

Chapter 5

All-in-one and one-for-all: Thetic structures in Buli grammar and discourse

Anne Schwarz

Universidad Regional Amazónica IKIAM

This paper describes the form and function of the thetic statement and its categorical counterpart in the north Ghanaian Gur language Buli. Instead of applying cross-linguistically well-known means to encode thetic statements, such as intonation and word order, a connective particle is positioned at the beginning of the verb phrase and followed by a dependent verb form or a noun in predicative function. This thetic construction may contain a prosodic break between the strictly preverbal subject constituent and the predicate. Interestingly, the structural features of thetic statements are also encountered in various embedded clause types and give way to different stacking configurations. In terms of function, it is argued that the internal lack of information structure renders the thetic statement compatible with a number of different pragmatic situations and thus contributed to its wide distribution in Buli grammar.

1 Introduction

The thetic/categorical opposition is a useful information-structural distinction in some languages, in particular in those that have developed special grammatical features for pragmatically marked and unmarked topics. However, despite its information-structural potential, the thetic/categorical distinction is only rarely considered in the functional-typological literature concerning African languages (but cf. Güldemann 1996; 2010; 2012; Fiedler et al. 2010: 250-253; Schwarz 2010a; beyond Africa, see in particular Kuroda 1972; Sasse 1987; 1995; Lambrecht 1987; 2000; Ulrich 1988).

The thetic/categorical distinction in linguistics was borrowed from philosophy in the 19th century when Brentano and his student Marty proposed a fundamental dichotomy between two different basic types of logical statements (Sasse 1987: 535). Categorical statements are those with a classical bipartite subject-predicate structure in the Aristotelian sense. These statements involve two successive acts: naming an entity and then



making a statement about it. Thethetic statement, on the other hand, is internally unstructured and expresses an event, a state or a situation without the concept of a topic that is commented about. Kuroda (1972) was one of the scholars who applied this statement classification in linguistics in order to describe the distribution of certain Japanese morphemes. Ulrich (1988: 388) states that a distinction between categorical andthetic constructions is universal, but that these constructions are expressed in different ways by different languages. Since the recognition of thethetic/categorical opposition provides new insights into the intricate distribution of certain sentence construction types in Buli, the distinction between statements that are mentally construed as unstructured (thetic), versus structured into topic and comment (categorical), has also been adopted here (Table 1, cf. also Ulrich 1988).

Table 1: Thethetic/categorical distinction

	Categorical	Thetic
Philosophical level	gives a statement about an argument, predicative linkage	presents a fact as an inseparable unit
Pragmatic level	topic-comment structure	unstructured (without the notion of a topic to be commented on)

Transferred to the pragmatic level, the bipartite structure of categorical utterances consists of a topic and of a comment that predicates something about it. In thethetic statement, on the other hand, the propositional information is conceptualized as a single unstructured package. Unlike categorical statements,thetic statements have no internal information structure and have been described as monomial predications where “the entire situation, including all of its participants, is asserted as a unitary whole” (Sasse 1995: 4f).

The term TOPIC is used here to refer to “what the utterance is about,” thinking of the topic as a textual location or storage address for subsequent information. Reinhart (1982) uses a file card metaphor according to which the topic corresponds to the title of a file card to be followed by the comment. Functioning as a storage address, it is a characteristic feature of topics to be established early in the discourse to be then commented on. Accordingly, the subject often has the (unmarked) topic role. In coherent discourse, topics tend to display certain continuity and thus often represent information already shared between the interlocutors. Languages may therefore provide particular (pronominal) morphology or allow even zero forms for encoding such topic continuity. Accordingly, it is possible to encounter categorical statements that have a conceptual topic about which the speaker comments, even though the topic itself is not mentioned in the sentence.

In intonation languages, such as English, the formal distinction between thetic and categorical statements often relies on accentuation. This is exemplified in (1)-(2). In the unstructured thetic utterance there is only one accent (marked in small capital letters) which is placed on the subject. In the pragmatically bipartite categorical utterance, on the contrary, both subject and predicate are accented.

- (1) English thetic utterance (unstructured):
The CHICKEN's burnt.
As a reply to 'What the hell is this stench?' (Sasse 1987: 529)
- (2) English categorical utterance (bipartite topic-comment structure):
The CHICKEN's BURNT.
In the context of 'Where is the roast chicken I expected, why do I have a hamburger on my plate?' (Sasse 1987: 529)

Another well-known linguistic means for the expression of the thetic/categorical distinction is word order change. In Modern Greek, for instance, thetic statements display subject inversion. The non-topical subject in (3) follows the verb rather than preceding it; the latter is, however, the case in the categorical statement in (4).

- (3) Modern Greek thetic utterance:
Xtipise to telefono.
rang ART phone
'The PHONE rang.' (thetic) (Sasse 1987: 536)
- (4) Modern Greek categorical utterance:
To telefono xtipise.
ART phone rang
'The PHONE RANG.' (categorical) (Sasse 1987: 536)

In terms of discourse functions, it is obvious that the categorical organization of the utterance particularly serves incremental information transfer, as the conceptual presence of a topic provides a cognitively useful dedicated storage address for any new information. Categorical statements seem to represent the dominant pattern in several communicative situations in coherent discourse, and are therefore overall pragmatically less marked than the unstructured thetic organization. Thetic statements, on the contrary, can be used to deliberately deviate from discourse coherence, for instance, by presenting an unexpected event in episode-initial position. Cross-linguistic studies (Sasse 1995) have also confirmed that certain lexical classes of predicates are typically not used to comment on topics across languages. Weather verbs are particularly well-known in this respect. Cross-linguistically we can therefore expect categorical statements to represent the basic communicative model in most coherent discourse sections, while thetic statements might be more appropriate at scene-initial points and with particular predicate types. Following Güldemann (2010: 86), we will suggest that when the speaker has the choice and applies the thetic rather than the categorical configuration, she deliberately

uses this “grammatical device to CANCEL the sentence-internal information structure that is induced by the morphosyntax of an unmarked sentence.”

The remainder of the paper describes the basics of the thetic/categorical distinction in Buli. §2 provides a brief introduction to the language. §3 deals with information structure in Buli and first outlines the language-specific distinction between the two configurations (§3.1), followed by more detailed structural accounts of categorical (§3.2) and thetic statements (§3.3). §3.4 deals with the pragmatic range of situations for the application of thetic statements and §3.5 discusses the embedding of corresponding structures. A short summary is provided in §4.

2 The language

Buli belongs to the Gur language family and is spoken by approximately 150,000 people (Lewis, Simons & Fennig 2015) in northern Ghana. There are three dialectal zones, two of them with further minor internal variations. The data used for illustration here all stem from the central zone and were recorded with speakers from Sandema and Wiaga between 1996-2009. Within the Gur family, Buli is classified as a member of the Oti-Volta branch, in which it forms, together with its closest relative and south-western neighbour Kɔnni, the Buli/Kɔnni subgroup (Manessy 1979; Naden 1989). Buli shares several linguistic features with other Gur languages, but also reveals various peculiarities. Like most Gur languages, it is a tone language and has a three-tone contrast (High, Mid, Low). Deviating from several relatives, though not from its sister language Kɔnni (Cahill 2007), Buli verbs have no lexical tone and the grammatical relevance of tone is fairly high. Both lexical and grammatical tones are active beyond their primary tone bearing unit (TBU): High tones extend their domain in certain environments, and in most dialects, Low tones spread rightward into High TBU's, but spare Mid ones (Schwarz 2003; 2007). Grammatical affixes and clitics are attested both with and without inherent tone. Tone alignment patterns are complex and depend on lexical, grammatical, and sometimes syntactic variables. In addition to transparent tone patterns which only change in their surface realization, there are also some nouns and verb forms which display an “unstable” rising tone pattern, realized as LM on bimoraic segments in phrase-final position. However, the final rise often gets completely lost, unlike the High component in other tone patterns which remains traceable even under tone spreading conditions. This and other phenomena point to intonational influences in grammatical and lexical tone.

Buli is a head-marking language with a morphological type that is mildly agglutinating, predominantly with suffixes in the nominal domain and with mostly proclitics in agreement. The basic word order is SV(O)(OTHER). The verb system is aspectually organized, with the perfective representing the unmarked aspect for dynamic verbs. Serial verb constructions are common and fulfil voice as well as particular tense-aspect-mood-specifications. The nominal domain is characterized by a fully operational and mostly overt noun class system. Nouns display suffixes and agreement operates primarily phrase-externally employing proclitics or prefixes. Two nominal forms with different suffix sets exist. In singular, only the definite forms have fully overt noun class suffixes,

while the indefinite forms often display just suffix traces (see Schwarz 2007 for more details on the noun class morphology).

3 Information structure in Buli

Buli grammar closely interacts with information-structural categories. Under certain conditions, for instance, it seems that the morphological marking of a focal constituent within the sentence is grammatically required, i.e. its lack would yield an ungrammatical sentence. This situation and the dominance of morphosyntactic means for information-structural encoding is reminiscent of other African languages for which grammaticalized focus systems are reported (Hyman & Watters 1984; Robert 2000; Wolff 2005).

With respect to the information-structurally controlled morphosyntactic variation in Buli, we observe major encoding differences in terms of different sentence constructions. Long-lasting studies concerning this variation, taking data into account from both natural discourse and more controlled elicitation tasks (as those by Skopeteas et al. 2006 and Schwarz 2007), suggest that it is not just sentence-internal information-structural categories such as topic and focus that are systematically morphosyntactically encoded, but rather the distinction between thetic and categorical statements in order to structure and organize discourse above the simple clause-level.

3.1 The thetic/categorical distinction in Buli

In Buli, the propositional content of an utterance can be expressed in at least two formal ways: either in a simple categorical form as in (5), or in a thetic statement as in (6). Note that the different subject forms in the sentences used here for illustration – a definite subject noun in the categorical statement and an indefinite subject noun in the thetic one – are not mandatory but reflect some typical features to be discussed further below.¹

(5) Categorical (simple):

Nípōōwá ɲòb kà túé.
 nípōk=wá ɲòb ká túé
 woman=DEF1 eat KA bean.NC6

‘The woman ate beans.’ [BL (1a-trans) 2004: 28.01]

¹ The representation of examples contain a first data line that follows common orthographic practice to some degree, but also marks surface tone. Most examples contain a second data line that provides morpheme segmentation and underlying tones (wherever these differ from orthographic practice) and uses IPA symbols with those consonants that could otherwise be misinterpreted. Unpublished data is from my fieldwork. The examples stem from audio-recorded narratives, conversations and contextualized elicitations and translations. The source provided in square brackets at the end of an example free translation represents the name and number of my Toolbox record under which they are stored.

- (6) Thetic:
 Nípōk lē ηðb tùē.
 nípōk lē ηðb túē
 woman.NC1 CON eat bean.NC6
 ‘A woman ate beans.’ [BL (2) 2005: 454-2]

Both sentences display SVO order, but they differ in two principled ways: In simple (e.g., monoclausal) categorical statements, the verb is subject to regular inflection including tonal subject agreement where applicable. In the affirmative perfective, the inflection includes tonal differences on the verb depending on whether the subject is a speech act participant, or not (Schwarz 2007). Furthermore, in simple categorical statements in the affirmative, there is usually a postverbal particle *ká*. This particle marks the left edge of a postverbal constituent, either one immediately after the verb or one more distantly following the verb. In (5), *ká* precedes the object noun and is subject to Low-tone spreading that extends from the grammatical Low tone of the immediately preceding verb.

In the thetic statement on the other hand, the subject is followed by a predicate-initial connective particle *lē* (allomorph *nē*) after which the content verb follows. Perfective content verbs do not display tonal subject agreement but are characterized by the unstable rising tone. Since the verb in (6) is followed by an object noun, it remains without trace of the final rise. The Low tone of the verb form (*ηðb*) spreads right onto a following High TBU, just as would happen from any regular Low tone verb. Note also that in the thetic statement, the postverbal particle *ká* cannot occur.

The formal dichotomy between simple categorical and thetic statements is schematically outlined in Table 2. The properties of both configurations are discussed in more detail in §3.2–§3.3.

Table 2: Basic categorical/thetic distinction for a Buli SVO sentence in the affirmative perfective

Categorical (simple):	S (* <i>lē</i>)	V (subject agreement by tone, if applicable)	<i>ká</i>	nominal O
Thetic:	S <i>lē</i>	V (no subject agreement; unstable rising tone)	(* <i>ká</i>)	nominal O

3.2 Simple categorical encoding

The unmarked sentence construction in Buli usually correlates with a categorical reading in which the subject has the unmarked topic role about which the predicate (with its further arguments and adjuncts) comments. In the absence of subject nouns or disjunctive

provides circumstantial information in the postverbal domain by appearing at the left edge of this constituent and setting it literally apart from the preceding information. There seems to be a preference for marking the sentence-final constituent, but if there is more than one, the particle remains flexible with respect to which postverbal constituent it selects. In the case of multiple postverbal constituents, the particle is thus reminiscent of an argument focus marker. On the other hand, if the lexical verb or the truth value of the sentence is in focus, the particle *ká* still appears in postverbal position and not before the verb that represents (part of) the focal information. In such predicate-centered focus conditions, the particle combines with an extension (*-mā*) which emphatically confirms the state of affair (Schwarz 2010b). Example (9) illustrates the use of particle *ká* with the extension, here following an intransitive stative verb.

- (9) Jāāmú wōŋā kāmā pāā.
 dʒāa-mú wōŋa ká-mā pāa
 thing-DEF14 be.long KA-CONF much
 ‘This story is really long!’ [Jein 43]

Categorical encoding in Buli also applies to nominal predicates (cf. also Schwarz 2009b). The predicate in (10) consists of the nominal compound *mí-wòŋ* ‘long rope’ which comments on a topical subject noun. The categorical configuration is expressed by the particle *ká* marking the left edge of the constituent in predicative function.

- (10) Mííkádé ká mí-wòŋ.
 míi-ká-dé ká mí-wòŋ
 rope-DEF12-PROX KA rope-long.NC12
 ‘This rope is long.’ (Schwarz 2009a: 267)

Some aspects of the categorical encoding with nominal predicates are remarkable. In the absence of a verb, the invariable particle resembles a copula element that signals the predicative function of the nominal it precedes. Unlike verbs, however, the particle *ká* does not occur with a prosodic linker (*à=*). Example (11) illustrates that even after a dedicated pause (indicated by the comma), the linker does not occur.

- (11) Nípōómá mééná, ká wà lèébà.
 nípōk-má méena ká wà=lée-ba
 woman-DEF2 all KA NC1=daughter-NC2
 ‘All the women are his daughters.’ (Schwarz 2009a: 268)

Unlike categorical statements with verbal predicates, categorical configurations with nominal predicates in the function of the comment do not even require a bound subject pronoun to express the topic within the same clause. The topics of the categorical statements in (12a)-(12b) are only contextually retrievable; within the statements they are not expressed.

- (12) a. *Ká mí.*
 ka 1s.D
 ‘It’s me.’ (e.g., reply to ‘Who’s that?’ after knocking) (Schwarz 2009a: 269)
- b. *Ká mí-wɔŋ.*
 ka rope-long.NC12
 ‘It [a rope] is long.’ [den noff4: 10]

Note that the configuration described here operates within a clause. Buli has another categorical configuration with a comment that is expressed by an entire clause rather than the verb phrase. The clausal comment is particularly suitable in situations in which the sentence-initial topic does not have subject function (13). Such a complex categorical statement contains the clausal conjunction *tè* at the beginning of the comment and prohibits the use of particle *ká* in postverbal position. For reasons of space, the complex categorical construction is not further discussed here.

- (13) *nídɔ́wá tè nípṓwá nàk.*
 nídɔ́a-wá tè nípṓk-wá nàg
 man-DEF1 CNJ woman-DEF1 hit
 ‘The man, he was hit by a woman.’ (lit. ‘The man, the woman hit him.’) [BL
 (1a-trans) 2004: 330]

3.3 Thetic encoding

The thetic configuration is marked by a predicate-initial marker *lē* (allomorph *nē*) and a dependent verb form. In the perfective, the verb (including suffixes) is characterized by the unstable rising tone, rather than the inflectional verb tone pattern with tonal subject agreement as present in categorical statements. Example (14) illustrates this with a perfective predicate to which an object pronoun binds.

- (14) *Síri lē dòm wā.*
 síri lē dòm-wā
 bee.NC5 CON bite-NC1
 ‘A bee stung her.’ (Reply to: ‘What happened?’) [BL (3) 2005: 444-2]

The thetic statement in (15) contains an imperfective predicate (stem *vī* ‘follow’). The imperfective verb is encoded like the dependent, non-initial verb in a series, as evidenced by the High-toned imperfective marker *á* which elsewhere precedes dependent verbs. The imperfective marker here binds to the preceding connective particle *lē* and the resulting form is often subject to considerable and irregular surface tone changes (Schwarz 2007).

- (15) Bīāk làà vī bíík, níðààbìlì.
 bīāk lē-á vī bíík níðàa-bìlì
 dog.NC12 CON-IPF follow child.NC12 man-small.type.NC5
 ‘A dog is chasing a child – a boy.’ (Reply to: ‘What is going on?’) [BL (3) 2005: 440-1]

Parallel to verbs and clausal connectives, the predicate-initial marker *lē* can also host the prosodic linking vowel *à*= (16).

- (16) Gbáŋ àlē d̀̀à t̀̀éb̀̀ùlk̀̀ù z̀̀úk.
 gbáŋ à=lē d̀̀à t̀̀éb̀̀ùl-k̀̀ú z̀̀úk
 book.NC12 &=CON lie table-DEF15 top.NC15
 ‘There is a book on the table.’ (Schwarz 2009a: 271)

Thetic configurations also extend into the domain of nominal predicates (17). The particle *lē* here directly precedes the predicative nominal, complementing the constructions with particle *ká* encoding categorical nominal predicates. Note that while nominal predicates in categorical statements (where *ká* precedes the nominal predicate) are only occasionally accompanied by a topical subject expression, the non-topical subject of thetic statements is always expressed. In categorical nominal predication, the semantic reading depends on the nature and encoding of the predicative noun. Available readings are PROPER INCLUSION where “a specific entity is asserted to be among the class of items specified in the nominal predicate”, or EQUATION where the entity referred to by the subject is “identical to the entity specified in the predicate nominal” (Payne 1997: 114). Thetic nominal predication, on the contrary, always serves identification. Accordingly, the thetic configuration is particularly common when identifying an individual by name, as illustrated in (17).

- (17) j̀̀wàlá yúé lē Àk̀̀ànchiēm-á̀̀ŋm̀̀í.
 ñ-wá-lá júe lē A.
 DEM-NC1-DET name.NC6 CON A.
 ‘The other one’s name is Akanchiem-anyami.’ [BL (1b-sess) 2004: 222]

There is also a subtype of thetic nominal predication which uses a demonstrative identifier for identification and serves the presentation of entities “out of the blue” (18)-(19). The subject can thus be introduced into the discourse without anchoring it via situational deixis or other contextual information. This is achieved by way of a nominal dummy predicate following the connective particle *lē*. The dummy predicate, a syllabic nasal to which the determiner *lá* suffixes, functions like a demonstrative identifier that provides no lexical information but rather refers back to the initial subject noun.

- (18) Nààwā lē ñnā.
 nàa-wá lē ñ-lá
 chief-DEF1 CON DEM-DET
 ‘This is the chief.’ (Schwarz 2007: 88)

- (19) \dot{M} bisáɲá wàɲɪ àlē ònā.
 m=bí-sa-ɲá wà=ɲɪ à=lē ò-lá
 1s=child-NC12-DEF6 NC1=one &=CON DEM-DET
 ‘This is one of my children.’ (Schwarz 2007: 88)

It is noteworthy that in the absence of a content verb in nominal predication, the connective particle *lē* resembles a copula verb. By hosting the prosodic linker (and even proclitic subject pronouns in embedded sentences, see §3.5) it is also more similar to other verbs (and some connectives) than is its postverbal counterpart *ká* of the categorical construction. Would the particle *lē* therefore be better analyzed as a defective copula verb that is also used in clefts for the thetic encoding? Schwarz (2009b) argues that synchronically neither *lē* nor *ká* are copulas, even though there could be a historical relation between the connective particle *lē* (*nē*) and a widely attested old Niger-Congo copula that surfaces as *ni* or in a somewhat similar form in various languages. Interestingly, the only preposition in Buli, *lè* (allomorph *nè*) ‘with, and’, differs only in tone from the predicate-initial connective particle found in thetic statements (cf. also Schwarz 2010a). Hence there is a tonal and a syntactic differentiation between two linking lexemes having the segmental structure *le* (*ne*) (Table 3).

Table 3: Comparison of linking morphemes with segmental structure *le*

<i>lè</i> (<i>nè</i>) ‘with, and’	preposition (NP/argument-initial) in comitative and coordinative function	[wá lè nààwā] ... ‘he and the chief ...’
<i>lē</i> (<i>nē</i>) (equative ‘be’)	connective particle (VP/predicate-initial) in thetic statements	wá lē nààwā ‘he is the chief’

Both elements share connecting/coordinating semantics and both have prepositional properties, but they differ with respect to their tones, their syntactic domains and consequently also their predicative force. Despite these differences, it is very likely that they are historically related.

So far we have seen that there is a systematic formal distinction between categorical and thetic statements in Buli, which applies to sentences with verbal and nominal predicates alike. Simple categorical statements contain the particle *ká* in affirmation, the position of which is restricted to the postverbal domain (if there is a verbal predicate) and seems to be able to single out focal constituents to some degree. Subject and verb are in such a close relationship that the verb even displays subject agreement tone in the perfective. Thetic statements, on the other hand, contain the connective predicate-initial particle *lē* and never the postverbal particle *ká*. In the perfective, there is no tonal subject agreement at the verb, and the verb in the imperfective resembles the dependent, non-

initial verb in a series. The verbal predicate of thethetic statement is obviously less finite than that of the categorical statement, even though it still functions as predicate. Subject and the dependent verb form are markedly concatenated by means of the particle *lē*.

In the following, I outline the major pragmatic and syntactic conditions under which this marked subject-verb concatenation is very common in Buli.

3.4 The pragmatics ofthetic statements

Thetic statements have no internal information structure. Semantically and pragmatically, thethetic configuration responds to discourse conditions in which a sentence topic is (a) either not available or (b) not wanted for discourse-structuring purposes. For example, thethetic encoding is used whenever the speaker wants to deliberately interrupt the coherence of the discourse and set the following paragraph apart from the previous text. This is most typically the case when new scenes, major participants, and unexpected events are presented.

We find thethetic encoding very often with indefinite subjects, while such subjects are rather rare in simple categorical statements. Thethetic encoding with indefinite subjects is characteristic for existential clauses, such as in (20). Here the existence or presence of the entities referred to by the subject noun is asserted with the help of a posture verb.

- (20) Yíé ηàyè lē zà.
 jíte ηà=jè lē zà
 house.NC6 NC6=two CON stand
 ‘There are two houses.’ [BL (1b-sess) 2004: 218]

The high frequency of indefinite subjects inthetic statements is a result of the scene-setting function of thethetic configuration, among others, with respect to the presentation of major discourse participants. Example (21) illustrates the introduction of the two protagonists at the beginning of a story, immediately after the time frame has been set.

- (21) núr nè wà d5ā àlē jàm bōrós.
 núr nè wà=d5a à=lē dʒàm bō-rós
 person.NC1 with NC1=friend.NC1 &=CON PAST exist-LOC
 ‘there was a man and his friend.’ [Azuima 2]

The scene-setting function ofthetic statements often also consists of reporting particular events. Such reports are typically about sudden changes and include a predicate in the perfective (22); but if the situation persists, imperfective encoding may also be encountered (23).

- (22) Nááb lē biàg biik.
 náab lē biàg biik
 cow.NC14 CON give.birth child.NC12
 ‘A calf has been born.’ [Heiss 392 PA]

- (23) *M̄* biik làā yūāk.
 m̄=biik lē-á jūag
 1s=child.NC12 CON-IPF become.sick
 ‘My child got sick (and still is).’ [BL (2) 2005: 0450-2]

Examples (21)-(23) show that the structural thetic properties in Buli remain the same – the predicate-initial connective particle *lē* and a less finite, dependent verb form – whether the thetic statement states the existence of an entity (21) or that of an event (22)-(23). The two subtypes were identified by Sasse (1987: 526f.) as “entity-central” and “event-central” thetic statements and are only in some languages formally differentiated. Pragmatically, both subtypes lack any internal structuring, but semantically, they focus on different ontological categories.

The formal properties of thetic statements can also be regularly observed in the responses to questions of the ‘What happened?’ type, in particular if the reply does not include impersonal pronouns but nouns in subject function (24). While some scholars interpret such responses as sentence focus occurrences (using *Wh*-questions as focus diagnostics; cf. also Fiedler et al. 2010), in line with Sasse (1987: 572f.) and Güldemann (2010: 86) we can readily identify them as event-central thetic statements.

- (24) Gálási àlē lò tēŋ à mòb.
 galasi à=lē lò tēŋ à=mòb
 glass.NC1 &=CON fall ground.NC12 &=break
 ‘A glass has fallen down and got broken.’ (reply to: ‘What happened?’) [BL (3) 2005: 0448-1]

The formal properties of thetic statements are also regularly used when the focus of assertion is on the subject constituent, as in (25). The utterance represents the direct response to an information question about the identity of the agent (which demands subject function). The utterance can be interpreted as an entity-central thetic statement which states the identity of the subject referent in respect to the rest of the proposition. Regardless of the particular subtype, it can be concluded that the thetic encoding is chosen in order to revoke the categorical interpretation of the simple clause; since otherwise the subject would be regarded the unmarked topic and the focus would be restricted to information pertaining to the comment.

- (25) (Ká) Mary àlē ŋǒbī.
 ká Mary à=lē ŋǒbī
 KA M. &=CON eat.ASS
 ‘Mary ate them.’ (reply to: ‘Who ate the beans?’) (Fiedler et al. 2010: 246)

As indicated by the particle in parenthesis in (25) and by (26), thetic statements of both subtypes can also begin with the sentence-initial particle *ká*.

- (26) *Ká bòlím nē chògsi yèrí á dē.*
ká bòlím lē chògsi jéri á dē
 KA fire.NC14 CON catch house.NC5 IPF eat
 ‘A house has caught fire.’ (reply to: ‘What happened?’) [BL (1b-sess) 2004: 414.15]

The presence of the particle *ká* is grammatically optional and seems to be pragmatically determined. From the data at hand it appears that the sentence-initial particle has scope over the wholethetic construction and is overall more frequent inthetic statements that occur in dialogues and respond to immediate verbal contexts. I therefore hypothesize that the particle *ká* both supports discourse cohesion and is part of a comment that may also include an embeddedthetic statement (27). The topic of this comment is not expressed in the sentence itself, but is provided in the previous discourse context.³

- (27) [Categorical [TOPIC Ø-expression] [COMMENT *ká* [Thetic ... *lē* ...]]]

3.5 Embedding withthetic encoding

One of the striking features ofthetic encoding in Buli is the frequency with which the formal properties occur in various grammatical functions. In the following, I present some common cases of embedded clauses that are based on the same morphosyntax asthetic statements.

One of the two relative clause strategies in Buli, the head-internal relative clause which relativizes subjects and sometimes the objects of the matrix clause (cf. also Hiraiwa 2003), employs the structural means used forthetic statements. The head-internal relative clause in (28) refers to the object the interlocutor sees, and her perception represents the object of the matrix clause, a request to provide a thorough description. In general, relativized head nouns in Buli are marked by an indefinite noun class pronoun or, in case of a headless clause as in (28), often by an indefinite pronoun of noun class 5. The predicate of the head-internal relative clause always begins with the connective particle *lē*. In addition to the basicthetic encoding, relative clauses as well as other embedded clauses typically end with the clause-final determiner *lá*, as is the case in the following example.

- (28) *Māgsī [fi làā nā dīi lá].*
māgsi fi=lē-á nā dīi lá
 describe 2S=CON-IPF see IND5 DET
 ‘Describe what you see!’ [BL (1b-sess) 2004: 106]

Other embedded clause types regularly using thethetic encoding plus the clause-final enclitic determiner *lá* are sentence-initial clauses with adverbial functions as well as what are called TAIL-HEAD LINKAGE clauses. The latter term describes a “discourse pattern which consists in repeating, at the beginning of a new sentence, the main verb of the

³ The etymology of the particle *ká* is not clear, but it is possible that it originated from a proform of noun class 12 (anaphoric pronoun *kà*=).

preceding sentence for discourse cohesion” (Guillaume 2011: 109). The linkage clause in (29) repeats propositional content of the immediately preceding sentence.

(29) ‘... and he caned me.’

Wà nē mii mì lá, mí yàà kàláá kūm ...
 wà=lē mii mí=lá mí jàa kàli-á kūm
 NC1=CON cane 1S.D=DET 1S.D TENSE sit-IPF cry

‘He caned me/when he had caned me, and I was then sitting crying ...’

[Karichiwade 14-15]

Note that the connective particle *lē* can host proclitic subject pronouns (29)-(29), as if it were a (stative⁴) verb. Bound subject pronouns at the connective particle are not uncommon in adverbial and linkage clauses with thetic structures and they can also be found in head-internal relative clauses, the object of which is relativized. Bound subject pronouns are absent, however, from entity-central thetic statements (or subject focus occurrences), where nominal or disjunctive emphatic subject pronouns prevail for semantic/pragmatic reasons.

Summarizing a particular discourse topic for the interlocutor is another common situation where thetic structures occur. Speakers often present the title of their monologue at the beginning or end of their speech, as illustrated by the stacked thetic construction in (30). The initial embedded thetic structure provides the content description in form of a subject clause. The matrix clause presents the title to the listeners by means of a matrix thetic structure that contains the demonstrative identifier (*n̄nā*). The speaker summarizes the topic of her speech by means of a headless relative clause ‘how we celebrate our Fiok festival in Buluk’. Instead of a lexical head, it contains an indefinite pronoun in noun class 5 (*d̄i*) which contributes to the circumstantial manner interpretation. The predicate in the relative clause includes the connective particle *lē*, here with the prosodic linker (*à=*). The verb *dē* ‘eat’, here with the sense ‘celebrate’, comes in a dependent verb form marked for the imperfective aspect by preceding morpheme (*á*). The embedded clause ends with the determiner *=lá* and represents the clausal subject in a thetic matrix clause. The connective particle *lē* of the matrix clause is followed by the demonstrative identifier and fulfils the presentational function.

(30) [Tàmā bùlidē ànàā dē tì fiòkū d̄i lá] àlē n̄nā.
 tàmá bùli-de à=lē-á dē tì=fìo-kú d̄i=lá à-lē n̄-lá
 1P.D Balsa-here &=CON-IPF eat 1P=harvest-DEF17 IND5=DET &-CON DEM-DET

‘This is how our fiok festival is celebrated in Buluk.’ [Fiok 1]

Another example with a clausal subject and the demonstrative identifier as nominal predicate is given in (31). The speaker announces her future career aspirations as the topic of her upcoming talk. In contrast to (30), she here also adds the sentence-initial particle *ká* which explicitly connects her presentation to my earlier request for her speech.

⁴ The tonal properties of the connective particle with the subject proclitic differ from those with dynamic verbs, but resemble those with the irregular stative verb class.

- (31) Ká [mí nàà yāā mí chúmdīpō bōká à jē dīi lá]
 ká mí lē-á jāa mí chúm dīi-pō bō-ká à=jē dīi=lá
 KA 1S.D CON-IPF want 1S.D tomorrow IND5-in exist-DEF12 &=do IND5=DET
 àlē ñnā.
 à=lē ñ-lá
 &=CON DEM-DET

‘This is how I want my future career to be.’ [Chumdiipo 21]

These common embeddings of thetic structures are schematically summarized in (32). Scheme A in (32) reflects the thetic encoding in sentence-initial dependent clauses which have adverbial or discourse linkage function, as illustrated in (29). Scheme B refers to the stacking of two thetic constructions, as illustrated by the presentation of the discourse title in (30). Scheme C is an extension of the stacked thetic structures. The embedding thetic structure gets itself embedded in a categorical statement, as indicated by the sentence-initial particle *ká* and illustrated in (31). A simpler version of the scheme in C with just one thetic structure embedded in a categorical statement is given in (27) above.

- (32) A. [_{Thetic} ... *lē* ...]_{ADVERBIAL/LINKAGE} + Matrix clause
 B. [_{Thetic} [_{Thetic} ...*lē*...] ...*lē* ...]
 C. [_{Categorical} [TOPIC Ø] [COMMENT *ká* [_{Thetic} [_{Thetic} ...*lē*...] ...*lē* ...]]

4 Summary

This paper has described the structural and pragmatic features of the thetic statement and its (simple) categorical counterpart in Buli. Buli grammar applies a connective particle with prepositional properties for the encoding of thetic statements. The connective particle (*lē*) differs only tonally from the comitative and the NP coordinating preposition. In the thetic statement, it is followed either by a predicative noun or by a dependent verb form, as evidenced by the tone of the perfective verb and the tone of the preverbal imperfective marker, respectively. While the connective particle *lē* literally *connects* the predicate to a non-topical subject, both elements are often prosodically separated. The optional presence of the predicate-initial prosodic linker, which appears verb-phrase initially after a pause, indicates that subject and predicate are not necessarily within the same intonational phrase in thetic statements in Buli.

Thetic statements are used to present entities and report events for various discourse structuring reasons. Thetic encoding is also used in response to information questions concerning the referent of a subject constituent or an event. The function of thetic encoding consists in such cases in the cancelation of the prototypical categorical interpretation in favor of a single information chunk. The application of a simple concatenating morphosyntactic device to this effect, which can be easily embedded, has further contributed to the success of thetic encoding in Buli grammar and the apparent polyfunctionality of thetic structures.

Acknowledgements

Funding of the Buli research within the Gur research projects at Humboldt University Berlin (1996-2001) and the Research projects B1, D2, B7 of the Collaborative Research Centre on Information structure (SFB 632, 2003-2009) by the German Research Foundation (DFG) is gratefully acknowledged. I thank the editors and two anonymous reviewers for their very helpful and constructive comments which significantly contributed to improving the quality of the paper. I am also grateful to the language consultants who provided data, transcriptions and translations, ideas and thoughts about the grammatical structure and use of their language.

Abbreviations

ASS	assertive	IPF	imperfective
CNJ	conjunction	LOC	locative
CON	connective	NC	noun class
CONF	confirmation	P	plural
D	disjunctive pronoun	PROX	proximal
DEF	definite	Q	question marker
DEM	demonstrative	S	singular
DET	determiner	TBU	tone bearing unit
IND	indefinite pronoun	&	prosodic linker

References

- Cahill, Michael C. 2007. *Aspects of the morphology and phonology of Kɔnni*. Dallas: SIL International.
- Fiedler, Ines, Katharina Hartmann, Brigitte Reineke, Anne Schwarz & Malte Zimmermann. 2010. Subject focus in West African languages. In Malte Zimmermann & Caroline Féry (eds.), *Information structure: Theoretical, typological, and experimental perspectives*, 234–257. Oxford: Oxford University Press.
- Guillaume, Antoine. 2011. Subordinate clauses, switch-reference, and tail-head linkage in Cavineña narratives. In Rik van Gijn, Katharina Haude & Pieter Muysken (eds.), *Subordination in native South American languages*, 109–139. Amsterdam: John Benjamins.
- Güldemann, Tom. 1996. *Verbalmorphologie und Nebenprädikationen im Bantu – Eine Studie zur funktional Motivierten Genese eines konjugationalen Subsystems*. Bochum: Universitätsverlag Dr. N. Brockmeyer.
- Güldemann, Tom. 2010. The relation between focus and theticity in the Tuu family. In Ines Fiedler & Anne Schwarz (eds.), *The expression of information structure – a documentation of its diversity across Africa*, 69–93. Amsterdam: John Benjamins.

- Güldemann, Tom. 2012. Thetic speaker-instantiating quotative indexes as cross-linguistic type. In Isabelle Buchstaller & Ingrid Van Alphen (eds.), *Quotation: Cross-linguistic and cross-disciplinary perspectives*, 117–142. Amsterdam: John Benjamins.
- Hiraiwa, Ken. 2003. Relativization in Buli. In George Akanlig-Pare & Michael Kenstowicz (eds.), *Studies in Buli grammar* (Working Papers on Endangered and Less Familiar Languages 4), 45–84. Cambridge, MA: MIT.
- Hyman, Larry M. & John R. Watters. 1984. Auxiliary focus. *Studies in African Linguistics, Supplement* 15(3). 233–273.
- Kuroda, Shige-Yuki. 1972. The categorial and the thetic judgement: Evidence from Japanese syntax. *Foundations of Language* 9(2). 153–185.
- Lambrecht, Knud. 1987. Sentence focus, information structure, and the thetic-categorial distinction. In Jon Aske, Natasha Beery, Laura A. Michaelis & Hana Filip (eds.), *Proceedings of the 13th Annual Meeting of the Berkeley Linguistics Society – General Session and Parasession on Grammar and Cognition*, 366–382. Berkeley, California: Berkeley Linguistics Society.
- Lambrecht, Knud. 2000. When subjects behave like objects: An analysis of the merging of S and O in sentence-focus constructions across languages. *Studies in Language* 24(3). 611–682.
- Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). 2015. *Ethnologue: Languages of the world*. 18th edn. Dallas, TX: SIL International. <http://www.ethnologue.com>.
- Manessy, Gabriel. 1979. *Contribution à la classification généalogique des langues voltaïques*. Paris: Centre National de la Recherche Scientifique; SELAF.
- Naden, Anthony J. 1989. Gur. In John Bendor-Samuel & Rhonda L. Hartell (eds.), *The Niger-Congo languages*, 140–168. New York: University Press of America.
- Payne, Thomas E. 1997. *Describing morphosyntax: A guide for field linguists*. Cambridge: Cambridge University Press.
- Reinhart, Tanya. 1982. Pragmatics and linguistics. An analysis of sentence topics. *Philosophica* 27. 53–94.
- Robert, Stéphane. 2000. Le verbe wolof ou la grammaticalisation du focus. In Bernard Caron (ed.), *Topicalisation et focalisation dans les langues africaines*, 229–267. Louvain – Paris: Peeters.
- Sasse, Hans-Jürgen. 1987. The thetic/categorial distinction revisited. *Linguistics* 25(3). 511–580.
- Sasse, Hans-Jürgen. 1995. ‘Theticity’ and VS order: A case study. In Yaron Matras & Hans-Jürgen Sasse (eds.), *Verb-subject order and theticity in European languages*, 3–31. Berlin: De Gruyter.
- Schwarz, Anne. 2003. Low tone spreading in Buli. *Cahiers Voltaïques / Gur Papers* 6. 121–130.
- Schwarz, Anne. 2007. *Aspekte der Morphosyntax und Tonologie im Buli – Mit Schwerpunkt auf dem Buli von Wiaga*. Berlin: Humboldt-Universität zu Berlin. (Doctoral dissertation.) www.urn:nbn:de:kobv:11-10085362.
- Schwarz, Anne. 2009a. To be or not to be? About the copula system in Buli (Gur). In Margarida Petter & Ronald Beline Mendes (eds.), *Proceedings of the Special World Congress*

5 *All-in-one and one-for-all: Thetic structures in Buli grammar and discourse*

of *African Linguistics – São Paulo 2008: Exploring the African language connection in the Americas*, 263–278. São Paulo: Humanitas.

- Schwarz, Anne. 2009b. Tonal focus reflections in Buli and some Gur relatives. *Lingua* 119. 950–972.
- Schwarz, Anne. 2010a. Discourse principles in grammar: The thetic/categorical dichotomy. In *Proceedings of the Inaugural Tropics of the Imagination Conference, James Cook University Cairns, 2-3 November 2009* (Etropic 9). <https://journals.jcu.edu.au/etropic/article/view/3425/3363>.
- Schwarz, Anne. 2010b. Verb-and-predication focus markers in Gur. In Ines Fiedler & Anne Schwarz (eds.), *The expression of information structure. A documentation of its diversity across Africa*, 287–314. Amsterdam: John Benjamins.
- Skopeteas, Stavros, Ines Fiedler, Samantha Hellmuth, Anne Schwarz, Ruben Stoel, Gisbert Fanselow, Caroline Féry & Manfred Krifka. 2006. *Questionnaire on information structure (QUIS)* (Interdisciplinary Studies on Information Structure 4). Potsdam: University of Potsdam.
- Ulrich, Miorita. 1988. Thetisch vs. Kategorisch und Informationsstruktur. In John Albrecht, Harald Thun & Jens Lüdtke (eds.), *Energieia und Ergon: Sprachliche Variation – Sprachgeschichte – Sprachtypologie*, 387–399. Tübingen: Narr.
- Wolff, H. Ekkehard. 2005. Intrinsic focus and focus control in two varieties of Hausa. In F. K. Erhard Voeltz (ed.), *Studies in African linguistic typology*, 397–415. Amsterdam: John Benjamins.