Investigating Enggano morphosyntax using data from historical corpora and language documentation

Charlotte Hemmings, University of Oxford

Research Seminar on the documentation and lexicography of the Enggano language, Udayana University, Monday 18th March 2024

Introduction

- In this seminar, I will present some findings from our research into the **morphosyntax of Enggano**, an Austronesian language spoken on Enggano Island, Sumatra, Indonesia.

- I will outline how we collected the data used in our description of Enggano morphosyntax by analysing published **historical materials** and comparing with contemporary materials collected via **language documentation**.

- In doing so, I will comment on the **challenges** faced when using such corpora for language description, as well as the **opportunities** that were presented to us by including both historical and contemporary perspectives in our work.

Roadmap

- Background on Enggano

- Contemporary Enggano Morphosyntax

- Historical Change

- Conclusions

Background on Enggano
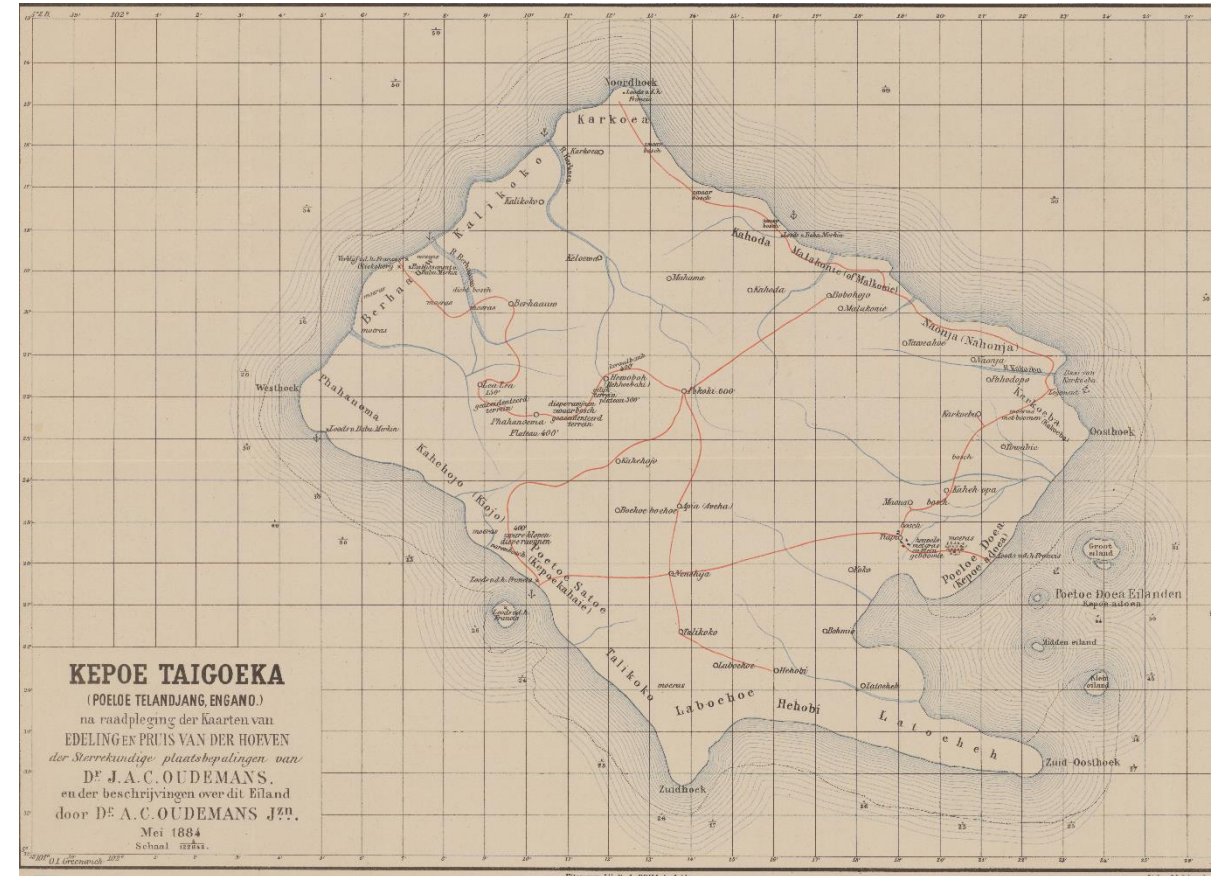
Background on Enggano
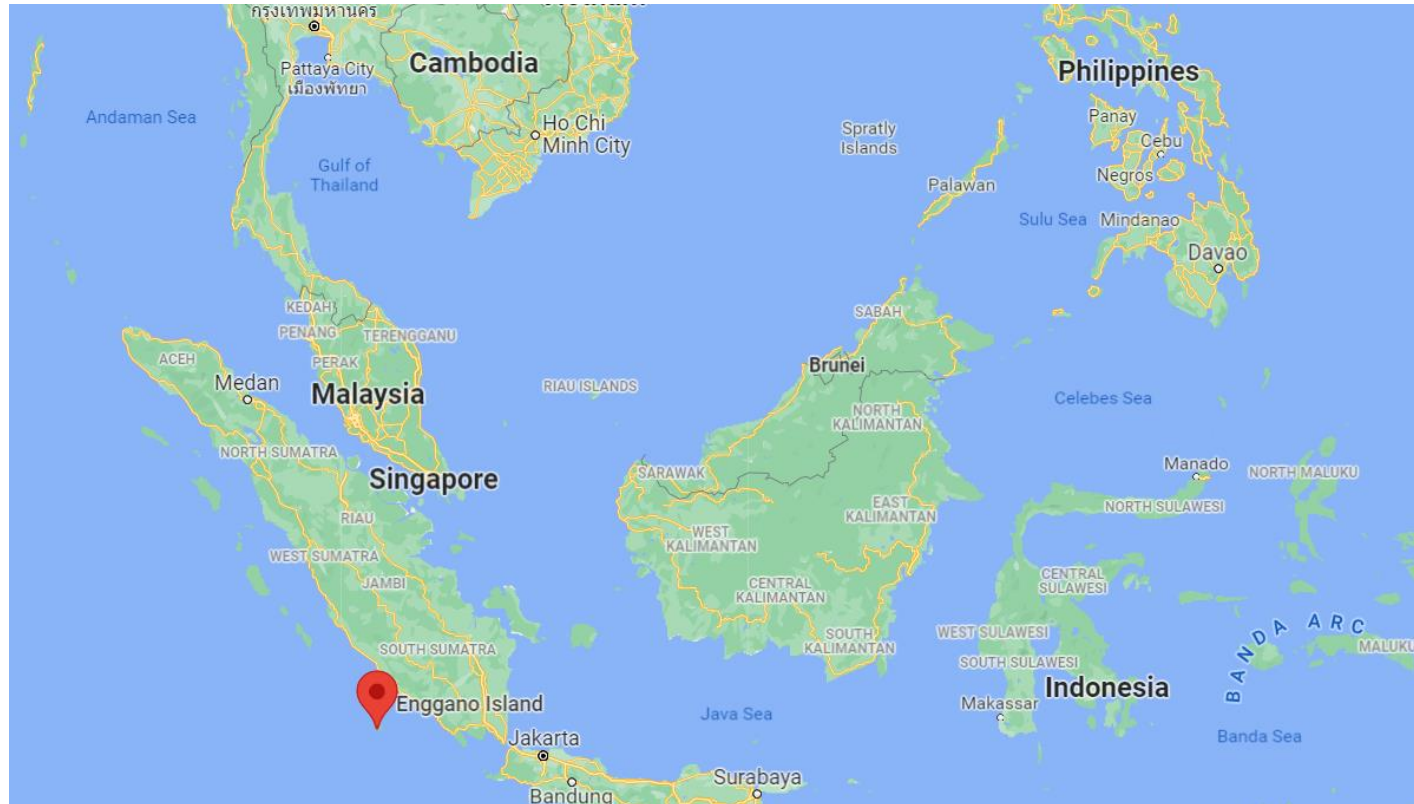
- Enggano is spoken by approx 1,500 speakers on Enggano Island, Sumatra, Indonesia

- Historically debated, but most people now argue that Enggano is Austronesian (Dyen 1965, Nothofer 1986, Edwards 2015, Smith 2017, 2020, McDonnell & Billings 2022)

Background on Enggano

map from ter Keurs (2006: 134)

map from Oudemans (1884)

Previous Documentation

- Enggano has a relatively long history of documentation, starting with wordlists from the 19th Century

Author	Year of publication	Stay in Enggano	Number of words
Djoeragan Boewang	1854	one year	21
H. von Rosenberg	1855	10-24 September, 1852	154
J. van der Straaten and P. Severijn	1855	10 June-2 July 1854	201
J. Walland	1864	10 June-2 July 1854	250
Oscar Louis Helfrich and J.A.J.C. Pieters	1891	10 Nov-3 Dec 1885, 13 December 1891, inter alia	1012
Oscar Louis Helfrich	1916		1100

Previous Documentation

Walland 1864

Hooge taal.		Lage taal.
De zon.	bakka–kaha.	bakka–moenai.
" maan.	kanoeah.	moena.
" sterren.	peloa.	kenoeaij–noekie (kleine maan).
" zee.	oeweh.	praboh.
" aarde.	loppoh.	
Water.	é–boh.	lebo–leboh.
Steenen.	ééah.	pakoe–pakoe.
Vuur.	apoeah.	obie.
Rivier.	baka–oeweh.	beloea.
Berg.	koh.	pawah.
Bosch.	howdah.	koë–koë.
Wit.	kha–oedaija.	
Zwart.	khah–khoh.	kahèlo–èloh.
Rood.	kanoeaij.	kai–koh–koh.
Groen.	ka–akiema.	

Helfrich 1916

A. WOORDENLIJST.

a

a ä ä o e; oudere broeder of zuster.

a ä p o e (w) a (q) [a ä f o e (w) a (q)]; neersmijten, neersmakken.

a b a; in, te, op.

a b a i e k a h a i e q [a b a i e k a h a è q]; negen.

a b a k o e (w) a (q) [a b a k o (w) a; a b a k o (w) ě]; geleiden, begeleiden.

a b a o e (w) a i e (j) o n i e (j) a; last van iemand die heengaat.

a b a p o e (w) a p a o e (w) a; onderzoek, navraag.

a b è h a [w a b è h a]; bamboeriet.

a b è o; braken.

a b o e a q ă; aandoen, aanhebben, gebruiken, zooals kleederen.

a b ö h a; afgedaan, klaar, reeds, voltooid, gereed, ten einde, in orde gekomen uit een verwarden toestand, gedaan hebben met iets, tot rust gekomen.

a d è q [a r è q]; herwaarts, hierheen.

a d i e [a r i e]; antwoord, antwoorden, antwoord geven.

a d i e b a [a r i e b a; a l i e b ě]; vijf; a d i e b a h i e a d o e (w) a [a r i e b a h i e a d o e (w) a; a l i e b ě h i e a d o e (w) ě], zeven.

a d o; kwaad doen, kwaad stichten.

a d o e (w) a [a r o e (w) a; a d o e (w) ě]; twee.

a h a d i e; roepen, bij den naam roepen, noemen.

a h a d o; zich oprichten, opstaan, overeind zetten.

a h a p i e o r i e; met de hand drukkend over een lichaamsdeel

Previous Documentation

D. VERHALEN (Ekoedaäjo).

I. Ekoedaäjo ekoanoe kipahóna hi ekaka.
Tjeritera tikoes kawin sama orang.

| E. | Kikia | hale | ilopo | eana | aroea | epoinamo | moö, |
| M. | Adalah | doeloe | dinegrie | itoe | doea | gadis | jang |

| E. | káoea | | epoeahhadia, | kamoephi | ipahona | hi | | epo- |
| M. | amat elok | roepanja | | tiada maoe | kawin | | dengan | boe- |

| E. | nanami | kidara | ilopo | | eana, ke | | kikia | pe ilopo |
| M. | djang | semoea | di negrie | itoe | tetapi | adalah | di negrie |

| E. | eana | kahai | ekaka, | moö | kioha | | ikarara | epoeahhadia, |
| M. | itoe | satoe | orang | jang | boeroek | sekali | roepanja |

| E. | kia | kamoephi | | boekokoi | dakioho, | iarakoea |
| M. | dia | tiada maoe | | berdjalan | siang | | soepaja djangan |

| E. | kiepöa | kia | ekaka, | be | | mania | kia, ke | | kaoea |
| M. | di lihat | dia | orang | djangan | maloe | dia | tetapi | bagoes |

| E. | ikarara | eohoradia, | kia | kaha | | dakóahia | ioeba |
| M. | sekali | njanjiannja | dia | berdjalan | malam | | diroemah |

| E. | kapoinamo | aroea | nèèni, ke | | keaba | iahapi | kia | boekoi |
| M. | gadis gadis | doea | tahadi | tetapi | tiada | maoe | dia | boeka |

| E. | eoeba | jonia, | | kaeanaha | diboehohora | ekaka | kioha |
| M. | roema | sama | dia | djadi | menjanjilah | orang | boeroek |

I. Verhaal van eene rat, die met twee meisjes geslachtelijke gemeenschap had.

In eene zekere landstreek waren twee beeldschoone meisjes, die niet wilden huwen met de jongelingen uit het land. Aldaar woonde ook een jongeling, die, van wege zijn leelijk uiterlijk, zich schaamde om zich over dag op straat te vertoonen; hij zong echter prachtig. Op een avond begaf hij zich naar de woning van die meisjes, om hun een bezoek te brengen; zij wilden hem echter niet in huis laten. De jongeling hief een gezang aan, waarop de meisjes genegenheid voor hem opvatten en hem lieten binnenkomen. In de woning brandde geen licht.

Previous Documentation

Challenges

❖ **Accessibility:** the wordlists and texts are mainly translated into Dutch, generally no morphological gloss is given

❖ **Comparability:** Transcription systems differ and there is often no information given on what dialect is represented or what sound is represented

Previous Documentation

- The most important materials collected to date were collected by the German Scholar Hans Kähler who spent 7 months on Enggano Island between July 1937 and January 1938:

Sketch Grammar	Kähler 1940
Text Collection	Kähler 1955, 1957, 1958, 1960a, 1960b, 1961, 1962, 1964, 1975
Dictionary	Kähler 1987

Previous Documentation

Kähler 1940

§ 2. Numerusbildung beim Substantivum.

a) 1. *eke'epa e'ana ka'au* jener Vogel stirbt = jene Vögel sterben.
   2. *e'ana euba'au* jenes ist mein Haus = jene sind meine Häuser.
   3a. *ehuda kia* sie ist eine Frau.
    b. *kahauda ki* sie sind Frauen.
   4a. *kia kipudu epunānāmī e'ana* er tötet jenen Jüngling.
    b. *kia kipudu kapunānāmī e'ana* er tötet jene Jünglinge.

Wie bereits in § 1 erwähnt ist, macht das Enggano bei Substantiven einen scharfen Unterschied zwischen solchen, die eine «Gattung» und solchen, die ein «Einzelwesen» bezeichnen. Unter den Begriff

Previous Documentation

Kähler 1955

Kähler 1975

EIN TEXT VON DER INSEL ENGGANO (WESTKÜSTE VON SUMATRA).

Aufgenommen und bearbeitet von Hans Kähler.

I. ekēda'ayɔ ukaka[1]) halɛɛ[2]) 'akɔrupada[3]).

1. ki kahɔɛ ipiada[4]), kahaɛa ba'apia[5]). kabia e'ana ea'ahāɛāmōhā[6]) ukɔ[7]) hɛ̄mɔ'ɔ kaniā[8]) epākā[9]) hɛ̄mɔ'ɔ ediɔɔ ba'ubuɔda'a[10]) ukaka bakaɔhēa[11]). — kɛ'anaha dibahaɛha ipia e'ana ki 'akɔrupara. yaba'apia nāɛnā[12]) upiada hɔɔ 'apūkā[13]). kamɔhɔmāhā[14]) dadudia[15]) kabahɔɛ iyapahalipa[16]) 'ɔupākā[17]) e'ana. kabuda'aɔha[18]) kahɔi'i e'ūmāɔ[19]), kamūnā'ā yai ipia hɛ̄mɔ'ɔ eiñāhā[20]) honānīā[21]) hii earara[22]). kɛ earara e'ana ku'uɔhɔ[23]), bɛ epanūkɔ[24]) i'iɔɔnīā, kia kipiaka[25]) ekarahadia[26]).

Dorfleben.

1. (Heno') kikia eeheku uɔya halee ua'aka'udara ukaka e'ana kino'oi'ie euku'ania e'ana : kikia kahai'i ekaka kaniā Ka'arubi ika'udara kahabauba e'ana. kabia kahai'i emani 'oukaka Ka'aropo e'ana, kamahapi kia kahēadia kia ea'ahaona. kabahaɛha ika'udara ukaka kaniā Ka'ahēao yamahaona ebuda it'ita Ka'ahēao e'ana, ina'aɛna itopodia umahau aika'udaradia.

Afrika und Übersee, Band XXXIX   91

legen. Zu der Zeit trug der epăkă genannte Baum Früchte, den die Leute zu benutzen gewohnt waren. — Deshalb gingen sie zu dritt mit ihrem Kinde nach der Pflanzung. Die Mutter arbeitete in der Pflanzung, die bereits überwuchert war. Auch ihr Mann machte sich auf den Weg, er fing mit der Schlinge (Vögel) bei jenem epăkă-Baum. Er fing eine Wildtaube, die er mit nach der Pflanzung nahm, die der Aufenthaltsort seiner Frau mit ihrem Kinde war. Aber ihr Kind schlief, denn Müdigkeit hatte es überfallen, und sein Körper war matt.

2. Sein Vater machte sich auf den Weg und kam dort an. (Da) sprach er zu seiner Frau: "Brate eine Wildtaube als Nahrung für uns hier!" "Ja!", erwiderte seine Frau als Antwort auf die Frage ihres Mannes. Alsdann briet sie die Wildtaube. "Nun, sie ist gar!" Darauf sprach die Frau folgendermassen zu ihrem Manne: "Lass uns nur essen! Los, wecke unser Kind, lass uns essen!"

Dorfleben.

1. Folgendermassen verhielt es sich mit der früheren Regelung des Dorf-Gründens der Leute : In jenem Dorfverband gab es einen Menschen (d.h. eine Frau)(namens, d.h. hier:) von der Sippe Ka'arubi[1]). Und es gab einen Mann von den Ka'aropo-Leuten (d.h. von der Sippe Ka'aropo), der eine Heirat plante. Er geht ins Dorf der Ka'ahēao genannten Leute, damit er dort eine Frau (aus) der Ka'ahēao (-Sippe) heiratet; er bringt sie in sein Land oder in sein Dorf.

Previous Documentation

- Kähler (1987) was published posthumously by former students. It includes root words, inflected forms and information on cognates/sources of borrowings

```
eānŭmãẽ          = eʔanĭmãẽ     Pandanusart

°añãmŭ           (ꞌ ML ayam) eañãmŭ ehuda    Huhn

                 eañãmŭ emãnĭ   Hahn

ʔao             kaʔao    sterben, tot/verstorben sein, (Pflanzen usw)

                 absterben, eingehen, (Feuer) erlöschen

                 hō baʔao    bereits gestorben

                 kabaʔao    starb

                 kaʔaoa    wird sterben

                 kaʔaəa (lo8z34)    am Sterben sein

                 hō baʔaoa    im Begriff sein zu sterben

                 eaʔao = e[ʔ]abaʔáo    Tod, Todesfall, das Sterben,

                 Erlöschen (von einem Feuer)
```

Previous Documentation

Challenges

❖ **Accessibility**: translation into German, no glossing, no original fieldnotes

❖ **Comparability**: Each of the resources uses a different transcription system

Meeting the Challenges

- To make the materials accessible and comparable we developed a **unified transcription system** and had the materials translated from Dutch/German into English and Indonesian

- Kähler, Hans. Texte von der Insel Enggano [Texts from the Island of Enggano/*Teks dari Pulau Enggano*].

  Enggano and German transcriptions by Barnaby Burleigh (University of Oxford) and Michele Bianconi (University of Oxford), checked by Sam Corner (University of Oxford). Alignment of Enggano and German transcriptions by Barnaby Burleigh. English translations by Barnaby Burleigh. Indonesian translations by I Komang Sumaryana Putra (Udayana University, Bali, Indonesia). FLEx analyses by Mary Dalrymple and Charlotte Hemmings.
  Transkripsi Enggano dan Jerman oleh Barnaby Burleigh (University of Oxford) dan Michele Bianconi (University of Oxford), dan diperiksa oleh Sam Corner (University of Oxford). Penylarasan transkripsi Enggano dan Jerman oleh Barnaby Burleigh. Terjemahan bahasa Inggris oleh Barnaby Burleigh. Terjemahan bahasa Indonesia oleh I Komang Sumaryana Putra (Universitas Udayana, Bali, Indonesia).

  - *Afrika und Übersee* 39 (1955): 89--94. FLEx analysis.
  - *Afrika und Übersee* 41 (1957): 153--156. FLEx analysis.
  - *Afrika und Übersee* 42 (1958): 179--189. FLEx analysis.
  - *Afrika und Übersee* 43 (1960): 299--312. FLEx analysis.
  - *Afrika und Übersee* 44 (1960): 41--71. FLEx analysis.
  - *Afrika und Übersee* 45 (1961): 205--225. FLEx analysis.
  - *Afrika und Übersee* 46 (1962): 127--145
  - *Afrika und Übersee* 48 (1964): 46--58

Meeting the Challenges

- We subsequently used SIL's Fieldworks Language Technology (FLEx) to add interlinear gloss

1.2	Word	ekėda'ayo			ukaka		halEE	'akorupada	
	Morphemes	e-	kėda'a	-yo	u-	kaka	halEE	'akodu	-pada
	Lex. Entries	e-₁	kėda'a	-o	u-₂	kaka	halEE	'akodu	-pada
	Lex. Gloss	DIR	tell	NOM.patient	OBL	person	in.earlier.times	three	together
	Lex. Gram. Info.	v:VNomArt	v1 (Class1)	v:(NOMsuff)	n:Art/OblArt	subs	adv	num	num>adv

Free Eng I. Tale of People in the old Days who were together (three) with their Child

Ind I. Kisah Orang pada Zaman dahulu (berjumlah tiga) bersama dengan Anak mereka

Ger I. Erzählung von Leuten in alter Zeit, die zu dritt mit ihrem Kinde waren

Mid-Summary

- By transcribing materials like the Helfrich (1916) texts and the Kähler (1955-1975) texts in a common transcription system and glossing in FLEx, we now have searchable corpora that can be used to look at structures in discourse.

Helfrich Corpus [late 19th century]	Approx 2,000 words	Approx 450 clauses
Kähler Corpus [1930s]	Approx. 40,000 words	Approx 5,000 clauses

The Enggano Project

- From data collected after Kähler (e.g. Nothofer 1986, 1992, Yoder 2011), it is clear that contemporary Enggano has undergone quite dramatic changes from the variety documented by Helfrich and Kähler:

Helfrich Form	Kähler Form	Contemporary Form	Meaning
lopo	dopo	dop	land
moö	mõ'õ	mė'	relativiser
kie-pöa	ki-pəa	ki-pú	see
e-poea-hha=dia	e-pəa-ha=dia	pủa-h=de	his face

- These changes, together with the sociolinguistic context, provided the motivation for our ongoing AHRC-funded documentation project ("The Enggano Project").

The Enggano Project

- The documentation project began in 2018 with two **pilot visits** to Enggano Island

- During the early stages of the project, Mary and Wayan established links with the Enggano community and began to **train local consultants**.

- Subsequently an AHRC grant was awarded to Mary Dalrymple, I Wayan Arka and Bernd Nothofer

The Enggano Project

- The aims of the Enggano project:

  ❖ Collect an archivable **documentary corpus** of contemporary Enggano

  ❖ Produce a **grammar** based on a **FLEX database** of glossed texts and lexicon

  ❖ Produce **teaching materials** for use in the community

  ❖ Assess the position of Enggano in the Austronesian family – both in terms of **subgrouping** and **typology**.

The Enggano Project

- The aim of language documentation today is to compile "**a representative and lasting multipurpose record of a natural language or one of its varieties**" (Himmelmann 2006)

- This involves collecting a corpus of **audio and video recordings** and "adding-value" with time-aligned **transcription**, **annotation** and **translation** into a language of wider communication (see Austin 2010, 2016).

- This is accompanied by systematic recording of **metadata** and **archiving in language archives** to make the materials understandable, findable, preservable and useable (Austin 2016).

The Enggano Documentation Project

- So far, we have collected a corpus of over 200 **audio and video recordings** covering a range of genres.

- These include wordlists, structural elicitation, stories, songs, descriptions of cultural practices & texts for the teaching materials
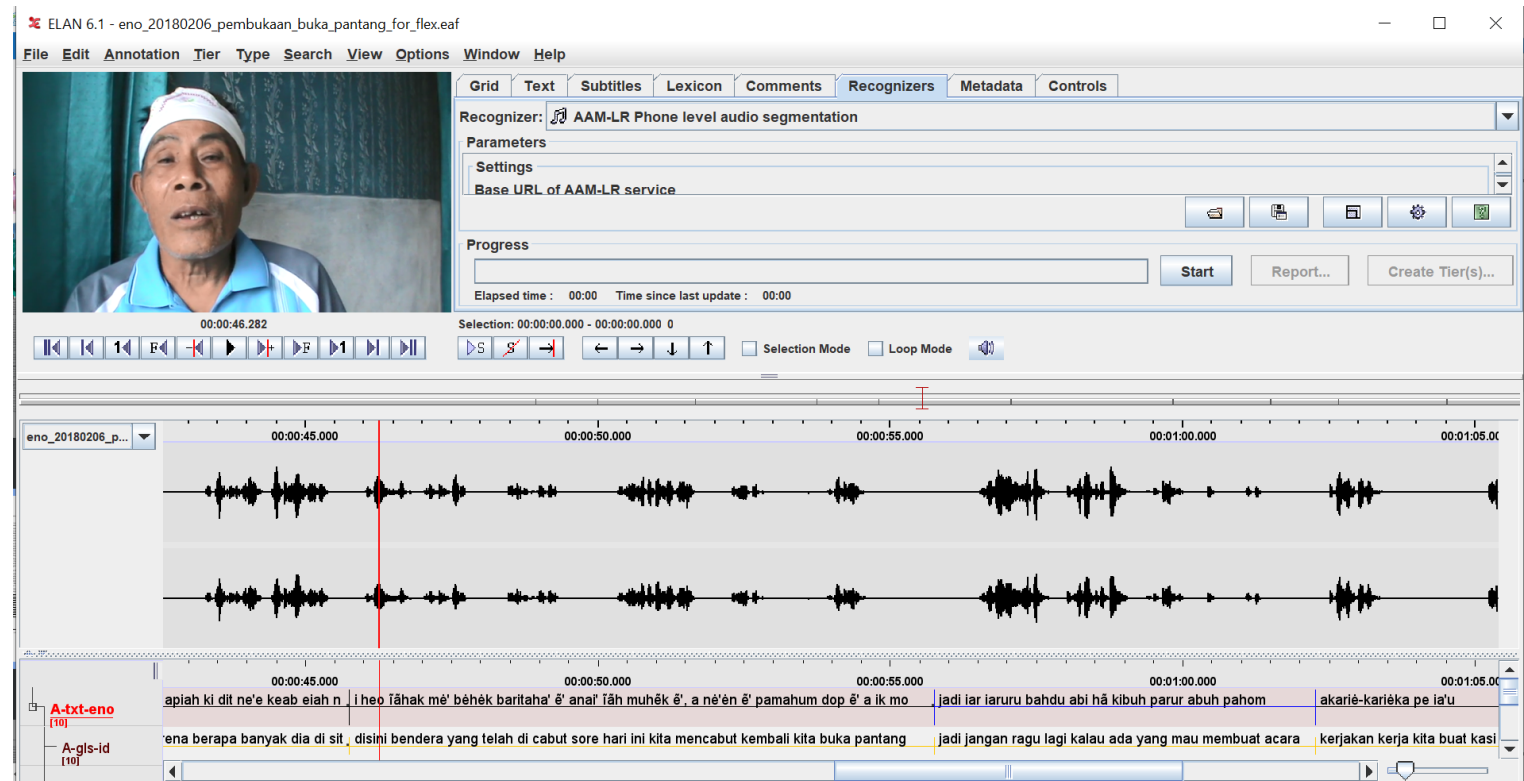
- We store **metadata** using the SayMore software.

The Enggano Project

Challenges

❖ **Accessibility**: audio and video recordings need to be transcribed, translated and glossed

❖ **Comparability**: we need to develop a common transcription system and gloss consistently!

The Enggano Project

- The recordings were initially **transcribed** and **translated into Indonesian** by Engga Zakaria Sanggian using ELAN software.

- The Indonesian was further **translated into English** by I Komang Sumaryana Putra and I Kadek Dony Suantika, Udayana University.
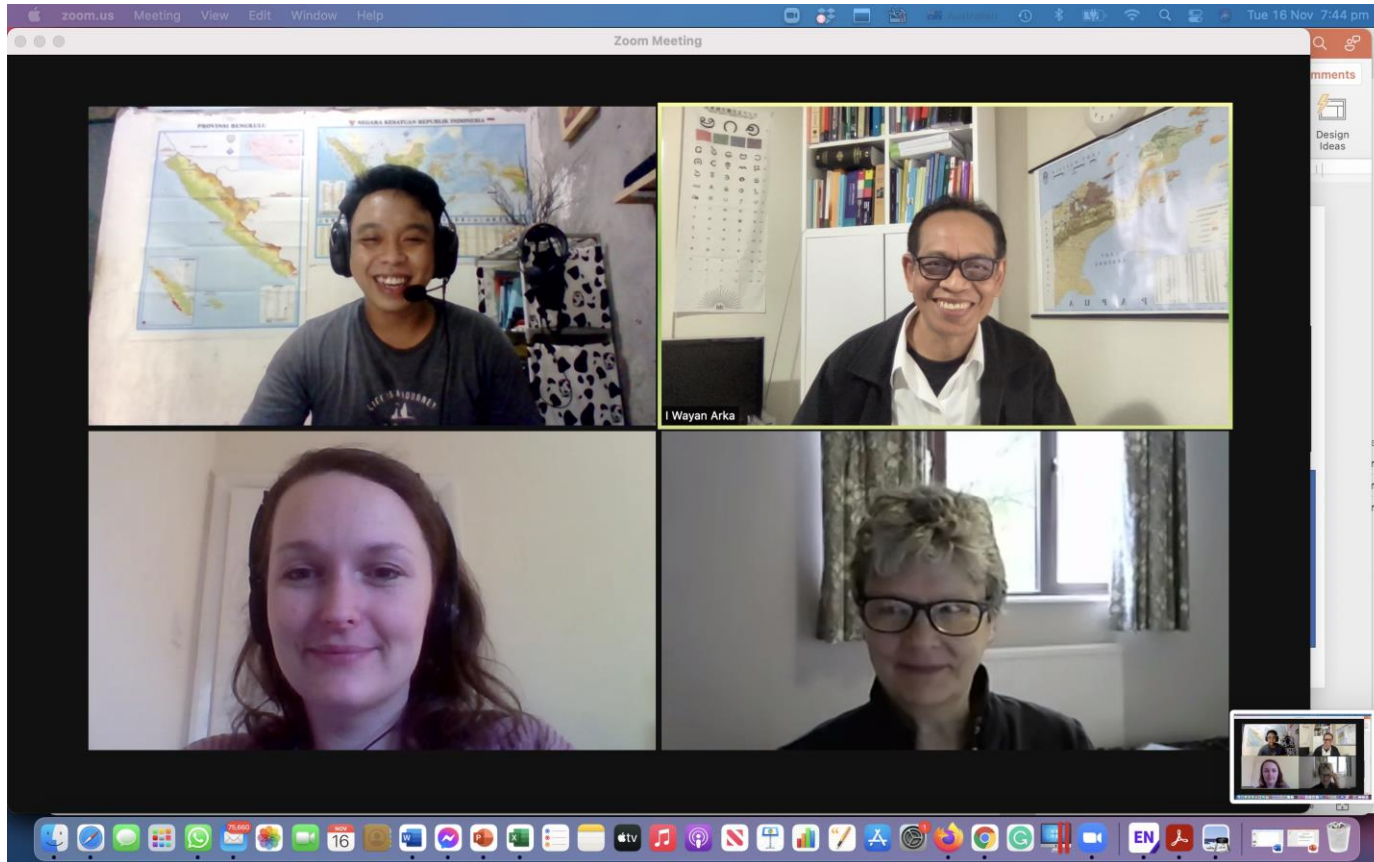
The Enggano Project

- Since the COVID pandemic, we have been having **regular meetings on Zoom**.

- We use these sessions to go through the ELAN files and discuss **issues in transcription** and **word glossing**

- In consultation with the Enggano community, we developed a **standard orthography** (Hemmings et al 2023)

The Enggano Project

- Once we had checked the segmentation, transcription and translation in ELAN, we imported the transcripts into FLEX for **interlinear glossing**

The Enggano Project

- This allows us to build a corpus of glossed texts in FLEX that is both **searchable** and **broadly comparable** to the historical materials

1 Word	né'én	u	kibúh	búkúda'ai	ekak	mé'	ka'pan	tikus
Morphemes	né'én	u	ki- búh	bú- kúda'a -i	e- kak	mé'	k- a'- pan	tikus
Lex. Gloss	now	1SG	KI want	BU tell FUT	NM person	REL	KI VBLZ become	mouse
Word Gloss Eng	now	I	want	tell	person	which	become	mouse
Word Gloss Ind	sekarang	saya	mau	beritahu	orang	yang	menjadi	tikus
Word Cat.	adv	pro	aux	v	n	rel	v	n

Free Eng Now I will tell you a story about man who turned into a mouse
Ind Sekarang aku mau kasih tau orang yang menjadi tikus

Contemporary Corpus [2018-2024]	Approx. 30,000 words	Approx 6,500 clauses (of which 2,000 from naturalistic texts)

Summary

- The Enggano language has a **long history of documentation** from early wordlists, to the materials collected by Kähler and our ongoing documentation project.

- However, both the published historical materials and the audio and video recordings collected as part of language documentation need to be further processed (via transcription, translation, glossing) in order to be **accessible** and **comparable.**

- This is very **time-consuming** and requires experts in many different languages (and linguistics). However, it results in **parallel corpora** that can be used for comprehensive description, based on naturalistic language use.

Contemporary Enggano Morphosyntax

Enggano Morphosyntax

- Using the contemporary Enggano **FLEx database** (which includes naturalistic texts and elicited sentences) we are working towards a description of **Enggano morphosyntax**

- There is a major distinction between **nouns** and **verbs** in terms of the morphosyntax

Nominal Morphosyntax

- Nouns in Enggano can be divided into two classes on the basis of how number is marked:

<table>
<tr><td>

<u>Non-human Nouns</u>
not distinguished for
singular and plural

</td><td>

<u>Human Nouns</u>
take a *ka-* prefix in the
plural

</td></tr>
</table>

e.g.   *koi* 'pig/pigs'   e.g.   *pa* 'child' vs *kapa* 'children'

Nominal Morphosyntax

- Nouns also differ from verbs in that they can optionally take a **nominal marker *e-***

(1) **e-pa**   ean   mė'   ku-'ueh   ean
     NM-child   DEM   REL   KI-sleep   DEM
     'the child who was sleeping' (Kähler 1955 retelling)

(2) ka-m-kahnėn   **pa**   ean
     3-BU-fall.asleep   child   DEM
     'The child fell asleep' (eno_20181213_burung_hantu)

Nominal Morphosyntax

- This is inherited from the direct case marker *e-* that contrasted with the *u-* oblique marker in Old Enggano.

(3) ka-bu-pu̇a=da'a   e-dahao=dia   e-ka'a:i'io
3-BU-see=EMPH   DIR-niece=3SG.POSS   DIR-spear
'His niece saw the spear' (Kähler 1975: Krieg, 29.18)

(4) kamõhõ   epaE   e'ana ka-bu-'ua   nõ'õːĩ̃'ĩ̃ẽ i'ioo uko'E'E
3-different   DIR-child   DEM   3-BU-say   like.this OBL   OBL-devil
'Then the child spoke to the devil as follows' (Kähler 1955)

Nominal Morphosyntax

- However, *e-* no longer marks **core arguments** in contemporary Enggano but seems to emphasize the status of a word as nominal.

(5) jadi da-pu̇-pu̇ tingkah-tingkah **e-ke'ep**
    so 3PL-REDUP-see behaviour-FULL.REDUP NM-bird
    'Then they observed the birds behaviour' (eno_20181213_burung_hantu)

- The distinction between core arguments and oblique is now marked via the generalised **preposition** *o* or *ė* (< *i'ioo*)

(6) kinė'ah pe u ki-'iu [o ka-pa ė']
    like.that PT 1SG KI-say PREP KA-child DEM
    'That is what I say to this child' (eno_20181213_pidahrumah)

Nominal Morphosyntax

- Finally, nouns differ from verbs in that they can be possessed.

- When the possessor is nominal, it simply appears after the possessed item:

(7)   yub   Ali
        house   Ali
        'Ali's house/rumah Ali'

(8)   e-uba   u-ko'e'e
        DIR-house   OBL-devil
        'the devil's house' (Kähler 1975)

Nominal Morphosyntax

- When the possessor is pronominal, either a free pronoun can be placed after the possessed noun…

> (7) yub   u
>   house 1SG
>   'my house/rumah saya'

- … or suffixal/enclitic pronouns can be used to express the possessor:

(8a) yuba-'	(8b) yuba-b	(8c) yub=de
house-1SG	house-2SG	house=3SG
'my house/	'your house/	'his house/
rumahku'	rumahmu'	rumahnya'

Nominal Morphosyntax

- The suffixal/enclitic possessor strategy is the only strategy attested in the historiacal corpora

1SG	-'(V)u	euba**'au**	My house/ rumahku
2SG	-bu	euba**bu**	Your house/ rumahmu
3SG	=dia	euba**dia**	His house/rumahnya
1PL.INCL	-ka	euba**ka**	Our house/rumah kita
1PL.EXCL	=dai	euba**dai**	Our house/rumah kami
2PL	=du	euba**du**	Your house/rumah kalian
3PL	-da	euba**da**	Their house/rumah mereka

Nominal Morphosyntax

- The contemporary pattern is no longer regular:

1SG	-'	yuba'	My house/ rumahku
2SG	-b	yubab	Your house/ rumahmu
3SG	=de	yubde/yumde	His house/rumahnya
1PL.INCL	-k	yubak	Our house/rumah kita
1PL.EXCL	=da	yubda/yumda	Our house/rumah kami
2PL	=du	yubdu/yumdu	Your house/rumah kalian
3PL	-r	yubar	Their house/rumah mereka

Nominal Morphosyntax

- The patterns of possession may be even more complicated if the possessed noun is multimorphemic

ka'nė' 'my friends'/'teman-teman saya'	ka- 'plural' + **an** 'friend' + -' '1SG.POSS'
hẽkũã' 'my seat'/'tempat duduk saya'	**hẽk** 'sit' + -a 'LOC.NOM' + -' '1SG.POSS'
tėhė' 'my drink'/'minuman saya'	**it** 'drink' + -o 'PAT.NOM' + -' '1SG.POSS'

Verbal Morphosyntax

- Verbs in Enggano are expressed in one of three basic constructions

(9a)	a=dʉhʉr		ean	**ka**-b-dėhė	bu- verb
	when=finished		DEM	3-BU-hear	
	'afterwards, he heard'				

(9b)	Ki	ke'	**i**-dėhė	bare verb
	3SG	NEG	3-hear	
	'he didn't hear'			

(9c)	**Ki**	ki-dėhė	ki- verb
	3SG	KI-hear	
	'he heard'		

Verbal Morphosyntax

- *Bu-* verbs and bare verbs also take different sets of verbal agreement markers which agree with the subject (S/A)

	Set 1 (bu-)	Set 2 (bare)
1SG	u-	u-
2SG	ė-	u-
3SG	ka-	i-
1PL.INCL	ka-	ka-
1PL.EXCL	u- -a	u- -a
2PL	ė- -a	u- -a
3PL	da-	da-

Verbal Morphosyntax

- *Ki-* verbs, in contrast, use free pronouns/NPs to express grammatical functions.

- These are the main structures used in **relative clauses**:

  (10)   ẽ'   pa   [mẻ'   ki-pu]
           DEM   child   REL   KI-run
           'that's that child that ran'

- We believe that *ki-* is related to the marker *si/si=* that is used to mark relative clauses in many languages of Sumatra (e.g. Nias, Mentawai, Karo Batak, Gayo)

- It may have been extended to main clauses via **reanalysis** of (pseudo-)cleft constructions

Verbal Morphosyntax

- In addition to bearing the markers *bu-* and *ki-* and the agreement prefixes that accompany *bu-* and bare verbs in main clauses, verbs may also express **tense** and **aspect** (see Wijaya 2018).

- **Perfective aspect/past tense** is expressed via the auxiliary *ho*. This is always followed by a verb root with an allomorph of *bu-* and NO agreement.

  (11)   u   **ho**-m-no
         1SG   PFV-BU-eat
         'I already ate'

- We believe that *bu-* is related to the PAN actor voice marker *-um-* and many languages also used AV morphology in dependent clauses (see e.g. McDonnell & Chen 2022)

Verbal Morphosyntax

- Progressive aspect is indicated (as in many Austronesian languages) via the use of reduplication:

(12)   u   ki-**no**-no

          1SG   KI-REDUP-eat

          'I am eating'

- Finally, future tense/volitionality can be expressed either using the auxiliary *buh* (an innovation) or the *–(h)a* suffix also attested in Kähler:

(13)   Ki   **buh**   Ø-puak

          3SG   FUT   BU-go.home

          'he will go home'

(14)   Ki   ki-poka-**h**

          3SG   KI-go.home-FUT

          'he will go home'

Verbal Morphosyntax

- There are several different **verb classes** in Enggano

- Firstly, there are verbs that **do not take an overt allomorph of *bu-*** even in the contexts where we expect it (e.g. with agreement in main clauses and following auxiliaries like *ho*)

(15) ka-Ø-puak
3-BU-go.home
'He went home/dia pulang'

(16) ki ho-Ø-puak
3SG PFV-BU-go.home
'He went home/dia sudah pulang'

(17) ka-Ø-pari' ya-kare
1PL-BU-make NMLZ-dance
'We perform a dance'
(Perkawinan)

(18) Be a ho-Ø-pari' yakoro
because 1PL PFV-BU-make peace
'because we made peace'
(Asal Mula Tarian Perang)

Verbal Morphosyntax

- Secondly there is a class of **stative** verbs. These tend to be formed from stative roots/nominal roots plus the verb stem forming prefixes *a-* and *a'-:*

ka'daih 'white'
ka'kĩh 'dry'
ka'man 'fragrant'
kanih 'fast'
kanik 'clear'

ki- + a'- + daih 'white'

ki- + a- + nik 'clear'

yub 'house' → ka'yub 'have a house'

- These have slightly different meanings when reduplicated

ka'daihdaih 'pale'
karoprop 'very dirty'

Verbal Morphosyntax

- Finally, there are distinctions between **transitive** and **intransitive** verbs. Both can occur in the three main verbal constructions. However, they can combine with different derivational prefixes.

- Only transitive verb roots can combine with the detransitivising voice prefixes *di-* 'passive' and *aH-* antipassives (cf. *di-* vs *meN-* in Indonesian):

(19) paic   ean   **di-pari'**   Engga
      knife   DEM   PASS-make   Engga
      'That knife was made by Engga'

(20) Engga **k-a-bari'**   paic
      Engga ki-antip-make knife
      'Engga makes knives'

Verbal Morphosyntax

- Intransitive verbs may combine with **causative** *pa-* or the **applicatives** -i and -a' to derive transitive verbs

uhur 'roll'

(21) u **ki-pa-hur** ie ean
1SG KI-CAUS-roll stone DEM
'I roll the stone'

ku 'enter'

(22) u **ki-kua-h** yub Engga
1SG KI-enter-APPL house Engga
'I enter Engga's house'

pu'ur 'fall'

(23) Engga **ki-pu'da-'** kara bak ẽ' ne'e
Engga KI-fall-APPL glasses DEM earlier
'Engga dropped his glasses earlier'

Verbal Morphosyntax

- The marker *pa-* can also mark **middle voice** (which is a common pattern of polysemy in Austronesian) with both oppositional verbs (where it marks reciprocity) and non-oppositional verbs:

(24) Ki   **ki-pa-pú-pú**   iẽn
      3PL   KI-RECIP-REDUP-see   3PL.OBL
      'They looked at each other'

(25) u   **ki-pa-ko-koeh**
      1SG   KI-MID-REDUP-squat
      'I squat'

Verbal Morphosyntax

- Finally, there is a marker *-aba-* which is used to mark **consecutive action** in clause chains:

(26)   ka-bu-pa-kahre   pa   ean   **y-ab-a'a**
       3-ʙᴜ-ᴘᴀ-kill   child   DEM   3-ᴀʙᴀ-die
       'He killed the child and he died' (Kähler 1955 retelling)

(27)   ka-b-ahaːE=ha   **y-aba-kiu**=ha   i-tita.
       3-ʙᴜ-go=ᴇᴍᴘʜ   3-ᴀʙᴀ-hide=ᴇᴍᴘʜ   ʟᴏᴄ-there
       'It went there and sought shelter there' (Kahler 1955, 12.16)

Summary

- Having access to the **historical corpora** gave us a guide for elicitation (e.g. it led us to expect that verbs would occur in three major structures and to look for agreement markers)

- We might not have identified these without this expectation, since there are lots of potentially **ambiguous** forms, e.g. *ka-* which can be a 3SG agreement marker with *bu-* verbs but also represent the combination of *ki- + a-*)

- It also gave us an explanation for why suffixed forms (like possessed nouns and applicativized verbs) seem to contain an extra vowel that is not predictable from the semantics, namely these are **thematic vowels** that have been regularly lost word-finally and in non-stressed positions.

Summary

- This all helps to produce a more **comprehensive description** of the grammar of the language which may otherwise have been difficult given the context of largescale shift to Indonesian (and contact-induced change) and variation among speakers.

- This description shows that Enggano has a number of affixes that are cognate with **Austronesian morphology** (e.g. *bu-, di-, aH-, pa-* etc.)

- We will now turn to look at one other advantage of using historical materials in our analysis: the ability to look at **historical changes**!

Historical Change

Using the Historical Materials

- There are many changes in the morphosyntax of contemporary Enggano compared with Enggano as documented in Helfrich and Kähler.

- We have already discussed the loss of **case-marking** in nouns, the development of a new **possession marking strategy**, and the change in the form of affixes due to morpho-phonological processes like **final-vowel deletion**.

- In this final section of the talk, I want to illustrate two **syntactic changes** that we can see when we compare how structures are used across the corpora:

    (1) word order
    (2) relative clauses

Word Order

- In the historical corpora, *bu-* and bare verb constructions have verb-initial order:

(28) **ka-bu-púa=da'a** e-dahao=dia   e-ka'a:i'io
3-BU-see=EMPH DIR-niece=3SG.POSS DIR-spear
'His niece saw the spear' (Kähler 1975: Krieg, 29.18)

(29) kaupE **i-paka'ã:ũã'ã** ka-po'inãmõ e-púaha u-kaka e'ana
not.yet 3-know   DIR.PL-maiden DIR-look OBL-person DEM
'The maidens didn't know yet the appearance of the person.'
(Kähler 1940 Grammar, 24.15)

Word Order

- The only time that a subject appears before the verb is when it is preceded by the word kamohoː

(30)	**ka-mõhõ**	**e-pae**	e'ana	**ka**-b-ia	i-dita
	3-different	DIR-child	DEM	3-BU-exist	LOC-there
	'and the child remained there.' (Kähler 1955)				

Word Order

- These could be considered a topicalisation structure as it tends to introduce contrastive topics:

(31) ke'anaha e-'aupaka-**ra**, **d**-aha:e i-uba,
then DIR-departure-3PL 3PL-go LOC-house
[discourse topic = the parents]

ka-mõhõ **e-pae** e'ana **ka**-b-ia i-dita
3-different DIR-child DEM 3-BU-exist LOC-there
[discourse topic = the child]
'And then their exit happened, they went home, and the child remained there.' (Kähler 1955)

Word Order

- In contemporary Enggano, however, *bu- verb* constructions can be **verb-initial** but also appear with **pre-verbal subjects** even when the subject is not a contrastive topic:

(32)   e-ko'oe'   ean   **ka-b-abe'**
      NM-devil   DEM   3-BU-stand
      'The devil stood up' (Kähler 1955 retelling)

- Consequently, we can talk about word order change!

Relative Clauses

- In the historical corpora relative clauses may be introduced with an optional relativizer *mõ'õ*:

Helfrich Corpus

(33a) ke   ano=nia   [hemo'o k-a'ahko] i-ab-ako   i-kaudara   kahai

and   friend=3SG.POSS REL   KI-swim   3-ABA-arrive LOC-village one

'And her friend who swam arrived in a village.'(Helfrich 1916, Rat 39)

(33b)   … e-kaka   [ki-la   e-ayo   eana]

DIR-person   KI-bring   DIR-fish DEM

'the people who brought the fish' (Helfrich 1916, Earthquake 16)

Relative Clauses in Enggano

- If the relative clause contains a verbal predicate this is marked with *ki-*:

Kähler Corpus

(34a) Ka-'ėdėha=ha   e-paE   [hẽmõ'õ   **ku-'uoho**]
3-startle=EMPH   DIR-child   REL.SG   KI-sleep
'The child, who was sleeping, was startled' (Kahler 1955, 6.2)

(34b) e=apama   u=kaka   [mo'o **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)

Relative Clauses in Enggano

- Relative clauses may also contain nominal predicates (marked with the direct case marker e-).

- The **passive nominalisation** is used as an alternative means of relativizing on P

Kähler Corpus

(35a) i'iaha   e-kude-a   u-mẽhẽ-nũ   [mõ'õ aruu   **ki-nõ-nõõ**]?
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)

(35b) e-huda   e'ana   [mo'o   **e-di-pėa**   ama-nai]
DIR-woman   DEM   REL   DIR-PASS-see   OBL.father-1PL.EXCL.POSS
'The woman who was seen by you ('our father')' (Kähler 1957: 153)

Relative Clauses in Enggano

- In contemporary Enggano there are several important changes.

- Firstly, the number of relative clauses with an overt relativizer gets progressively higher:

	Without Relativizer	With Relativizer	Total
Helfrich Corpus	30 (56%)	24 (44%)	55
Kähler Corpus	151 (29%)	370 (71%)	521
Contemporary Corpus	64 (15%)	372 (85%)	436

Relative Clauses in Enggano

- The percentage of **headless relative clauses** in the dataset is much higher. In Old Enggano only around 12% of relative clauses were headless. In Contemporary Enggano, 218/436 (or 50%) of relative clauses are headless:

(36a)   di'iėb   ho-mėk   [mė'   ki-dap]

    PASS-do   PFV-many   [REL   KI-fall]

    'because many have fallen' (Asal Mula Tarian Perang, 29)

(36b)   ean   [mė'   k-a'-kũhẽ]

    DEM   REL   KI-VBLZ-use

    'those are the useful (ones)' (Kehidupan dalam Masyarakat, 11)

Relative Clauses in Enggano

- Finally, it is not only possible to use *ki-* marking as a strategy within relative clauses but ***bu- verbs*** also occur in the naturalistic texts:

(37a) ean   [mė'   **da-bu-**'u   burung   hantu]

DEM   REL   3PL-BU-say   bird   ghost

'that's what they call burung hantu (owl)' (Burung Hantu)

(37b) [mė'   **u-**pakõ'õã'   ẽ']

REL   1SG-know   DEM

'What I know is…' (Malakoni)

Relative Clauses in Enggano

- We argue that this fits with the story that *ki-* originated as a dedicated relative clause marker (like *si-* in other languages of Sumatra) and was **extended into main clauses**

- This is facilitated by the fact that Enggano (unlike Nias for example) has an **overt relativizer** (mõ'õ or mė') which takes on the main role of marking relative clauses (as evidenced by its increased frequency)

- Once ki- is seen as just a verbal marker that can alternate with bu- in main clauses, **the use of bu- is extended to other contexts** where it wasn't originally used (e.g. relative clauses)

- This has the effect of **removing any differences** between relative clauses and main clauses, just has been argued for in the development of PAN voice (Kaufman 2018)

Summary

- Consequently, the historical materials not only help in the documentation and description of contemporary Enggano, but also enable analysis of **paths of historical change,** which isn't always possible with underdocumented languages (see e.g. Feist & Palancar 2021)

- These can have important implications for our wider understanding of linguistic **theory and typology**

- Hence, they are an important part of achieving the final goal of the project **to assess the place of Enggano within the larger Austronesian family**.

Conclusions

Conclusion

- Enggano has a rich history of documentation, and is in the relatively fortunate position of having parallel text corpora from C19, C20 and C21.

- The historical resources present a challenge in terms of their **accessibility** and **comparability** – especially since we do not really know what the data collection and publication process involved (see also Dobrin 2021).

- However, they give us an important window into a **morphologically complex language** – providing an explanation for puzzling synchronic patterns, a template for analysing new texts as we build up the corpus and a window into diachronic changes.

With thanks to…

The Enggano Community

With thanks to…

Dendi Wijaya, Kantor Bahasa Bengkulu
Engga Zakaria Sangian, Universitas Dehasen Bengkulu
Mary Dalrymple, University of Oxford
I Wayan Arka, Australian National University

With thanks to…

Bernd Nothofer, Goethe-Universität Frankfurt am Main
Daniel Krausse, Goethe-Universität Frankfurt am Main
Erik Zobel, Independent Researcher

I Komang Sumaryana Putra and Colleagues at Udayana University, Bali