

# Challenges in describing contact-induced language variation

Lidia Federica Mazzitelli

Universität zu Köln/ Australian National University

Paris, 3. December 2021



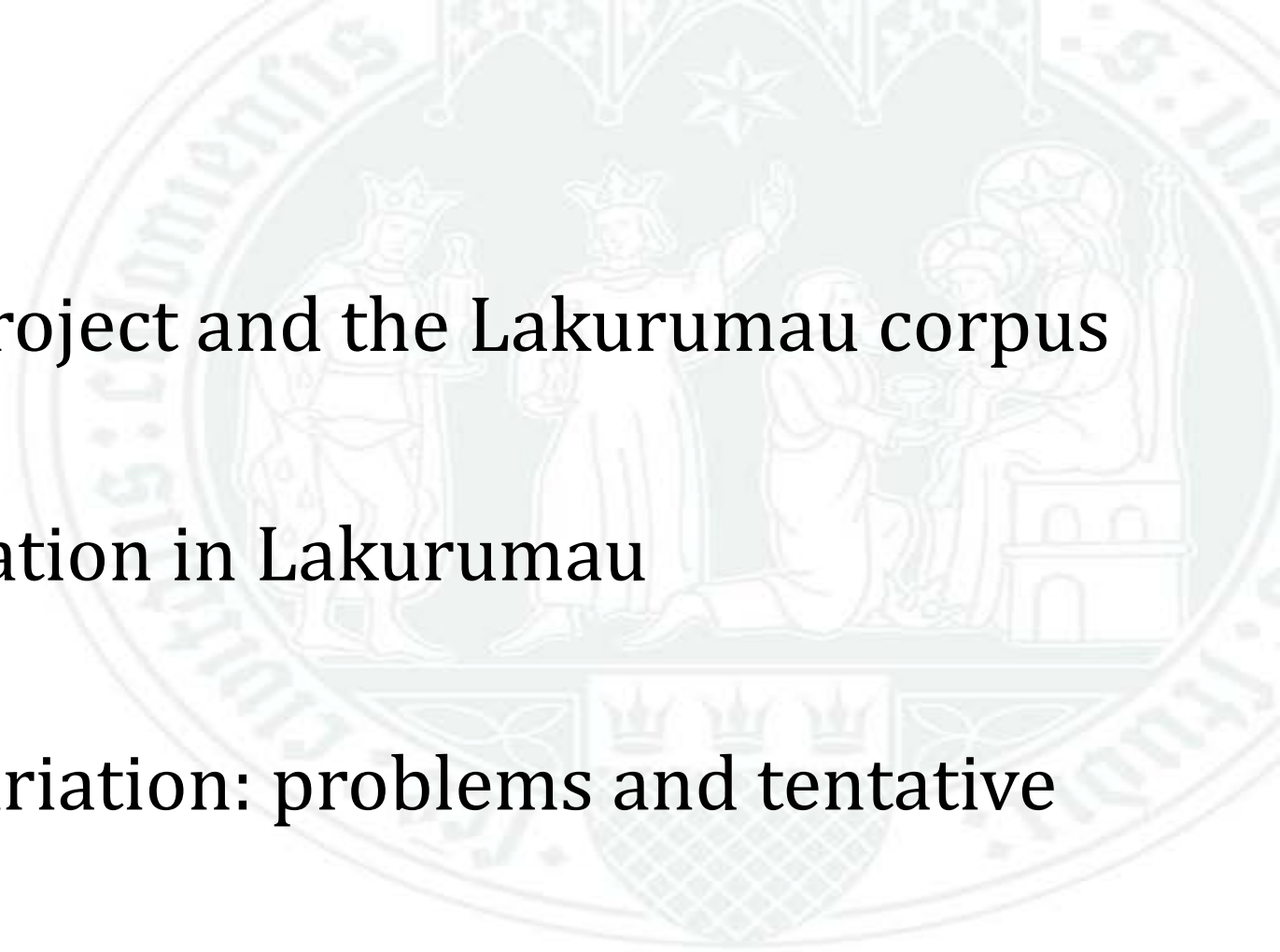
Australian  
National  
University

Universität zu Köln



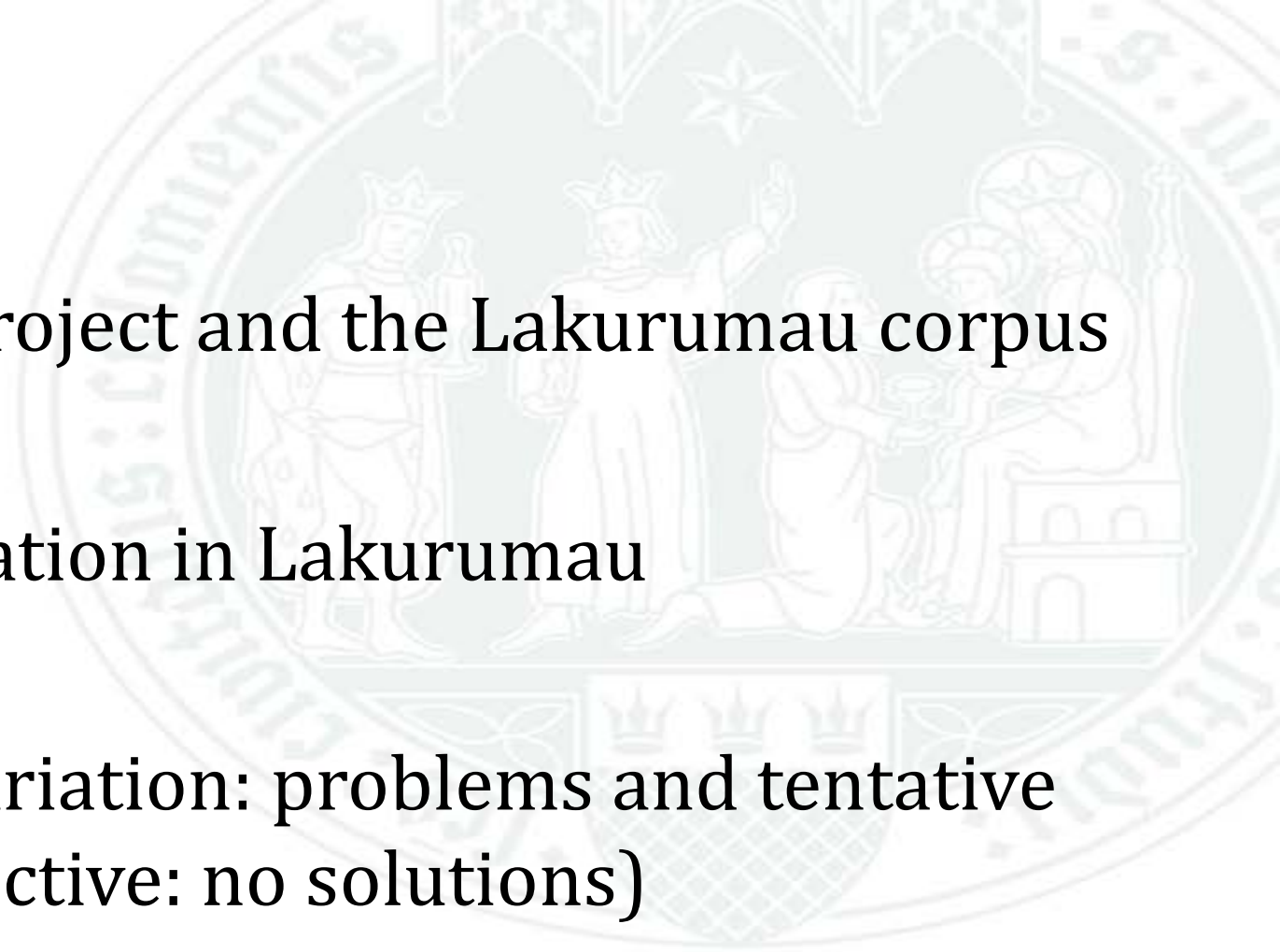
# Contents

- The Lakurumau description project and the Lakurumau corpus
- Types of contact-induced variation in Lakurumau
- Describing contact-induced variation: problems and tentative solutions



# Contents

- The Lakurumau description project and the Lakurumau corpus
- Types of contact-induced variation in Lakurumau
- Describing contact-induced variation: problems and tentative solutions (in an honest perspective: no solutions)



# The Lakurumau language and the Lakurumau description project



# Lakurumau

- Lakurumau is a Western Oceanic (Austronesian) language
- It is spoken in the Lakurumau village of New Ireland Province of Papua New Guinea by ca. 800 (?) people
- It is endangered → Lakurumau is still passed on to children in most households, but Tok Pisin is the dominant language of the community





# Lakurumau village

Lakurumau bay



Lakurumau main street



House of my host family





# The Lakurumau documentation project

- No previous published documentation or description work on the language, but Kara, the closest neighbouring language (with a high level of mutual intelligibility with Lakurumau) has a Bible translation and a grammar
- *Documentation and description of Lakurumau*, funded by ELDP
- 2017 – 2019; 3 fieldwork trips, for a total of 9 months

# The Lakurumau documentation project: Community wishes

- To have a linguist investigating Lakurumau and “certify” that it is indeed a language just as much as the neighbouring ones (Kara, Nalik)
- To have Lakurumau educational materials for schoolchildren
- Bible translation (but that was really beyond my means)
- Dictionary



# The Lakurumau documentation project: Community wishes

- The grammar and the documentation corpus were not immediate wishes of the community
- However, all Lakurumau speakers I interacted with were quite enthused to participate in the documentation project
- And they were also happy at the prospect of a Lakurumau grammar (in the form of a book)

# The Lakurumau documentation project: Outcomes (so far)

- Audio-video documentation corpus
  - › 22 h recorded
- Educational materials for the community
  - › Storybook and alphabet book for schoolchildren; storybook for adults; translation from Tok Pisin of an educational booklet for children
- Descriptive papers
  - › On valency-changing morphology; aspectual system; role of animacy
- Grammar
  - › Still to do!
  - › Reference grammar (at least at first)

# The Lakurumau corpus

- Archived with ELAR (<https://elar.soas.ac.uk/Collection/MPI1093372>)
- 22h recorded;
- 19h transcribed and translated into English
- Corpus sessions in this presentation are referred to with the header *lox001, lox002, ...*

The screenshot shows the ELAR digital archive interface for the Lakurumau corpus. At the top, there are navigation buttons for 'CATALOGUE', 'UP', and 'Advanced search'. Below this is a search bar with a dropdown menu for 'Language' and a search button 'ADD'. The main heading is 'lakurumau-mazzitelli-0475', with a breadcrumb trail 'Browse Archive > lakurumau-mazzitelli-0475'. A 'Sort by Type' dropdown is visible. The central area contains a table of metadata:

<b>lakurumau-mazzitelli-0475</b> lakurumau-mazzitelli-0475 <b>Collection</b> Collection Handle: <a href="http://hdl.handle...">http://hdl.handle...</a> Title: Documentation and Description... <b>Location</b> Continent: Oceania	<b>Country:</b> Papua New Guinea <b>Region:</b> New Ireland <b>Address:</b> Lakurumau <b>Latitude:</b> -2.886392 <b>Longitude:</b> 151.247898 <b>Languages</b>	<b>Language</b> Name: Lakurumau <b>Project</b> ID: IPFo270 Name: 0475-IPFo270 Description: Year: 2017Start_Date: En... <b>Project Information</b>	<b>Deposit Status:</b> Collection online <b>Funding Body:</b> ELDP <b>Actor</b> Name: Lidia Federica Mazzitelli Role: Depositor Affiliation: University of Cologne
--	---	---	---

Below the metadata is a 'REFINE YOUR SELECTION' sidebar with filters for Continent (Europe (1), Oceania (181)), Country (Germany (1), Papua New Guinea (181)), Language (English (175), Lakurumau (153), Tok Pisin (6)), and Person (Alex Tolivaina (5), Angela Junias (3)). The main content area is a grid of video thumbnails with titles: 'A traditional song', 'A trip to the yacht', 'A very big fish', 'A vision', 'Ananas', 'Bamboo', 'Betty Sikil informed consent', 'Biraau Dangui informed...', 'Bride price', 'Building a house', 'Cacao harvesting', 'Cannibalism', 'Childrens\_b...', 'Clans in Lakurumau', 'Colours, flowers and plants', 'REQUEST ACCESS', and 'Conversation about fishing'. A 'Search within' bar and a 'FOLDER' search button are also present.



# The Lakurumau corpus: Genres

Genres and sub-genres	Time (HH:MM:SS)
<b>Narrative</b>	
Personal narrative (eye witnessed)	04:00:30
Traditional narrative (non-eye witnessed)	01:56:40
Fiction	00:17:05
History	00:41:50
Exposition	01:59:46
Narrative stimuli	00:53:42
Reports	00:04:40
Tot.	09:54:13
<b>Dialogues</b>	
Spontaneous chats	01:45:55
Spontaneous discussions	01:30:56
Staged chat	00:06:47
Staged discussion	05:31:54
Tot.	08:55:32
<b>Procedural</b>	00:48:50
Observational filming	00:50:00
Elicitation	01:35:36
Music	00:06:34
Total	22:10:45

<b>Personal narrative</b>	description of one's life/family; the building of a house; hunting stories
<b>Traditional narrative</b>	traditional stories; origins of the clans,...
<b>Fiction</b>	two stories invented (and written!) by a consultant
<b>Oral history</b>	Second World War; the German colonization of New Ireland
<b>Exposition</b>	funerary and marriage rites; description of flora and fauna
<b>Narrative stimuli</b>	e.g. <i>The Frog Story</i>
<b>Spontaneous and staged chats</b>	chats with no predetermined topic
<b>Spontaneous and staged discussions</b>	chats with a predetermined topic (e.g. „the future of the Lakurumau language“)
<b>Procedural</b>	how to weave a basket; how to make sago; how to bake coconut cake

# The Lakurumau corpus: Genres



Dangui Mosley tells a story about the warrior Malaxon  
**Traditional narrative / Oral history**



Tolingare Tokulau and Konda Pikia talk about how to fish sharks  
**Staged discussion + exposition**



Talking while peeling sweet potatoes  
**Spontaneous chat**  
**!!! They knew the camera was on !!!**



Roberta Sarameli and Dinah Gurumang talk while sewing a blouse  
**Spontaneous chat + procedural**

# The Lakurumau corpus: The speakers

- 36 speakers: 33 adults (19 to 89) and 3 children (11 to 16)
- 19 women, 14 men
- Different education backgrounds (primary education only to college degree)
- Different level of proficiency in Lakurumau, even though all use it at home to some extent
- No monolingual speakers: all Lakurumau speakers are fluent in Tok Pisin; many also in other local languages
- Two non-native speakers (identified as such in the corpus)



# The Lakurumau corpus: Speakers

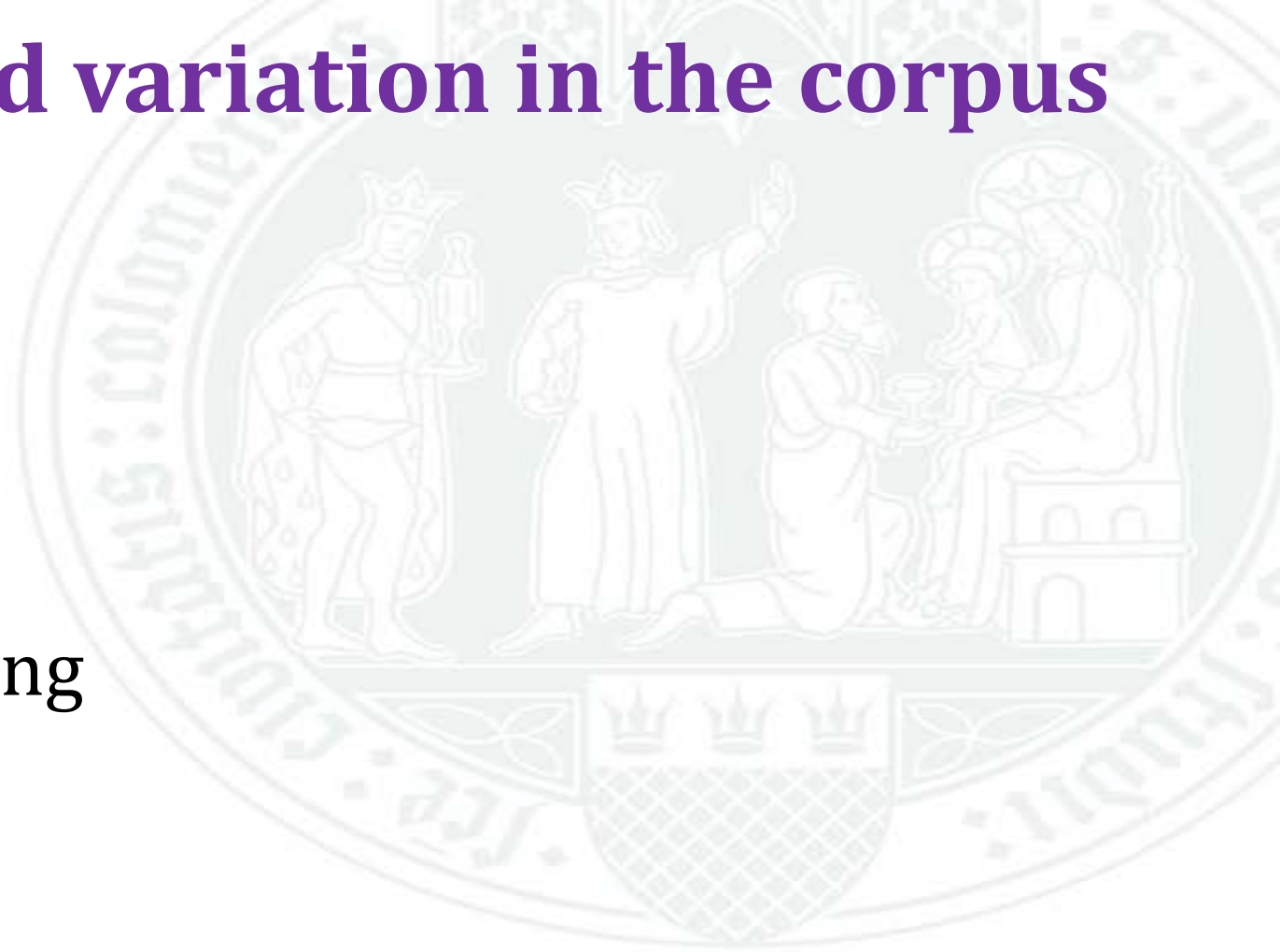
<b>Number of speakers</b>	<b>Year of birth</b>	<b>Age (at the time of recording)</b>	<b>Time (HH: MM)</b>
3	2008 – 2006	11 – 15	00: 18
2	1998 – 1996	20 – 22	
3	1980	37	01: 21
6	1973 – 1961	45 – 53	20: 59
16	1960 – 1950	58 – 68	
3	1949 – 1943	69 – 75	
3	1939 – 1933	79 – 85	

# Documenting and describing contact-induced variation in Lakurumau



# Types of contact-induced variation in the corpus

- Lexical borrowings
- Code-mixing and code-switching
- Grammatical calquing





# Types of contact-induced variation in Lakurumau

- Lexical borrowings
  - › Very frequent, in all speakers of all ages and level of proficiency
  - › Mostly for modern-age realia, but not only

66 years old speaker

*Dira daa **rekordim** azo ra mos dira daa fedawaai wana, singsaxaa naadi anaam... Usi lo **tok ples skul, tok ples skul** maandi **tok ples**, singsaxaa **tok ples** mon lo lena **klasrum**, aa lo yaan di fang seng, Tok Pisin malaam ket, Tok Pisin. **So** ka vit di roi vadikdik a vedawaai*

‘We will **record** the things we talk about, but them up there [at the school]... Like in the **tok ples school**, in **the tok ples school** we speak **tok ples** but we speak **tok ples** only in the **classroom** and when they go away [from school], they only speak Tok Pisin. **So** they don’t hold well onto the language’ (*lox024*)

# Types of contact-induced variation in Lakurumau

- Lexical borrowings
  - › Very frequent, in all speakers of all ages and level of proficiency
  - › Mostly for modern-age realia, but not only

59 years old speaker

*Okay maadi valaa asang. Maadi walongin asang anga wiu xa **baak**. Ga daa pet oy anga wiu xa xupkup, ah.*

‘Okay, the two of us went further. We then heard a dog barking [*baak*]. I will say that a dog was barking [*xupkup*], ah’. (lox020)

# Types of contact-induced variation in Lakurumau

- Code-mixing and code-switching
  - › Quite rare in my corpus, even in spontaneous dialogues
  - › Only found in dialogues when a speaker turns to someone with whom they usually use Tok Pisin or English, or when reporting someone else's words

56 years old speaker

[Context: She was illustrating me the uses of different plants in her garden]

*A ngona ayaam... **very nice, a sweet.**  
Dira daa yaan. **Okay. When it's ripe.***

'This fruit here... **very nice**, it's **sweet**. We usually eat it. **Okay**. [We eat it] **when it's ripe.**' (*lox071*)

# Types of contact-induced variation in Lakurumau

- Code-mixing and code-switching

- › Quite rare in my corpus, even in spontaneous dialogues
- › Only found in dialogues when a speaker turns to someone with whom they usually use Tok Pisin or English, or when reporting someone else's words

59 years old speaker

*Aa nam Naomi zi xa fai zik aa nam baat akam ka valik kus naan o xa daa valaa xus nam Naomi “**yu no ken toktok**”*

‘And Naomi instead went down there and that person went and told her... he went and told Naomi «**you cannot talk**» (lox023)



# Types of contact-induced variation in Lakurumau

- Grammatical calquing
  - › In some cases, difficult to tell, because Oceanic languages of New Ireland are the substrate for Tok Pisin, which is structurally very similar to them
  - › However, there are some clear cases where Lakurumau speakers use «wrong» Lakurumau constructions, which are correct in Tok Pisin
  - › In the corpus is most frequent in younger speakers or speakers who use predominantly Tok Pisin

# Types of contact-induced variation in Lakurumau

- Grammatical calquing from Tok Pisin

## Loss of in/alienable possession distinctions

13 years old speaker

*Di-na*                      *za-gu*                      instead of                      *Din-nangu*  
same.sex.sibling-3SG    ALIEN.POSS-1SG                      same.sex.sibling-1SG  
'My sibling' (*lox264*)

Tok Pisin: *Sista/brata bilong mi* 'sister/brother of me' no alienability distinction

# Types of contact-induced variation in Lakurumau

- Grammatical calquing from Tok Pisin

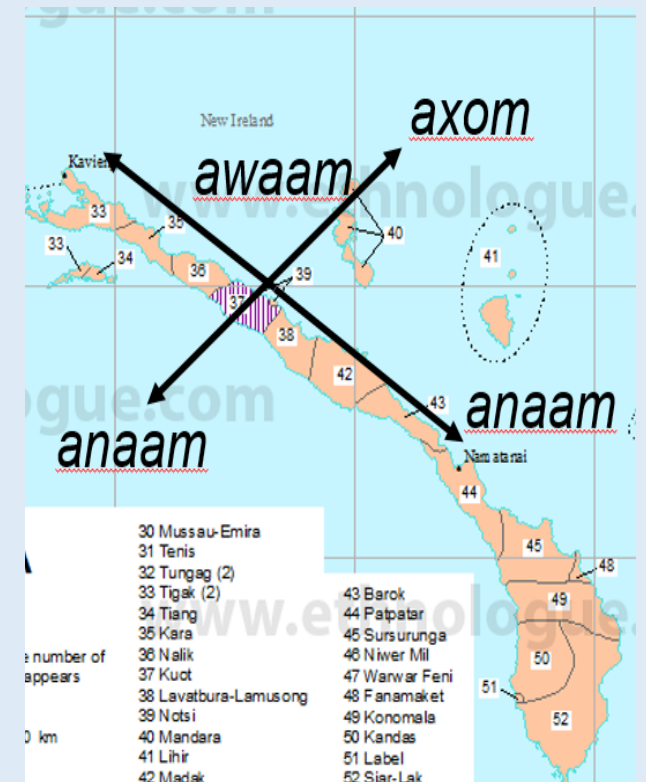
## Loss of spatial distinctions

37 years old speaker (uses predominantly Tok Pisin)

*Naan in **awaam** Kanada* instead of *Naan in **axom** Kanada*  
'He is from Canada' (*lox204*)

*Awaam*: on one's left facing the sea

*Axom*: on sea; away from New Ireland



# Types of contact-induced variation in Lakurumau

- Grammatical calquing from Tok Pisin

## Use of prepositions with spatial nouns

11 year old speaker

*Maandi= laak asang lo latan* instead of *Maandi laak asang latan*

1PL.EX.S= go.up IAM LOC bush

'We went up to the bush' (*lox250*)

Tok Pisin: *mipela go antap long bus* 'we went up into the bush'



# Types of contact-induced variation in Lakurumau

- Grammatical calquing from Tok Pisin

## Subject markers and numerals

15 years old speaker (NB: not very proficient; at home she speaks Tok Pisin and goes to boarding school in another village)

A *ralavaat* *revin* *malef lo* *zaxaa* *cube* instead of

ART four woman sleep LOC one cube

A *revin* *a=* *ralavaat* *u=* *maleflo* *cube* *a=* *zaxaa*

ART woman 3SG.S= four 3PL.S= sleep LOC cube 3SG.S= one

'Four women sleep in the same room' (*lox265*)

Tok Pisin: *foapela menmeri silip long wanpela cube*

# Types of contact-induced variation in Lakurumau

- Grammatical calquing from Tok Pisin

**Anteriority marker *bin* – also in older and very proficient speakers**

59 years old speaker

*Orait nanga= valaa sik ket a ron akam sa=gu*

alright 1SG.S= go.forth take again ART spear that AL.POSS-1SG

*nanga= **bin** poi naan paamua.*

1SG.S= ANT spear 3SG first

‘Alright I went to get back my spear (with which) I had speared him before’ (*lox020*)

instead of: *nanga woi naan paamua // nanga woi naan **tabung*** ‘in older times’

Tok Pisin: *mi **bin** sutim en*

# Judgements

- Lakurumau speakers define as «bad language» both lexical borrowings and grammatical calquing
- Lexical borrowings however are more accepted, even more so those for which no Lakurumau equivalent exist.
- Grammatical calquing on the other hand is considered as «bad language» - also because it is objectively «wrong»

# Documenting variation: problems

- Getting the data was not a problem: (at least I think that) the corpus is diverse and big enough to actually offer a wealth of variation examples
- But of course it will need to be expanded with many more recordings of younger speakers which at the moment are sorely lacking



# Documenting variation: problems

- I did receive occasional complaints as to some speakers I decided to include, who were seen as “bad speakers” by other community members
- However, my motivation of representing a wide range of Lakurumau speakers and varieties was eventually well understood and accepted (I fear also because I didn’t let them much choice – I just recorded anything and anyone willing to)
  - › But how much of this acceptance is just because the recordings are seen as a “less prestigious” form of language description with respect to the written standard (educational materials/grammar)?

# Describing variation: problems

- The description of (also contact-induced) variation is quite interesting for linguists and I think it would be good to have it in the reference grammar
- However, how can I describe contact-induced variation without flagging those speakers as «bad speakers»?
  - › The grammar will be cross-referenced with the corpus, and speaker information would be easily be retrievable
  - › Lakurumau speakers may not all read the reference grammar, but some may – it is a source of prestige to have a printed grammar of the language

# Solutions (?)

- Avoid using depreciative terms such as «mistakes» (it's obvious but still)



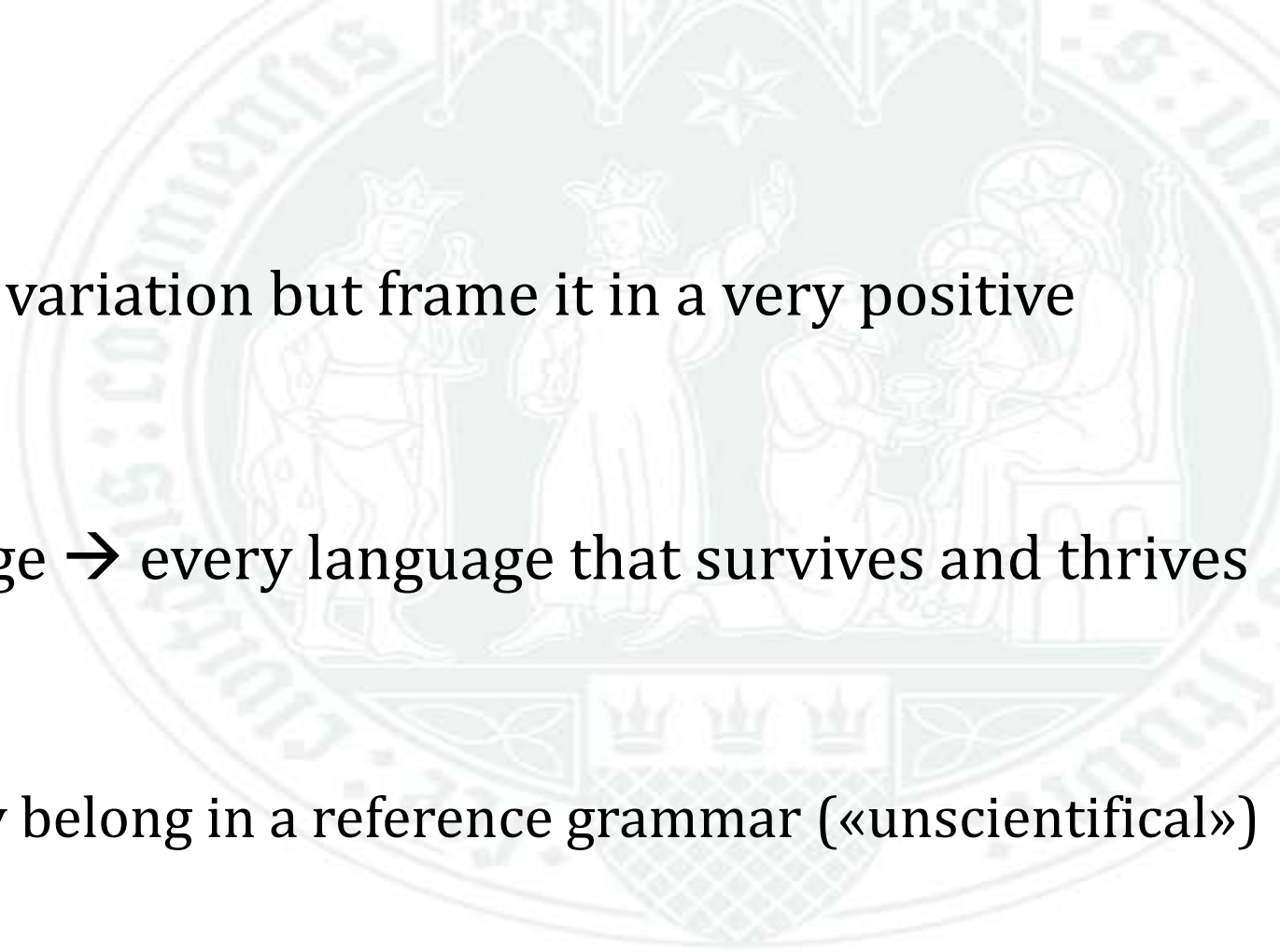
# Solution A

- Just for the chapter on variation, leaving speaker information and corpus cross-reference aside → present the examples, but give no source
  - › It is true that data retrievability is a tenant of corpus-based language description, and future linguists may want to see the examples in their context:
  - › however, such information may be given in the corpus metadata and willing linguists will be able to find it
  - › But community members may still be able to track them down to the source speakers



# Solution B

- Have speakers data in the chapter on variation but frame it in a very positive language
- Emphasis on the naturalness of change → every language that survives and thrives changes at some point (also
  - › Appreciative language does not really belong in a reference grammar («unscientific»)



# Solutions C

- Leave the description of variation aside altogether in the grammar, and publish it in a separate form (a paper)
  - › Academic papers may be less accessible to the community → hence speakers are more protected
  - › But:
    - » it doesn't feel very ethical to «hide» information from the community;
    - » papers might still be read by them;
    - » it makes the grammar less informative (?)

*Ka doxo marasaxaa - Thank you - Merci!  
(sorry I didn't provided any ideal solution)*



*Sunset in Lakurumau*

# Bonus track: Recognition (→ and funding)

- Grammars are usually well recognised as «good enough» outputs for Ph.D.
- But they rarely are further in the academic career → they often are not recognised as a «valuable monograph» (e.g. for German *Habilitation*, or for getting a tenure in the US system)
- Even worse is the situation with documentation corpora (at least so I experienced), which have very little «official academic» recognition
- I might be wrong and this may just be my impressionistic biased view, but I think it may be worth also discussing about this a bit