

A corpus-based analysis of grammatical relations in Enggano

Charlotte Hemmingsⁱ, Erik Zobel^j & Mary Dalrympleⁱ

University of Oxfordⁱ & Independent Researcher^j

GRELSPOC, INALCO 15th-16th June 2023

Introduction

- This paper presents a **diachronic study** on the usage of different constructions to express grammatical relations in Enggano, an Austronesian language of Sumatra, Indonesia.
- We will compare **verbal constructions** in a corpus collected by Hans Kähler in the 1930s (“Old Enggano”) with materials collected as part of an ongoing documentation project since 2018 (“Contemporary Enggano”).
- Enggano verbs occur in one of three forms: ***bu-***, **bare** and ***ki-***. *Bu-* and bare verbs express grammatical relations via verbal agreement, whilst *ki-* verbs do not take agreement and instead use overt NPs or pronouns for GRs.

Introduction

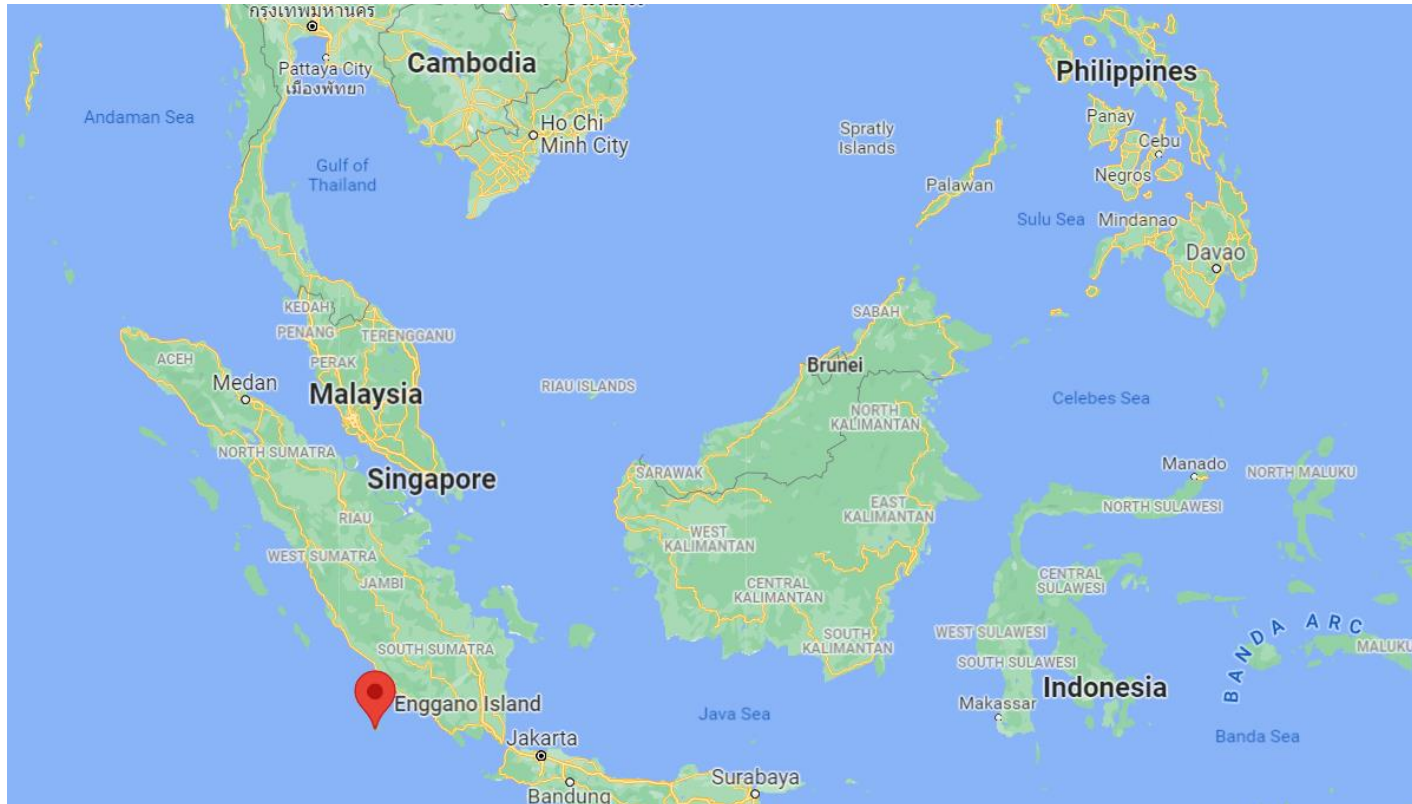
- The morphology of *bu-* and bare verbs is cognate with **symmetrical voice** morphology in other Austronesian languages (see e.g. Edwards 2015). However, they do not reflect a symmetrical voice alternation in Enggano and instead represent a **realis/irrealis split**.
- *Ki-* is an **innovative form**, most likely cognate with the relativiser *si=* in Nias (see Brown 2001). It may have started out as a relative clause marker but has been co-opted for SVO **main clauses**.
- The aim of the paper is to **compare** the three constructions in Old Enggano and Contemporary Enggano, and to see how they are used in discourse

Roadmap

- Background on Enggano
- The verbal constructions & expression of grammatical relations
- Comparing the corpora
- Conclusions

Background on Enggano

Background on Enggano



- Enggano is spoken by approx. 1,500 speakers on Enggano Island, Sumatra, Indonesia
- There is some debate around classification but most people now agree that Enggano is **Austronesian** (Dyen 1965, Nothofer 1986, Edwards 2015, Smith 2017, 2020, Billings & McDonnell 2022)

Background on Enggano



map from ter Keurs (2006: 134)

- Today, Enggano is spoken in six main villages across the island.
- There are also **non-Enggano populations** who migrated to the island. Inter-ethnic communication is in **Indonesian**.
- Enggano is considered **endangered** as speakers increasingly shift to Indonesian (see Arka et al. 2022).
- The language is most vital in the central villages of Meok, Apoho and Malakoni

Background on Enggano

1850-1900	Early Wordlists	von Rosenberg 1855, van der Straaten & Severijn 1855, Walland 1864, Oudemans 1879 Helfrich & Pieters 1891, Helfrich 1893, 1916
1930s	Hans Kähler	Grammar Sketch (Kähler 1940) Text Collection (Kähler 1955, 1957, 1958, 1960, 1961, 1962, 1964, 1975) Dictionary (Kähler 1987, published posthumously)
1980s-2020s	Recent Work	Nothofer (1986, 1992), Nikelas et al (1994), Yoder (2011) Wijaya (2018), Butters (2021) Riswari et al (2021)
2018-present	AHRC-funded documentation project	Corpus of audio and video recordings (Meok) Swadesh lists from across the villages Grammar, FLEX database of glossed texts and lexicon

Verbal Constructions & Expression of grammatical relations

Verbal Constructions in Old Enggano

- Old Enggano verbs occur in one of three forms:

- (1a) **ka-bu-pudu-ha** epaE e'ana bu-form
 3.NOM -**bu**-kill-EMPH child DEM.MED
 'and he killed the child' (Kähler 1955:90)
- (1b) kea-ba'a **i-pudu** e-koyo e'ana bare form
 NEG-INTENSIVE 3.ERG-kill DIR-pig DEM.MED
 'He didn't kill the pig' (Kähler 1940:101)
- (1c) e-kaka e'ana **ki-pudu** e-koyo ki-form
 DIR-person DEM ki-kill DIR-pig
 'That person killed a pig.' (Kähler 1940:108)

bu- form

- *bu-* is a reflex of PAN *-um- that elsewhere in the Austronesian world marks actor voice:

Allomorph	Example with 3sg agreement	Root
bu-	ka bu dodo ‘he grasps sth’	dodo
b-	ka b edo ‘he cries’	edo
mu-	ka mu noo ‘he eats sth’	noo
m-	ka m ãhãhõ ‘he fears sth/is afraid’	ãhãhõ
-ub-	ka u bore ‘he lifts sth’	kore
-um-	ka u mõõ ‘he tries’	kõõ
∅-	ka ∅ padi’o ‘he makes’	padi’o

bu- form

- Main clause *bu-* verbs take agreement markers that agree with S/A

(2a) **ka-bu-pudu** *kia*
 3-BU-kill 3SG
 ‘They fought against him’ (Kähler 1975: 61)

(2b) **ka-mu-na’a** [*kia*] [*e-aka’a:i*]
 3-BU-take 3SG DIR-headband
 ‘He takes his headband’ (Kähler 1975)

(2c) **ka-bu-péa=ha** [*e-ko’E’E*] [*e-hũã u-kanĩxõõ*]
 3-BU-see=EMPH DIR-devil DIR-fruit OBL-tree.species
 ‘(When) the devil saw the fruits of the ekanîxôô-tree’ (Kähler 1955)

1SG	‘u-
2SG	‘o-
3SG	ka-
1DU.INCL	ka-
1PL.INCL	ka- -a’a
1PL.EXCL	‘u- ‘ai
2PL	‘o- -a’a
3PL	da-, di-, ki-

bare form

- Bare forms most likely derive from a **dependent undergoer voice** form (that consisted of just the bare verb root) (see e.g. Zobel forthcoming)
- They also occur with **agreement** for S/A:

(3a) kau=pe **i-paka'aua'a** [ka=po'inamo] [e=puaha u=kaka e'ana]
 NEG=yet 3-know PL-maiden DIR=look OBL=person that
 'The maidens didn't know yet the appearance of the person.' (Kähler 1940:103)

(3b) ka kea=ba'a **y-a'u'ua** [e=kidei-da e'ana]
 and NEG=INTENSIVE 3-good DIR=belly-3s.GEN that
 'They did not agree' (lit. 'there belly was not good') (Kähler 1975:54)

bare form

- The agreement forms are different with bu- and bare verbs – which is again reminiscent of **symmetrical voice** and suggests a different origin (see Zobel forthcoming)

Gloss	Set 1 (bu- verbs)	Set 2 (bare verbs)	free pronouns
1SG	'u-	'u-	'ua
2SG	'o-	u-	'o'o
3SG	ka-	i-	kia
1DU.INCL	ka-	ka-	'ika
1PL.INCL	ka- -a'a	ka- -a'a	'ika'a
1PL.EXCL	'u- 'ai	'u- 'ai	'ai
2PL	'o- -a'a	u- -a'a	adiu/aduu/aruu
3PL	da-, di-, ki-	da-	ki

bu- vs bare form

- Importantly, however, the choice of *bu-* vs bare verb does **NOT** represent a symmetrical voice alternation in Enggano:
 - Agreement is with S/A in both cases
 - Enggano has no subject-only extraction restriction in relative clauses
 - There is evidence for an S/A pivot with both *bu-* verbs and bare verbs (see Zobel et al 2022)

bu- verbs	bare verbs
<ul style="list-style-type: none"> • realis main clauses • embedded clauses (without agreement) • Intransitive subordinate clauses 	<ul style="list-style-type: none"> • irrealis main clauses (negatives, hortatives and imperatives) • clause-chaining/purposive clauses • transitive subordinate clauses

bu- form

- *bu-* verbs are also used in dependent clauses **without agreement** following auxiliaries like *hoo* ‘perfective’ and *hii* ‘repeated action’:

(4a) 'ika hoo **mu-na'a** e-kixo kitera
 1pl.incl pfv bu-take dir-ant all
 ‘we have already taken all the ants’ (Kähler 1955)

(4b) kE'anaha ka-hii **b-a'ioi=xa** e-paE e'ana
 then 3-repeat bu-follow=emph dir-child dem
 ‘Then he chased the child again’ (Kähler 1955)

Bu- vs bare form

- Bare verbs (typically with the derivational marker *aba-*) mark **consecutive action/ purposive clauses** in clause-chaining contexts:

(5a) *ka-b-ai-xa* ama ka-pae e'ana [*y-aba-pea* ki]
 3-BU-come-EMPH father PL-child DEM 3-CNS-see 3PL
 'The father of the children came, in order to see them' (Kähler 1957: 154)

(5b) *ka-b-ahae-ha* [*y-aba-kiu-ha* i-tita]
 3-BU-go-EMPH 3-CNS-hide-EMPH LOC-there
 'It went there and sought shelter there' (Kähler 1955: 90)

Bu- vs bare form

- In subordinate clauses marked with the conjunctions *a*= ‘if/when’ and *be* ‘because’, transitive clauses are marked with bare verbs (plus agreement) whilst intransitive verbs occur with *bu-* (without agreement)

(6a) *a=da-dohoi* *e=di-'ua-dia*
 CONJ=3PL -hear DIR=PASS=say-3sg.GEN
 ‘when they heard what he had said’ (Kähler 1975:80)

(6b) *a=b-ai* *ki* *na'ani*
 CONJ=BU-come 3PL later
 ‘when they will come later’ (Kähler 1975:32)

- We treat this as a remnant of the conservative pattern of **ergative alignment** found in other languages of Sumatra & Sulawesi, e.g. Northern Nias (see Sundermann 1913)

Ki- form

- Finally, ki- is likely cognate with Nias si= which functions as a **relative clause marker** for relativisation on S/A (see Brown 2001)

Allomorph	Example	Root
ki-	kia ki dodo 'he grasps'	dodo
k-	kia k edo 'he cries'	edo
kV-	ki ku 'uoho 'they sleep'	'uoho

Ki- form

- It is the only verbal form used in **relative clauses**:

(7a) ka-'édèha=ha e-paE [hemo'ò **ku-'uoho**]
 3-startle=EMPH DIR-child REL.SG KI-sleep
 'The child, who was sleeping, was startled' (Kahler 1955)



(7b) e=apama u=kaka [mo'ò **ki-'ope** kia] e'ana
 DIR=number OBL=person REL FOC-ambush 3SG that
 'the number of the people who lay in ambush for him' (Kähler 1975:61)



(7c) i'iaha e-kude-a u-měhě-nu [mo'ò aruu **ki-no-noo**]?
 where DIR-originate-LOC.NOM OBL-food-2PL.POSS REL 2PL KI-REDUP-eat
 'Where does the food that you eat come from?' (Kähler 1957: 153)



Ki- form

- However, it is also used in main clauses. These have SVO order in contrast to the **verb-initial** (VSO) order seen with *bu-* and bare verbs:

(8) ki **k-aha:e** i-pia=da
 3PL KI-go LOC-garden=3PL.POSS
 ‘They set off for their plantation’ (Kahler 1955)

- Given the usage in relative clauses (and the existence of Nias *si=*) we analyse main clause *ki-* as the result of a process of **insubordination**
- A similar story is posited in the history of Austronesian voice morphology (see e.g. Kaufman 2018, Ross 2006)



Arts and
Humanities
Research Council



Summary

	Agreement	verb-initial word order
<i>bu-</i> verbs	Yes	Yes
bare verbs	Yes	Yes
<i>ki-</i> verbs	No	No

Verbal Constructions in Contemporary Enggano

- The morphosyntax of contemporary Kelabit has not changed that much in comparison with major **phonological changes**:

Change	Old Enggano form	Contemporary Enggano form
Loss of final vowels	noo 'eat' pia 'garden' hëkũ 'sit'	no pi hëk
/o/ → /ə/ /õ/ → /ã/	dodo 'grasp' kõõ 'try'	dèr kě
vowel elision	do'orao 'sand' kabudodo kamunoo	do'ra kabdèr kamno

Verbal Constructions in Contemporary Enggano

- It is still the case the verbs occur in *bu-*, bare and *ki-* forms:

(9a) adühür ean **ka-b-dèhè**
 CONJ=finish DEM 3-B U -hear
 ‘afterwards he heard’

(9b) ki ke’ **i-dèhè**
 3SG NEG 3-hear
 ‘he didn’t hear’

(9c) ki **ki-dèhè**
 3SG KI-hear
 ‘he heard’

Verbal Constructions in Contemporary Enggano

- And that bu- and bare verbs take agreement marking with S/A, whilst ki- verbs use free pronouns to express grammatical relations:

Gloss	Set 1 (bu- verbs)	Set 2 (bare verbs)	free pronouns
1SG	u-	u-	u
2SG	è-	u-	è'
3SG	ka-	i-	ki
1PL.INCL	ka-	ka-	ik
1PL.EXCL	u- a	u- a	a
2PL	è- a	u- a	ari
3PL	da-	da-	ki/hamè'

Verbal Constructions in Contemporary Enggano

- The major change is word order, specifically *bu-* and bare verbs can also occur in subject-initial orders:

(10a) [e-am-de] ka-puak
NM-father-3SG.POSS 3-go
'Its father started on his way' (Kähler 1955 retelling)

(10b) ki ke' i-dèhè
3SG NEG 3-hear
'he didn't hear' (elicitation)



Arts and
Humanities
Research Council



Verbal Constructions in Contemporary Enggano

- This may be the result of **contact** with predominantly SVO Indonesian and/or it may result from the **reanalysis of a topicalization construction** that is common in Old Enggano:

(11) ka-moho=maha dadu=dia ka-b-ahaꞤE
 3-different=also husband=3SG.POSS 3-BU-go/walk
 ‘Her husband also started on his way’ (Kahler 1955)

Summary

	Agreement	Verb-initial Word order
<i>bu-</i> verbs	Yes	?
bare verbs	Yes	?
<i>ki-</i> verbs	No	No

Comparing the corpora

Comparing Corpora

- We have seen that both Old Enggano and Contemporary Enggano allow for different **verbal constructions** that correspond to grammatical relations being expressed in different ways.
- To get an initial sense of how the constructions are used in discourse, we compared the three verbal constructions in an Old Enggano folk story (Kähler 1955) and a contemporary retelling of the story.

All finite clauses	bu-	bare	ki-	Total
Old Enggano	118 (c. 50%)	70 (c. 30%)	48 (c. 20%)	240
Contemporary Enggano	125 (c. 50%)	59 (c. 24 %)	63 (c. 26%)	247

Comparing Corpora

- In certain constructions (e.g. relative clauses, negation) the **choice of verbal construction is fixed**. If we exclude these cases then the results are as follows:

Main clauses	bu-	bare	ki-	Total
Old Enggano	106 (c. 58%)	55 (c. 30%)	21 (c. 12%)	182
Contemporary Enggano	118 (c. 62%)	42 (c. 22%)	31 (c. 16%)	191

- The results are very similar in terms of overall frequency – though in contemporary Enggano the percentage of *ki-* forms is slightly higher, and bare forms is lower.

Comparing Corpora

- It seems as though *bu-* verbs occur at **episode boundaries** as the first clause in a **new discourse unit** and then followed by bare clauses:

(12)	ka-moho	e-paE	e'ana	ka-bu-pua,	
	3.NOM-different	DIR-child	DEM.MED	3.NOM-BU-run	
	y-aha:E	y-aba-kiu	i-hoo	u-kãĩ	u-kepũũ
	3.ERG-go	3.ERG-MOTION-hide	LOC-inside	OBL-base	OBL-fern

'And the child fled, it went away and sought shelter on the inside of the foot of the fern.' (Kähler 1955).

- Thus, the choice of *bu-* vs bare verbs may be triggered by **similar factors to symmetrical voice** languages (see Riesberg et al 2022).

Comparing Corpora

- Ki- verbs in main clauses are typically used to **establish new topics** in the discourse, whilst *bu-* verbs reflect **continuing topics**:

(13a) ki **k-aha:e** i-pia-da
3PL KI-go loc-garden-3PL.POSS
'They started towards their plantation'

(13b) kE e-arara e'ana **ku-'uoho**
but dir-child dem ki-sleep
'but the child was sleeping' (Kahler 1955)

Comparing Corpora

- Similar patterns also occur in contemporary Enggano:

(14a) **ka-puak** **y-ah** b-ah-uar yãh be-ké'ém i heo
 3-leave 3-go cns-antip-find place cns-hide loc inside

yea ku pakis
 foot tree fern

'it went away and sought shelter on the inside of the foot of the fern.'

(14b) e-nai=de **k-a'-karié** i pi [...],
 nm-mother=3sg.poss ki-vblz-work loc garden

e-dar=de **ki-puak** **y-ah**
 nm-husband=3sg.poss ki-leave 3-go

'The mother worked on the plantation [...] The father started on his way'
 (Kahler 1955 retelling)

Comparing Corpora

- We also compared Old/Contemporary Enggano with the texts in Helfrich (1916) and in these texts **bare forms** seem to be more frequent:

Main clauses	bu-	bare	ki-	Total
Helfrich text	27 (c. 39%)	29 (c. 42%)	13 (c. 19%)	69
Kahler text	106 (c. 58%)	55 (c. 30%)	21 (c. 12%)	182
Contemporary text	118 (c. 62%)	42 (c. 22%)	31 (c. 16%)	191

- (15) moko hinua kia **i-kua** i-uba ka-poinamo eana,
 many times 3SG 3-enter LOC-house PL-girl DEM
 'The young man visited the girls multiple times' (rat story)

Conclusions

Conclusion

- Like other Austronesian languages, Enggano verbs can occur in **different morphological forms** that in turn have different forms for marking grammatical functions.
- However, unlike many Western Austronesian languages, these alternations do not equate to **symmetrical voice**.
- Instead Enggano appears to have straightforward **accusative alignment** in agreement and cross-clausal pivots, and only maintains something approaching **ergative alignment** in the context of *a*= subordinate clauses

Conclusion

- From symmetrical voice morphology, the language seems to have developed three separate **verbal constructions** that are used in different contexts and result from separate changes.
- In both Old Enggano and Contemporary Enggano, *bu-* verbs are the most **frequent in discourse**. These are often used at episode boundaries as the first clause in a new discourse unit and then followed by bare clauses.
- This is interesting as it suggests that the choice of *bu-* vs bare constructions in Enggano may be determined by **similar factors to symmetrical voice languages** (cf. Riesberg et al. 2022), even though the mapping to subject does not change.

Conclusion

- From looking at even older texts, it appears that the rise of *bu-* as the basic (most frequent) realis clause type may be a **new development** and that bare verbs may have been more frequent in the past (as is common in Sumatra and Sulawesi)
- In any case, the study reinforces the idea that **diachronic development** may provide explanations for **synchronic patterns** and we are fortunate enough to be able to explore this in Enggano given the rich documentation and corpora available.



Arts and
Humanities
Research Council



With thanks to...

The Enggano Community



With thanks to...

Dendi Wijaya, BRIN

Engga Zakaria Sangian, Universitas Dehasen Bengkulu

I Wayan Arka, Australian National University



With thanks to...

Milson Kaitora, Daniel Krausse, Sarah Ogilvie, Gede Primahadi Wijaya Rajeg, Jonas Tan and everyone involved in our recent project meeting in Oxford



With thanks to...



Bernd Nothofer, Goethe-Universität Frankfurt am Main

Colleagues at Udayana University, Bali

Arts and Humanities Research Council UK

The John Fell Fund, University of Oxford

The Endangered Language Fund

Workshop participants for their kind attention!