From Pragmatics to Sentence Type:
Non-topical S/A Arguments and Clause-second Particles in the Kalahari Basin

Tom Güldemann

Humboldt University of Berlin
Max Planck Institute for the
Science of Human History Jena

Lee J. Pratchett

Humboldt University
of Berlin

Alena Witzlack-Makarevich

Hebrew University of Jerusalem

Abstract: Languages of the Kalahari Basin contact area share a feature whereby a special type of particle occurs in clause-second position, often after the S/A constituent. Previous accounts have used a wide range of labels such as declarative, indicative, emphatic nominative, or topic, which point to a diverse but insufficiently understood functional array of this particle type. We address the phenomenon from a discourse-oriented and comparative perspective by exploring relevant cases in languages of three different families: Northern Khoekhoe of Khoe-Kwadi, N‖ng of Tuu, and Ju of Kx'a. We conclude that the particles are involved in a network of constructions spanning such diverse domains as nonverbal predication, focus, entity-central theticity, declarative, and possibly even differential S/A marking. The last two functions that relate to sentence types and grammatical relations, respectively, and (may) no longer display a marked information structure (IS) configuration, emerge from the overuse of thetic particle constructions and thus are the result of so-called "depragmaticization".*

Key words: Non-topical S/A argument, clause-second particle, Kalahari Basin

1. Introduction

The Kalahari Basin is a linguistic contact area first proposed by Güldemann (1998). It comprises the three language families formerly subsumed under the spurious concept of "Khoisan" (see Güldemann (2014) for current language classification). One of the features listed by Güldemann & Fehn (2017) in their recent survey of the Kalahari Basin is a widespread gram type, illustrated in (1) with |Xam of the Tuu family. The example exhibits an element *gnn* (provisionally glossed using a question mark) that occurs in clause-second position after the S/A constituent, separating the latter from the rest of the clause (see section 2 for more diverse constituent types preceding such an element in other languages).[1]

(1) *au*   *too=**gnn***   *n|e*   *!ii–ya*
     CONN   red.ochre=?   IPFV   be.red-STAT
     'But/and ochre is red.'   (Güldemann 2013b: 428)

Güldemann (2006: 119–122) originally proposed this syntactic feature as an areal isogloss of the smaller Cape linguistic area comprising !Ui languages of the Tuu family and Khoekhoe languages of the Khoe-Kwadi family. The following studies have since provided analyses of the phenomenon in natural discourse: Güldemann & Siegmund (2009), Güldemann (2010), Güldemann & Witzlack-Makarevich (2013), Güldemann & Pratchett (2014), and Pratchett (2017).

---

for his mother tongue expertise on Namibian Khoekhoe. All examples cited from published sources use their original orthography or transcription and may thus differ from the orthography we normally use here for the relevant language. Examples from N‖ng are mostly from Güldemann et al. (2007–2014) and are then accompanied by an identifier for the text and clause; the data corpus is accessible in the deposit under http://elar.soas.ac.uk/deposit/0089. Since this article deals with information structure, most examples are preceded by their discourse contexts in curly brackets. Elements focused on in the discussion are highlighted in bold. The following abbreviations are used in examples, figures and tables: A agent of transitive verb (semantic role), COM comitative, CONN discourse connective, COP copula, CSTR construct form, DECL declarative, DEM demonstrative, DIM diminutive, DIST distal, DU dual, ECT entity-central theticity, EMPH emphasis, F feminine, FG foreground, ID identificational, IPFV imperfective, IRR irrealis, IS information structure, LOC locative, M masculine, MPO multi-purpose oblique, NEG negation, P patient of transitive verb (semantic role), PASS passive, PL plural, PN personal name, PRO pronoun, PROG progressive, PROX proximal, PST past, Q question, RECP reciprocal, REL relative, RELV relevance, S single argument of intransitive verb (semantic role), SG singular, SUB subordination, TF term focus, V verb, VE valency external. The symbol # marks a pause.

[1] The status of the relevant |Xam element as an enclitic rather than a particle is unique across the relevant cases in the Kalahari Basin.

Table 1.   Clause-second particles in the Kalahari Basin (Güldemann & Fehn 2017)

Language (variety)	Family, branch	ISO	Form	Label	Source
ǀXam	Tuu, !Ui	xam	=ŋ	Emphatic nominative	Bleek (1928–1930: 87–88)
Nǁng (aka Nǀu)	Tuu, !Ui	ngh	*ke*	Declarative	Güldemann (2003), Collins & Namaseb (2011: 9)
East !Xoon of Taa	Tuu, Taa-Lower Nossob	nmn	*ń*	Indicative	Traill (1994: 193), Guldemann (2013a)
!Ora	Khoe-Kwadi, Khoekhoe	kqz	*tje*	Subjekt-Determinativ	Meinhof (1930: 49–50)
Standard Khoekhoe	Khoe-Kwadi, Khoekhoe	naq	*ge*	Declarative, Indicative	Hagman (1977: 108–113), Haacke (2013: 335)
Tsumkwe Juǀ'hoan	Kx'a, Ju	ktz	*ḿ*	Verb particle	Dickens (1994: 234, 2005: 44)
Ekoka !Xun	Kx'a, Ju	knw	*má*	Topic, Subject case	König (2006a, 2008a, b), Heine & König (2015)
N!aqriaxe	Kx'a, ǂ'Amkoe	huc	*ki*	—	Berthold & Gerlach (p.c.)

Table 1 gives a yet fuller survey of such clause-second particles in the Kalahari Basin. The column "Label" reflects the difficulties in their analysis in presenting the bewildering range of terms that span the marking of sentence type, grammatical relations, and information structure (henceforth IS) or are functionally indeterminate. In some languages, the particles are in complementary distribution with syntactically similar grams, notably those marking questions (*ba/xa* in ǀXam, *xa(e)* in Nǁng, *kha* in Standard Khoekhoe, *re* in Tsumkwe Juǀ'hoan), which lends support to the idea that they are related somehow to sentence type.

One of the primary aims of this article is to rectify the prevailing functional indeterminacy of these clause-second particles and to elucidate their relevance for IS research and other research fields. The issues we raise in this case study shed light on a more widespread problem in the IS domain, broaching a diverse array of functions including de-topicalization, focus, theticity, topic shift, discourse linkage, and others. We also draw attention to a possible development from a marked IS construction for non-topical S/A terms to an unmarked declarative aka indicative sentence.

With the above goals in mind, in section 2 we first survey the findings of the discourse analyses of clause-second elements in three language complexes, one from each of the lineages of the Kalahari Basin area. We start with *ge* in Northern Khoekhoe of the Khoe-Kwadi family focusing on Standard Namibian Khoekhoe and Richtersveld Nama (section 2.1), proceed with *ke* in Nǁng of the Tuu family (section 2.2), and conclude with *má* in North-central Ju of the Kx'a family (section 2.3). In section 3 we summarize the results and propose a unified functional and historical account of clause-second particles in the Kalahari Basin and discuss the implications for IS research in general.

2. Clause-second particles in the Kalahari Basin

2.1. Northern Khoekhoe *ge*

The cluster of Northern Khoekhoe (Khoe-Kwadi), spoken in large parts of Namibia and a small area in northwestern South Africa, possesses the particle *ge* (or *ke* and, in older sources, *gye*). Our point of departure is the most widely spoken member of this cluster, Standard Namibian Khoekhoe, in which *ge* is analyzed as a declarative aka indicative marker as shown in the following quote (cf. also Hagman (1977: 54) and Haacke (2013: 335)):[2]

> The subject is followed immediately by the particle *gye* ... It marks the subject as such and cannot be translated into German. This "subject particle" only occurs in declarative sentences, and there only in main clauses. It is lacking in questions, imperatives and dependent clauses. [our translation] (Dempwolff 1934/5: 36)

While the above and other characterizations depict a single element with a unitary syntactic position after a clause-initial slot, in fact, on closer inspection, they conceal the diverse structural and functional repertoire of *ge*. There are at least three distinct subtypes of *ge* in clause-second position alone, depending on the kind of constituent preceding it. Examples (2)–(4) illustrate these distinct patterns. To better tease apart the differences in terms of syntax and IS, the comparable constituents are aligned and annotated in a uniform way.

(2)   [FOCUS]   [A.TOPIC]   [BACKGROUND]

a. *ao–b*   ∅   *ge*   ∅   *tara–s–a*   *ra*   *mû*
    man-M.SG   "DECL"   woman-F.SG-P   IPFV   see
    'THE/A MAN is seeing the/a woman.'   (Haacke 2006: 114)

b. *tara–s–a*   *=b*   *gye*   *khoi–b–a*   ∅   *go*   *mû*
    woman-F.SG-P   =3M.SG.A "DECL"   person-M.SG-S   PST   see
    'The man saw the/a WOMAN.'   (Dempwolff 1927: 74)

c. *‖ari*   *=b*   *gye*   *khoi–b–a*   *tara–s–a*   *go*   *mû*
    yesterday   =3M.SG.A "DECL"   person-M.SG-S woman-F.SG-P   PST   see
    'The man saw the/a woman YESTERDAY.'   (Dempwolff 1927: 75)

(3)   [LINKER]   [A.TOPIC]   [FOCUS]

o   *=b*   *gye*   *khoi–b–a*   *tara–s–a*   *go*   *mû*
then   =3M.SG.A "DECL"   person-M.SG-S woman-F.SG-P   PST   see
'Then the man saw the/a woman.'   (Dempwolff 1927: 75)

(4)   [A.TOPIC]   [FOCUS]

khoi–b   ∅   *gye*   ∅   *tara–s–a*   *go*   *mû*
person-M.SG   "DECL"   woman-F.SG-P   PST   see
'The man saw the/a woman.'   (Dempwolff 1927: 73)

---

[2] Washburn (2001) is a study dedicated to the particle. However, it merely aims at a narrow theory-specific account and draws its data entirely from Hagman (1977), so that it does not contribute any new information to the present discussion.

Traditional approaches to Khoekhoe grammar focus on the common denominator in (2)–(4), whereby *ge* occurs in the Wackernagel position (marked in bold) that bisects the clause into a pre-field and a middle-field (the final and more rarely occupied post-field is after the last finite verb and not shown here). This basic clause structure is schematized in (5).

(5)   [PRE-FIELD] =**(S/A PRO)** *ge* [MIDDLE-FIELD] **PREDICATE** [POST-FIELD]

The clause-second *ge* is in complementary distribution with other particles such as *kha* which is used in emphatic questions and *kom* which marks certainty in conjunction with clause-final *o*. The syntactic position of *ge* aside, there are crucial differences between the three patterns in (2)–(4). First, the initial pre-field hosts constituents with very different IS roles which, in turn, has different implications for the middle field. Thus, the particle *ge* follows different types of fronted, mostly contrastive, focus constituents in (2), a discourse linker in (3), and a S/A topic in (4). (Similar to (3) is another pattern wherein conjunctions such as *o* 'then' have a functional parallel but structurally complex substitute in the form of an adverbial background clause.) Particularly relevant for the IS role of *ge* is the reversal of background and focus constituents between (2) on the one hand and (3) and (4) on the other. Also, if the pre-field constituent does not refer to the S/A, *ge* is preceded by an obligatory S/A pronoun, often an enclitic on the last pre-field element, which can be accompanied in addition by a full S/A noun phrase in the middle-field. This holds in (2b), (2c), and (3) with =*b* and *khoiba*. The fact that *ge* follows either a full noun phrase or a pronominal referring to the S/A argument has motivated occasional labels as "subject particle" and the like. This is, however, a problematic term because the elements before *ge* refer to a semantic role and not a structurally defined grammatical relation. A remarkable observation is that the structures in (2a) with A focus and (4) with A topic are not only segmentally identical but, according to S. Job (p.c.), cannot be distinguished prosodically either. This is surprising in view of the cross-linguistic generalization that an S/A-focus structure should, if anything, be more marked than other focus structures (cf., e.g., Fiedler et al. 2010). Therefore, we would expect a clearer distinction between (2a) and a categorical statement with an S/A topic, as in (4). This being said, the significance of such sentences with two simultaneously occurring noun phrases that mark the S/A in the pre-field and the P in the middle-field to the present discussion is currently unclear. While they are recurrently presented and discussed in descriptions of Khoekhoe and are judged by native speakers as grammatical, these sentences seem to be so rare in natural discourse that their potential for informing basic IS analyses in the language still needs to be determined.

The above subtypes of *ge* do not do justice to the full complexity of the particle: it is involved in at least two other constructions. These are rarely entertained in connection with declarative *ge* but are crucial for understanding its general functional profile. First, the particle is used as an identificational marker with scope over a preceding noun phrase, as in (6); this structure is, for example, a typical response to questions such as 'Who/what is this?'.

(6)   *gei   omi   **ge***
       big   house:M.S   ID
       'It is a BIG HOUSE.'   (Olpp 1963: 98)

In (7), *ge* also occurs as a sentence-final particle with scope over a constituent that looks like a noun preceded by a sentential relative-like modifier. While this appears to be similar to the structure in (6), the IS interpretation is quite different in that the noun before *ge* is typically excluded from the assertion domain, as shown in (7). To the best of our knowledge, the only substantial discussion of this construction is Haacke (2006: 114–116). Unfortunately, it lacks sufficient information on the structure's behavior in natural discourse. We are thus unable to characterize its function conclusively and leave that open.

(7)   a. *gei   a   omi   **ge***
          big   COP   house:M.S   ?
          'The house IS big./The house is BIG.'   (Olpp 1963: 98)
       b. *‖nā–ti   go   mâ   ‡hôa–s   **ge***
          this-like   PROX.PST   stand   news-F.S   ?
          'The story ENDS like this./THAT's how the story ends.'   (S. Job, field notes)

Ia	[Term	***ge***]			Identification
Ib	[REL.Clause Term	***ge***]			?
II	[Focus(=s/a)	***ge***]	[(S/A)	Other V]	"Declarative" after term focus
III	[Conj./Clause=s/a	***ge***]	[(S/A)	Other V]	"Declarative" after discourse linking
IV	[S/A	***ge***	∅	Other V]	"Declarative" after S/A topic

Figure 1.   Structural and functional profile of *ge*-constructions in Northern Khoekhoe

The five contexts of *ge* are summarized in Figure 1 by means of structure schemas that visualize their similarities and differences in terms of syntax and function. We argue that the constructions are related to each other, at least historically; hence, we are dealing with a single element *ge*. We defer the discussion on how these constructions can be accommodated within a semantic-map approach to section 3.1, once the largely similar profile of clause-second particles in other Kalahari Basin languages has been demonstrated. Presently, we only seek to evaluate the characterization of *ge* as a purported declarative (or indicative) marker in three of the five constructions in Figure 1, which should be done in view of the standard definition of "declarative":

> A term used in the grammatical classification of sentence types, and usually seen in contrast to imperative, interrogative, etc., moods. It refers to verb forms or sentence/clause types typically used in the expression of statements ... (Crystal 1997: 104)

If *ge* is indeed a canonical marker of this kind, we would expect it to occur in all relevant independent clauses. While this is implied in more contemporary

descriptions of Namibian Standard Khoekhoe by Hagman (1973: 106, 1977) and Haacke (2013), it is problematic for Northern Khoekhoe as a whole. That is, *ge* has a significantly different profile in varieties that were recorded prior to the emergence of the modern standard and/or where the geographic distance is such that they appear to have been uninfluenced by it. We are in a position to demonstrate this with data collected by the last author in 2003–2006 on Richtersveld Nama. This is the last surviving Khoekhoe variety of South Africa, spoken in the area whence Khoekhoe speakers originally migrated into modern Namibia, and which has been less affected by language standardization in the north.

Figure 2 illustrates the distribution of clause-second particles in a corpus of ten Pear Stories of variable length but with a comparable story line. The small corpus from ten different speakers shows considerable variation. Speakers A, B, and K use *ge* regularly befitting its declarative analysis, while speakers C, D, E and I use it only in some contexts; speaker F employs it only on special occasions and two speakers (G and H) do not use any clause-second particles. While it remains unclear how to best characterize *ge* in Richtersveld Nama, the data nevertheless suffice to conclude that it cannot mark canonical declaratives.
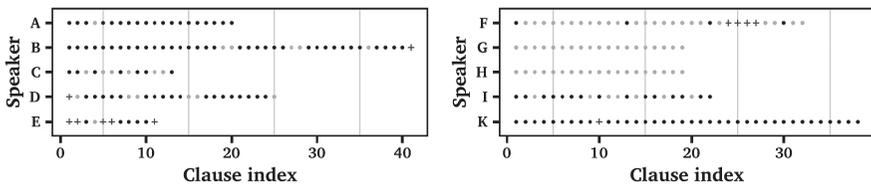

Figure 2.   Frequency of *ge* (and *kom*) in ten Pear Stories from Richtersveld Nama (black dot = *ge*-clause, cross = *kom*-clause, light grey dot = unmarked clause)

There is evidence to assume that the picture in this peripheral Northern Khoekhoe variety reflects the more conservative situation vis-à-vis modern Namibian Khoekhoe. For one thing, early assessments of *g(y)e* in Northern Khoekhoe up until as late as the first half of the 20th century regularly describe the particle as a facultative element without a discrete function (e.g. Hahn 1870: 40–42, Seidel 1891: 13–14, Planert 1905: 16, 20, Meinhof 1909: 53, Dempwolff 1934–1935: 37). As a matter of fact, the earlier the source, the more likely it is to find the most basic sentence pattern described without "declarative" *ge*, while the sentence pattern with *ge* is viewed as a free variant. Wallmann (1854: 26–30, 1857: 32–33), for example, views *ge* as a copula and suggests that *ge*-marked clauses reflect a "compound conjugation" of a vague indicative nature—a view repeated also in some later accounts. Moreover, this author draws attention to dialectal differences between distinct ethnic groups speaking Northern Khoekhoe today and describes the frequent use of *ge*-clauses as typical of the speech of the Nama native to Namibia rather than that of the Oorlam who immigrated from South Africa and whose original variety is thus likely to have been more like Khoekhoe spo-

ken there. Likewise, Vedder (1923: 159) reports that a declarative *ge* of the Nama type was not present in the early varieties of the Damara variety of Namibian Khoekhoe.

Finally, if we expand the survey to consider other related languages, we find further evidence that the declarative-like *ge* is a recent development that seems to be restricted to the prescriptive Namibian standard variety. Thus, in !Ora, the second-most well attested Khoekhoe language, the cognate marker *tje* (or *dje*) is not a grammaticalized declarative marker either. The best !Ora description by Meinhof (1930: 49–50, 59–60) merely reports that the particle occurs occasionally after the subject, as in the opening line to a story in (8), and regularly in association with *ni* in a presumably related construction expressing obligation, as in (9).

(8) |eie-b   ***tje*** *tje* |ui   tse   ḣei–b   !na ǂnoa |o:-s
     jackal-M.SG   ?   PST   one   day   tree-M.SG   in   sit   child-F.SG
     ḣoa
     find:RELV
     'One day the jackal had found a girl sitting in a tree.'

                                                              (Engelbrecht 1936: 230–231)

(9) ‖xara–e=b   ***tje*** *ni*
     punish-PASS=3M.SG.S   ?   IRR
     'er muß bestraft werden [he must be punished]'   (Meinhof 1930: 53)

2.2. N‖ng *ke*

The N‖ng language complex (often called N|uu after its most extensively documented northwestern dialect) is the last surviving and now moribund member of the !Ui branch of the Tuu family, once spoken in the South African part of the Kalahari (see Güldemann 2017 for more information). N‖ng possesses an element *ke* that is very similar to Standard Namibian Khoekhoe *ge*. In fact, the two elements have been linked historically in Güldemann (2006: 119–122) within the more general contact hypothesis in the Cape linguistic area between the Khoekhoe subgroup of Khoe-Kwadi and the !Ui branch of Tuu. The similarity of this marker to Khoekhoe *ge* has also been observed by Collins & Namaseb (2011: 9).

Our point of departure for analyzing N‖ng *ke* is the first extensive modern material on the northwestern N|uu variety, which is known under the doculect name N|huki and was collected in the 1960s by Westphal (n.d.). The unpublished field notes stored at the Bleek and Lloyd Archive in Cape Town include data that was elicited from a single speaker and has been analyzed comprehensively in Güldemann (2003).

The most important observation about *ke* in N|huki is its regular presence as a particle after the S/A argument in the basic clause illustrated in (10). Occurring in 77% of 184 relevant clause tokens of the corpus, this salient *ke*-construction seems to be so similar to *ge* in Standard Khoekhoe that Güldemann (2003, 2006) also analyzed it as a declarative marker. Another similarity of N|huki *ke* to Khoekhoe *ge* is its use as an identificational marker, as in (11).

(10) ǂoo   a   **ke**   ǂxoa   n|a   ng   anci
     man this "DECL" speak COM   1SG   father
     'this man speaks with my father'   (Westphal n.d.)

(11) *ng*   **ke**
     1SG   ID
     'it is I'   (Westphal n.d.)

In the late 1990s, encouraged by the revelation of a larger number of remnant speakers of N‖ng, various research initiatives resulted in a more extensive documentation. This has led to a reanalysis of the limited description provided by earlier studies. Modern morpho-syntactic data on N‖ng were collected in two research contexts, one of which resulted in the grammatical sketch by Collins & Namaseb (2011) describing *ke* very briefly as follows:

> *ke* only appears in declarative sentences, and never in questions, relative clauses, or imperatives. Furthermore, the declarative marker is mostly absent in complement clauses. […] When the subject is a pronoun, the declarative marker *ke* can be optionally replaced by *-a*. […] The declarative markers *ke* and *-a* are often dropped, for example with negation and also in stories. (Collins & Namaseb (2011: 9))

While the authors closely follow the preliminary analysis by Güldemann (2003, 2006), a brief inspection of their own text collection (2011: 71–114) casts doubt on their account because *ke* occurs only sporadically.

The second research team collected natural discourse data within the ELDP-funded project "A text documentation of N|uu" (Güldemann et al. 2007–2014) and also worked on the data collected in 1936 in connection with the Kalahari expedition of the University of the Witwatersrand (made publicly available by Traill 1997, see Güldemann 2017 for details). This second data set was recorded at a time when the language was still the means of communication in a relatively coherent speech community. The analysis of these discourse data result in a more complete profile of *ke*-constructions. First, the two structures with *ke* observed in the Westphal corpus are confirmed, namely its use as an identificational marker and as an apparent "declarative" marker, to be discussed in more detail below.

A new function not recognizable clearly in Westphal's N|huki data is its use in contexts of (mostly contrastive) term-focus. The question-answer context targeting the P argument in (12) is from an elicitation (see Güldemann (2010: 74-6, 89) for more examples).

(12) *tyui   xae Katarina aa |oba   i   ǂxanisi **ke** Katarina aa |oba   i*
     what   Q   PN   give child ?   book   TF   PN   give   child   ?
     'What does Katarina give the child?'   'Katarina gives the child a BOOK.'

If the initial focus constituent does not refer to the S/A argument, as in (12), the out-of-focus part is formally identical to a non-S/A relative clause except that it lacks a relative linker. That is, the clause after *ke* starts with an obligatory S/A

noun phrase and has a lexical "gap" coreferential with the clause-initial focus item. If this corresponds with an oblique constituent, a resumptive index *'n‖aa* occurs in its expected syntactic position. One is thus confronted with a cleft-like construction of an identificational clause, as in (11), followed by a background clause, whereby *ke* is the pivot between the initially exposed focus and the extra-focal rest of the sentence. For the later discussion, it is important to recognize that in the case of S/A focus there is no reference to the S/A argument after *ke*, so that this structure is segmentally indistinguishable from the apparent "declarative" construction shown in (10).

Yet another construction only identified in the text data exemplifies *ke* as part of phrasal discourse linkers. This pattern should be analyzed as a special case of the bisected cleft-like construction treated previously in that the initial focus position hosts a generic (pro)nominal anaphor referring back to preceding propositional content. The material in the initial slot varies; attested are *(ha) gao **ke**, (tya) gao **ke*** 'it is (that) thing …', *tya !'ama **ke*** 'it is that reason …', *tya **ke*** 'it is that …'. The common denominator is that the pattern creates phrasal discourse linkers for reason and consequence, rendering such English expressions as 'that is why' (as in (13) below), 'for that reason', 'therefore', 'accordingly', 'consequently', 'so', or simply 'then'.

(13) {Moon has chopped Hare's mouth crosswise}
       gao   **ke**   *n!au*   *tyuu*   #   *!oon*   *|kx'aba*
       thing   TF   hare   mouth   now   be.red
       'That's why the hare's mouth is now red.'   (NX360000-01_A.034–035)

The constructional range of *ke* in N‖ng is highly similar to that of Northern Khoekhoe *ge* in Figure 1 in that only the use under Ib is lacking. Recall, however, that the role of *ke* as a declarative marker was initially identified by Güldemann (2003) on account of Westphal's (n.d.) elicited data without any later substantiation by more robust evidence. We are in the position to assess this hypothesis and the regularity and frequency of *ke* in spoken discourse, as we possess a coherent text corpus of manageable size (Güldemann et al. 2007–2014). We annotated a sample of 21 dialogues and monologues by six speakers with more than 4500 clause-like pause units for the presence and type of *ke*-construction. (In natural discourse, there is often no direct correlation between a pause unit and a clause but we are confident for the present statistical evaluation about using these numbers as a proxy for the proportion of *ke* tokens per clause units.)

Figure 3, in which dots represent individual clauses of a text-(like) unit, illustrates the distribution of *ke*-marked clauses in a similar format as Figure 2 for Richtersveld Nama. Individual texts vary in length so that we truncated long texts to the first 150 clauses to facilitate a better comparison. The texts 1-21 are a representative selection of our modern data. The texts I–VI show the *ke*-occurrence in the naturalistic data collected during the Kalahari expedition of the University of the Witwatersrand in 1936. By comparison, corpus A is an elicitation session made during the same expedition, while B and C are two samples of 150 sentences

each from Westphal's (n.d.) elicitation, as annotated in Güldemann (2003). A first observation is that *ke*-occurrence varies considerably: some texts lack *ke* entirely or display it only very occasionally. This is surprising in view of its analysis as a plain declarative particle: of 4644 clauses annotated, only 254 (5.5%) are marked by *ke*. Another crucial result is that the presence of *ke* in potential declarative clauses is considerably elevated in the samples A—C, representing elicitation—an unnatural use of language.[3]
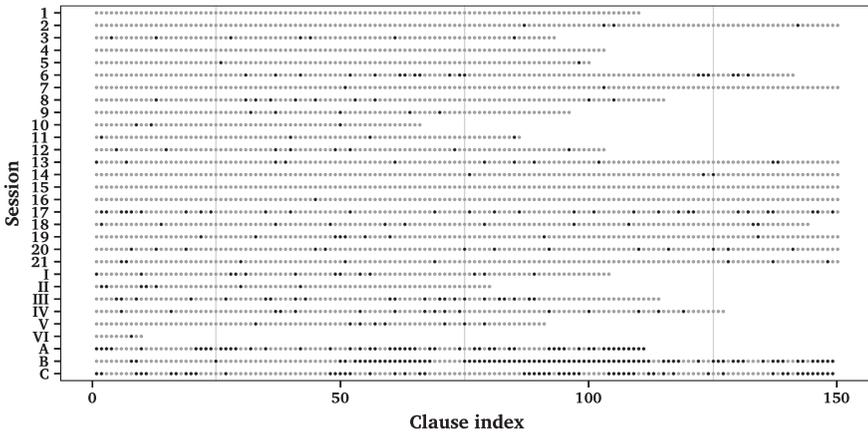
Figure 3.   Frequency of *ke* in N‖ng sources (1–21 = modern texts, I—VI = old texts, A—C = elicitations, black dot = *ke*-clause, light grey dot = unmarked clause)

Table 2.   Frequency of different *ke*-constructions

Construction type		Total	% of all clause-like units (4644)
I	Identification	83	1.79%
II	Term-focus	42	0.9%
III	Discourse linking	14	0.3%
IV	"Declarative"?	115	2.5%
	Total	254	5.5%

Table 2 summarizes the frequency of *ke* according to construction type, revealing a picture very similar to that of *ge* in Richtersveld Nama. Apart from the parallel syntactic restrictions and constructional contexts, the generally low discourse frequency means that it cannot be a declarative marker. A dedicated analysis of the apparent "declarative" *ke* in textual data, recorded in Table 2 under

---

[3] The comparison between the modern data and those from 1936 show a decrease in the frequency of *ke*: In contrast to 5.5% of marked clauses today, 14.6% of sentences (67 of 459 sentences) are so marked in the data from 1936. We have no explanation for this difference.

IV, has been undertaken by Güldemann & Siegmund (2009) and Güldemann & Witzlack-Makarevich (2013). Their results show that the analysis proposed by Güldemann (2010) for syntactically bisected cleft-like constructions in the other Tuu languages |Xam and Taa also hold for N‖ng in that the relevant *ke*-clauses should be associated with the concept of "theticity". This has been studied by only few scholars intensively, notably Sasse (e.g., 1987, 2006) and Lambrecht (e.g., 1994, 2000, using the label "sentence focus") but it is crucial for the IS domain in general. We largely follow the theoretical account and terminological usage of Sasse (1987) but we should point out that we divert from Sasse in viewing theticity to be primarily IS-related, in line with his own apt characterization of thetic statements, which are opposed to categorical statements with an internal foreground-background structure:

> The thetic statement forms a unit with respect to what it contributes to the discourse at a given point. It expresses a pragmatically unanalyzed state of affairs and presents it as a piece of complex information. … Thetic statements are thus uttered at those points of the discourse when compact information is required. This is not the case with the categorical statement. It presents a state of affairs as something analyzed, dissected into different information units. It selects one of the participants of the state of affairs in order to present it as a predication base and arranges the rest in such a way that it forms the predication about the selected predication base. We thus utter categorical statements at those points of the discourse when information is built up in successive bits. (Sasse 1987: 558)

Concretely, the bisected non-focus *ke*-clauses in N‖ng are used for the introduction of or shift to new participants; in exclamations of surprise; heightened assertion (including the expression of obligation, very similar to (9) of !Ora); the provision of settings, explanations, and other types of background information; and, finally, in weather and season expressions. This range of contexts matches very closely the profile of theticity identified by Sasse (1987: 566–567). "Declarative" *ke* thus instantiates a subtype of the cleft-like structure conveying a S/A-central thetic statement. We come back to the relation between cleft-like clauses and theticity in section 3.1. For the time being, we diagnose that *ke* in N‖ng can be characterized in a unified way by having scope over diverse types of preceding constituents. Its constructional profile is summarized in Figure 4 (cf. the numbering that parallels Figure 1 for Northern Khoekhoe *ge*, structure IIb of N‖ng corresponds to IV of Khoekhoe).

I	[Term	*ke*]			Identification
IIa	[Focus	*ke*]	[(S/A)	V Other]	Term focus
IIb	[S/A	*ke*]	[Ø	V Other]	Entity-central theticity (earlier IV)
III	[Discourse anaphor	*ke*]	[S/A	V Other]	Discourse linking

Figure 4.   Structural and functional profile of *ke*-constructions in N‖ng

2.3. North-central Ju *má*

We conclude section 2 with an analysis of the clause-second particle *má* in Ekoka !Xun, a variety of the North-central subcluster of the language complex Ju, spoken in southern Angola, across northern Namibia, and north-western Botswana and representing one of two branches of the Kx'a family.[4] In previous treatments, the particle *má* is described as a topic marker that is "obligatory for subjects in declarative main clauses" (Heine & König 2015: 260, cf. also König 2008a, 2008b: 273–276). This appears to be the case in example (14).

(14) *Càālò* **má** *ā*   *cŋ̀-ā*   *g‖ú*   *kē*   *‖xūm̀* *khùyā*
      Calo ?   PROG   drink-VE   water MPO   river   place:CSTR
      'Calo is drinking water at the river.'   (König 2008a: 70)

König (2008a: 255) offers a number of other formal attributes of sentences exhibiting the particle. While most are compatible with the constraints applicable to Khoekhoe *ge* and N‖ng *ke*—for example, being excluded from interrogative clauses—some, like the incompatibility with sentences starting with a coordinator word, are open to question (cf. Pratchett 2017: 230). In the following, we apply a similar corpus-based approach to the analysis of *má* using only naturalistic Ekoka !Xun texts from König & Heine (2001), namely four short narratives by two speakers (we adapt the glossing in all examples to the conventions used elsewhere in this article).

The first use of *má* in Ekoka !Xun is that of a predicator in an identificational clause, as shown in (15) and (16). In the second example, whose context is the explanation of how a hunter should approach his prey depending on whether it can sense him by using the wind, the clause connector *tà* attests to the clause-final position of *má*.

(15) {Where are they? They are right there.}
      tū'm-cí   **má**
      be.near-NOM   ID
      'It is a nearby place. [our translation]'   (König & Heine 2001: 179)

(16) {Then you see it [small animal] and just hit it.}
      hà   *ŋ̀ŋ*   *n!hàù*   **má** *tà*   *òhā*   *hǎ-è*   *|ōā*
      PRO1   PROX   rabbit.1   ID   CONN   COP   PRO1-REL   NEG
      ‖òhlĭ-yà
      sense.wind-VE
      'It is a rabbit and [it] doesn't sense the wind. [our translation]'
                                              (König & Heine 2001: 184)

The second *má*-construction identified in the corpus takes the form of a cleft-like sentence and functions primarily to express contrastive term focus. As with

---

[4] Other varieties in the Ju language complex have similar but partly different clause-second particles. See, for example, Güldemann & Pratchett (2014) regarding the Southeastern Ju subcluster, where the phenomenon even exhibits a complex case of historical layering.

N‖ng, there is no overt subordination in the cleft-like sentence and the fronted focal referent is not resumed anaphorically in the background clause.

(17) {You are the big one, pull!}

à	hŋ́	**má**	̀ohā	n‖à–n‖à
2SG	EMPH	TF	COP	big

'No, [it is] YOU [who is] are the big one!'   (König & Heine 2001: 168)

While we assume that different participants can be focused, the only example in the available text corpus is (17) with contrastive S focus. Heine & König (2015: 262) suggest that the emphatic particle *hŋ́* is the focus marker and *má* "is already grammaticalized in its new function as a subject case marker". The inconclusive case hypothesis aside, this interpretation does not take the real possibility of a cleft-like structure into account: in the event of the omission of the background clause an identificational clause, as in (15) and (16), remains.

The third construction with *má*, exemplified in (18), is also analyzed as a cleft-like sentence. Here, a fronted deictic *kā–ndò'à* 'that' (or *kū–ndò'à* 'there', as in (25) below) refers anaphorically to preceding propositional content and emphasizes the temporal immediacy and/or causality holding between two states of affairs. The entire pattern can be paraphrased as 'that is when/why'. This structure is very akin to construction III in N‖ng (cf. (13) and Figure 4 of section 2.2), where the particle *ke* is part of phrasal discourse linkers.

(18) {The horse kicked the hyena and the jackal repetitively}

kā–ndò'à	**má**	ǀāàlè	xāŋ́	kē	ǃàhà
PRO3-DIST	TF	jackal.1	then	PST	run.away

'(It is) THEN [that] the jackal ran away.'   (König & Heine 2001: 169)

The next construction involves a sentence-initial temporal *kā*-clause that is offset from the following main clause by *má*. We argue that the meaning of (19) is that the main clause event is immediately precipitated by the dependent event. This is the result of the *kā*-clause being in the scope of *má* and its associated function of raising the discourse saliency of preceding constituents. The expression as a whole foregrounds the dependent clause and focusses on the tight temporal relation between two events (cf. Güldemann (2002: 262–268) for the similar effect of a foregrounding auxiliary in Shona). In English, this is best conveyed by introducing the dependent clause with conjunctions like 'whenever' or 'as soon as'.

(19) {Jackal takes hyena's tail and horse's tail and ties them together}

kā	hà̀	‖'ŋ̀–‖x'āē	cā	**má**	hà̀	kū–ndò'à	kwèé
SUB	PRO1	tie-together	3DU	FG	PRO1	LOC-DIST	say

xāŋ́	ǃ̀ò'è
then	pull

'When [as soon as] he had tied the two together, he said: "Pull!"'

(König & Heine 2001: 167)

As mentioned above, examples of term focus in the form of a cleft are sparse

in the corpus. There are, however, ample tokens of *má*-clefts marking entity-central theticity. We provide here four typical cases: the beginning of a story in (20), the presentation of a participant in (21), and two exclamations in direct reported speech in (22) and (23).

(20) {Opening line in a story}
　　　kwā　ŋ̀ŋ̀　　**má**　*gè-ā*　　*|xūúnnū cā*　　*‖āè*
　　　LOC DEM　ECT　exist-VE crocodile 3DU　monkey
　　　'There were crocodiles and a monkey.'　　　　　(König & Heine 2001: 173)

(21) {One day when the monkeys came again to drink [at the river]}
　　　|xūúnnū-mà　**má**　*hà̀*　*t̀àqè*　　*kē*　*‖x'òà　t-ē*　　　*kwèé*
　　　crocodile-DIM ECT　PRO1 mother.1 PST send　CONN-PST say
　　　'[there was a] small crocodile, his mother sent [him] telling [him]…' [or: 'there was a small crocodile whose mother had sent him saying...']
　　　　　　　　　　　　　　　　　　　　　　　　　（König & Heine 2001: 174）

(22) {And the jackal says he will lie to the lion and say:}
　　　kū-ndò'à n!ún　**má**　*gè-ā*　　　*tcàò*　　*cā*　　*gàò*
　　　right:there　　　ECT　exist-VE　　hunger　3DU　thirst
　　　'There is hunger and thirst!'　　　　（König & Heine 2001: 179）

(23) {The animals know that lion is around and jackal goes to elephant and pro-claims:}
　　　è　　　*!h̀àwà*　**má**　*ā*　　　　*tōān*
　　　1PL.EX　family　ECT PROG　finish'
　　　'Our family is being finished!' (i.e. the lion is eating them)
　　　　　　　　　　　　　　　　　　　　　　　　　（König & Heine 2001: 178）

When compared to the constructional counterparts in the languages/dialects treated earlier, the Ekoka !Xun examples are interesting in that the constituents before *má* are recurrently not S/A referents. In (20) and (22) it may be argued that this is due to an idiosyncratic argument structure of the verb *gè* with the valency-sensitive suffix -*a*. In (21), however, the best analysis seems to be that the initial noun is the possessor of the A in the following clause. Nevertheless, clear S/A-oriented cases of the construction type do exist, as in (23). Entity-central thetic sentences with *má* also serve other typical functions attested elsewhere, for example, topic shift toward an activated (or accessible) referent, as in (24).

(24) {the **jackal**ᵢ beat the horseⱼ to stand up,}
　　　|òh̀è　**má**　*xāŋ́*　*kē*　*!!'h̀è*　*g|ùĩ*
　　　horse　ECT　then　PST　kick　hyena
　　　'And the **horse**ⱼ then kicked the hyenaₖ'　　　（König & Heine 2001: 168）

In Ekoka !Xun, more than one instance of *má* can appear in a sentence, as illustrated in (25), in which the first token involves a deictic discourse anaphor with *kū-ndò'à* 'there' and the second token encodes topic shift. The deictic *kū-ndò'à* 'that place' could be interpreted either as in the locative sense 'from there' or simply as 'then'.

(25) {I never want to see the jackal again!}

kū–ndò'à	**má**	*cā*	**má**	*xāŋ́*	*wòhà*	*‖àqŋ̄–ā*	*kòè*
LOC-DIST	FG	3DU	ECT	then	just	separate-VE	RECP

'From then on the two separated forever.'   (König & Heine 2001: 169)

Finally, there are 25 *má*-tokens, representing about a fifth of the total, that still defy a robust classification based on the previous comparative or language-internal evidence. We discuss two patterns that are recurrent enough to merit treatment as subtypes (*má* is glossed accordingly as ?).

(26) {Jackal and hyena both agree that neither of them will go walking on the next day}

tà–kā	*n!āē‖àŋ*	**má**	*|àālè*	*n|è*	*m–ē*	*!!'hàò*	*cū*
CONN	next.day	?	jackal	alone	ECT-PST	walk	around

'But [it is] the next day [that], jackal alone walked around.'

(König & Heine 2001: 165)

(27) {The lion usually catches and eats the animals [at the water hole]}

gāō	*hà̄–ē*	*kē*	**má**	*‖hā–m̀hè*		*kū–ndò'à*	*kwèé*
day	PRO1-?	?	?	animal.3-DIM.PL		LOC-DIST	say

'One day, the animals said [why is our family being eaten up by the lion?]'

(König & Heine 2001: 177)

One type with seven tokens, illustrated in (26) and (27), appears to be related to the structure in (18) and (25) where *má* follows a deictic, because the elements preceding it are also time adverbs. While the case in (26) may still be viewed as focusing on the immediacy of the event with respect to the previous state of affairs, this seems less likely in (27). Both examples do not convey a shift of participant topic but rather of the narrative scene as a whole. While this partly betrays the relation to the original core domain of theticity, it is conceivable that such cases increase in frequency and by virtue of their overuse become a structure for plain frame-setting topics. This pattern comes close to unmarked discourse linkage with *ge* in Khoekhoe.

Another subtype in this "Other" category, shown in (28), has eight tokens; they are found exclusively in a procedural text with this particular semanto-syntactic configuration.

(28) a.

kā	*bà*	*‖áúl̃e‖hā–mà*	*bà*		**má**	*gù*	*à*	*n‖à̀ò* [...]
SUB	2SG	hunt animal-DIM	2SG		?	take	2SG	bow

'If/when you hunt an animal, you take your bow [and arrow and go into the bush].'

b.

kā	*bà*	*‖àbà*	*!'ō*	*bà*	**má**	*xāŋ́*	*n!hún–ā*	*tî*
SUB	2SG	go.in	bush	2SG	?	then	walk.SG-?	search

kā
PRO3

'When you go into the bush, you walk and look for it.'

c. *kā   bà   n!hún-ā tī   kā   bà   **má** xāŋ́ n!hún-ā n!ùm̀*
  SUB 2SG walk-? search PRO3 2SG ? then walk-? creep.up
kx'ōān kā
look.at PRO3
'when you walk and look for it, you then walk and creep up on it.'
(König & Heine 2001: 181–182)

All sentences in (28) show an initial subordinate *ka*-clause followed by the main clause with an initial second-person pronoun in the scope of *má*.[5] As the particle does not follow the dependent clause directly, this is only superficially similar to the pattern exemplified in (19). Such sentences also do not instantiate a topic shift. One can conceive of various hypotheses to explain the structure functionally and embed it within the overall profile of *má*. With the restricted data at hand, we refrain from this but note that such cases seem to represent another context for *má* moving away from its original domain of highlighting a preceding constituent.

I	[Term	**má**]		Identification
IIa	[Focus	**má**]	[(S/A) V Other]	Term focus
IIb	[Term	**má**]	[(S/A) V Other]	Entity-central theticity
IIIa	[Discourse anaphor	**má**]	[S/A V Other]	Foreground discourse linking
IIIb	[[*ká* Clause]	**má**]	[S/A V Other]	Foreground discourse linking
IIIc	[[*ká* Clause] S/A	**má**]	[∅ V Other]	?

Figure 5.   Structural and functional profile of *má*-constructions in Ekoka !Xun

The survey of diverse *má*-constructions is summarized in the structural and functional profile of Figure 5. Again, it reveals considerable similarities with that in Figure 1 of Khoekhoe and Figure 4 of N‖ng, as can be discerned from the parallel numbering of structures.
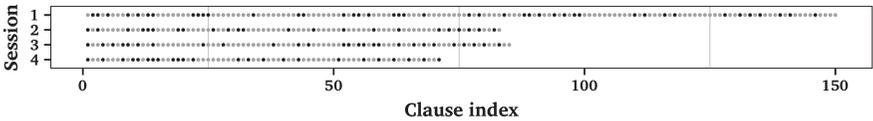

Figure 6.   Frequency of *má* in Ekoka !Xun sources (black dot = *má*-clause, light grey dot = unmarked clause)

We proceed with the quantitative analysis of the corpus. The four narratives contain 349 declarative clauses displaying 122 instances of *má*. This figure goes against Heine & König's (2015: 263) and König's (2008b: 273–276) claim that "on average, every declarative main sentence [in Ekoka !Xun] contains at least one topic marker which has acquired properties of a subject marker". We suspect

---

[5] Incidentally, this pattern looks very similar to the Khoekhoe pattern [[BG clause]=S/A pronoun *ge* [clause]] in Figure 1 of section 2.1. For now, we abstain from speculating about any possible significance of this superficial isomorphism.

that their generalization is based largely on elicited data that we find has a greater affinity for this gram type (see section 3.2).

Table 3.   Frequency of different *má*-constructions

Construction type	Total	% of clause total (349)
I Identification	7	2.0%
IIa Term focus	1	0.3%
IIb Entity-central theticity (including topic shift)	63	18.0%
IIIa Discourse linking with deictic anaphor	8	2.3%
IIIb Discourse linking with *ka*-clause	18	5.2%
Other (including IIIc)	25	7.2%
Total	**122**	**35.0%**

Table 3 presents the frequency of the different *má*-constructions, few of which are instances of topic marking, making such an analysis very unlikely. The data also show that focus clefts of type IIa need not be more frequent in discourse than thetic clefts of type IIb.

3. Discussion
In this section, we offer a unified functional account of clause-second particles in the Kalahari Basin (section 3.1), discuss some implications of our results for IS research in general (section 3.2), and close with a few remarks on the repercussions these have for typological research on grammatical relations (section 3.3).

3.1. A unified semantic map
We have argued above that most clause-second particles dealt with have their central functions in the IS domain but also involve other uses related from a grammaticalization perspective. Since they establish a "family" of constructions with a diverse functional profile, we propose a unitary analysis within the semantic-map framework shown in Figure 7 and, in the following, discuss the major formal and functional shifts that are central to the present topic. (The poorly understood IS structure Ib occurs only in Khoekhoe and we do not feel it contributes to our core subject matter, so that it is not treated further.)

IIb: S/A-central thetic statement	IIb: S/A-central topic shift/contrast	IV: S/A-topical declarative

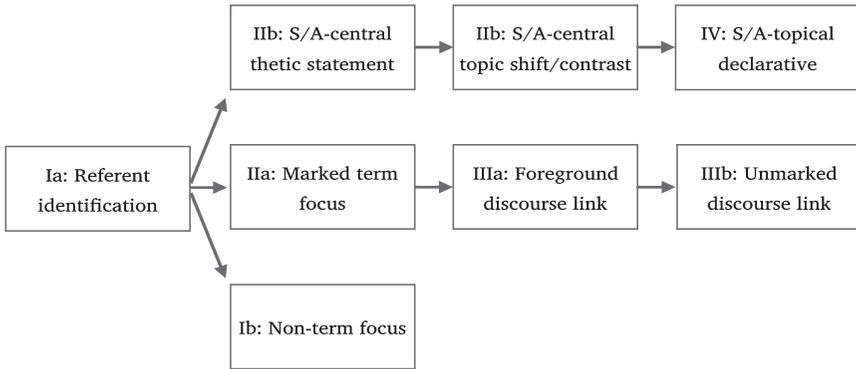| Ia: Referent identification	IIa: Marked term focus	IIIa: Foreground discourse link	IIIb: Unmarked discourse link

Ib: Non-term focus

Figure 7.   Semantic map of the uses of clause-second particles in the Kalahari Basin

The crucial initial development in Figure 7 from function Ia to II is a cross-linguistically recurrent phenomenon: a non-verbal identificational (or presentational) sentence where a particle (or some other element) instantiates the predication is used as the nucleus of a complex cleft-(like) or pseudo-cleft-(like) sentence that syntactically exposes a term against a background clause. The analysis of this construction has traditionally concentrated on its recurrent role as a focus structure (see Schachter (1973) as an early influential study in a row of numerous others). With respect to this change from Ia to IIa, we have nothing to add here except for the observation that it is attested in all the languages surveyed above.

The same syntactic structure is, however, also recruited for a very different IS function, represented in Figure 7 as IIb, attesting to the absence of a one-to-one mapping between form and function in IS and beyond. As already outlined in section 2.2, Sasse (e.g., 1987) argues persuasively for a recurrent distinction between two types of sentence organization, namely an unmarked "categorical statement" involving a sentence-internal IS profile and a marked "thetic statement" wherein any unbalanced configuration is neutralized toward a compact information unit. He (1987: 526) classifies thetic statements further into two subtypes called "entity-central" and "event-central". Viewed normally in the context of term focus, the structural properties of the bisected cleft-(like) structures are particularly adept for encoding entity-central theticty, as amply shown by Sasse (1987) and for Tuu languages of the Kalahari Basin specifically by Güldemann (2010) (see also Tosco (2002) for similar findings in Somali). This is because such split sentences are associated with intercepting the syntactic subject-predicate relation of a categorical statement, on the one hand, and cancelling the assertivity of the central state-of-affairs expression, thereby backgrounding it, on the other hand. In Sasse's (1987: 542) words, the special effect of cleft-like structures is to:

> disrupt the direct connection of the entity and the event by first introducing the entity by an existential or copular clause and then, quasi appositionally,

making a statement about it. This 'making a statement about', however, differs radically from the normal theme-rheme or topic-comment structure. The assertion of the event is subordinated pragmatically, semantically, and syntactically to the existential expression.
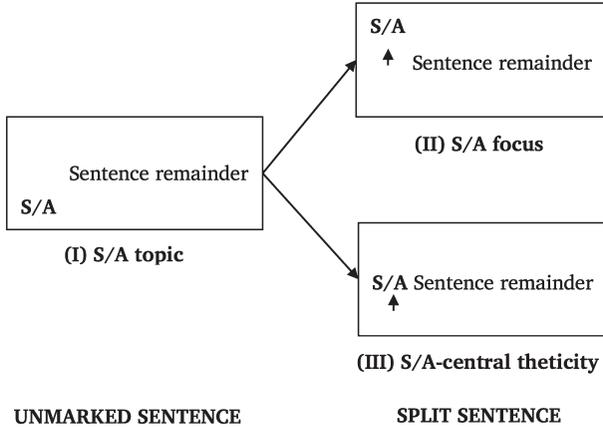

Figure 8.   Three salience relations between S/A argument and its sentence

The mechanism underlying the double IS function inherent to a syntactically clefted structure is schematized in Figure 8. The unmarked categorical sentence shown in panel (I) on the left is associated with the default IS role for a S/A argument as topic (cf. e.g., Güldemann, Zerbian & Zimmermann 2015: 159–160), which is symbolized by the higher focal salience of (parts of) the sentence remainder. The split sentence that is oriented toward the S/A argument by exposing it syntactically is represented in the two panels on the right, the upward arrow symbolizing its increased syntactic prominence and the resulting dissociation from its default IS role as topic. The focus construction in panel (II) elevates the S/A argument toward a marked focus role that is more salient than the rest of the sentence—the background clause (cf. e.g., Fiedler et al. (2010) for the specific relationship of S/A focus and formally marked IS structures). The cleft-construction with the thetic function in panel (III) also elevates the S/A argument but merely toward the salience level of the remaining part of the sentence. In doing so, it cancels a possible hierarchy between the different information units, which results in an entity-central thetic statement in the above sense. This phenomenon holds for N‖ng and Ju and, we argue, should be reconstructed for Khoekhoe.

Bisected cleft-like structures thus tend to be insufficiently characterized functionally, which is due to the accidental history of science. Since Schachter's (1973) cross-linguistic treatment of expressions that resemble English clefts in formal and functional terms and the incorporation of its findings in theories of transformational syntax and "movement", such structures are well researched regarding their focus use. Entity-central theticity, the recurrent second function is, however,

still underestimated. This is in spite of the fact that it has been documented cross-linguistically in sufficient detail by Sasse (1987) and later works and can in some languages be a very salient or even the more frequent function according to fuller analyses based on natural discourse corpora, as evident from the above data or similar findings in other languages (cf. Tosco (2012) on Somali).

A second development recorded in the semantic map of Figure 7 involves the split structure under IIa coming to create different types of discourse linkage. Here the initial constituents in the scope of the relevant particles, referring largely to time and reason, are of diverse syntactic complexity and pragmatic status. One extreme is represented by Ekoka !Xun and other Ju dialects where the particles follow a fully elaborate dependent clause that, in line with the central use of the particle, is foregrounded within a marked type of clause linkage recorded under IIIa. The other extreme is found in Northern Khoekhoe where the conjunction words *ō*, *tsî*, and *hîa* or adverbial subordinate clauses before *ge* merely serve as signals of discourse cohesion recorded under IIIb. Other cases in |Xam (see Güldemann 2013b: 424), N‖ng, and Ekoka !Xun fall between these two extremes. This suggests that the situation in Khoekhoe is the endpoint of a process of grammaticalization in which the earlier status of the pre-particle material as foreground has been lost completely.

This "re-pragmaticization" of material in the initial position of a complex focus structure from IIIa to IIIb anticipates what is also relevant for the development from IIb to IV and which is central to our discussion of the syntactic and pragmatic role of sentence-initial S/A arguments. Regarding Khoekhoe *ge* we argue for a functional change in the Namibian standard variety from an entity-central thetic structure oriented to a S/A-referring constituent to a plain declarative sentence. The crucial phenomenon here concerns the shift from a syntactically exposed non-topical referent within a marked thetic statement to a topical one in an unmarked categorical statement. The exact mechanism cannot be reconstructed transparently from Khoekhoe-internal data and in section 3.2 we entertain a partly non-functional explanation for the historical change in this language.

At this point, we refer to the data provided in section 2 on the comparable particles in N‖ng and even more so in North-central Ju because these allow a glimpse at how a construction dedicated originally to focal or at least non-topical S/A arguments potentially encroaches on the topic domain. For one thing, we have shown contexts where they encode S/A topic shift and contrastive selection out of established referent sets that due to their accessibility may be difficult to distinguish from unmarked topics. A pattern not treated above but highly relevant in this context has been observed in |Xam (Güldemann 2010: 80) and ǂKx'ao‖'ae (Güldemann & Pratchett 2014), namely paired contrasts between both S/A and P referents in which the former end up as sentence-initial contrastive topics marked by cleft-like structures, as demonstrated with clause-second *kòm* from ǂKx'ao‖'ae.

(29) {If a young woman and a young man like each other,}

    a. *tè*   *dshàú-mà*   **kòm**   *gè–à*   *|xòà* *hǎ*   *màmà*   #
        CONN   woman.1-DIM ECT   stay-VE COM PRO1 granny.1

        sà   *kxàè*   *g!òmtjù*
        3DU   have   house.3

        'the GIRL stays with her grandmother and the two have a G!OMTJU,

    b. *tè*   *!árîkxàò*   **kòm**   *kxàè*   *!'hǎàn*
        CONN   young.man.1 ECT   have   house.3

        'while/but/and the BOY has a !'HAAN.'

In |Xam, this encoding of contrastive topics is otherwise still far from marking anything like a topic, which indicates that this is a generally latent use of entity-central thetic structures of the split type. Finally, in connection with Ekoka !Xun *má* of section 2.3, we hinted at a couple of structural patterns in which a potential shift toward topic marking is conceivable. These uses of the relevant particles—and the possibly other still undiagnosed cases—provide potential bridging contexts for their shift toward topic markers, or as a related context, their use in plain declarative sentences with S/A topics.

Sasse, when fleshing out his concept of theticity, has himself (1987: 560) entertained that a categorical reading of a sentence tends to be generally preferred as soon as it contains an expression of an entity as well as one of an event related to the respective referent:

> When an event, part of which is an entity, is stated, the problem arises that the entity is a possible candidate for a predication base, and the event is a possible candidate for a predicate. This problem is due to the fact that if a linguistic expression denoting an event and a linguistic expression denoting an entity are combined in a sentence, the most normal reading is that the relation between the two is a predicative one.

(30) a.   [S/A TOPIC - PREDICATE - ...]

      b. [S/A NON-TOPIC - **PARTICLE**] - [S/A GAP   - PREDICATE - ...]

The mechanism of the implied change also has a concrete structural aspect if, as is the case in the above languages, the constituent exposed in the clause-initial position refers predominantly to the S/A. The two relevant sentence types are schematized structurally in (30), the unmarked categorical sentence in (30a) and the originally split thetic sentence in (30b) (syntactically identical positions are aligned). While syntax and pragmatics of the two sentence patterns are quite distinct, the differences on the surface are minimal in the languages concerned, as the S/A-oriented cleft in (30b) has a gap in the subject position of the out-of-focus clause. The only overt difference of the two is thus the absence/presence of the particle between the initial S/A constituent and the final predicate. It is thus conceivable that, in analogy to the pattern in (30a), the pattern in (30b) invites an interpretation of the S/A argument as the predication base in line with Sasse's above idea. The IS behavior of the pattern in (30b) discussed above also makes a

change in the pragmatic configuration likely.

Over time, the original thetic function in bisected sentences with exposed S/A arguments can erode in that this constituent is reinterpreted as a canonical subject-topic in the grammatical-relation sense and the particle between it and the rest of the sentence merely comes to mark the status of the clause as asserted and independent, or of the S/A constituent as subject. This change may well be accelerated by an overuse of the thetic structure, decreasing its pragmatic markedness. It is relevant in this respect that theticity appears to be inherently subject to considerable variation in frequency when it comes to internal text dynamics, discourse type, idiolect, dialect, and possibly other factors, which would seem to provide ample opportunity for change. This can result in the complete change from function IIb to IV in Figure 7—a reanalysis from a S/A-central thetic statement to a normal categorical "declarative" with a topical S/A subject.

3.2. The challenge of theticity, or why discourse analysis matters

A recurring theme throughout the above discussion is relevant for the treatment of IS in general. In all languages treated in section 2 we challenge previous analyses of the clause-second particles. They have been analyzed variably as markers of declarative~indicative sentence type (Northern Khoekhoe *ge*, N‖ng *ke*), subject (Northern Khoekhoe *ge*, Ekoka !Xun *mâ*), or topic (Ekoka !Xun *mâ*). These accounts focus too much, we argue, on superficial syntactic distributions. The declarative characterization reflects the exclusion of the particle construction from specific contexts like subordinate clauses etc. and thus interprets a negative distribution by means of an apparent positive function. The characterizations as subject or topic marker capture the phenomenon that the element follows the clause-initial S/A argument from which it is concluded that it also exclusively relates functionally to this constituent. Both lines of analysis do not account for many other characteristics of the particle use and are also contradicted by data that would be relevant for the respective functional account. We argue that this is due to the lack of assessing the particles within their constructional context by means of natural discourse data, particularly so for thetic structures that are notoriously ambiguous within the IS domain, as we discuss below. Looking at the empirical basis of the earlier proposals for N‖ng and Ekoka !Xun, the absence of a more systematic discourse-based approach is self-evident. (The discussion of the case of *ge* in Northern Khoekhoe is deferred to the end of this section.)

Collins & Namaseb's (2011: 9) "declarative" analysis of N‖ng *ke* is, as observed above, contradicted by their own data. Their stance was possibly influenced by two circumstances. The second author is a native speaker of Namibian Khoekhoe and may well have been biased toward the account of the very similar particle *ge* in his own mother tongue. The first author started his fieldwork with access to Güldemann's (2003) draft sketch, which contained a precipitant analysis (later qualified by Güldemann & Siegmund 2009 and Güldemann 2010).

As for Ekoka !Xun *mâ*, we assume that Heine & König's (2015: 260–263) characterization as a topic marker that is "obligatory for subjects in declarative

main clauses" is skewed by an analysis of examples that are disconnected from discourse contexts, giving the impression of a high incidence of *má* after a sentence-initial S/A constituent. Some of the central aspects of their structural and functional assessment are not corroborated by their own textual data.

Our alternative proposal takes a fuller range of data into account and approaches the functionally problematic use of the particle in terms of Sasse's (1987) theticity concept in the sense of presenting information in a compact form that cancels a normal foreground-background profile. The above discussion should have made clear that theticity is an inherently versatile phenomenon, subject to an enormous amount of variation according to such factors as textual dynamics, discourse type, idiolect, dialect and presumably others.

In this respect, it is important to recall that the different analyses of N‖ng *ke* hinge on the distinction between elicitation and coherent texts. Güldemann's (2003) "declarative" account was based originally on Westphal's corpus while our present analysis in terms of theticity is grounded in a survey of a larger amount of natural discourse. The different frequency and distribution of *ke* in these two corpora is not coincidence. Elicitation is a pragmatically impoverished type of language skewed toward theticity vis-à-vis normal speech, even if one tries to provide context by means of questions and the like. In particular, responses to conventional non-contextualized elicitation tasks share with thetic statements the character of an "out-of-the-blue" context that lack sufficiently accessible participants. This results in a higher occurrence of structures that are normally used for the IS function of theticity—this in spite of the fact that it is often quite difficult to identify them by means of targeted IS elicitation. This, we assume, is behind the skewed distribution of the relevant particle structures between elicitation and natural speech.

Table 4.   IS configurations across discourse genres

Discourse	Typical (though not exclusive) IS dynamics	Typical IS configuration
Conversation	Contrastive discontinuity	Marked focus
Elicitation	"Out-of-the-blue" utterance	Theticity
Narrative	Topic discontinuity	
	Topic continuity	Categorical declarative

We hypothesize in general that the incidence of different IS configurations and their associated constructions differ considerably across various discourse genres, as presented in a preliminary and simplified fashion in Table 4. What follows from this is clear: language-specific structures produced in isolated elicitation contexts are potentially opaque with respect to their conventional IS configuration and thus remain functionally indeterminate and, more crucially, are misleading for linguistic analysis. Thus, a construction that is repeatedly elicited may in fact be marked and, after a thorough corpus analysis, turn out to be rarer in actual language use. Hence, without an analysis of a comprehensive spoken language corpus, the functional interpretation of elicited data runs the risk of being partly or com-

pletely inadequate. Put more succinctly, an elicited single-sentence construction may not express declarative, subject, topic, and the like but rather be thetic.

More generally, elicitation, while indispensable for linguistic research, provides a data type that is notoriously difficult to interpret. The limitations of elicited data, in particular for IS research, are beginning to become clear (cf., e.g., Apel, Jacob & Tadesse (2015) and Fiedler (2017) for some discussion of the different analytical results from elicitation and more natural discourse data). However, the caveat is worth reiterating in view of the long tradition of elicitation for the evaluation of grammatical structures in general and of IS configurations in particular, especially when they are morpho-syntactically marked.

We conclude this part by addressing the case of Northern Khoekhoe, which is not the same as in N‖ng and Ekoka !Xun. Certain historical and dialectal evidence cast doubt on the declarative analysis of the particle *ge*. In fact, such an account is only given in grammatical descriptions that reflect Standard Namibian Khoekhoe. The modern colloquial variety also seems to have *ge* as a declarative marker according to field notes and information furnished by the mother tongue linguist S. Job, although a systematic analysis is pending. This would mean that in one case in the Kalahari Basin grammaticalization has gone full circle. At the same time, we lack sufficient historical data to flesh out this development with empirical facts.

It is also crucial that we discuss a sociolinguistic aspect of the development. It is without doubt that modern spoken Khoekhoe in Namibia has been shaped for several generations by the spread of the prestigious standard variety, notably in the church as well as the formal education system and its application of prescriptive language norms. As mentioned above, the public use and teaching of the standard grammar communicate a clear employment of *ge* as a declarative marker. This variety, however, is the result of a complex codification process (see Haacke 1989 for details). This was initiated in the 19th century by missionaries (cf. Tindall 1857, Wallmann 1857) and with their active participation was later continued by the German and South African colonial authorities in order to create a standard language (cf. Seidel 1891, Planert 1905, Meinhof 1909, Dempwolff 1934/5, Olpp 1963). Thus, the primary agents in this process were non-native speakers who were faced with enormous problems deciphering the use of thetic *ge*. The following quotes about the so-called "*ge* subjectivus" from Olpp (1963; our translation), a major source for Khoekhoe teaching at the time, amply testifies to this:

> This particle is used in DECLARATIVE MAIN CLAUSES [sic] without necessarily occurring in every main clause. However, it must not appear in subordinate, interrogative (direct or indirect) or imperative clauses.
> Note: The beginner is advised at first to use the *ge* subjectivus in ALL [sic] main clauses, which shall also be practiced in the following exercises and examples. In colloquial speech, in conversation etc. it is often dropped at the will of the speaker. (Olpp 1963: 10)
> Since the *ge* subjectivus, heretofore used regularly in declarative sentences, need not always occur, it shall be omitted where dispensable. [Its] retention

or omission is frequently determined simply by euphony of speech for which rules are impossible to establish. (Olpp 1963: 30)

With such vague but nevertheless authoritative guidance, the lesson for learners, including future school teachers, would have amounted to what is suggested to beginners—to use *ge* throughout. This implicit prescription reflects an impoverished linguistic intuition that did not spring from constant exposure to and habit of daily language practice. Hence, the use of Northern Khoekhoe *ge*, differing between native natural discourse in Richtersveld Nama and other "conservative" varieties on the one hand and the standard and modern mainstream variety of Namibia on the other, can be seen as a sociolinguistically mediated case of the above difference between spontaneous spoken discourse and elicitation.

3.3. Non-topical S/A arguments, clefts, and main clauses: a typological perspective

The emergence of an unmarked declarative sentence out of a cleft-like structure that exposes, among other things, the S/A argument is the endpoint of a complex process of grammaticalization. It represents a full cycle of an earlier formally unmarked structure being replaced by a formally marked one, the former ending up in "peripheral" contexts like other more conservative clause types, fixed expressions, etc. As can be expected, this phenomenon is not unique to the Kalahari Basin and has been observed before. For example, Rapold reports such a case in Benchnon, a Ta-Ne language of Omotic, writing (2007: 82):

Cleft sentences, whose typical function is commonly seen as focus on the extra-posed constituent ..., may develop into pragmatically unmarked, default constructions through the loss of markedness that accompanies the drastic extension of their functional domain. In the process, the amount of focality needed to trigger a cleft construction is progressively diminished, until cleft constructions are so frequent that they become the norm.

We have tried here to flesh out this development by means of a detailed analysis of discourse data from several languages. In line with Güldemann, Zerbian & Zimmermann (2015: 159–160) we conclude that this process is tied to a more general semantic-functional phenomenon, namely the default association of an S/A referent with the pragmatic status as topic and the fact that from a cross-linguistic perspective such a pattern is frequently fixed in a language-specific grammatical relation in the sense of Dryer (1997). As soon as there is a dissociation between semantic role and canonical pragmatic status, many languages resort to a marked linguistic expression—in the present areal context, split cleft-like sentences.

It is important to point out in this connection that the grammaticalization outlined in the semantic map of Figure 7 and its possible historical repercussions do not only concern sentence type marking. We observed above that the late stages of the grammaticalization cycle have also received other labels, such as

marker of "emphatic nominative" in |Xam or "subject" in Northern Khoekhoe and Ekoka !Xun. Such terms clearly refer to a quite different domain, namely grammatical relations. At the same time, looking at the relevant cases, it is clear that the above terms do not refer to a genuine grammatical relation but rather to the set of semantic S/A roles in a situation where, at least originally, it does not coincide with its default pragmatic role as topic. The result is an instance of differential argument marking – something that has become an increasingly discussed topic in the more recent literature.

Our present discussion relates specifically to what is called "differential subject marking" (see, e.g., De Hoop & De Swart 2008) or what we call differential S/A marking. In the above cases, differential S/A marking is motivated by the contrast between the non-topical and topical status of S/A arguments. It goes without saying that the wider phenomenon is not tied in structural terms to cleft-like sentences dealt with here, given that entity-central thetic statements can be conveyed by a diverse range of constructions (see already Sasse (1987) for some discussion). For example, the particles in the Kalahari Basin can be viewed as segmental analogues to so-called "subject-accented (in our terms S/A-accented) sentences" in English.

Casting the net wider, differential S/A marking is in turn related to so-called "marked nominative" systems (cf. Handschuh (2014) for a global perspective and König (2006b, 2008b) and Dimmendaal (2014) for African surveys). This holds insofar as marked nominative systems can conceivably emerge from the generalization of constructional patterns of the above type for S/A constituents irrespective of their IS status. Likewise, a conceptual link should be made to ergative-absolutive case systems as the possible result of a generalization of the relevant constructional pattern for A constituents alone, for which it would be particularly interesting to investigate the IS profile of so-called optional ergative marking (see, e.g., McGregor 2006, McGregor & Verstraete 2010).

All these phenomena together relate to some form of structural markedness of S/A constituents. A similar default association can be stated for P(atient)s and assertive focus: as soon as this default pattern does not hold, namely in the context with a topical P, marked sentence structures tend to be employed (Güldemann, Zerbian & Zimmermann 2015: 160). This is one motivation of the well-known phenomenon of differential object or P marking, which has been discussed explicitly in IS-related terms already in Bossong (1985) and with increasing frequency more recently, for example, in Nikolaeva (2001) and Dalrymple & Nikolaeva (2011).

In general, the research on differential argument marking is increasingly showing that this phenomenon is intimately linked to IS and it cannot be accounted for without considering IS-related constructions (see, e.g., Bickel, Witzlack-Makarevich & Zakharko (2015), or the contributions in Seržant & Witzlack-Makarevich (2018) on diachrony). Hence, dedicated research on IS constructions and their potential historical development promises to provide major explanatory power for elucidating the theory of grammatical relations. One may in

fact wish to investigate more systematically whether grammatical relations in some languages are not only shaped by the IS status of a certain constituent but rather reflect it directly—this in addition to or even more strongly than its semantic role.

4. Conclusions

In this paper, we scrutinized three clause-second particles in languages from the distinct lineages forming the Kalahari Basin area. We leave them with structural and functional profiles more complex than how we found them, with implications for the study of language far beyond the Kalahari Basin. We first presented evidence for the claim that, while very versatile, their functions primarily lie in the IS domain. We also proposed that these clause-second particles establish a family of constructions emerging from the initial employment of an identificational clause for a complex cleft-like sentence. The structural properties of the latter are particularly adept for encoding non-topical S/A constituents within entity-central thetic statements in terms of Sasse (1987). The overuse of this construction can eventually give rise to an expression that no longer displays a marked IS configuration but merely conveys a declarative sentence, potentially entering a system of differential S/A marking. Last but not least we advance that IS configurations and the associated constructions differ considerably across distinct discourse genres, specifically between naturalistic texts and elicitations. The last method of data collection in particular does not give full access to certain grammatical structures and, when exclusively used for linguistic analysis, may even lead in the worst-case scenario to drawing the wrong conclusions.

References

Apel, Viktoria, Peggy Jacob, and Henok Wondimu Tadesse (2015) A discourse-corpus approach to focus constructions in three African languages (Pular, Kenga, and Gamo). Paper presented at the Final Conference of the SFB 632 "Advances in information structure research 2003–2015", Humboldt University of Berlin, 9 May 2015.

Bickel, Balthasar, Alena Witzlack-Makarevich, and Taras Zakharko (2015) Typological evidence against universal effects of referential scales on case alignment. In: Ina Bornkessel-Schlesewsky, Andrej L. Malchukov, and Marc D. Richards (eds.) *Scales and hierarchies: A cross-disciplinary perspective*, 7–43. Berlin: De Gruyter Mouton.

Bleek, Dorothea F. (1928–30) Bushman grammar: A grammatical sketch of the language of the |Xam-ka-!k'e. *Zeitschrift für Eingeborenen-Sprachen* 19: 81–98/ 20: 161–174.

Bossong, Georg (1985) *Empirische Universalienforschung: Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr.

Collins, Chris and Levi Namaseb (2011) *A grammatical sketch of N|uuki with stories*. Cologne: Rüdiger Köppe.

Crystal, David (1997) *A dictionary of linguistics and phonetics*. London: Blackwell.

Dalrymple, Mary and Irina Nikolaeva (2011) *Objects and information structure*. Cambridge: Cambridge University Press.

De Hoop, Helen and Peter de Swart (eds.) (2008) *Differential subject marking*. Dordrecht: Springer.

Dempwolff, Otto (1927) Die Hervorhebung von Satzteilen als Anlass zur Verwendung besonderer Wortformen. In: Franz Boas et al. (eds.) *Festschrift Meinhof*, 73–79. Glück-

stadt/Hamburg: J. J. Augustin.

Dempwolff, O. (1934/5) Einführung in die Sprache der Nama-Hottentotten. *Zeitschrift für Eingeborenen-Sprachen* 25: 30–66, 89–134, 188–229.

Dickens, Patrick J. (1994) *English-Juǀ'hoan/Juǀ'hoan-English dictionary*. Cologne: Rüdiger Köppe.

Dickens, Patrick J. (2005) *A concise grammar of Juǀ'hoan with a Juǀ'hoan-English glossary and a subject index*. Cologne: Rüdiger Köppe.

Dimmendaal, Gerrit J. (2014) Marked nominative systems in Eastern Sudanic and their historical origin. *Afrikanistik online* 11, urn:nbn:de:0009-10-38596.

Dryer, Matthew S. (1997) Are grammatical relations universal? In: Joan Bybee, John Haiman and Sandra A. Thompson (eds.) *Essays on language function and language type: Dedicated to T. Givon*, 115–143. Amsterdam: John Benjamins.

Engelbrecht, Jan A. (1936) *The Korana: An account of their customs and their history with texts*. Cape Town: Maskew Miller.

Fiedler, Ines (2017) The structuring of information in natural discourse in Ama (Nyimang). In: Raija Kramer and Roland Kiessling (eds.) *Mechthildian approaches to Afrikanistik: advances in language based research in Africa, Festschrift für Mechthild Reh*, 45–61. Cologne: Rüdiger Köppe.

Fiedler, Ines, Katharina Hartmann, Brigitte Reineke, Anne Schwarz, and Malte Zimmermann (2010) Subject focus in West African languages. In: Malte Zimmermann and Caroline Féry (eds.) *Information structure: Theoretical, typological and experimental perspectives*, 234–257. Oxford: Oxford University Press.

Güldemann, Tom (1998) The Kalahari Basin as an object of areal typology – a first approach. In: Mathias Schladt (ed.) *Language, identity, and conceptualization among the Khoisan*, 137–196. Cologne: Rüdiger Köppe.

Güldemann, Tom (2002) When 'say' is not say: The functional versatility of the Bantu quotative marker *ti* with special reference to Shona. In: Tom Güldemann and Manfred von Roncador (eds.) *Reported discourse: A meeting ground for different linguistic domains*, 253–287. Amsterdam: John Benjamins.

Güldemann, Tom (2003) Westphal's field notes on Nǀhuki with a linguistic analysis. unpublished manuscript.

Güldemann, Tom (2006) Structural isoglosses between Khoekhoe and Tuu: the Cape as a linguistic area. In: Yaron Matras, April McMahon, and Nigel Vincent (eds.) *Linguistic areas: Convergence in historical and typological perspective*, 99–134. Hampshire: Palgrave Macmillan.

Güldemann, Tom (2010) The relation between focus and theticity in the Tuu family. In: Ines Fiedler and Anne Schwarz (eds.) *The expression of information structure: A documentation of its diversity across Africa*, 69–93. Amsterdam: John Benjamins.

Güldemann, Tom (2013a) Syntax: Taa (East !Xoon dialect). In: Vossen (ed.), 408–419.

Güldemann, Tom (2013b) Syntax: ǀXam. In: Vossen (ed.), 419–431.

Güldemann, Tom (2014) "Khoisan" linguistic classification today. In: Tom Güldemann and Anne-Maria Fehn (eds.) *Beyond 'Khoisan': Historical relations in the Kalahari Basin*, 1–41. Amsterdam: John Benjamins.

Güldemann, Tom (2017) Casting a wider net over Nǁng: The older archival resources. *Anthropological Linguistics* 59(1): 1–34.

Güldemann, Tom and Anne-Maria Fehn (2017) The Kalahari Basin area as a "Sprachbund" before the Bantu expansion. In: Raymond Hickey (ed.) *The Cambridge handbook of areal linguistics*, 500–526. Cambridge: Cambridge University Press.

Güldemann, Tom, Martina Ernszt, Sven Siegmund, and Alena Witzlack-Makarevich (2007–2014) *A text documentation of N|uu*. London: SOAS, Endangered Languages Archive, ELAR. URL: https://elar.soas.ac.uk/Collection/MPI194591

Güldemann, Tom and Lee J. Pratchett (2014) KoMmenting on ǂKx'ao‖'ae. Paper presented at the 5th International Symposium on Khoisan Languages and Linguistics, Riezlern, 14–16 July 2014.

Güldemann, Tom and Sven Siegmund (2009) The marker *ke* in N|uu—declarative or not? Paper presented at the 6th World Congress of African Linguistics (WOCAL), University of Cologne, 17–21 August 2009.

Güldemann, Tom and Alena Witzlack-Makarevich (2013) The risks of analysis without spoken language corpora: clause-second *ke* in Richtersveld Nama and N‖ng. Paper presented at the International Workshop "Information Structure in Spoken Language Corpora (ISSLaC)", University of Bielefeld, 10–12 June 2013.

Güldemann, Tom, Sabine Zerbian, and Malte Zimmermann (2015) Variation in information structure with special reference to Africa. *Annual Review of Linguistics* 1: 155–178.

Haacke, W. H. G. (1989) Nama: Survival through standardization. In: István Fodor, and Claude Hagège (eds.) *Language reform: History and future* 4, 397–429. Hamburg: Buske.

Haacke, W. H. G. (2006) Syntactic focus marking in Khoekhoe ("Nama/Damara"). In: Ines Fiedler and Anne Schwarz (eds.), *Papers on information structure in African languages*, 105–127. Berlin: Zentrum für Allgemeine Sprachwissenschaft, Typologie und Universalienforschung.

Haacke, W. H. G. (2013) Syntax: Namibian Khoekhoe (Nama/Damara). In: Vossen (ed.), 325–340.

Hagman, Roy S. (1973) Nama Hottentot grammar. Ph.D. thesis: Columbia University New York.

Hagman, Roy S. (1977) *Nama Hottentot grammar*. Bloomington: Indiana University Publications.

Hahn, Theophilus (1870) *Die Sprache der Nama, nebst einem Anhange enthaltend Sprachproben aus dem Munde des Volkes*. Leipzig: J. A. Barth.

Handschuh, Corinna (2014) *A typology of marked-S languages*. Berlin: Language Science Press.

Heine, Bernd and Christa König (2015) *The !Xun language: A dialect grammar of Northern Khoisan*. Cologne: Rüdiger Köppe.

König, Christa (2006a) Focus in !Xun. *Frankfurter Afrikanistische Blätter* 18: 91–99.

König, Christa (2006b). Marked nominative in Africa. *Studies in Language* 30(4): 655–732.

König, Christa (2008a) Topic in !Xun. In Ermisch, S. (ed.), *Khoisan languages and linguistics: proceedings of the 2nd International Symposium, January 8–12, 2006, Riezlern/Kleinwalsertal*, 249–263. Cologne: Rüdiger Köppe.

König, Christa (2008b) *Case in Africa*. Oxford University Press.

König, Christa and B. Heine (2001) *The !Xun of Ekoka: A Demographic and Linguistic Report*. University of Cologne.

Lambrecht, Knud (1994) *Information structure and sentence form. Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.

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.

McGregor, William B. (2006) Focal and optional ergative marking in Warrwa (Kimberley, Western Australia). *Lingua* 116(4): 393–423.

McGregor, William B. and Jean-Christophe Verstraete (eds.) (2010) Optional ergative marking. *Lingua* 120(7).

Meinhof, Carl (1909) *Lehrbuch der Nama-Sprache*. Berlin: Reimer.

Meinhof, Carl (1930) Der Koranadialekt des Hottentottischen. *Zeitschrift für Eingeborenen-Sprachen*, Beiheft 12.

Nikolaeva, Irina (2001) Secondary topic as a relation in information structure. *Linguistics* 39(1): 1–49.

Olpp, J. [ed. by W. Moritz] (1963) *Einführung in die Namasprache*. Karibib: Rheinische Missionsgesellschaft.

Planert, Wilhelm (1905) *Handbuch der Nama-Sprache in Deutsch-Südwestafrika*. Berlin: Reimer.

Pratchett, Lee J. (2017) Dialectal diversity in Southeastern Ju (Kx'a) and a documentation of Groot Laagte ǂKx'ao‖'ae. Unpublished doctoral dissertation. Humboldt-Universität zu Berlin.

Rapold, Christian J. (2007) From demonstratives to verb agreement in Benchnon: A diachronic perspective. In: Azeb Amha, Maarten Mous and Graziano Savà (eds.) *Omotic and Cushitic studies: papers from the Fourth Cushitic-Omotic Conference, Leiden, 10–12 April 2003*, 69–88. Cologne: Rüdiger Köppe.

Sasse, Hans-Jürgen (1987) The thetic/categorical distinction revisited. *Linguistics* 25(3): 511–580.

Sasse, Hans-Jürgen (2006) Theticity. In: Giuliano Bernini, and Marcia L. Schwartz (eds.) *Pragmatic organization of discourse in the languages of Europe*, 255–308. Berlin: Mouton de Gruyter.

Schachter, Paul (1973) Focus and relativization. *Language* 49(1): 19–46.

Seidel, A. (1891) *Praktische Grammatiken der Hauptsprachen Deutsch-Südwestafrikas*. Wien/Pest/Leipzig: A. Hartleben.

Seržant, Ilja A. and Alena Witzlack-Makarevich (2018) *Diachrony of differential argument marking*. Berlin: Language Science Press.

Tindall, Henry (1857) *A grammar and vocabulary of the Namaqua-Hottentot language*. Cape Town.

Tosco, Mauro (2002) A whole lotta focusin' goin' on: information packaging in Somali texts. *Studies in African Linguistics* 31/1-2: 27–53.

Tosco, Mauro (2012) "Subject focus" and theticity. Paper presented at the Workshop 'Mise en relief et mise en retrait: Le marquage morphologique de la hiérarchie discursive', CNRS Paris, 19–20 October 2012.

Traill, Anthony (1994) *A !Xóõ dictionary*. Cologne: Rüdiger Köppe.

Traill, Anthony (1997) *Extinct: South African Khoisan languages (CD and booklet)*. Johannesburg: University of the Witwatersrand.

Vedder, von H. (1923) *Die Bergdama*, 2 vols. Hamburg: L. Friederichsen.

Vossen, Rainer (ed.) (2013) *The Khoesan languages*. London and New York: Routledge.

Wallmann, Johann C. (1854) *Vokabular der Namaqua-Sprache nebst einem Abrisse der Formenlehre derselben*. Halle: J. Fricke.

Wallmann, Johann C. (1857) *Die Formenlehre der Namaquasprache*. Berlin: W. Hertz.

Washburn, Paul (2001) A minimalist approach to Khoekhoe declaratives. In: Arthur Bell, and Paul Washburn (eds.) *Khoisan: Syntax, phonetics, phonology, and contact*, 28–56. Ithaca: Cornell University.

Westphal, Ernest Oswald Johannes (n.d.) Unpublished field notes on N|huki. Manuscripts and Archives Department, University of Cape Town.

Authors' contact information:
Tom Güldemann
Humboldt-Universität zu Berlin
Kultur-, Sozial- und Bildungswissenschaftliche Fakultät
Institut für Asien und Afrikawissenschaften
Unter den Linden 6
10099 Berlin
Germany

【要　旨】

語用論から文タイプへ
——カラハリ盆地諸言語における非主題S/A項と節2番目の小詞——

Tom Güldemann　　　　　　Lee J. Pratchett　　　　　Alena Witzlack-Makarevich

フンボルト大学（ベルリン）　　　フンボルト大学（ベルリン）　　　エルサレム・ヘブライ大学

マックスプランク研究所

　　カラハリ盆地言語帯には，節内の 2 番目の位置（多くの場合，その節の S/A 要素の後）に，ある特別な小詞類が生起するという特徴を共有する言語が存在する。この小詞類に対するこれまでの説明は，平叙文，直接法，強調された主格，主題といったさまざまなラベルを用いてなされており，この小詞類が，多様ではあるが充分に理解されていない一連の機能群を有していることを示している。本稿は，ディスコースに重点を置き，比較の観点から 3 つの異なる語族の言語，すなわちコエ‐クワディ語族の北コエコエ，トゥー語族のヌン，カー語族のジューに見られる関連する事例を探求する。そして，当該の小詞が，非動詞叙述，焦点，entity-central theticity，平叙文，さらにおそらく示唆的 S/A 標示までにも及びうるほどの多様な領域に広がる構文のネットワーク形成に関わっていると結論づける。この最後の 2 つの機能（それぞれ文タイプと文法関係に関連するものであって，もはや情報構造上有標である配置を示すものではない）は，thetic 小詞構文の過度の使用から出現するものであり，いわゆる「脱語用標識化」の結果である。