

## Chapter 21

# Optional ergativity and information structure in Beria

Andrew Wolfe

SIL International

Tajeldin Abdalla Adam

Ergativity in Africa is rare; König (2008: 95–96) lists only twelve African languages that have been described as exhibiting ergative phenomena. Even more rarely does OPTIONAL ERGATIVE MARKING (OEM) appear, in which the use of an ergative marker may depend on information-structural or discourse-pragmatic considerations. McGregor (2010: 1631) lists a sole instance in Africa.

Previous literature on Beria, a Saharan language, describes a system of focus marking that shows ergative alignment, wherein one marker, =*gu*, focuses transitive actors (A) and another, =*di*, focuses either transitive patients (P) or intransitive single arguments (S) (Jakobi & Crass 2004: 151–154; Jakobi 2006).

Based on new data from texts and judicious elicitation, we suggest that this =*gu* functions as an optional ergative marker which speakers employ in diverse pragmatic and syntactic contexts, not simply for assigning argument focus to the A term (Lambrecht 1994) but also for identifying an A when the P term is in focus; for highlighting brand new A participants (Prince 1981) in sentence focus contexts; for disambiguating grammatical roles; and for marking the A of quotative constructions and embedded relative clauses.

We reanalyze =*di* as a specificational copula (Mikkelsen 2005), contrasting with the predicational copula =*i*. Speakers may use either of these copulas in cleft constructions in order to focus constituents other than transitive actors.

## 1 Introduction

### 1.1 Optional ergative marking in Africa

The term ERGATIVE describes any linguistic system in which the single or principal argument of an intransitive verb (S) patterns with the patient (P) of a transitive verb rather than with the actor (A) of the transitive verb. See Figure 1.



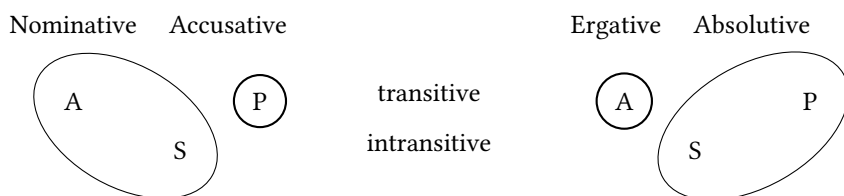


Figure 1: Ergative-absolutive vs. nominative-accusative (Dixon 1994; Payne 1997: 129–168)

In OPTIONAL ERGATIVE MARKING (or OEM; McGregor 2009; 2010; McGregor & Verstraete 2010), use or non-use of the ergative marker is conditioned by principles of information structure, discourse, or semantics. In the analysis below we follow the definition in McGregor (2009: 493):

Optional ergative case-marking refers to the situation in which the ergative marker may be present or absent from the Agent NP without affecting the grammaticality or interpretation of the clause in terms of who is doing what to who. The term ‘optional case-marking’, like ‘free variation’, is potentially misleading, and [...] does not mean that the marker is used randomly.

Only rarely does ergativity appear in descriptions of African languages. König (2008: 95–96) lists twelve languages with potential ergative phenomena, clustered in four families: West Nilotic, Kordofanian, Mande, and Chadic. McGregor (2010: 1631) offers only one of the West Nilotic languages from König’s list, Shilluk, as the lone example of OEM from the over 2,000 languages in Africa.

## 1.2 Beria

Using new data from texts and elicitation, we propose adding Beria, a Saharan language of eastern Chad and western Darfur, to this short list of African languages known to exhibit OEM.

In Jakobi & Crass’s (2004) grammar of the Kube dialect of Beria (henceforth J&C, 2004: 151–154; cf. Jakobi 2006), they present two enclitic focus markers, =*gu* and =*di*. They argue that these two focus markers follow an ergative distribution pattern: speakers employ =*gu* to mark focused A terms and =*di* to mark focused P and S terms. A third focus-marking strategy they characterize as a cleft construction, as it uses the copula =*i*. All three of these, =*gu*, =*di* and =*i*, encliticize to lexical NPs or free pronominals, though most participant reference tracking in natural discourse with unmarked pragmatics appears only in Beria’s verbal cross-reference system, wherein suffixes indicate the person and number of A (agent / actor) arguments and prefixes indicate the person and number of P (patient / object) arguments.

The present study seeks to extend and refine Jakobi & Crass’s analysis of =*gu*, =*di* and =*i* by examining contextualized discourse data with the tools of information structure

(especially Prince 1981 and Lambrecht 1994). In a preliminary review of two texts from a local news broadcast genre, we describe a previously unrecognized usage of =*gu* in contexts with *sentence focus* (Lambrecht 1994: 233–235), i.e., where the entire sentence consists of new information. Specifically, we propose that in our data, this optional use of =*gu* only appears with *brand new* transitive-agent referents (Prince 1981: 233–237). We also confirm that =*gu* may also occur in contexts with *argument focus* (Lambrecht 1994: 228–233) on A.

In addition to cases where the A term features more prominently in its context, we delineate three specific constructions that feature a *less* prominent A argument, also marked with =*gu*. All of these occur in contexts where the speaker accords higher prominence to the P argument. This prominence on the P may be pragmatic, as with contexts where P is in argument focus or serves as the subject of a pseudo-passive, or it may be syntactic, as the head of a transitive relative clause. We also identify a quotative use of =*gu*.

Turning from =*gu*, we reanalyze =*di* and =*i* as two different copulas with distinct semantics, and recast each of their argument-focus usages as variations on the same cleft construction. Copular =*i*, which J&C consider a “copula of identification,” we rebrand as a *predicational copula* (Mikkelsen 2005), and we call =*di* a *specificational copula*. In one documented difference between the two cleft constructions, OEM marker =*gu* appears on the A term in =*di* clefts but not in =*i* clefts. We suggest that this difference owes to the higher transitivity of the =*di* cleft, where the clause satisfies Hopper & Thompson’s (1980: 252) high-transitivity criterion of the *individuation of O*.

## 2 Overview of key concepts and text sources

### 2.1 Information structure

Space considerations preclude an extensive review of the literature on information structure and related subdisciplines, where terminological choices and definitions vary widely. For present purposes it is helpful to highlight Lambrecht’s (1994) definitions of focus as well as Prince’s (1981) taxonomy of new information, both of which figure into the analysis below.

Lambrecht (1994: 211) offers a shorthand definition of *focus* as that which is “unpredictable” or “non-recoverable” in any utterance. As a corollary, every sentence has some element in focus. He delineates three domains of focus: predicate focus, argument focus, and sentence focus.

The unmarked subject-predicate (topic-comment) sentence type [...], in which the predicate is in the focus and in which the subject (plus any other topical elements) is in the presupposition, will be said to have PREDICATE-FOCUS STRUCTURE; the identificational type [...] in which the focus identifies the missing argument in a presupposed open proposition will be said to have ARGUMENT-FOCUS STRUCTURE; and the event-reporting or presentational sentence type, in which the focus extends

over both the subject and the predicate (minus any topical non-subject elements), will be said to have SENTENCE-FOCUS STRUCTURE. (Lambrecht 1994: 222).

Prince (1981) sketches out a taxonomy of discourse referents in terms of newness and givenness. New discourse referents may be either BRAND NEW, where the speaker asks the hearer to create a new entity in her mental representation; or UNUSED, where the speaker invokes a referent new to the discourse but assumed to be familiar to the hearer. Given discourse referents may be TEXTUALLY EVOKED, referred to previously in the discourse itself; or SITUATIONALLY EVOKED, which Prince primarily uses for the discourse participants themselves (first and second person). Between these new and given categories lie INFERABLE referents, whose existence the hearer is able to deduce by logical or cultural implication, as the presence of a “bus” in the discourse implies the existence of a “driver.”

## **2.2 Text sources**

New data analyzed below come from publicly available Radio Dabanga newscasts from March 18, 2013 and May 1, 2013, downloaded from the Radio Dabanga website (<http://www.radiodabanga.org>) the day following their broadcast. Based in the Netherlands, Radio Dabanga (henceforth RD) broadcasts daily Darfur-related news in Arabic and various languages of Darfur. The news anchor for the Beria portion of the broadcast, Tayiba Abdelkarim Abdul, is a well-respected native speaker of the Kube dialect. For the purposes of the broadcast, she translates from Arabic. Having been given the text well in advance of her studio time, she considers beforehand how best to render the meaning of each news piece as a whole into Beria. The final recording undergoes a light editing process before broadcast.

In general, the paragraph-by-paragraph translation method and the linguistic reputation of the translator notwithstanding, some interference from the source language cannot be ruled out. For present purposes, these concerns are mitigated by two factors: Arabic – even Darfuri Arabic – has no ergative case marking; and the second author of this study, as a native Kube Beria speaker, can critically evaluate the naturalness of the data in what concerns the conclusions below. Nonetheless, future studies will need to confirm the present findings with a more extensive data set that has no foreign-language origin. Accordingly, we qualify our conclusions below as preliminary while we work to assemble a larger, more diverse corpus.

Elicited data from the present study’s second author appear marked with “TAA” below. In our discussions, elicitation never involved direct translation from English. Rather, we would discuss a pragmatic communicative context and ask what a native Beria speaker might say. Additionally, while recognizing the limitations of ungrammaticality judgments as linguistic evidence, in limited contexts below we deemed evaluations of infelicity important enough to the argument to include them.

### 3 Uses of =gu

#### 3.1 A-argument focus constructions with =gu

Based on data like that given in (1)<sup>1</sup>, J&C argue that =gu and =di are focus markers with complementary distribution, whose usage follows an ergative-absolutive alignment (2004: 151–154):

- (1) a. (Focus markers with ergative alignment Jakobi 2006)  
 bágvɔ=ɔgvɔ=gvɔ Ø:kú-gú-ɾ-í. focused A: =gu  
 wife-POSS.3SG=FOC<sub>ERG</sub> OJ:3:PFV:3-call-SJ:3-PFV  
 ‘It’s his wife who called him.’ (Jakobi 2006: 136)
- b. náá=drí nɛ-gɛr-g-ɪ. focused P: =di  
 PP:2SG=FOC<sub>ABS</sub> OJ:2-look.for-SJ:1SG-IPV  
 ‘It’s you I’m looking for.’ (Jakobi 2006: 137)
- c. sʉltán=drí Ø:nɪ-Ø-ɪ. focused S: =di  
 sultan=FOC<sub>ABS</sub> OJ:3:die-SJ:3-PFV  
 ‘It’s the sultan who has died.’ (Jakobi 2006: 137)

The enclitic =gu only appears with transitive actors. When speakers want to focus patients or intransitive arguments, they invariably cliticize =di or =i to the focused noun phrase. In light of Beria’s split-S verbal system (Mithun 1991; Dixon 1994: Chapter 4; see Jakobi 2011 for detailed analysis of the Beria facts), it bears noting that it is ungrammatical to use =gu with the S of even active intransitive verbs such as ‘run’ or ‘leave,’ which Beria codes with the A suffixes:<sup>2</sup>

- (2) \*Ai=gu suk=tu hiri-g-í.  
 1S=ERG market=DAT run-1SG.A-PFV  
 ‘It was I who ran to the market.’ (TAA)

cf. *Ai-di suk-tu hirigí.*

- (3) \*Ai=gu sur-g-í.  
 1S=ERG exit-1SG.A-PFV  
 ‘It was I who left.’ (TAA)

cf. *Ai-di surgí.*

<sup>1</sup>We have reproduced J&C’s examples as-is, with translation when in French. Note that they wrote before a working orthography was in place, so their examples use IPA transcription. Abbreviations from their work that do not follow the Leipzig Glossing Rules are listed in the Abbreviations section.

<sup>2</sup>New Beria examples are written in the working orthography approved by local administration (Faris 2006). High, [+ATR] vowels are represented with a circumflex diacritic: *û, î*. Mid, [+ATR] vowels remain unmarked, predictable from the presence of [+ATR] high vowels within the harmonic domain. Lexical tone is not written. Pluralization of both nouns and verbs uses the simplified orthographic convention of doubling the final letter of the word, capturing what is in actuality a complex system of tonal patterns. Imperfective verbs end in *-i* and perfective verbs in *-í*, reducing grammatical tone phenomena to an iconic visual form. Both rhotics are written *r*. Note that we have maintained a single representation for enclitics like =gu even though they harmonize for ATR with their host.

Because the label of ergative properly applies at the level of the particular grammatical construction, not of whole languages (Croft 2001: 132–171), it is unsurprising to find languages like Beria with both split-S and ergative alignments in operation in different components of the grammar. Note also that since =*gu* does not appear with S arguments, it indeed follows an ergative pattern and not a *marked nominative* one (König 2008: 138–203).

In fact, not even all clauses with two core arguments allow =*gu* on the more agentive argument. Certain low-transitive (Hopper & Thompson 1980) bivalent verbs such as ‘have,’ ‘learn’ and ‘know’ do not accept =*gu* but may use =*di* instead to focus the more agentive argument:

- (4) \*Ber=*gu* arabie kidí.                      A arab tigo.  
 3S=ERG car 3.A:have:PFV:SG 1S car have:1S.A:NEG  
 ‘It is he who has a car. I don’t have a car.’ (TAA)  
 cf. Ber-*di* arabie kidí. Ai arabie tigo.
- (5) jàmâl=*dɪ* àrmá Ø:áwáá-r-ɪ.  
 Jamal=FOC<sub>ABS</sub> Arabic OJ:3:learn-SJ:3-PFV  
 ‘It’s Jamal who has learnt Arabic.’ (Jakobi 2006: 139)
- (6) \*Hawa=*gu* tir=*egî*.                      ege=*gini*.                      Maha ege=*gino*.  
 Hawa=ERG name=1SG.POSS know=3A.PFV:AUX Maha know=3A.PFV:AUX:NEG  
 ‘It was Hawa who learned (lit., ‘knew’) my name. Maha didn’t learn it.’ (TAA)  
 cf. Hawa-*di* tir-*egî* ege-*gini*. Maha ege-*gino*.

Turning to information structure, among Lambrecht’s three focus categories (see §2.1), J&C have already established the use of =*gu* for *argument focus*, with focus on the A term.<sup>3</sup> Within this usage, argument focus on the A can accomplish at least two distinct purposes: asserting the *identity* of the A term, or asserting its *role*. This becomes clear in negative assertions as in (7–8). In (7), the identity of the agent is corrected, whereas in (8) it is the clause’s role assignments that are contradicted. Both contexts use =*gu*.

- (7) Bur=*do=gu* kana tene sai=*gi-n-o*.                      O                      gie=*i*.  
 boy=that=ERG NEG.PFV girl hit=3.A.PFV-AUX-NEG person other=COP<sub>PR</sub>  
 ‘It wasn’t *that boy* that hit the girl. It was someone else.’ (TAA)
- (8) Bur=*gu* kana tene sai=*gi-n-o=ru*,                      tene=*gu*.  
 boy=ERG NEG.PFV girl hit=3.A.PFV-AUX-NEG=CONJ girl=ERG  
 ‘It wasn’t that the *boy* that hit the *girl*. The *girl* hit the *boy*.’ (TAA)

<sup>3</sup>They do not use Lambrecht’s terms, but argument focus is clearly indicated in their work by cleft translations (Jakobi & Crass 2004: 151–153, examples 232–236 and 248; and in Jakobi 2006: 136, example 19). In the other four examples given of =*gu* (2004: 152, example 237; 2006: 137–138, examples 20, 21, and 26), it is unclear, based on the example and translation given, what the nature of the “focus” is. Some of these are treated below.

### 3.2 Sentence focus, brand new A construction

Not all uses of =*gu* entail argument focus, however. In both RD texts, the speaker uses =*gu* in contexts such as (9) where the remainder of the sentence is being asserted as new, not presupposed.

- (9) *barûgû* *yom\_al\_ahad=tu, oo* *toûra kuni=gu, arabie*  
 A.few.days.ago Sunday=on person:PL armed INDF.PL=ERG vehicle  
*tijari ni genî Gireda=re ji-e, je Jokhana=ru*  
 commercial INDF village Gireda=ABL COP<sub>LOC</sub>.3-CVB PROG Jokhana=DAT  
*si-r-i=gi, toûû ki-si-n-e=ge, bodo ajas=te*  
 go.IPFV.3-3A-IPFV=REL gun:PL 3.A.PFV-shoot-3.A-CVB=RSN that reason=ABL  
*oo wetti ku-nu-e, ere oo sogodî gie noko=ru*  
 person:PL three 3A.PFV-kill:P.PL-CVB again person:PL ten and one=DAT  
*hiara=gine=ii ginîi.*  
 wounded=3A.PFV:AUX:CVB=COP<sub>PR</sub>:PL 3A.PFV:say: PL

‘Last Sunday some armed people(=ERG) opened fire on a commercial vehicle that was from the town of Gireda, going to Jokhana. Because of that they killed three people, and 11 people have been injured, they said.’ [RD, 3/18/14, 5.1]

The news anchor does not expect the hearer to know there were people who opened fire on a commercial vehicle on the road from Gireda to Jokhana, nor does she merely assert that it was certain unnamed armed men who committed this deed; this is a news broadcast, and the entire utterance is unpredictable.

In the immediately following sentence, she restates and elaborates on the news just announced.

- (10) *oo Ø-nuî=gi araa=gi ni Radio Dabanga kerigi er*  
 person:PL 3.P-die:P.PL:PFV=these person:PL INDF Radio Dabanga within REL.PR  
*k-î-î=gi oo toûra malisha hakuma=ru tabi=gi*  
 3.A.PFV-say-PFV=REL person:PL armed militia government=DAT affiliated=REL  
*dî=ra hirdee=ra k-ori-e=gine=re*  
 camel:PL=CONJ horse:PL=CONJ 3.A.PFV-mount-CVB=3A.PFV:AUX:CVB=NFIN  
*arabie tijari yom\_al\_ahad tû=gi gardi genî Jokhana=re*  
 vehicle commercial Sunday evening=this road village Jokhana=ABL  
*ji-e=re je Gireda=r ke-si-r-î, toûû*  
 COP<sub>LOC</sub>.3-CVB=NFIN PROG Gireda=DAT VEN-go.IPFV.3-3A-IPFV gun:PL  
*ki-si-n-e=ge, bodo ajas=te oo wetti*  
 3.A.PFV-shoot-3.A-CVB=RSN that reason=ABL person:PL three  
*ku-nu-e, ere oo sogodî gie noko=ru*  
 3A.PFV-kill:P.PL-CVB again person:PL ten and one=DAT  
*hiara=ginîi ginîi.*  
 wounded=3A.PFV:AUX:PL 3A.PFV:say: PL

‘One of the relatives of the deceased told Radio Dabanga that the armed,

government-affiliated militia people mounted on camels and horses and opened fire on a commercial vehicle on Sunday evening on the road coming from the village of Jokhana to Gireda. Because of that they killed three people and 11 people were injured, they said.’ [3/18, 5.2]

This example well illustrates the optionality of the ergative marking as defined by McGregor and cited in §1.1: a “situation in which the ergative marker may be present or absent from the Agent NP without affecting the grammaticality or interpretation of the clause in terms of who is doing what to who” (McGregor 2009: 493). On the level of grammatical relations – and, in fact, even of sentence focus vs. argument focus – nothing changes between (9) and (10), yet =*gu* appears in (9) and not in (10). In both, ‘armed people’ (*oo toûra*) ‘shot’ (*toûû kisine*) a ‘commercial vehicle’ (*arabie tijari*) and ‘killed’ (*kunue*) people. In neither sentence is any one argument in focus. Everything that changes is irrelevant to the syntax of core arguments: the speaker adds the source of the information (‘one of the relatives of the deceased’), expands on her description of the transitive agent (characterizing them as ‘government-affiliated militia’), notes attendant action (transitive: they ‘mounted camels and horses’), and shifts some minor details (mentioning the ‘road’ and changing the aspect of ‘have been injured’ to ‘were injured’).

No salient grammatical details change, but on a second repetition the once BRAND NEW identity and role of the transitive agent have become TEXTUALLY (Prince 1981: 233–237). All four of the non-quotative tokens of sentence-focus =*gu* occur with *brand new* transitive-agent participants.

### 3.3 Constructions with backgrounded A=*gu*, foregrounded P

Another distinct pragmatically marked construction with =*gu* appears in one token in the RD texts, which fits the pattern for neither A-argument focus nor sentence focus constructions:

- (11) *genîa ha Sharq al Jabal hie=ru Yara=ra Nimra=ra Dalma=ra*  
*villages Mt. Sharq al Jebel direction=LOC Yara=CONJ Nimra=CONJ Dalma=CONJ*  
*genîa=kî=ra, oo toûra malishat hakuma=ru tabi*  
*villages=these=CONJ person:PL armed militia government=DAT affiliated*  
*begî=gu a-we gami=giniî.*  
*this=ERG go.3A.PL-CVB ambush=3A.PFV:AUX:PL*

‘The villages around Mt. Sharq al Jebel – Yara, Nimra, and Dalma – these villages too, these armed government-affiliated militiamen(=ERG) went and ambushed.’  
 [3/18, 3.1]

Here the A is not new material but *textually evoked*, as shown by the deictic *begî* ‘this, above.’ In context this sentence immediately follows the description of another ambush by these militia, in the ‘villages that lie to the northeast of Kutum’ (line 2.1). The speaker marks continuity with the previous episode here not only through anaphoric participant reference (*begî*) but also with the lexical repetition of *gami-giniî* ‘they ambushed’ and



*genîa* ‘villages.’ In other words, everything here is textually evoked except the names of the ambushed villages and the occurrence of a new incident in ‘these villages, too’ *genîa-kî-ra*.

The new village names added to the list of ‘villages ambushed by the Janjaweed’ are the unrecoverable content in sentence (11). The speaker fronts the brand new P argument and demotes the textually evoked A argument in a sort of pseudo-passive construction. As noted by Zakaria Fadoul Khidir (2005: 80), Beria has no exact equivalent to the French passive construction because the presence of passive verbal morphology in Beria requires the total omission of an agentive lexical noun phrase. To express a demoted agent, the verbal morphology must remain the same but the constituent order changes to PAV and A gets marked with =*gu*.

Sentence (11) also exemplifies a phenomenon known to the literature as *ERGATIVE HOPPING* (Haviland 1979: 155, cited in Rumsey 2010: 1657) whereby an ergative-marked A term may simultaneously function as the S argument of an intransitive verb (here ‘went and’) that intervenes between the ergative-marked noun phrase and the bivalent verb that licenses it. Three tokens of =*gu* exhibit ergative hopping.

One of J&C’s examples also falls into the pattern where P is promoted and constituent order is reversed:

- (12) bɪɛ kî=dî ábā égi=gú Ø:sí-é-ɾ-î.  
 house this=FOC<sub>ABS</sub> father my=FOC<sub>ERG</sub> OJ:3:build-PFV:3-SJ:3-PFV  
 ‘It’s this house that my father has built.’ (Jakobi 2006: 138)

Although the situational context for utterance (12) is not given, it seems likely that the A term here is part of a presupposition ‘my father built a house’ or ‘my father built something,’ putting the A ‘my father’ in the pragmatic background (i.e., not in the focus domain) and focusing on the clefted P term ‘this house.’

### 3.4 Quotative construction

Two separate quotative constructions account for the remaining tokens of =*gu* in the RD texts. In one of them, which is used in our data exclusively for *direct* reported speech, no verb of saying appears between the quoted agent and the quotation itself. The speaker uses =*gu* to signal to the hearer that she is transitioning from the A argument to a direct quotation, as in (13):

- (13) ...genî=gî kerigi oo jii=**gu**, “Ta-rdasin-e  
 village=this within person:PL COP<sub>LOC</sub>:3:PL=ERG 1PL.P-come.together-CVB  
 baa=ru ere je korekk=tu ba=gî je oû-d-í,”  
 hand:PL=INS again PROG spade:PL=INS mine=this PROG dig-1PL.A-PFV  
 gi-n-e kííí.  
 3.A.PFV-say-CVB 3.A.PFV-say-PFV:PL  
 ‘... the people who are in this village said, “We are coming together and we are digging this mine out with our hands, and again with spades.” [5/1, 4.3]

This “A – direct Quotation – quotative Verb” pattern satisfies Beria’s default APV word order, with the quotation serving as the P. Quotative OEM constructions are well documented for at least two other languages (Rumsey 2010): Ku Waru (Trans-New Guinea) and Bunuba (non-Pama-Nyungan).

The existence of a quotative =*gu* construction sheds light on one of J&C’s examples as well:

- (14) bágâ tɔ-gʊ-ε ʃɪI gɪŋ, bɛr=**gʊ** tɛbi  
 millet:PL RV-scatter-3:SJ:CVB<sub>1</sub> LC:3:AFF:P SUB PP:3=FOC<sub>1</sub> take:IMP  
 gɪ-n-I.  
 PRF:3-say-3:SJ:AFF:PRF

‘Since the millet had been scattered, he said: “Take it!”’ (Jakobi & Crass 2004: 152)

Without knowing the communicative context we cannot know for sure, but a direct quotative =*gu* reading here seems to fit more naturally than an argument focus reading.

The quotative construction appears in information-structurally diverse contexts: In (13) and the other RD token, the quoted speaker is a brand new referent and the utterance displays sentence focus whereas in (14), the quoted speaker is presumably textually evoked (by virtue of its pronominal reference) and the utterance thus has predicate focus.

### 3.5 Subordinate constructions with =*gu*

In the other quotative construction the quotation is introduced with a relative clause ‘what A said.’ In this relative clause, the A argument appears marked with =*gu*, as in (15; quotation omitted for space):

- (15) o wakil\_amin\_al-am iga Umam\_al\_Muttahida hifz\_al\_salaam  
 person Undersecretary.General mission United.Nations peacekeeping  
 kî tir-ogo Hervé Ladsous gine\_îrî, [o=kî=**gu** er  
 GEN name-3SG.POSS Hervé Ladsous is.called person=this=ERG REL.PR  
 k-î-î]=gi, [...] gine\_kîi.  
 3A.PFV-say-PFV.SG=REL [...] he.said

‘The UN Under-Secretary General for Peacekeeping, whose name is Hervé Ladsous, what this person said: [“”], he said.’ [5/1, 5.1]

In both such tokens in our data, a brand new participant enjoys an elaborated introduction from the speaker, which is topicalized at the beginning of the utterance. The speaker then proceeds to preface the participant’s quoted speech act with the short relative clause ‘what he said.’ This relative clause uses a reduced noun phrase to refer to the quoted participant: ‘this person’ *o-kî* in (16), and in the other example even more simply ‘he’ *ber*. Such a reduced form of reference signals that the quoted participant has become an *evoked*, less prominent entity; critically it allows the speaker to pivot from the communicative purpose of introducing the participant to that of telling the hearer what it was that the participant said.

The ergative marking of A constituents embedded within relative clauses extends beyond quotative uses. In (16), *tene* ‘girl’ functions as both the patient of the verb *kidigari* ‘loves’ in the embedded relative clause and as the single argument of the verb *kari* ‘came’ in the matrix clause. As the pivot, or shared argument between the two clauses, the P argument enjoys greater syntactic prominence than the A term.

- (16) Tene [bur=*gu* ki-dig-a-r-*i*]=do ka-r-*i*.  
 girl boy=ERG PFV.3-love-PFV.3-3.SG-PFV:SG=that come-3.A-PFV:SG  
 ‘The girl [that the boy loves] came.’ (TAA)

### 3.6 Summary

Table 1 shows a taxonomy of =*gu* constructions identified to this point. Two of these constructions, shown at bottom – the disambiguation of roles shown in ‘it wasn’t that the *boy* hit the *girl*; the *girl* hit the *boy*’ and the signaling of transition from quoted speaker to quoted speech – have no discernible function in assigning prominence, whether information-structural prominence or syntactic prominence. Rather, their only function is to disambiguate roles and constituents. This in itself provides the strongest argument for characterizing =*gu* as primarily a marker of ergativity and not as a marker of focus.

Common to the remaining constructions is the assignment of *marked prominence* to the A term of a transitive verb, whether higher prominence, as in the cases of sentence focus with a brand new A and of argument focus on A, or lower prominence, as in the cases of A embedded in relative clauses or of PAV word order (due to argument focus on P or to pseudo-passivization).

Table 1: Summary explanatory matrix of constructions in which A is marked with =*gu*

assigning prominence	higher prominence of A	main	argument focus on A (§3.1, (7)) sentence focus on brand new A (§3.2)
	lower prominence of A	main (PAV)	clefted P with argument focus (=di), A in presupposed material (§3.3, (12)) P as subject of pseudo-passive, evoked, demoted A (§3.3, (11))
disambiguating	disambiguating roles (§3.1, (8)) quotative, signaling transition to speech complement (§3.4)	relative	P as pivot of clause linkage, A within relative clause (§3.5)

This schema provides a clear account of all data points in our RD texts as well as elicited data, and may also elucidate examples given in previous literature that did not seem to express argument focus on the A term. One significant line of evidence remains unresolved, however: native speaker intuitions as to the function of =*gu*. The next section briefly turns to this before moving on to =*di* in §4.

### 3.7 Native speaker intuition: =*gu* as *deixis of role*

When Zakaria Fadoul Khidir, a native speaker of Kube Beria, discusses =*gu* under the category of “passive voice” (2005: 80; see §3.3 above), he glosses it as a “deictic” marker. While at first glance this does not seem to integrate easily with the analysis above, it well captures native speakers’ intuitions about =*gu*, as the second author of this study attests and as is further confirmed by Amir Libiss (p.c.), a third Beria speaker.

Beria already has two sets of deictic markers, proximal =*kî* / =*gi* and distal =*to* / =*do* (Jakobi & Crass 2004: 126), both with wide syntactic distribution. These markers can even co-occur with =*gu*, as in the noun phrase *o-kî-gu* ‘this person=ERG’ in (15) above. Clearly =*gu* does not encode mere referential deixis.

At the same time, a deictic reading of =*gu* accords well with the broader social and cognitive purpose of deixis and demonstrative marking. If, as Diessel (2006: 463) argues, “demonstratives function to coordinate the interlocutors’ joint focus of attention,” =*gu* could be understood not perhaps as pointing at a referential entity per se, but as pointing at that entity’s grammatical relation within the transitive clause. By using =*gu*, a speaker “points” and invites his hearer to focus attention on the agentive grammatical role played by the indicated referent. Deixis and argument focus perform related social and cognitive functions in terms of coordinating interlocutors’ attention.

Of course, as argued at length above, =*gu* appears in more diverse constructions than just argument focus on the A term. Nevertheless this does not conflict with the singling out of A-argument focus as the prototypical use of =*gu*. By way of comparison, if asked to define the word ‘that,’ most English speakers would likely point – yet it is no less true that English ‘that’ has also been grammaticalized to serve in discourse-deictic functions and even as a relativizer. In fact, one would *expect* the evolution and grammaticalization of a deictic marker into such varied additional constructions (Diessel 2006, §4.2).

With this foundation laid, we propose DEIXIS OF ROLE as a descriptive term for the prototypical function of optional ergative marking in Beria. This not only incorporates native speaker insights about how Beria OEM works but it may provide fruitful directions for future OEM theorizing and research.

To fully comprehend the function of Beria’s =*gu*, it is also necessary to understand other focus markers in its environment, especially =*di*, which was previously described as an absolutive focus marker (Jakobi & Crass 2004: 151–154; Jakobi 2006). §4 turns to this.

## 4 Copulas and clefts with =*di* and =*i*

J&C characterize =*di* as an absolutive focus marker because of data such as in (1) above. To assign argument focus to the A term, Beria speakers use =*gu* but to focus P or S, they select =*di*. We reframe =*di*, however, as a previously unrecognized copula, which can then be employed in a cleft construction.

The strongest rationale for this shift is that Jakobi recognizes the existence of a “non-verbal predication marker” (2006: 138) =*di* sharing a presumed “common origin” with =*di* the absolutive focus marker. These two are in fact formally identical, as shown in (17-18, cf. (1b) above):

- (17) O      *kese-r-ì=gi*                      *Tayiba Abdelkarim Abdul=di.*  
 person speak-3A-IPFV.SG=REL *Tayiba Abdelkarim Abdul=COP<sub>SP</sub>-3*  
 ‘The person who is speaking is Tayiba Abdelkarim Abdul.’ [5/1, 2.2]
- (18) *ár=dì.*  
 PP:1SG=PRED  
 ‘It’s me.’ (Jakobi 2006: 138)

Furthermore, in argument focus constructions, =*di* shares common distribution with =*i*. Both =*di* and =*i* may focus S and P arguments but not A arguments. On the other hand, in our data, neither of these markers shows the complex and varied distribution summarized for =*gu* in Table 1 above.

As copular verbs, both =*di* and =*i* share a negative form, =*do*, while =*gu* has no negative form (cf. 7-8):

- (19) *ǵim=d-ō.*  
 owl=IC-NEG  
 ‘It’s not an owl.’ (J&C 2004: 101) (compare 23)
- (20) O      *arabie kidí=do*                      *ber=d-o*                      *kire=ego=di.*  
 person car      3.A:have:PFV:SG=REL 3S=COP<sub>SP</sub>-NEG brother=3SG.POSS=COP<sub>SP</sub>  
 ‘The person who has a car is not him; it is his older brother.’ (TAA)

Semantically, a clean division of labor exists between =*i* and =*di*. The first classifies or ascribes an attribute to the topic: X displays the characteristic Y (21-22), or X is a member of the set Y (23-24):

- (21) *m̄sā h̄err=i.*  
 pot full=IC:3:AFF  
 ‘The pot is full.’ (J&C 2004: 100)
- (22) *b̄ɔɔ=éǵí                      tákk̄ɔ̀r̄e=i.*  
 ram:PL=1S.POSS very.fat=IC:3:AFF  
 ‘My rams are very fat.’ (J&C 2004: 100)

- (23) g̃imm=i.  
 owl=IC:3:AFF  
 ‘It’s an owl.’ (J&C 2004: 100)
- (24) tàmarā=i.  
 Tama.person=IC:3:AFF  
 ‘It’s a Tama.’ (J&C 2004: 100)

The second marker predicates co-extensive reference between the two terms: X is fully described by Y and there are no other members in the set; see (17-18) above. Accordingly, following Mikkelsen (2005), we label =i a predicational copula, and =di a specificational copula. As she explains (2005: 1), predicational copulas “tell us something about the referent of the subject” whereas specificational copulas “says who or what the referent is.”<sup>4</sup>

These semantics for =di and =i carry over into their corresponding versions of the cleft construction. In (12) above, for instance, the speaker presupposes ‘a certain house exists which my father built’ and, through the specificational cleft, he predicates that ‘this is that house.’ The two referents of ‘this house’ and ‘the house that my father built’ are co-extensive. In (25), by contrast, the speaker assumes ‘a certain substance that the mother gave to the child’ and then through an predicational cleft, clarifies the nature of that substance. What sort of thing is that which the mother has given the child? It is of the class of ‘paste.’

- (25) g̃uɔ=i                      r̃à            jàá=r            Ø:kɛkk-i.  
 paste=COP.SG.ABS mother child=ADV OJ:3:give:SJ:3-PFV  
 ‘It’s paste that the mother has given to the child.’ (Jakobi 2006: 139)

One complicating factor in this rendering of the facts is the presence of =gu in the =di cleft in (14) but not in the =i cleft in (25). We propose to resolve this complication with reference to Hopper & Thompson’s (1980) scalar transitivity criteria. In §3.1, scalar transitivity already helped explain why low-transitive verbs ‘have,’ ‘learn’ and ‘know’ cannot occur with ergative-marked agents. If =gu correlates with high transitivity, then Hopper & Thompson’s criterion of the individuation of O (1980: 252–253) may motivate the use or non-use of =gu in these two sub-constructions of the Beria cleft. The focused element of a =di cleft is invariably a referential, definite expression like ‘this house’ whereas in =i clefts it is a non-referential, indefinite class or attribute like ‘paste.’ In other words, the fronted P or O in a =di cleft is *individuated* whereas in a =i it is not.<sup>5</sup>

Distributional, morphological and semantic lines of evidence converge to support the assertion that =di is a copula. The only counterevidence we are aware of finds natural explanation in the dynamics of scalar transitivity.

<sup>4</sup>Jakobi & Crass (2004: 102–106) catalogue two other Beria copulas: locative/existential *ji* and comitative *bei*.

<sup>5</sup>In fact, the *individuation of O* may also be relevant to the quotative =gu construction (§3.4). Rumsey (2010) theorizes that the tendency of quotative OEM marking to apply more to direct reported speech than to indirect stems from this same *individuation of O* criterion, since the shift of voice and perspective involved in a direct speech act sets it off more sharply from the framing speech act; it is more clearly individuated than indirect reported speech.

## 5 Directions for further research

Directions for future research on OEM phenomena in Beria abound. As discussed in §2.2, the present study, which represents the first results of an ongoing program of Beria-language corpus collection and discourse analysis, necessarily offers only preliminary conclusions due to the small corpus size and its foreign-language source material. With a larger annotated corpus, a higher-resolution picture will emerge of what additional factors may influence speakers' choices of when to use =*gu*, potentially including animacy, discourse macrostructure, activation status, unexpectedness of agency, aktionsart, and zero anaphora of other core constituents, among others. More robust explorations of relative clauses, quotative constructions, topic chains, and ergative hopping will complement this fuller analysis.

Another open question is why Jakobi & Crass's two folk tale texts (2004: 185–192) include no instances of =*gu*, even in contexts with transitive predicates and newly introduced participants, such as in (26):

- (26) sàgʊr tɛnɛ tɛbr-ɛ-ɾ-ɛ  
 jackal girl take-PRF:3-3: SJ-CVB<sub>1</sub>  
 'a jackal took a girl, then...' (Jakobi & Crass 2004: 185)

We suggest that the lack of =*gu* here owes in part to the hearers' assumed familiarity with stock folk tale participants, which among the Beria speech community would perhaps better be classified as *unused* participants than *brand new* ones. This is consistent with the fact that J&C's storyteller omits, throughout the story, lexical noun phrases that would clarify the identity of referents for the benefit of uninitiated readers. J&C supply these identities in square brackets in their French translation – 21 times in the first four-page text alone. Considerations of genre and register may also be playing a role here: we expect to see =*gu* especially in other formal settings calling for precise speech, such as testimony in traditional courtroom proceedings or recitation of cultural history unfamiliar to hearers. At least, if anything, the lack of =*gu* in J&C's texts would seem to *strengthen* the *optional* ergativity analysis of =*gu* given above.

The diachronic origins of =*gu* also remain opaque. Descriptions of case marking for Western Saharan (WS) languages (surveyed in König 2008: 38–57) demonstrate similar optional case marking to Beria's: Core arguments go unmarked in default APV constituent order, and the presence of case marking is influenced by "syntactic, semantic and pragmatic factors" that are "yet to be adequately described" (Hutchison 1986: 193). While WS's "nominative" marker, *ye* in Kanuri-Kanembu and *i* in Teda-Daza, bears little formal resemblance to Beria's =*gu*, its usage as an optional marker on A terms is not unlike =*gu*. In fact, it is not certain that *ye* / *i* indeed functions as a nominative marker uniformly across WS. Hutchison (1986: 203–205) does establish that in some dialects of Kanuri, *ye* may mark an S term – though rarely, and only in narrative discourse when the subject switches – but in Keshirda Dazaga, for instance, Josiah Walters (p.c.) has yet found no tokens of an S marked with *i*. Possibly OEM is not uncommon in the Saharan family.

## 6 Conclusions

Close scrutiny of =*gu* data in Beria’s Kube dialect yields much of interest to the typologist, discourse researcher, and Africanist alike. In the analysis of two radio news broadcasts and a handful of carefully elicited examples, we have teased apart as many as seven distinct constructions in which =*gu* marks only the A term of transitive clauses, and never P or S terms. Speakers employ =*gu* in these constructions to various communicative ends, including the disambiguation of roles and the assigning of either higher or lower prominence to a transitive agent. What brings unity to this diversity is the ergative nature of =*gu*: It is indeed foremost a case marker and not a focus marker. At the same time, it is essentially an *optional* ergative marker, the use of which varies according to syntactic and discourse-pragmatic principles. This makes it only the second such system described as such on the African continent.

Furthermore, we offered multiple lines of evidence showing that =*di*, which has been described as an absolutive focus marker, is in fact a previously unrecognized copula. We contrasted this specificational copula =*di* with the predicational copula =*i*. Both of these may be used in a clefted construction to focus an S or P argument. This reconfiguration may be summarized as in Table 2.

Table 2: Summary of reanalysis of Jakobi 2006

Jakobi & Crass 2004; Jakobi 2006			
= <i>gu</i>	<b>focus</b>	focus marker, ergative	optional ergative marker
= <i>di</i>	<b>focus</b>	focus marker, absolutive	specificational cleft
		non-verbal predication marker	specificational copula
= <i>i</i>	<b>focus</b>	cleft construction	predicational cleft
		copula of identification	predicational copula

In addition, this analysis lends further support to multiple components of Rumsey’s (2010) account of OEM phenomena in Ku Waru (Trans-New Guinea) and Bunuba (Non-Pama-Nyungan), especially in its reliance on scalar transitivity, particularly the *individuation of O*, in the description of a quotative OEM construction, and in the recognition of *ergative hopping* phenomena, in a language far removed from Ku Waru and Bunuba both geographically and genetically.

Finally, drawing from Khidir (2005) and Diessel (2006) we explored reframing Beria’s OEM as a variety of deixis we called *deixis of role*, because in its prototypical usage =*gu* “points” to the grammatical role of the ergative-marked noun phrase, directing a listener’s focus and attention to that agent role.

Much work remains to be done to understand Beria information structure as a whole, beginning with the assembly of a significantly larger and more diverse corpus, but the present study represents some basic steps forward toward this end. McGregor (2009: 1626) laments that the study of optional case marking suffers from “not enough data, not



enough ideas.” It is our hope that the present study makes a modest contribution toward addressing these lacunae.

## Acknowledgements

We gratefully acknowledge insightful responses given after our initial presentation at ACAL, especially from Colleen Ahland, Michael Ahland and Malte Zimmermann, who pointed us in helpful directions for both clearer presentation and further literature review and engagement. Thanks also to Ryan Pennington, who tipped us off to a number of particularly valuable references. Special thanks to Josiah Walters who shared unpublished data from Keshirda Dazaga. We are grateful to Amir Libiss, another Kube speaker who reviewed our analysis and offered insights. Thanks finally to Deborah Morton for reviewing an early draft and offering a multitude of incisive comments. All mistakes are, of course, ours.

## Abbreviations

Abbreviations follow the Leipzig Glossing Rules, except for the following:

A	agent of transitive verb or agentive single argument of intransitive verb	NFIN	non-final verb form
A <sub>IMPERS</sub>	impersonal agent	O	object
ADV	adverbializer	OJ	object marker
AFF	affirmative	P	Patient of transitive verb or patientive single argument of intransitive verb
CONJ	conjunctive	PP	person pronoun
COP <sub>PR</sub>	predicational copula	PRED	verbless predicate
COP <sub>SP</sub>	specificational copula	REL	relativizer
COP <sub>LOC</sub>	locative copula	REL.PR	relative pronoun
IC	identificational copula	RSN	reason clause
IPV	imperfective	RV	reduced valence
LC	locative copula	SA	agent-like subject
LM	lexical morpheme	SJ	subject marker
MAN	manner	SP	patient-like subject
MED	marker of medium verbs	SUB	subordinator
MID	middle voice	V	verb
		VEN	venitive

## References

- Croft, William. 2001. *Radical construction grammar*. Oxford: Oxford University Press.
- Diessel, Holger. 2006. Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics* 17(4). 463–489.

- Dixon, R. M. W. 1994. *Ergativity* (Cambridge Studies in Linguistics 69). Cambridge: Cambridge University Press.
- Faris, David. 2006. *Beria-English English-Beria dictionary*. 2nd edn. Iriba, Chad: Association pour le Développement Économique et Social du département de Kobé (ADESK).
- Haviland, John Beard. 1979. Guugu Yimidhurr. In R. M. W. Dixon & Barry J. Blake (eds.), *Handbook of Australian languages*, 27–180. Canberra: ANU Press.
- Hopper, Paul J. & Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56(2). 251–99.
- Hutchison, John P. 1986. Major constituent case marking in Kanuri. *Current approaches to African linguistics* 3. 191–208.
- Jakobi, Angelika. 2006. Focus in an active/agentive alignment system – the case of Beria (Saharan). *ZAS Papers in Linguistics* 46. 129–42.
- Jakobi, Angelika. 2011. Split-S in Beria. In Doris Löhler, Eva Rothmaler & Georg Ziegelmeyer (eds.), *Kanuri, Borno and beyond: Current studies on the Lake Chad region* (Topics in Interdisciplinary African Studies 22), 91–116. Cologne: Rüdiger Köppe Verlag.
- Jakobi, Angelika & Joachim Crass. 2004. *Grammaire du beria (langue saharienne)* (Nilo-Saharan: Linguistic Analyses and Documentation 18). Cologne: Rüdiger Köppe Verlag.
- Khidir, Zakaria Fadoul. 2005. *Bases et radicaux verbaux: Déverbatifs et déverbaux du beria (langue saharienne)*. Köln: Rüdiger Köppe.
- König, Christa. 2008. *Case in Africa*. Oxford: Oxford University Press.
- Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus, and the mental representation of discourse referents*. Cambridge: Cambridge University Press.
- McGregor, William B. 2009. Typology of ergativity. *Language and Linguistics Compass* 3(1). 480–508.
- McGregor, William B. 2010. Optional ergative case marking systems in a typological-semiotic perspective. *Lingua* 120. 1610–36.
- McGregor, William B. & Jean-Christophe Verstraete. 2010. Optional ergative marking and its implications for linguistic theory. *Lingua* 120. 1607–1609.
- Mikkelsen, Line. 2005. *Copular clauses: Specification, predication and equation*. Amsterdam: John Benjamins.
- Mithun, Marianne. 1991. Active/agentive case marking. *Language* 67(3). 510–46.
- Payne, Thomas Edward. 1997. *Describing morphosyntax*. Cambridge: Cambridge University Press.
- Prince, Ellen. 1981. Toward a taxonomy of given-new information. In Peter Cole (ed.), *Radical pragmatics*, 223–55. New York: Academic Press.
- Rumsey, Alan. 2010. ‘optional’ ergativity and the framing of reported speech. *Lingua* 120. 1652–76.