantic contrast obtains between (6a) and (7a).

3. Focus Construction.

A focus construction in Akan has a "point of prominence" within it (Boadi 1974) where contrastive information (of exclusivity) is intentionally placed for the purpose of emphasis. A constituent is contrastively focused in Akan when it is fronted in its extra-sentential projection of focus phrase (FocP). The (fronted) constituent in focus is also immediately followed by the Foc, *nà*. Specifically, the Foc appears at the head position of the projected FocP, as shown in (8). Also, observe in (8b) that when the sentential head is rather the constituent in focus, the same form of the verb-stem remains *in situ*. This is unlike the case of a questioned sentential head where *yɛ* is, instead, introduced in the canonical base position; see (2d).

(8) a. Kòfí rè-bòá Á'má
Kofi PROG-help Ama
'Kofi is helping Ama.'

b. [$_{FocP}$ Bòá$_i$ *nà* [$_{IP}$ Kòfí ré-bóá$_i$ Á'má]]
help Foc Kofi PROG-help Ama
'It is help (that) Kofi is helping Ama.'

c. [$_{FocP}$ Á'má$_i$ *nà* [$_{IP}$ Kòfí ré-bóá nó$_i$]]
Ama Foc Kofi PROG-help 3SG
'It is Ama (that) Kofi is helping.'

Boadi (1974: 7) explains that, in focus constructions, the Foc has the function of narrowing down the referential range of its host, the focus constituent. The function of the Foc in focus constructions, therefore, is a semantic one. That is, it has discourse information alteration significance and, for that matter, it induces semantic contrast. As will become evident, clear semantic contrast is realized between a focus construction and its canonical clause counterpart.

Boadi (1974) notes that *dèɛ̀*, which occurs in the same syntactic position as *nà*, also plays the role of a focus marker; for instance, *Á'má$_i$ dèɛ̀ Kòfí rébóá nó$_i$*

'as for Ama, Kofi is helping her' (cf. (8c)). As he finally asserts, however, let us note that *dèɛ̀* does not define the concept of contrastive information in definite terms. Unlike *nà*, it does not induce an exclusive focus on a fronted constituent. The non-exclusivity of *dèɛ̀*-focus construction is clearer in terms of contrastive account of focus. Supposing a statement is made with regards to an event, but a constituent in that statement (e.g., the subject or the object) is contrary to the truth of the event, in the correction of the statement by focus, the *nà*-focus construction gives the appropriate contrastive account. For instance, observe in (9) below that, in giving a contrastive object to the one in the declarative statement, cohesion (indicated by a continuous arrow) attains between the declarative statement and the *nà*-focus construction in (9b). On the other hand, we realize that the *dèɛ̀*-focus construction in (9c) does not logically/coherently follow from the declarative statement. This buttresses the point that *dèɛ̀* does not have the same focus marking function as *nà*. In other words, it does not induce an exclusive focus.

(9) a. Kòfí rè-bòá Yàẁ
Kofi PROG-help Yaw
'Kofi is helping Yaw.'

b. Dààbí! Á!má$_i$ *nà* Kòfí ré-bóá nó$_i$
no! Ama FOC Kofi PROG-help 3SG
'No! It is Ama that Kofi is helping (her).'

c. Dààbí! Á!má$_i$ *dèɛ̀* Kòfí rè-bòá nó$_i$
no! Ama FOC Kofi PROG-help 3SG
'No! As for Ama, Kofi is helping her.'

Also, unlike *nà*, *dèɛ̀* cannot come after a fronted Q-word, and thus using *dèɛ̀* in **Hwáń*$_i$ *dèɛ̀ Kòfí rébóá nó*$_i$ is ungrammatical. Therefore, aside from the fact that we do not consider *dèɛ̀* as a true FOC, it also falls outside the scope of this paper.

Indeed, there are other ways of putting a constituent in focus (specifically, in prominence) that do not involve constituent fronting: for instance, the use of intonation, as shown in (10a), and the use of inherently focus-marked words like 'only', as shown in (10b). However, it is important to note that a constituent that has been focused *in situ* (as shown in (10a-b)) could still be fronted for the purpose of contrastive information realization in Akan. As shown in (10c), for example, the subject is fronted along with the inherently focus-marked word, *ńkóáá* (cf. (10b)).

(10) a. [$_{IP}$ *Kòfí* [$_{VP}$ á-ꜜbóá Áꜜmá]] 'Kofi has helped Ama.'
Kofi PRF-help Ama

b. [$_{IP}$ *Kòfí ńkóáá* [$_{VP}$ á-ꜜbóá Áꜜmá]] 'Only *Kofi* has helped Ama.'
Kofi only PRF-help Ama

c. [$_{FocP}$ *Kòfí ńkóáá* nà [$_{IP}$ ɔ́-[$_{VP}$ á-bóá Áꜜmá]]]
Kofi only FOC 3SG- PRF-help Ama
'It is only *Kofi* who has helped Ama.'

In the light of the realization of focus examplified in (10c), we particularly draw attention to the fact that, in this paper, we are concentrating on focus marking that involves not only prominence, but new/contrastive information as well. That is to say, contrastive focus is only realized through constituent fronting in Akan.

A constituent cannot be contrastively focused *in situ* in Akan because the FOC cannot be used in the canonical clause. As noted earlier, the FOC is only introduced at the head position of the extra-sentential projected FocP. This explains the ungrammaticality of the constructions in (11); i.e., the introduction of the FOC in the canonical clause is unattested.

(11) a. *[$_{IP}$ Kòfí ré-bóá Áꜜmá *nà*] 'It is Ama (that) Kofi is helping.'
Kofi PROG-help Ama FOC

b. *[$_{IP}$ Kòfí *nà* ré-bóá Áꜜmá] 'It is Kofi who is helping Ama.'
Kofi FOC PROG-help Ama

It is important to note that the focus construction is related to the Q-word fronting construction in Akan with regards to constituent fronting and the use of the FOC at the head position of a projected functional phrase. Besides these two phrase structure facts, another connection between the two constructions is that a focus construction is more or less an answer to a Q-word fronting construction in a question-answer pair (Boadi 1974). Therefore, as exemplified with the subject in (12) below, the answer constituent to the Q-word in the Q-word fronting construction corresponds to the constituent in focus in the focus construction. Perhaps, this correspondence contributed to Saah's (1988) suggestion that a fronted Q-word is more emphatic, as compared to an *in situ* counterpart.

(12) *Question*

[$_{\text{FocP}}$ *Hwán*$_i$ nà [$_{\text{IP}}$ ɔ$_i$-ré-sómá àbɔ́frá nó]]
who FOC 3SG-PROG-send child DEF
'Who is sending the child?'

Answer/Focus

[$_{\text{FocP}}$ *Kùsí*$_i$ nà [$_{\text{IP}}$ ɔ$_i$-ré-sómá àbɔ́frá nó]]
Kusi FOC 3SG-PROG-send child DEF
'It is Kusi who is sending the child.'

4. More on Q-Word Fronting and Focus Constructions.

We have noted constituent fronting in Q-word and focus constructions in Akan. Current research in LFG (e.g., Berman 1997, Bresnan 2000, 2001) describes constructions exhibiting this phenomenon as forms with discourse function (DF), projected to absorb the fronted constituent. Observe in (12) above that, in light of the structural hierarchy at c-structure, the fronted constituents in Spec-FOCP show an iconic structural precedence and dominance over other constituents in both constructions. We have also observed that FOC appears at the head position of the projected DF (FOCP) in both constructions, as in (12) and other data already given.

One other feature, which both Q-word fronting and focus constructions exhibit and is worth noting in the light of LFG, is the presence of a resumptive pronoun (henceforth, RPro) in the canonical clause position of a fronted constituent (i.e., the Spec-FOCP constituent). This RPro agrees in number and in person with the Spec-FOCP constituent, as can be seen in (13-15) below. Observe in (13) that, with their appearance in Spec-FOCP, the plural subjects (in focus or in question) are replaced in the canonical position (i.e., Spec-IP) with the third person plural pronoun, *wɔ́(ń)*. The singular subjects (in focus or in question) in (14) and (15) are also replaced in the canonical position with the third person singular pronouns; i.e., ɔ̀-(*nó*) for an animate subject and ɛ̀-(*nó*) for inanimate subject. The RPros then refer back to the Spec-FOCP constituents, hence the co-indexing of Spec-FOCP and Spec-IP.

(13) [$_{\text{FocP}}$ *M̀máá$_i$* nà [$_{\text{IP}}$ wɔ̀$_i$-hwɛ́ mm̀ɔ́frá]]
female.PL FOC 3PL-look.HAB child.PL
'It is females who take care of children.'

[$_{\text{FocP}}$ *Hwáánóḿ$_i$* nà [$_{\text{IP}}$ wɔ̀$_i$-hwɛ́ mm̀ɔ́frá]]
who.PL FOC 3PL-look.HAB child.PL
'Which people take care of children?'

(14) a. [$_{\text{FocP}}$ *Ɔ̀báá$_i$* nà [$_{\text{IP}}$ ɔ̀$_i$-hwɛ́ mm̀ɔ́frá]]
female.SG FOC 3SG-look.HAB child.PL
'It is a female who takes care of children.'

b. [$_{\text{FocP}}$ *Hwáń$_i$* nà [$_{\text{IP}}$ ɔ̀$_i$-hwɛ́ mm̀ɔ́frá]]
who FOC 3SG-look.HAB child.PL
'Who takes care of children?'

(15) [$_{\text{FocP}}$ *Ɛ̀tòá$_i$* nà [$_{\text{IP}}$ ɛ̀$_i$-bɔ́-ɪ́]] 'It is a bottle that broke.'
bottle FOC it-break-PAST

[$_{\text{FocP}}$ *Déɛ́ń$_i$* nà [$_{\text{IP}}$ ɛ̀$_i$-bɔ́-ɪ́]] 'What broke?'
what FOC it-break-PAST

It is important to note that the RPros are essentially pronouns of Akan. As presented in (16) below, therefore, we call attention to the fact that an RPro is not just an agreement marker, but a pronoun (in position) that has to agree in person and number with a fronted argument function. Note also that only the non-bracketed syllables of the pronouns are normally said in the appropriate argument positions, as observed in (13-15) above.

(16) *Pronouns of Akan*

	Subject		Object	
Person	Singular	Plural	Singular	Plural
1st	*mé* 'I'	*yɛ́(ń)* 'we'	*mè* 'me'	*yɛ́ń* 'us'
2nd	*wó* 'you'	*mó* 'you'	*wò* 'you'	*mò* 'you'
3rd	*ɔ̀(nó)* 'she/he'; *ɛ̀(nó)* 'it'	*wɔ́(ń)* 'they'	*(ɔ̀)nó* 'her/him'; *(ɛ̀)nó* 'it'	*wɔ̀n*`'them'

Going back to the data in (13-15), observe that the RPros are in the third person. This does not mean that an RPro should always be in the third person. The RPros are in the third person in (13-15) because third person pronouns are the right (pronominal) substitutes for full noun argument functions. As shown in (17a), where the argument function in focus is a first person singular pronoun for instance, its RPro should also be in the first person singular – i.e., the same pronoun. Otherwise, as also shown in (17b), the construction is ungrammatical. The focus construction in (17a), however, shares a common corresponding Q-word fronting construction with (14a); i.e., *Hwáńi nà ɔ̀i-hwɛ́ m̀mɔ̀fŕa* 'who takes care of children?' in (14b).

(17) a. [$_{FocP}$ *Méi* nà [$_{IP}$ *mèi*-hwɛ́ m̀mɔ̀fŕá]]
1SG FOC 1SG-look.HAB child.PL
'It is me who takes care of children.'

b. *[$_{FocP}$ *Méi* nà [$_{IP}$ ɔ̀i-hwɛ́ m̀mɔ̀fŕá]]
1SG FOC 3SG-look.HAB child.PL
'It is I who takes care of children.'

As noted by Saah (1988: 24) referring to Stewart (1963: 149), unlike in the subject position, the occurrence of RPro is restricted in the object position (and other post-verbal environments). This restriction has to do with the feature specification of animacy; i.e., [±animate]. Specifically, if a fronted object or object-in-question is animate ([+animate]), its canonical base position is filled with the appropriate RPro, as shown in (18a) and (19a). Lack of an RPro for a fronted animate object or object-in-question renders a construction ungrammatical, as shown in (18c) and (19c).

(18) *Focus:*
a. [$_{FocP}$ À̀tó*i* nà [$_{IP}$ ɔ̀báá nó [$_{VP}$ ré-bóá [$_{NP}$ nó*i*]]]]
Ato FOC lady DEF PROG-help 3SG
'It is Ato that the lady is helping.'

b. [$_{FocP}$ ɛ̀móó*i* nà [$_{IP}$ ɔ̀báá nó [$_{VP}$ nóá [$_{NP}$ Ø*i*]]]]
rice FOC lady DEF cook.HAB *e*
'It is rice (that) the lady cooks.'

c. *[$_{FocP}$ Àtó$_i$ nà [$_{IP}$ ɔbáá nó [$_{VP}$ ré-bóá [$_{NP}$ Ø$_i$]]]]
Ato FOC lady DEF PROG-help *e*
'It is Ato that the lady is helping.'

d. *[$_{FocP}$ ɛ̀móó$_i$ nà [$_{IP}$ ɔbáá nó [$_{VP}$ nóá [$_{NP}$ nó$_i$]]]]
rice FOC lady DEF cook.HAB 3SG
'It is rice (that) the lady cooks.'

(19) *Q-word fronting:*

a. [$_{FocP}$ Hwáń$_i$ nà [$_{IP}$ ɔbáá nó [$_{VP}$ ré-bóá [$_{NP}$ nó$_i$]]]]
who FOC lady DEF PROG-help 3SG
'Who is the lady helping?'

b. [$_{FocP}$ Déɛ́ń$_i$ nà [$_{IP}$ ɔbáá nó [$_{VP}$ nóá [$_{NP}$ Ø$_i$]]]]
what FOC lady DEF cook.HAB *e*
'What does the lady cook?'

c. *[$_{FocP}$ Hwáń$_i$ nà [$_{IP}$ ɔbáá nó [$_{VP}$ ré-bóá [$_{NP}$ Ø$_i$]]]]
Who FOC lady DEF PROG-help *e*
'Who is the lady helping?'

d. *[$_{FocP}$ Déɛ́ń$_i$ nà [$_{IP}$ ɔbáá nó [$_{VP}$ nóá [$_{NP}$ nó$_i$]]]]
rice FOC lady DEF cook.HAB 3SG
'What does the lady cook?'

Conversely, where the fronted object or object-in-question is inanimate ([-animate]), the RPro is covertly represented, as in (18b) and (19b). As shown in (18d) and (19d), an overt RPro for a fronted inanimate object or object-in-question renders a construction ungrammatical. Saah (1992: 221) refers to the lack of overt RPro in the inanimate situation as an "empty category" (*EC*) situation in Akan.

Regarding the animacy of an object and whether or not it is human (i.e., [±human]), as shown in (20a) and (20b) respectively, it is important to note that both [+animate; +human] and [+animate; -human] objects have the same RPro in the canonical position. However, as shown in (20c), an RPro could be optional in the case of [+animate; -human] objects.

(20) a. [$_{FocP}$ Àtó$_{[+animate, +human]i}$ nà [$_{IP}$ ɔ̀báá nó [$_{VP}$ ré-bóá [$_{NP}$ nó$_i$]]]]
Ato FOC lady DEF PROG-help 3SG
'It is Ato that the lady is helping.'

b. [$_{FocP}$ Àkókɔ́$_{[+animate, -human]i}$ nà [$_{IP}$ ɔ̀báá nó [$_{VP}$ ré-yɛ́ǹ [$_{NP}$ nó$_i$]]]]
fowl.SG FOC lady DEF PROG-help 3SG
'It is a fowl that the lady is rearing.'

c. [$_{FocP}$ Àkókɔ́$_{[+animate, -human]i}$ nà [$_{IP}$ ɔ̀báá nó [$_{VP}$ ré-yɛ́ǹ [$_{NP}$ Ø$_i$]]]]
fowl.SG FOC lady DEF PROG-help 3SG
'It is a fowl that the lady is rearing.'

In the *wh*-construction, while *hwáń* replaces [+animate; +human] objects (see (19a)), [+animate; -human] objects are replaced by *déέń* 'what' (or *déèǹ àbóá / àbóá béń* 'what animal'); e.g., *Déèǹ àbóá nà ɔ̀báá nó réyéǹ nó* 'what animal is the lady rearing?' (cf. (19a)).

Where there is a need to show in the c-/f-structures that the inanimate object is covertly represented, some versions of LFG account for the phenomenon through the Principle for Identifying Gaps (Bresnan 2001: 181) provided in (21). The principle is necessary in the linking up of such an *EC* to the Spec-DF (FOCP) constituent, thus enabling the integration of Spec-DF constituent (a non-argument) in the argument structure in f-structure.

(21) *Principle for Identifying Gaps:*
Associate XP $\rightarrow$ *e* with (($x\uparrow$) DF)= $\uparrow$

Through the Principle for Identifying Gaps, the violation of the Economy of Expression principle (e.g., Bresnan 2001) by having an *EC* in the c-structure is bypassed. The Economy of Expression principle states that all syntactic phrase structure nodes are optional and use of any of them is prohibited unless independent principles demand it.

Perhaps, the animacy restriction on objects, and not on subjects, also emphasizes the Subject Condition (SC) which LFG stipulates. SC requires every predicate to have a subject, but not necessarily an object. Based on the inspiration of SC, we posit the condition, Strict Phonetic Subject (SPS), stated in (22), for extra-sentential clauses in Akan (in this paper, Q-word fronting and focus constructions).

(22) *Strict Phonetic Subject:*
Every predicator in the embedded clause of an extra-sentential clause must have a phonetic subject.

SPS is motivated against a possible proposal that a fronted subject (in a focus or *wh-* construction) does not need RPro in the canonical clause, since it is still the most prominent in the relational hierarchy and is the default DF. In this sense, SPS is not merely a stipulation. In fact, it has to be satisfied in other extra-sentential constructions in Akan as well; e.g., topic constructions and relative clauses. SPS explains the grammaticality and ungrammaticality of the focus constructions in (23b) and (23c) respectively.

(23) a. [$_{IP}$ Kùsí rè-sòmá mé]
Kusi PROG-send 1SG
'Kusi is sending me.'

b. [$_{FocP}$ *Kùsí$_i$* nà [$_{IP}$ ɔ$_i$-ré-sómá mé]]
Kusi FOC 3SG-PROG-send 1SG
'It is Kusi who is sending me.'

c. *[$_{FocP}$ *Kùsí$_i$* nà [$_{IP}$ Ø$_i$-ré-sómá mé]]
Kusi FOC __-PROG-send 1SG
'It is Kusi who is sending me.'

5. Distinction: Discourse-Contextual Information.

So far, it has been made clear that both Q-word fronting and focus constructions essentially share a common marked (or extra-sentential) phrase structure configuration; i.e., [$_{FocP}$ *XP nà* [$_{IP}$ ...]]. However, considering the individual discourse-contextual information that is realized in the s-structure of each of them, through the i-structure (Vallduví 1992, Lambrecht 1994), as compared to the discourse-contextual information expressed in the s-structure of the respective *in situ* and canonical clause counterparts, we explain in this section that semantic contrast is clearly evident in focus constructions.

In exploring the different types of discourse-contextual information that obtain in Q-word fronting and focus constructions, let us assume that the semantic content in each of the constructions particularly has to do with (or is tied to) the obligatory occurrence of the FOC, besides the constituent fronting. With this as-

sumption, as already noted in section 3, we suggest that, unlike in focus constructions, the occurrence of the FOC in Q-word fronting constructions does not invoke any information of semantic significance in the discourse other than what obtains in the discourse of related Q-word *in situ* counterparts. In other words, in Akan, Q-word fronting does not alter the s-structure of the interrogative in any way.

We do not dispute the fact that, in some languages, Q-word fronting may invoke contrastive information (as against the discourse-contextual information expressed in a related *in situ* construction). Mutaka & Tamanji (2000: 221) note that in Bafut, a Grassfields Bantu language spoken in Cameroon, a Q-word fronting construction encodes more emphasis than its *in situ* counterpart, although both of them ask virtually the same question. The study clams that a Q-word fronting construction expresses a high degree of concern on the part of the speaker. Thus, as shown in (24) for example from Mutaka & Tamanji (2000: 221), the speaker is insisting to know the specific "thing" (the object) that *Suh* killed in (24c), but the speaker shows no such insistence in (24b).

(24) a. Sùù kɨ́ kó nô 'Suh killed a snake.'
Suh PAST catch snake

b. Sùù kɨ́ kó àkɔ̀ 'What did Suh kill?'
Suh PAST catch what

c. à nɨ́ àkɔ̀ mɔ̀ Sùù kɨ́ kô 'What is it that Suh killed.'
it be what that Suh PAST catch

Perhaps, this "speaker-intention" argument could be made in Akan as well. However, we contend that it has no semantic relevance in the discourse. That is, semantic contrast does not obtain between a Q-word fronting construction and its *in situ* counterpart in Akan as it does between a focus construction and a related canonical construction.

Boadi (1990: 78) suggests that the lack of semantic contrast in a Q-word fronting construction in Akan, as compared to a related Q-word *in situ* construction, is due to the semantic fact that Q-words are actually inherently focus-marked. Accordingly, they do not need any special reference. We further claim in this paper that a Q-word holds the core of the information profile of a construction within which it appears (i.e., the expression of interrogative). Thus, a Q-word does not need any further semantic buffer, in this case the FOC, to complete what it already inherently establishes. Sabel (2000: 430), citing Hovarth (1986: 118),

explains that focus is a syntactic feature that is assigned to a non-echo *wh*-phrase. This strengthens our position that Q-words are inherently focus-marked; so, they do not need to be assigned further focus. As Boadi (1990) notes, Q-word fronting is only an alternative, optional representation.

A test for ascertaining that Q-words are inherently focus-marked is that, following a previous discourse, only a Q-word could be used to represent the whole of a construction within which it occurs. Accordingly, in (25) below, the whole of (25b) can be replaced by (25c), drawing directly from (25a). In other words, (25c) is a follow up to (25a), just as (25b) is. On the contrary, observe that (25e) cannot be a follow up to (25a) and so, it cannot represent the whole of (25d). That is, as a non-Q-word, the word in (25e), *Kofi*, which is also the Spec-FocP constituent in (25d), is not inherently focus-marked. So, it can only be contrastively focused in the focus-presupposition structure, as shown in (25d).

(25) a. *Kòfí* bé-!dúá àbá nó 'Kofi will sow the seed.'
Kofi FUT-sow seed DET

b. *Hwáń*$_i$ nà ò$_i$-bé-!dúá àbá nó? 'Who will sow the seed?'
who FOC 3SG-FUT-sow seed DET

c. *Hwáń*? 'Who?' (appropriate alternative to b.)
who

d. *Kòfí*$_i$ nà ò$_i$-bé-!dúá àbá nó 'It is Kofi who will sow the seed.'
Kofi FOC 3SG-FUT-sow seed DET

e. *Kòfí* 'Kofi' (not an appropriate alternative to d.)

As noted earlier, contrary to the stance taken in this paper, Saah (1988: 19) claims that, as a motivation for the constituent left-periphery dislocation, extra-sentential clause-initial Q-word occurrence is more emphatic, as compared to the *in situ* counterpart. The question, however, is to what extent is a fronted Q-word more emphatic? With regards to discourse-contextual information, what can we draw from its i-structure (which is accessible to the s-structure, as noted earlier) that is different from what is obtained in the i-structure of a related Q-word *in situ* construction? Seemingly emphasized as a fronted Q-word in Akan is, it is actually vacuous in terms of semantic contrast to a related Q-word *in situ* construction. As

explained in section 2.1 (see and contrast (6a) and (7a)), Q-word fronting (with the employment of FOC) induces nothing more into its i-structure other than what is in the i-structure of the *in situ* construction (i.e., the general interrogative expression of the Q-words).

On the other hand, the identification of a semantic contrast in the i-structure of a focus (of extraction) construction, as compared to that of a related canonical construction, is indisputable and readily perceptible. Contrastive information is attained in focus construction, particularly relating to the constituent in focus. In this case, among all the constituents in the construction, the one in focus is highlighted as the point of contrastive discourse information (of exclusivity) in the construction; hence, its constitution as the "point of prominence" (Boadi 1974). For instance, the focus construction in (25d), *Kòfí nà òbé!dúá àbá nó* 'it is Kofi who will sow the seed', is interpreted as 'it is Kofi and only Kofi (within a discourse-relevant subset of individuals) who will sow the seed', and not just as 'Kofi will sow the seed'. With the latter interpretation, none of the constituents is identified as prominent (or new) information. Accordingly, other people besides *Kofi* might sow the seed as well; hence, the contrast between it and the former interpretation of focus. Kiss (1995) puts the interpretation of focus as follows: "the focus operator serves to express identification" (Kiss 1995: 212). In the focus construction in (25d), for instance, constituent fronting and the use of the FOC identify *Kofi*, and only *Kofi*, as the one who is sowing the seed.

We realize that a focus construction differs in semantic content from a related canonical clause when we put both constructions in *yes-no* question. In answering the question, with the focus construction, the constituent in focus alone could be retrieved into the answer, as can be observed in (26a).[3] On the other hand, with the canonical clause, the whole construction has to be repeated in the answer, as shown in (26d). The retrieval of any particular constituent into the answer results in question-answer incoherence, as shown in (26e). This is because, unlike in focus constructions, no particular constituent is put in (contrastive) focus in the canonical clause.

[3] In answering (26a), the whole focus construction could also be retrieved; i.e., *Àáné, Àdú$_i$ nà Kòfí fréè nó$_i$*. Having the canonical sentence, *Àáné, Kòfí fréè Àdú*, as an answer to (26a) sounds funny.

(26) a. Àdú$_i$ *nà* Kòfí fɾɛ́-ɛ̀ nó$_i$? 'Is it Adu who Kofi called?'
Adu FOC Kofi call-PAST 3SG

b. Àáné, Àdú (á) 'Yes, it is Adu.'
yes, Adu

c. Kòfí fɾɛ́-ɛ̀ Àdú? 'Did Kofi call Adu?'
Kofi call-PAST Adu

d. Àáné, Kòfí fɾɛ́-ɛ̀ Àdú 'Yes, Kofi called Adu.'
yes, Kofi call-PAST Adu

e. Àáné, Àdú 'Yes, it is Adu.'
yes, Adu

Despite the semantic distinction made between Q-word fronting and focus constructions in relation to their non-extracting clause counterparts, it is important to note that "focus-presupposition" information structure is reflected in both constructions, which goes to prove that both Q-word and focus express prominent new information. With the manifestation of "focus-presupposition" structure in Q-word fronting constructions, one cannot deny the fact that they involve some sort of focusing. Kroeger (2004: 139) notes that "the question word bears a pragmatic focus, since it specifies the crucial piece of new information which is required; the rest of the question is part of presupposition". Sabel (2000: 430) also puts it this way: "the *wh*-phrase designates what is not presupposed as known". Now, since a Q-word constitutes a linguistic device for the identification of a specific piece of prominent new information, it should be recognized as prominent new information as well. As shown in (27) below, we observe that it is from the questioning in (27a) that *Kusi* realizes as prominent new information in (27b) and, for that matter, the focus.

(27) a. *Question:* *Hwáń$_i$* nà ɔ̀$_i$-ré-sómá àbɔ̀fɾá nó?
who FOC 3SG-PROG-send child DEF
'Who is sending the child?'

b. *Focus:* *Kùsí$_i$* nà ɔ̀$_i$-ré-sómá àbɔ̀fɾá nó
Kusi FOC 3SG-PROG-send child DEF
'It is Kusi who is sending the child.'

Following the feature-based i-structure (Choi 1999, 2001, Lee 2001), which we extend here to include Q-words, Q-words and focused constituents in Akan would therefore show identical information profiles on discourse NEW(ness) and PROM(inence), as shown in (28).

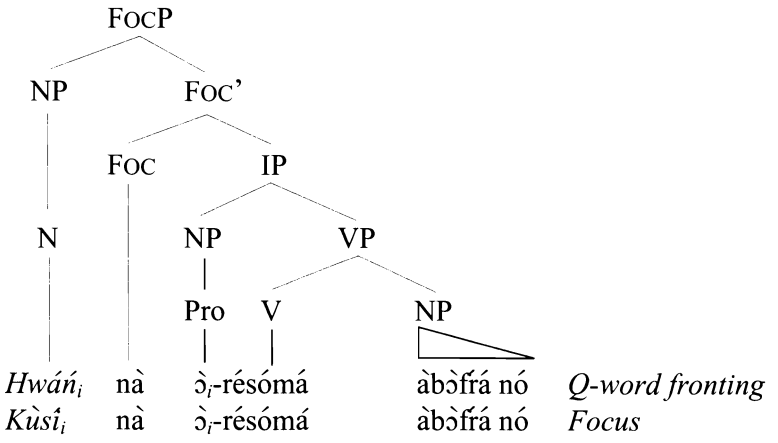
(28) *Information profile of focus and Q-word*

Focus $\begin{bmatrix} \text{NEW} + \\ \text{PROM} + \end{bmatrix}$ Q-word $\begin{bmatrix} \text{NEW} + \\ \text{PROM} + \end{bmatrix}$

Going back to Q-word fronting and focus constructions in relation to their non-extracting clause counterparts, however, it has been noted that, unlike in Q-word fronting constructions, the FOC has a semantic function in focus constructions; i.e., it alters the default discourse-contextual information of a related canonical clause. We refer to this semantic function of the FOC in focus constructions as discourse-contrast, since it results in contrastive information (of exclusivity; i.e., '*X* and only *X*') that characterizes focus constructions in Akan. Conversely, discourse-neutral (Lee 2001) is obtained with the occurrence of the FOC in Q-word fronting constructions, since the same discourse-contextual information expressed in related Q-word *in situ* constructions is expressed in them. It logically follows then that "Q-word fronting in Akan is only an optional representation" (Boadi 1990: 78) and the obligatory occurrence of the FOC with it is only a general syntactic restriction. In line with structural markedness, we refer to the FOC in Q-word fronting constructions as configurational focus, since its occurrence contributes to the marking of the whole phrase structure configuration. Recall that Q-word fronting and focus constructions are noted as marked sentence-types.

Having identified and explained the common information profile, defining pragmatic focus, of Q-words and focused constituents and their different semantic interpretations in a construction, we now present a common c-structure and individual f- and i-structures of the Q-word fronting and focus constructions in (29) below. In the individual i-structures in (29c) in particular, we show the different realizations of the common information profile (presented in (28)) in the interpretation of Q-word fronting and focus constructions relative to the discourse-contextual information that obtain in related non-extracting constructions – i.e., the realizations of discourse-neutral of Q-words and discourse-contrast of focus.

(29) a. C-structure (*for both Q-word fronting and focus constructions*)



b. *F-structures*

Q-word fronting

$$\begin{bmatrix} \text{PRED} & \text{'soma ⟨SUBJ, OBJ⟩'} \\ \text{ASP} & \text{PROG} \\ \text{DF} & [\ ] \\ \text{SUBJ} & \begin{bmatrix} \text{NUM} & \text{SG} \\ \text{PRED} & \text{'Pro'} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{NUM} & \text{SG} \\ \text{DEF} & + \\ \text{PRED} & \text{'abɔfra'} \end{bmatrix} \end{bmatrix}$$

Focus

$$\begin{bmatrix} \text{PRED} & \text{'soma ⟨SUBJ, OBJ⟩'} \\ \text{ASP} & \text{PROG} \\ \text{DF} & [\ ] \\ \text{SUBJ} & \begin{bmatrix} \text{NUM} & \text{SG} \\ \text{PRED} & \text{'Kusi'} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{NUM} & \text{SG} \\ \text{DEF} & + \\ \text{PRED} & \text{'abɔfra'} \end{bmatrix} \end{bmatrix}$$

c. *I-structures*

Q-word fronting

$$\begin{bmatrix} \text{FOCUS} & \begin{bmatrix} \text{F-TYPE} & \text{NEUTRAL} \\ \text{I-PRED} & \text{'Hwan'} \end{bmatrix} \\ \text{BGD} & \begin{bmatrix} \text{ɔrésòmá} \\ \text{àbɔ́frá nó} \end{bmatrix} \end{bmatrix}$$

Focus

$$\begin{bmatrix} \text{FOCUS} & \begin{bmatrix} \text{F-TYPE} & \text{CONTRASTIVE} \\ \text{I-PRED} & \text{'Kusi'} \end{bmatrix} \\ \text{BGD} & \begin{bmatrix} \text{ɔrésòmá} \\ \text{àbɔ́frá} \end{bmatrix} \end{bmatrix}$$

We have already discussed how the common c-structure is realized in section 4. The argument functions subcategorized by the verb, *sòmà*, in both constructions are also encoded in the individual f-structures. Also encoded in the f-

structures is the identification of the projected DF with an argument function, the subject. The individual semantic significance in the discourse of Q-word fronting and focus constructions is also given in the separate i-structures. Here, the focus type (F-TYPE) of the Q-word, *hwáń* (noted as information predicate (I-PRED)) is given as neutral following FOC function as discourse-neutral in Q-word fronting construction, while that of the focused constituent, *Kùsí*, is given as contrastive following FOC function as discourse-contrast in focus constructions. The rest of both constructions are given as presupposition/background information (BGD). I-structure is observed here as distinct structure from the f-structure projected off the c-structure (King 1997, Butt & King 1998). Recall that it is accessible to the s-structure for semantic interpretation.

Since Q-words have been noted as inherently focus-marked in Akan, finally, it is important to note that a Q-word fronting construction is distinguished from its *in situ* counterpart only on the basis of c-structure configurational markedness. As noted on several occasions, with respect to discourse-contextual information realization, both representations are essentially the same.

6. Constraining the Constructions: OT-LFG.

With a recast of LFG within Optimality Theory (OT-LFG) (Bresnan 2000, Choi 1999, Kuhn 2001), the common c-structure configuration of Q-word fronting and focus constructions is further established in this section. We also illustrate and constrain harmonic alignment (Aissen 1999, Bresnan 2000, Choi 2001, Lee 2001) between the common c-structure and the i-structure of a particular construction. Each of the parallel structures of LFG defines prominence in a hierarchical fashion. The matching of prominence definition in one structure to that in another structure constitutes a harmonic alignment.

6.1 Categorial representation. Two conflicting constraints readily come to mind concerning constituent fronting in Q-word fronting and focus constructions. These are OP-SPEC, motivated by the presence of a syntactic operator (Grimshaw 1997, Bresnan 2000, Kuhn 2001) and recast in expression as *operator in specifier of functional projection*, and *DISLOC, proposed in this paper on the inspiration of the economy principle and expressed as *don't dislocate*. As stated in (30), while OP-SPEC favors functional projection and the appearance of a constituent in question/focus in Spec-DF, *DISLOC stands to block such a categorial representation. For a Q-word fronting or focus construction word order to prevail, therefore, OP-SPEC must crucially outrank *DISLOC.

(30) *OP-SPEC:*
An operator (i.e., a constituent in focus/question) must be in the specifier position of its functional projection.

**DISLOC:*
Don't dislocate; the canonical phrase structure must not be altered.

The other typological traits of Q-word fronting and focus constructions noted earlier also need to be recast and explained in terms of constraints, if alternative categorial representations are to be properly rejected. It has been noted that the projected phrase of the operator function has to be headed by the FOC, *nà*. Also noted is the fact that an argument function that appears at the specifier position of the projected functional phrase has to be replaced in the embedded canonical clause position by an RPro. The appropriate constraints we employ to demand these representations are OB-HD/fp (Bresnan 2000, Choi 2001, Kuhn 2001) and PARSE/gf, proposed here on the motivation of SPS; (see (22)).[4] Respectively expressed as *obligatory head* and *parse argument functions*, OB-HD/fp and PARSE/gf are also stated in (31) below. In the constraint ranking, we assume a dominance of PARSE/gf between the two. However, both constraints should dominate *DISLOC and should be dominated by OP-SPEC (see Tableau I).

(31) *OB-HB/fp:*
The head position of a functional projection must be filled.

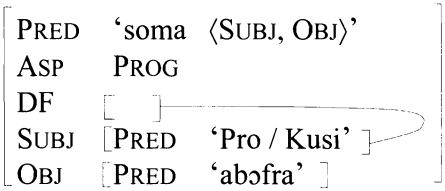
PARSE/gf:
Left dislocated argument function should be phonetically represented in the canonical clause position.

The f-structure in (32), a merged f-structure of both constructions in (29), is employed as the working input. The attribute-value matrix (AVM) of the operation and other features underscored in the individual constructions are not indicated in the input f-structure of the two constructions, since they do not undermine the c-

[4] An alternative view is that SPS should be kept in the constraint formulation, but that would restrict pronoun resumption to only the subject position. That is, considering the fact that fronted/focused animate objects also have to be resumed, PARSE/gf better captures the phenomenon.

structure configuration in any way. Tableau (I) also explains that, among the candidate set of (a), (b), (c), and (d), the optimal candidate is the one whose c-/f-structures best relate to this input. Both Q-word and focused constituents are represented in Spec-FocP as NP in the tableaux.

(32) *Input f-structure:* Hwáńi / Kùsíi nà ɔ̀i-résómá àbɔ́frá nó



(I) *OP-SPEC » PARSE/gf » OB-HB/fp » *DISLOC*

		Matrix Q-word fronting/focus	OP-SPEC	PARSE	OB-HB	*DISLOC
☞	a.	[FocP NPi *nà* [IP Proi [VP V NP]]]				*
	b.	[IP NP [VP V NP]]]	*!*			
	c.	[FocP NPi *nà* [IP ei [VP V NP]]]		*!		*
	d.	[FocP NPi *e* [IP Proi [VP V NP]]]			*!	*

In Tableau (I), candidate (a) outperforms the other candidates as follows: Candidate (b) is taken out (on two counts) for not having a functional projection, let alone a constituent in question/focus occurring in its specifier position. Candidate (c) is also ruled out on PARSE/gf for violating the requirement of having an RPro in place of the fronted argument function (in the present case, the subject function) in the embedded canonical clause. Candidate (d) is also taken out of consideration for its violation of OB-HB/fp, which ensures functional projection headedness. Consequently, the grammatical c-/f-structure of candidate (a) prevails as the optimal candidate. Note that the input f-structure in (32) essentially doubles as f-structure of candidate (a).

6.2 Information correspondence: alignment. We have noted that Q-word fronting and focus constructions share a common information profile in the i-structure as regards NEW and PROM. Choi (2001: 34) proposes i-/c-structure cor-

respondence/alignment constraints based on NEW and PROM that yield informationally-motivated marked c-structure. Relevant among these constraints in the present cases of Q-word fronting and focus constructions are NEW-L and PROM-L recast in (33) below.

(33) *NEW-L:*
[+NEW] aligns left in the construction of occurrence.

PROM-L:
[+PROM] aligns left in the construction of occurrence.

Since both Q-word and constituent in focus are noted as '[+PROM]; [+NEW]' in the feature-based i-structure, and each of them sits at Spec-FOCP, at present the most prominent position in the *structural hierarchy* at c-structure, it is obvious that the i-/c-structure correspondence constraints in (33) will be satisfied in both constructions (see Tableau II). Comparing their individual discourse-contextual information to the information that obtains in their respective Q-word *in situ* construction and canonical clause counterparts, however, Q-word fronting and focus constructions have been set apart in the s-structure through the projected i-structure (see (29c)) as "discourse-neutral" and "discourse-contrast" respectively. These separate semantic orientations of Q-word fronting and focus are expressed in constraint terms in (34) below, following Choi's (2001) NEW-L and PROM-L proposals.

(34) *NEUT-L:*
[+NEUT] aligns left in the construction of occurrence.

CONST-L:
[+CONST] aligns left in the construction of occurrence.

With the present constraints in the constraint set, Tableau II below show that CONST-L must crucially outrank NEUT-L where there is a need to establish i-/c-structure harmonic alignment in a focus construction (i.e., a correspondence between a constituent in focus and the Spec-FOCP position, as against harmonic alignment between a fronted Q-word and the Spec-FOCP position). Observe in the tableau that, unlike the ranking of CONST-L against NEUT-L, the ranking between CONST-L and NEW-L/PROM-L in the Tableau is hardly crucial and, for that matter, has little or no impact at all in the i-/c-structure correspondence. As noted earlier,

this is because both fronted Q-word and focus constituent sit at Spec-FOCP and specify for [+NEW]/[+PROM].

(II) *NEW-L » PROM-L » CONST-L » NEUT-L*

		[$_{FocP}$ NP$_i$ *nà* [$_{IP}$ Pro$_i$ [$_{VP}$ V NP]]]	NEW-L	PROM-L	CONST-L	NEUT-L
☞	a.	[$_{FocP}$ Papa$_{[+CONST, +NEW, +PROM]i}$ *nà* [$_{IP}$ Pro$_i$ [$_{VP}$ V NP]]]				*
	b.	[$_{FocP}$ Hwan$_{i[+NEUT, +NEW, +PROM]i}$ *nà* [$_{IP}$ Pro$_i$ [$_{VP}$ V NP]]]			*!	

It is important to note that CONST-L and NEUT-L are only necessary constraints motivated on individual semantic content to draw attention to the s-structure distinction between Q-word fronting and focus constructions. Thus, the fact that the focus construction outperforms the fronted Q-word construction in Tableau II does not mean that the Q-word fronting construction is ungrammatical. As has already been mentioned in previous sections, it only explains that, unlike in a focus construction, no semantic contrast is realized in the i-structure of a Q-word fronting construction, as compared to that of related *in situ* construction - i.e., the discourse-contextual information is not altered. Ranking NEUT-L over CONST-L will also select i-/c-structure correspondence in Q-word fronting construction.

7. Conclusion.

It has been shown in this paper that Q-word fronting (in *wh*-questions) and focus constructions in Akan essentially share the same phrase structure configuration, which involves constituent left dislocation, introduction of the focus marker (FOC), *nà*, and insertion of a resumptive pronoun (RPro) for a dislocated argument function. Further, it has also been illustrated, using the OT-LFG framework, that the same c-/f-structure constraints and their rankings essentially ensure the configuration of both constructions.

Through the attribute-value matrix (AVM)-based i-structure, however, we have drawn attention to the individual semantic content of Q-word fronting and focus constructions based on the individual discourse-contextual information that obtains in them in comparison to discourse-contextual information that obtains in

respective *in situ* construction and canonical clause counterparts. It has been explained that the occurrence of the FOC, along with constituent left-periphery dislocation in a Q-word fronting construction does not result in semantic contrast because the discourse-contextual information expressed in it is the same that obtains in an *in situ* counterpart. On the other hand, constituent left-dislocation and the occurrence of the FOC in a focus construction do bring into play semantic contrast. In other words, a constituent is highlighted among others as an obvious point of contrastive information in the information profile of a focus construction. Using OT-LFG, we have stressed this semantic information distinction between the two constructions, which further shows the optimization of a particular i-/c-structure alignment in the grammar.

REFERENCES

Aissen, Judith. 1999. "Markedness and subject choice in Optimality Theory". *Natural Language and Linguistics Theory* 17: 673-711.

Berman, Judith. 1997. "Empty categories in LFG". *On-line Proceedings of the LFG97 conference, University of California, San Diego*, ed. by Miriam Butt & Tracy Hollaway King. Stanford: CSLI Publications.
URL: http://csli-publications.stanford.edu/LFG/2/lfg97-toc.html.

Boadi, Lawrence A. 1974. "Focus-marking in Akan". *Linguistics* 140: 5-57.

Boadi, Lawrence A. 1990. "Questions in Akan". *Frankfurter Afrikanistische Blätter* 2: 70-92.

Bresnan, Joan. 2000. "Optimal syntax". *Optimality Theory: Phonology, Syntax and Acquisition*, ed. by Joost Dekkers, Frank van der Leeuw & Jeroen van de Weijer. Oxford: Oxford University Press. Pp. 334-385.

Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell.

Butt, Miriam & Tracy Holloway King. 1998. "Interfacing phonology with LFG". *On-line Proceedings of the LFG98 Conference, University of Queensland, Brisbane*, ed. by Miriam Butt & Tracy Holloway King. Stanford: CSLI Publications.
URL: http://csli-publications.stanford.edu/LFG/3/lfg98.html.

Choi, Hye-Won. 1999. *Optimizing Structure in Context: Scrambling and Information Structure*. Stanford, CA: CSLI Publication.

Choi, Hye-Won. 2001. "Resolution of mismatches". *Formal and Empirical Issues in Optimality-Theoretic Syntax; Studies in Constraint-based Lexicalism 5*, ed. by Peter Sells. Stanford, CA: CSLI Publications. Pp. 17-62.

Dalrymple, Mary. 2001. *Lexical-Functional Grammar*: Syntax and Semantics 34. New York: Academic Press.

Falk, Yehuda. 2001. *Lexical-Functional Grammar*: *An Introduction to Constraint-Based Syntax*. Stanford: CSLI Publications.

Grimshaw, Jane. 1997. "Projection, heads, and optimality". *Linguistic Inquiry* 28: 373-422.

Horvath, Julia. 1986. *FOCUS in the Theory of Grammar and the Syntax of Hungarian*. Dordrecht: Foris.

Kaplan, Ronald M. & Joan Bresnan. 1982. "Lexical-functional grammar: a formal system for grammatical representation". *Formal Issues in Lexical-Functional Grammar*, ed. by Mary Dalrymple et al., 1995. Stanford: CSLI Publications. Pp. 29-130.

King, Tracy Holaway. 1997. "Focus domain and information-structure". *On-line Proceedings of the LFG97 Conference, University of California, San Diego*, ed. by Miriam Butt & Tracy Holloway King. Stanford: CSLI Publications. URL: http://csli-publications.stanford.edu/LFG/2/lfg97-toc.html.

Kiss, Katalin É. 1995. "NP movement, operator movement, and scrambling in Hungarian". *Discourse Configurational Languages*, ed. by Katalin É. Kiss. Oxford Studies in Comparative Syntax, Oxford University Press. Pp. 207-243.

Kroeger, Paul R. 2004. *Analyzing Syntax: a Lexical-Functional Approach*. Cambridge University Press.

Kuhn, Jonas. 2001. "Generation and parsing in Optimality Theoretic syntax". *Formal and Empirical Issues in Optimality-Theoretic Syntax*: *Studies in Constraint-Based Lexicalism* 5, ed. by Peter Sells. Stanford: CSLI Publications. Pp. 313-366.

Lambrecht, Knud. 1994. *Information Structure and Sentence Form*. Cambridge: Cambridge University Press.

Lee, Hanjung. 2001. "Markedness and word order freezing". *Formal and Empirical Issues in Optimality-Theoretic Syntax*: *Studies in Constraint-Based Lexicalism* 5, ed. by Peter Sells. Stanford: CSLI Publications. Pp. 17-62.

Mutaka, Ngessimo N. & P. Ngwe Tamanji. 2000. *An Introduction to African Linguistics*. Munich: Lincom Europa.

Saah, Kofi K. 1988. "*Wh*-questions in Akan". *Journal of West African Languages* 18, 1: 17-28.

Saah, Kofi K. 1992. "Null object constructions in Akan". *Proceedings of the Kwa Comparative Syntax Workshop*. MIT Working Papers in Linguistics. Vol. 17: 219-44.

Sabel, Joachim. 2000. "Partial wh-movement and the typology of wh-questions". *Wh-Scope Marking*: *Linguistik Aktuell/Linguistics Today* 37, ed. by Uli Lutz, Gereon Muller & Arnim von Stechow. Amsterdam: John Benjamins. Pp. 409-476.

Stewart, John M. 1963. "Some restrictions on the object in Twi". *Journal of African Languages* 2: 145-149.

Vallduví, Enric. 1992. *The Informational Component*. New York: Garland.

Department of Linguistics
University of Hong Kong
HKU-Main Building
Pok Fu Lam Road
Hong Kong, S.A.R China
cmarfo@graduate.hku.hk
abbodomo@hku.hk

[received December 18, 2004
accepted September 9, 2005]

Studies in African Linguistics
Volume 34, Number 2, 2005

PUBLICATIONS RECEIVED

Cammenga, Jelle. *Igikuria Phonology and Morphology: A Bantu Language of South-West Kenya and North-West Tanzania.* (East African Languages and Dialects, vol. 15). Cologne: Rüdiger Köppe Verlag. 2004. Pp. 351. ISBN 3-89645-029-6. Paperback. €64.00.

The author's 1994 dissertation on Kikuria, defended at the Free University of Amsterdam, is revised in this volume. The first chapter, "Introduction" [15-29], situates the Kuria people in Kenya and Tanzania, outlines the historical classification of the language within Bantu, and summarizes previous research on the language. Chapter 2, "Igikuria Phonology" [30-182], begins with the segmental inventory and phonetic realizations, and gives theoretical representations of these sounds from an underspecified and autosegmental perspective. The chapter also outlines numerous prosodic shapes of verb roots, which significantly influence the surface tone pattern, and after giving examples of nearly five dozen prosodic patterns of verbs, the author lists 11 proposed rules to account for tone in infinitives. The application of these rules is then exemplified with ample derivations. Syllable structure is also discussed, assuming a X-bar model of the syllable, as are lexical rules of voicing dissimilation, post-glide lengthening, prenasalization, 'r-formation', processes of vowel coalescence, and various rules of vowel height harmony.

Chapter 3, "Igikuria Morphology" [183-334], describes the noun class system including prefix allomorphy, the nominal augment, various agreement morphemes, pronouns and determiners, and the number system. This chapter also covers verb inflection and derivation, including subject and object prefixes, and the more than four dozen tense-aspect categories of the verb.

Crass, Joachim. *Das K'abeena. Deskriptive Grammatik einer Hochlandost-kuschitischen Sprache.* (Cushitic Language Studies, vol. 23). Cologne: Rüdiger Köppe Verlag. Pp. viii, 383. 1 map, 52 tables, 1 illustration. ISBN 3-89645-480-3. Paperback. €60.00.

The Highland East Cushitic language K'abeena is described in this grammar. Chapter 1, "Einleitung" [1-14], situates the language, spoken in Wolkite, 160 km southwest of Addis Ababa, and relates it to other Cushitic languages. Chapter 2, "Phonologie [15-60], identifies the phonemes and presents minimal pairs. The vowel inventory includes a series of 5 short voiceless

allophones which have achieved the status of being quasi-surface-contrastive, and play a role in case marking, due to shifts in accent. This chapter also describes the complex system of inflectional classes and accentual alternations associated with case and other morphological categories. There are a number of phonological processes operative in the language, including vowel epenthesis which eliminates certain consonant sequences, gemination, metathesis and vowel harmony. Chapter 3, "Nomen" [61-121], describes the inflection of nouns for case, number, gender, possession and definiteness. Nouns may be singular or "transnumeral" in their citation form, and there are five inflectional groups in terms of patterns of singular, transnumeral and plural marking, all encoded by numerous suffixes. The chapter also described many noun derivational processes and the seven case forms whose variations give rise to 13 inflectional paradigms. Possessive pronouns are suffixed to nouns, creating seven additional inflectional categories.

Chapter 4, "Pronomen" [123-134], covers the various forms of personal pronouns, demonstratives and interrogatives. The fifth chapter, "Verb" [135-198], describes derivation (causative, middle, passive, reciprocal, repetitive, as well as denominal derivations, and combinations of derivational categories). Inflectional categories include perfective vs. imperfective, indicative and imperative, as well as "converbs", a kind of dependent verb form. Subject agreement reflects eight person/number distinctions as well as aspect, thus the morphological potential of a verb is considerable. Chapter 6, "Andere Wortklassen" [199-237], covers adjectives, numerals, quantifiers, relational nouns, ideophones and interjections, and chapter 7, "Adverbien und Zeitbezeichnungen" [239-243], describes adverbs and time phrases. Chapter 8, "Weitere Suffixe sowie Fokus- und Emphasemorpheme" [245-261], presents various focal and emphatic suffixes, and Chapter 9, "Syntax" [263-330], lays out basic the syntactic structure of sentences, covering both simple and complement clauses. The last chapter provides texts [331-342], and the book ends with 44 verb paradigms of 80 inflectional categories, bibliography, and numerous indices.

Dunham, Margaret. *Éléments de Description du Langi, Langue Bantu F33 de Tanzanie: Phonologie - Grammaire - Lexique*. (Langues et Littératures de l'Afrique Noire 13, SELAF n° 413). Leuven/Paris/Dudley, MA: Éditions Peeters. 2005. Pp. vi, 335. ISBN 90-429-1573-0. Paperback. €58.00

The little-known Bantu language Langi, spoken in north-central Tanzania, is described in this revision of the author's thesis at Université de la Sorbonne Nouvelle. Chapter 1, "Introduction" [6-16], describes the ethnohistory of the Langi people, the position of the language in Bantu, and the author's field research on the language. Chapter 2, "La phonologie" [18-72], covers phonemc constrasts. Section 2.2 lists phones and phonemes, consonant minimal pairs in word- and stem-initial positions and intervocalically, and gives consonant frequency tables. Section 2.3 treats the 20 phonetic vowels — length contrasts plus 10 surface vowel places, reduced to 7 phonemic vowels by eliminating [e,o] which are allophonically conditioned by high tense vowels and back [ɑ] which is conditioned by a preceding back consonant. Section 2.4 then treats syllable structure. Tone is not marked in the book, but is mentioned: there is significant final

neutralization, and there are no lexical tone contrasts in verbs. The chapter ends with examples of vowel harmony and post-nasal hardening alternations.

Chapter 3, "Le système nominal" [76-127], covers noun classes, modifier types (adjectives, demonstratives, quantifiers, pronouns, numerals), and noun class agreement. Chapter 4, "Le système verbal et prédicatif" [130-171], presents the subject and object prefixes, derivational extensions and their combinations, and gives a detailed account of the tense-aspect-mood conjugations of the language. Chapter 5, "Les fonctions syntaxiques" [174-190], covers different elements in their various syntactic functions such as subject or object; infinitives as arguments; copular constructions; adverbs; coordination and subordination; and negation. The final chapter "Annexes" gives a Langi-English-French lexicon [192-231] followed by analogous ones for French-Langi-English [232-268] and English-Langi-French [269-305], ending with a list of tree and shub names including Langi, Latin scientific names, and equivalents in various languages of Tanzania.

Kastenholz, Raimund & Anne Storch (eds.). *Sprache und Wissen in Africa. Beiträge zum 15. Afrikanistentag, Frankfurt am Main und Mainz, 30. September — 2. Oktober 2002*. (Afrikanistentage vol 8). Cologne: Rüdiger Köppe Verlag. 2004. Pp. 301. ISBN 3-89645-406-4. Paperback. €29.80

Fifteen papers from the 15th Afrikanistentag in Frankfurt are presented in this volume. Following introductory remarks by Hermann Jungraithmayr [11-16], contributions cover (Anyawu) the phonology of the Jukunoid language Wapã [17-34], (Becher) codeswitching and language creation in Dakar [35-56], (Beck & Wittman) African media and culture [57-85], (Diallo) social mobility and linguistic dynamism in Guinée [87-109], (Fleisch) the expression of emotion in Zulu and Ndebele [111-137], and (Ala-Gerull) Austronesian and African language contact in the Indian Ocean [139-158]. Further papers include (Kießling) causative and pluractional derivations in western Ring (Grassfields Bantu) languages [159-181], (Meißner) gender and number in Maa [183-195], (Möhlig) methodological issues in the use of historical texts [197-211], (Mous) properties of the causative suffix in the Cushitic language Konso [213-228], (Nanfah) language and national unity in Cameroun [229-244], (Rothmaler) a historical investigation of Borno toponymy [245-262], (Storch) word-class typology in Jukunoid Hɔne [263-278] and (Tröbs) perfective marking in Manding [279-301].

Möhlig, Wilheml J. G & Karl Peter Shiyaka-Mbereme. *A Dictionary of the Rumanyo Language*. (Southern African Languages and Cultures, vol. 2). Cologne: Rüdiger Köppe Verlag. 2005. Pp. 459. ISBN 3-89645-601-6. Paperback. €52.80

This dictionary has three major parts. The first is an overview which includes a map indicating were Rumanyo is spoken in Namibia, a list of abbreviations, a guide to using the dictionary, references, and a grammatical sketch [23-77] which outlines the phonemes, phonological processes, noun classes, noun derivation, nominal modifiers, verb agreement and tense-aspect-

mood inflection, and derivational extensions. The next part (identified as Part 1) is the Rumanyo-English dictionary [81-278], containing around 10,000 entries. Verbs are entered in stem form, whereas nouns are given in their noun class prefixed form. All data are tone marked, and entries contain noun class indications, grammatical category information, and cross-references to related entries. The final part [283-459] is the English-Rumanyo dictionary.

Storch, Anne. *The Noun Morphology of Western Nilotic.* (Nilo-Saharan Linguistic Analyses and Documentation, vol. 21). Cologne: Rüdiger Köppe Verlag. 2005. Pp. 448; 1 colored map, 3 b/w maps, numerous tables and charts. ISBN 3-89645-139-1. Paperback. €64.00.

Comparative noun morphology of Western Nilotic is studied in this book. Chapter 1, "Western Nilotic Overview" [17-28], gives an overview of the language group, listing all of the languages along with numbers of speakers, location of the language and dialect information. Chapter 2, "Documentation and Research" [29-39], gives extensive bibliographic information about each of the Western Nilotic languages, and also describes the author's fieldwork on Western Nilotic which provides the basis for the present study. The third chapter, "Theoretical and Methodological Framework" [40-58], surveys previous theories of noun classification in Nilotic and areal theories relevant to Nilotic, especially the T/K and N/K linguistic areas. Chapter 4, "Western Nilotic Comparative Phonology" [59-96], outlines aspects of the phonology of individual Western Nilotic languages and reconstructed Western Nilotic, especially giving details on the complex realization of vowels.

The core of the volume is contained in chapter 5, "Western Nilotic Noun Morphology" [97-379], which provides a detailed description and analysis of noun morphology. The subsections of this chapter concentrate on number, noun classification and derivation, the most salient morphological categories in this family. In-depth descriptions of salient features of noun morphology, constituting mini-grammars, are given for Mayak, Mabaan, Jumjum, Dinka, Nuer, Anywa, Päri, Shilluk, Luwo, Thuri, Belanda Bor, Southern Lwoo and Labwor. Chapter 6, "Emergence and Decay of Classifiers" [380-416], summarizes and synthesizes the preceding chapter, showing recurring patterns of singular and plural suffixation, linker morphemes, gender and certain derivational prefixes, and discusses sources of nominal categorization. A striking feature of these languages is the centrality of semantic classifications of nouns. The final chapter, "Ethnogrammatical and Cognitive Implications, or What May Be Classified" [418-433], discusses the relevance of cultural factors in shaping conceptualization and how this affects grammaticalization of noun classifiers.

Please also note the following online publication, *Papers in Bantu Grammar and Description*, Volume 43, 4: 2006, edited by Laura J. Downing, Lutz Marten & Sabine Zerbian. The volume can be found at:

http://www.zas.gwz-berlin.de/index.html?publications_zaspil