

Examples: how to get them, how they are used by others

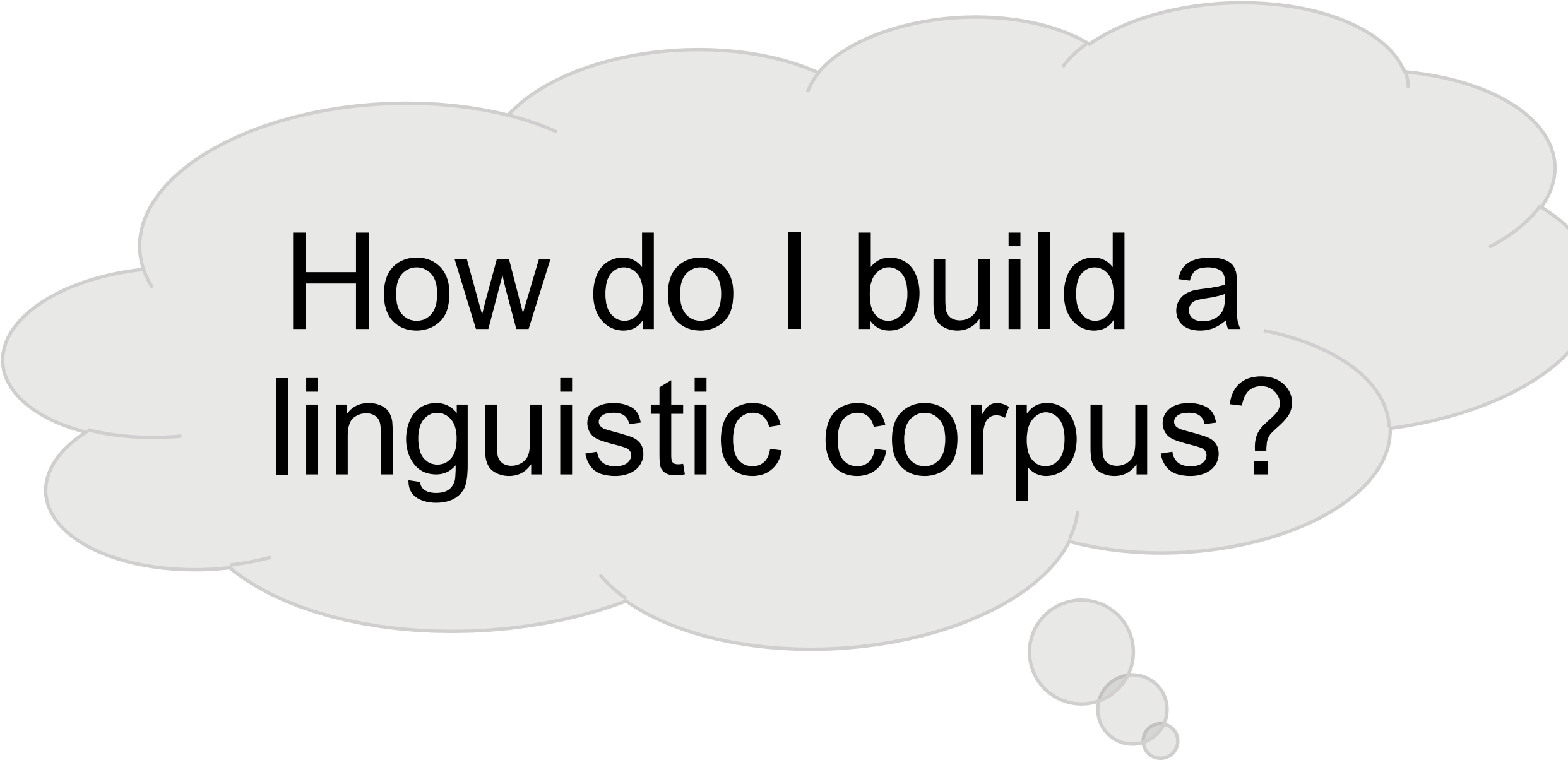
Eline Visser [ei'linə'fɪsəɹ]



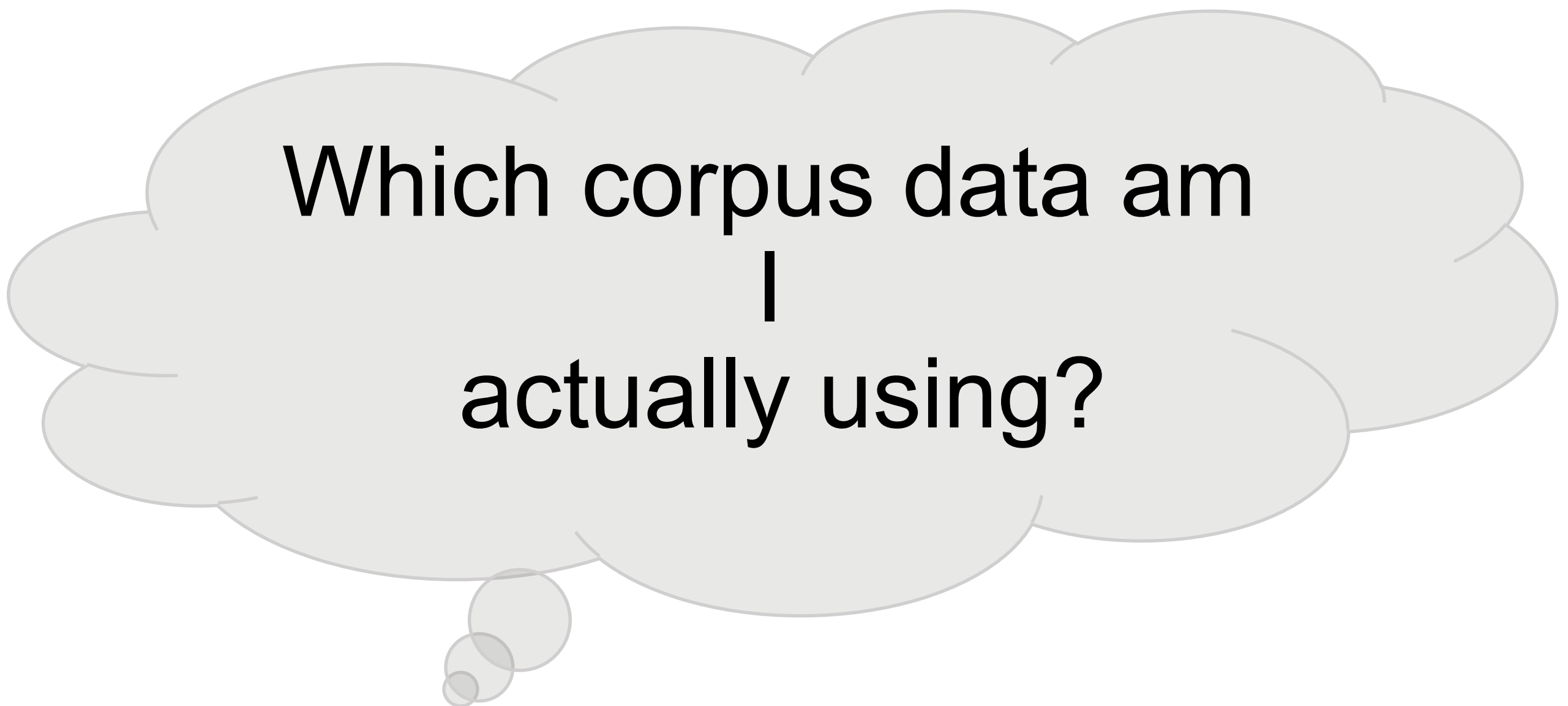
Wenner-Gren Stiftelserna
Wenner-Gren Foundations



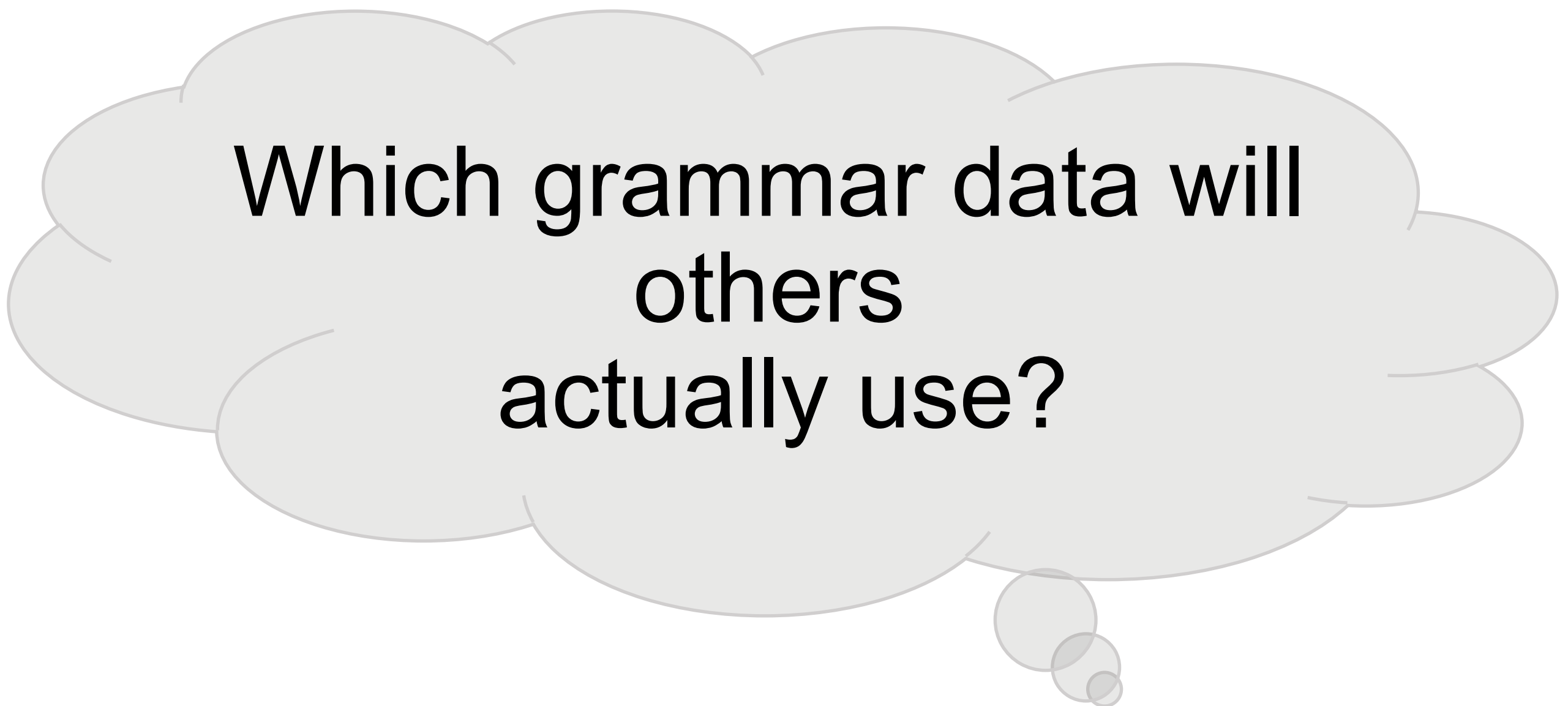
UiO : **University of Oslo**



**How do I build a
linguistic corpus?**



Which corpus data am
I
actually using?

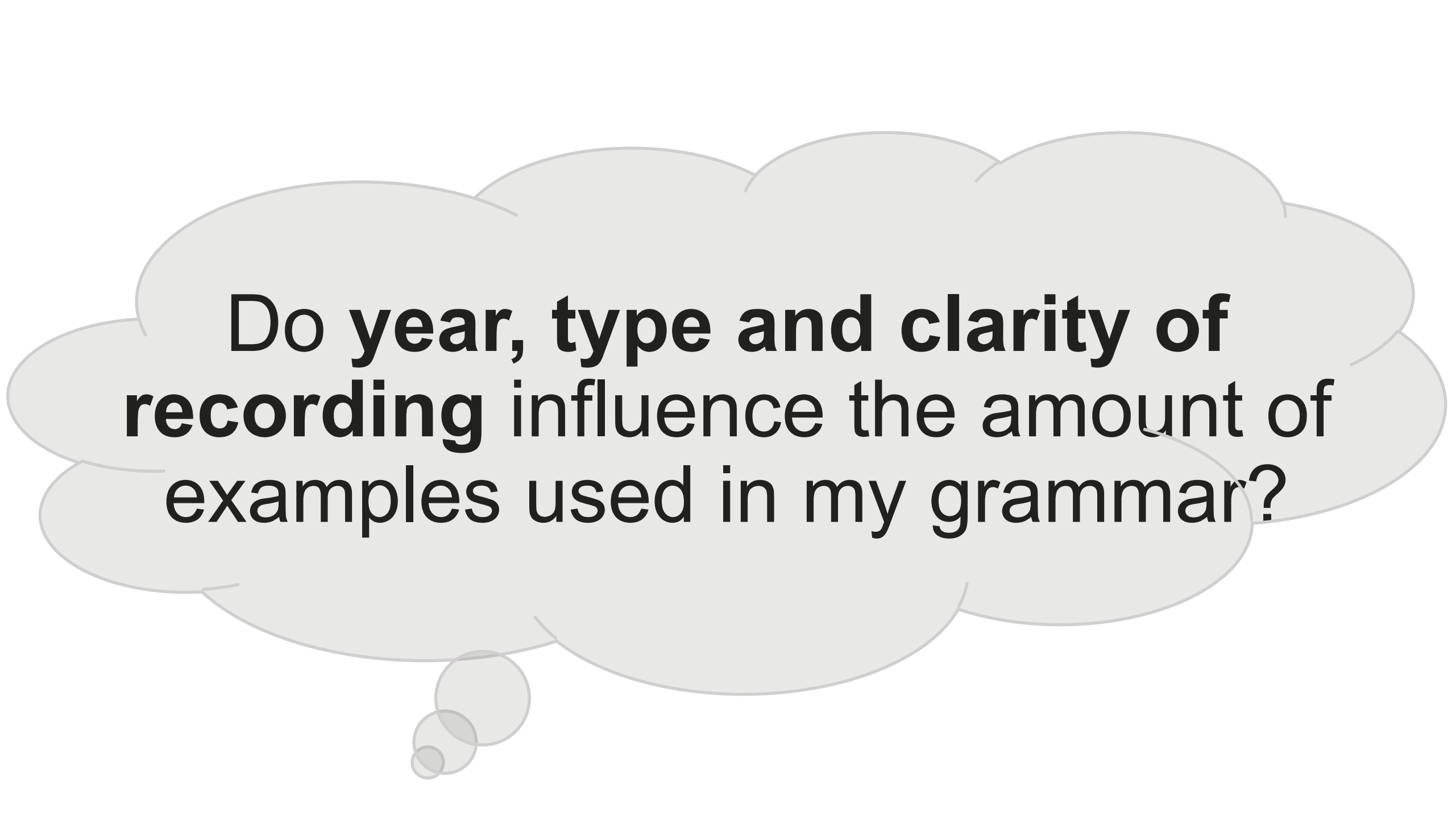


**Which grammar data will
others
actually use?**

Overview

- Grammar writer's perspective
 - stats from 4 grammars
 - year
 - type
 - clarity
 - Grammar user's perspective
 - typologists questionnaire
 - five principles for good examples
- } affect number of used examples

Grammar writer's perspective



Do year, type and clarity of recording influence the amount of examples used in my grammar?



A grammar of Palula

Henrik Liljegren



Attributive constructions in North-Eastern Neo-Aramaic

Ariel Gutman



A grammar of Yauyos Quechua

Aviva Shimelman



Advances in the study of Siouan languages and linguistics

Catherine Rudin (ed), Bryan James Gordon (ed)



Tone in Yongning Na: Lexical tones and morphotonology

Alexis Michaud



A grammar of Rapa Nui

Paulus Kieviet

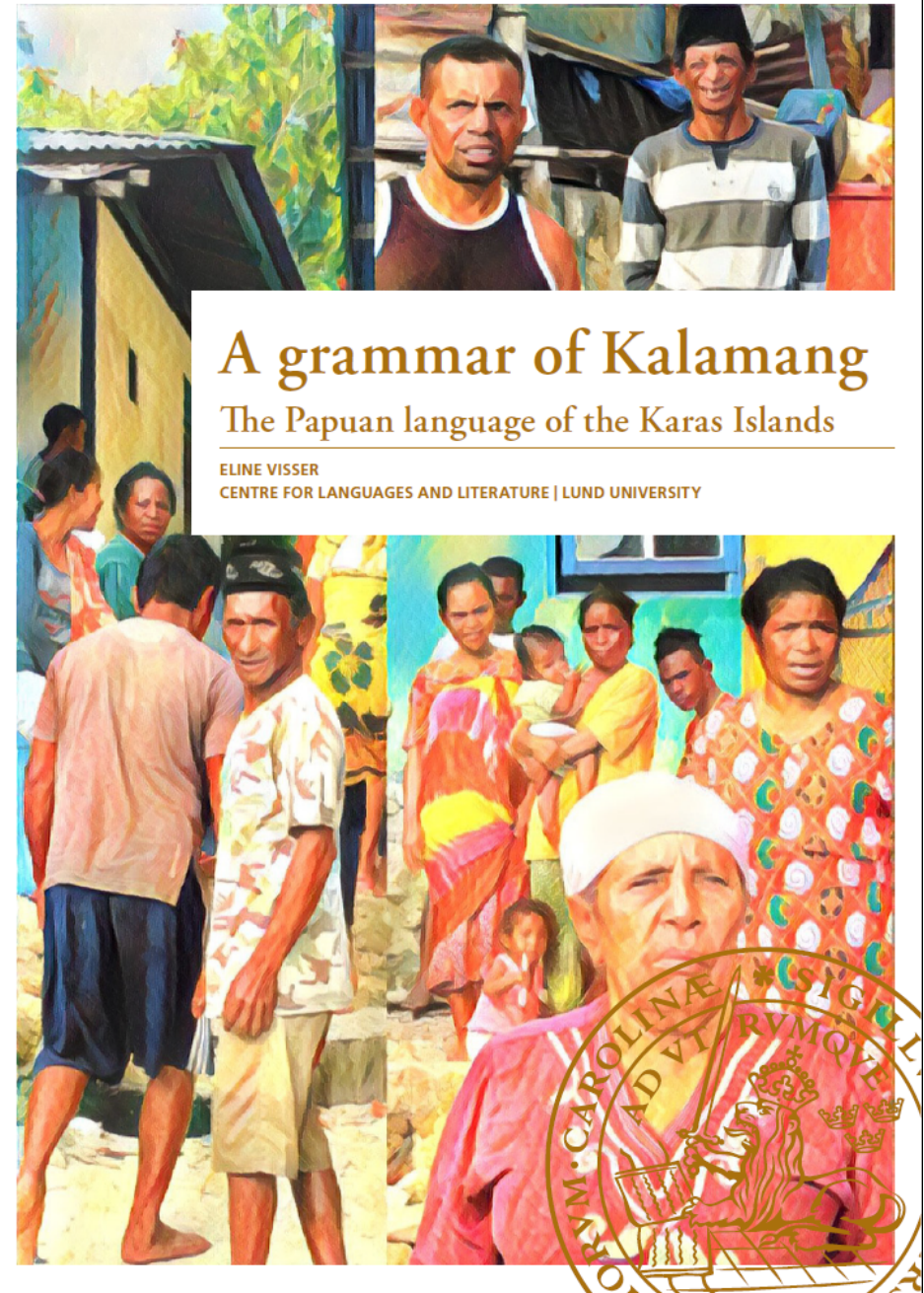
- Variables

- Results

- all independent variables have an effect on dependent variable

- year, type and clarity all influence the amount of examples that are used

Kalamang

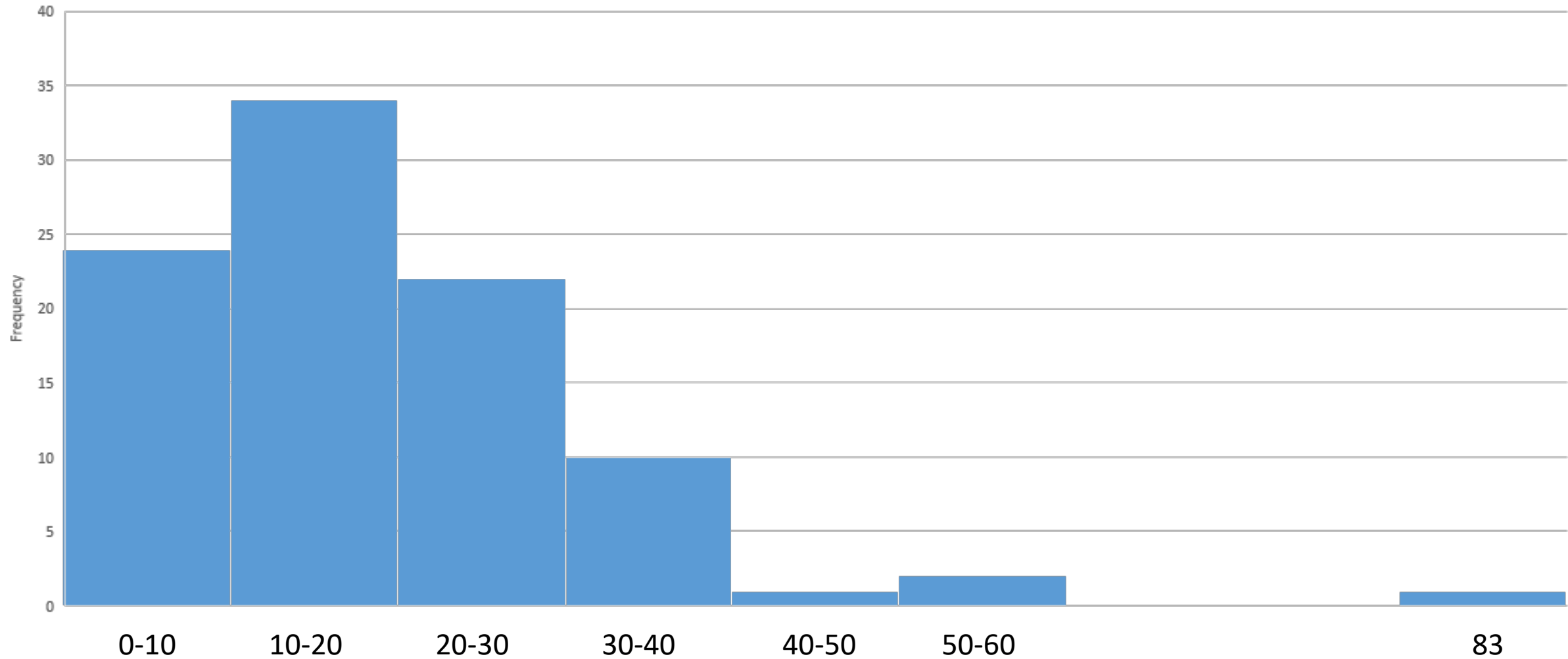


- 107 texts
- 70.000 words (avg 670 per text)
- 1050 examples

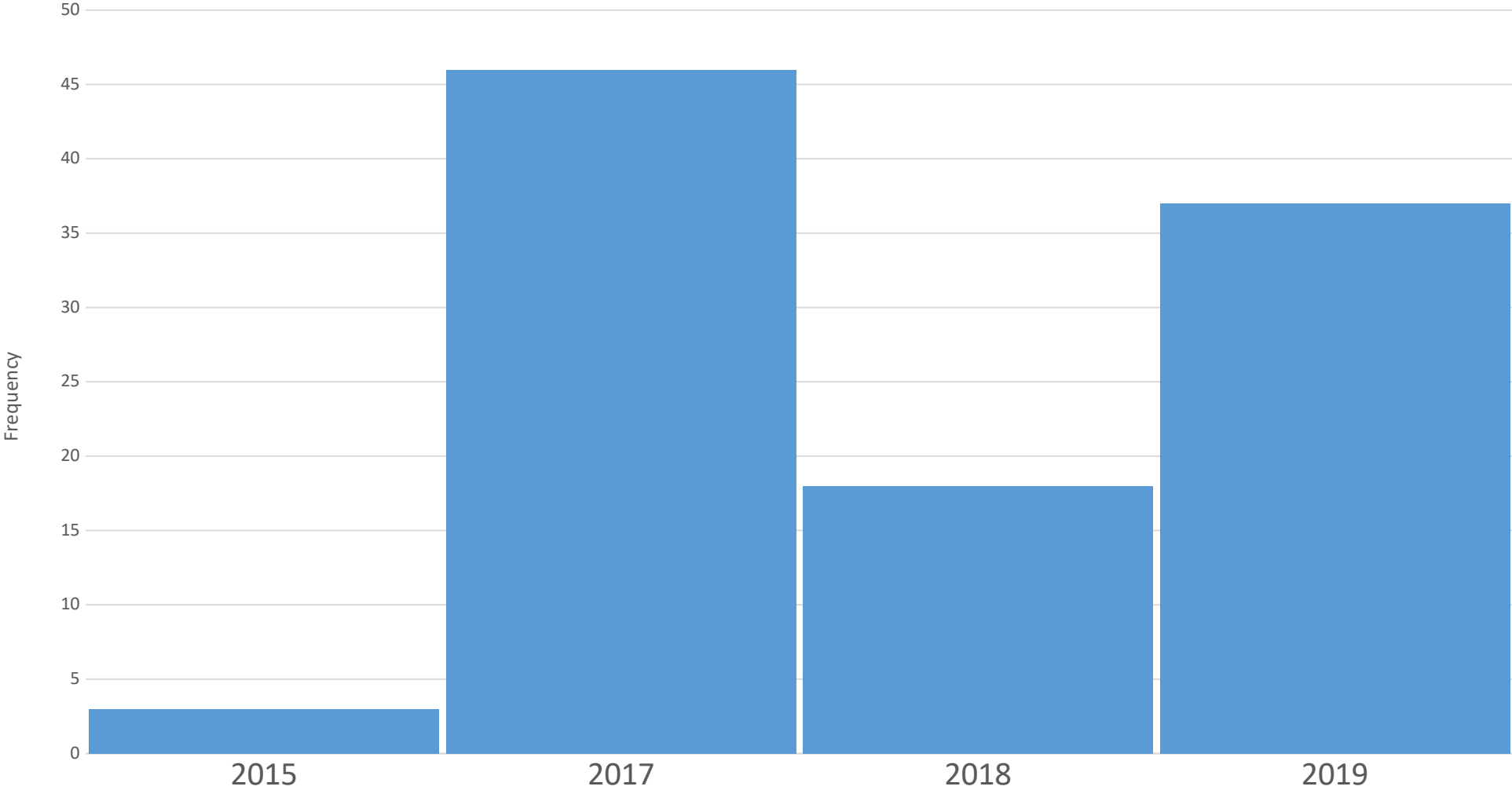
examples per 1000 words

- range: 0 – 50 (one outlier)
- average 16,4

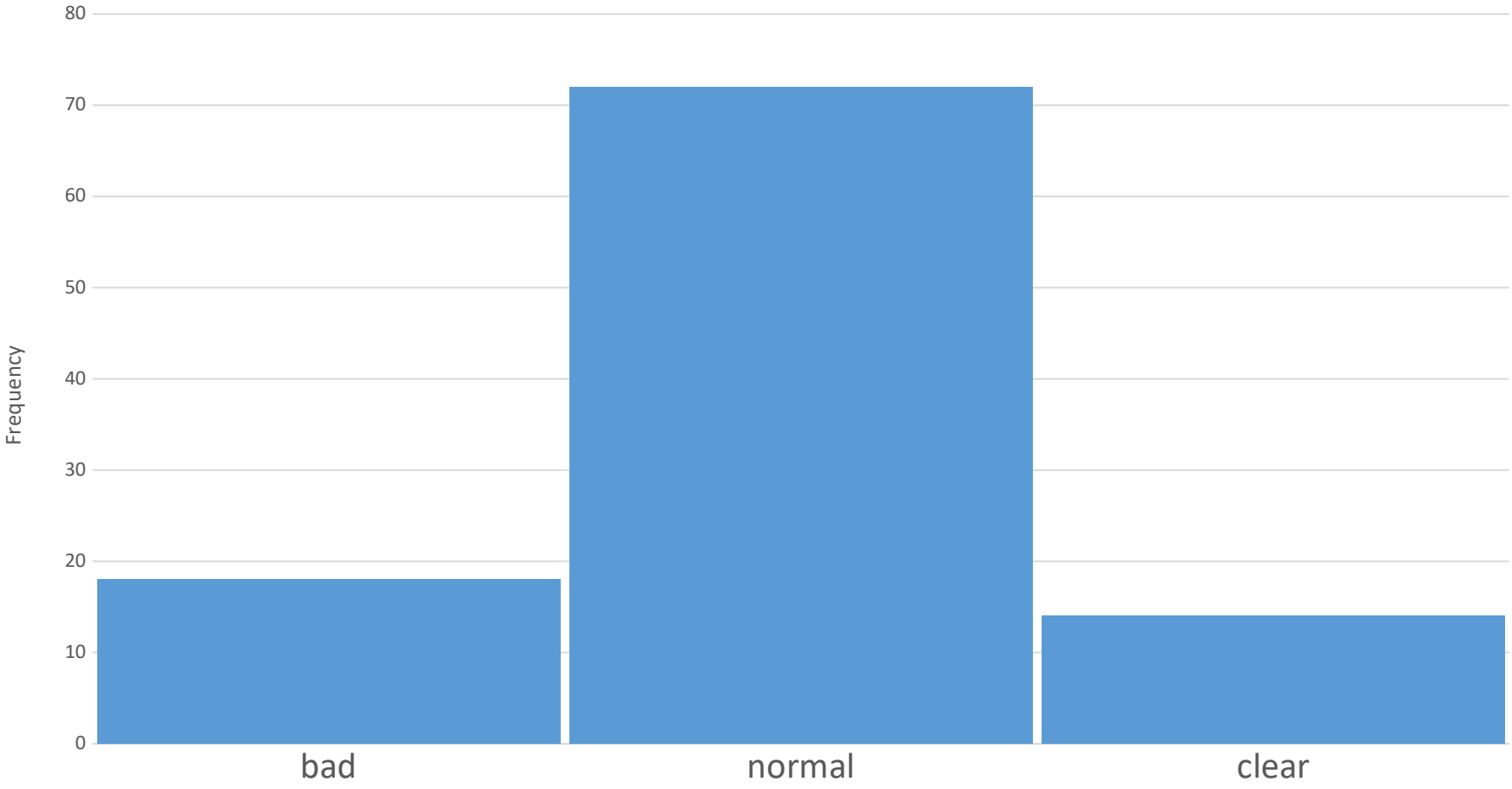
amount of used examples



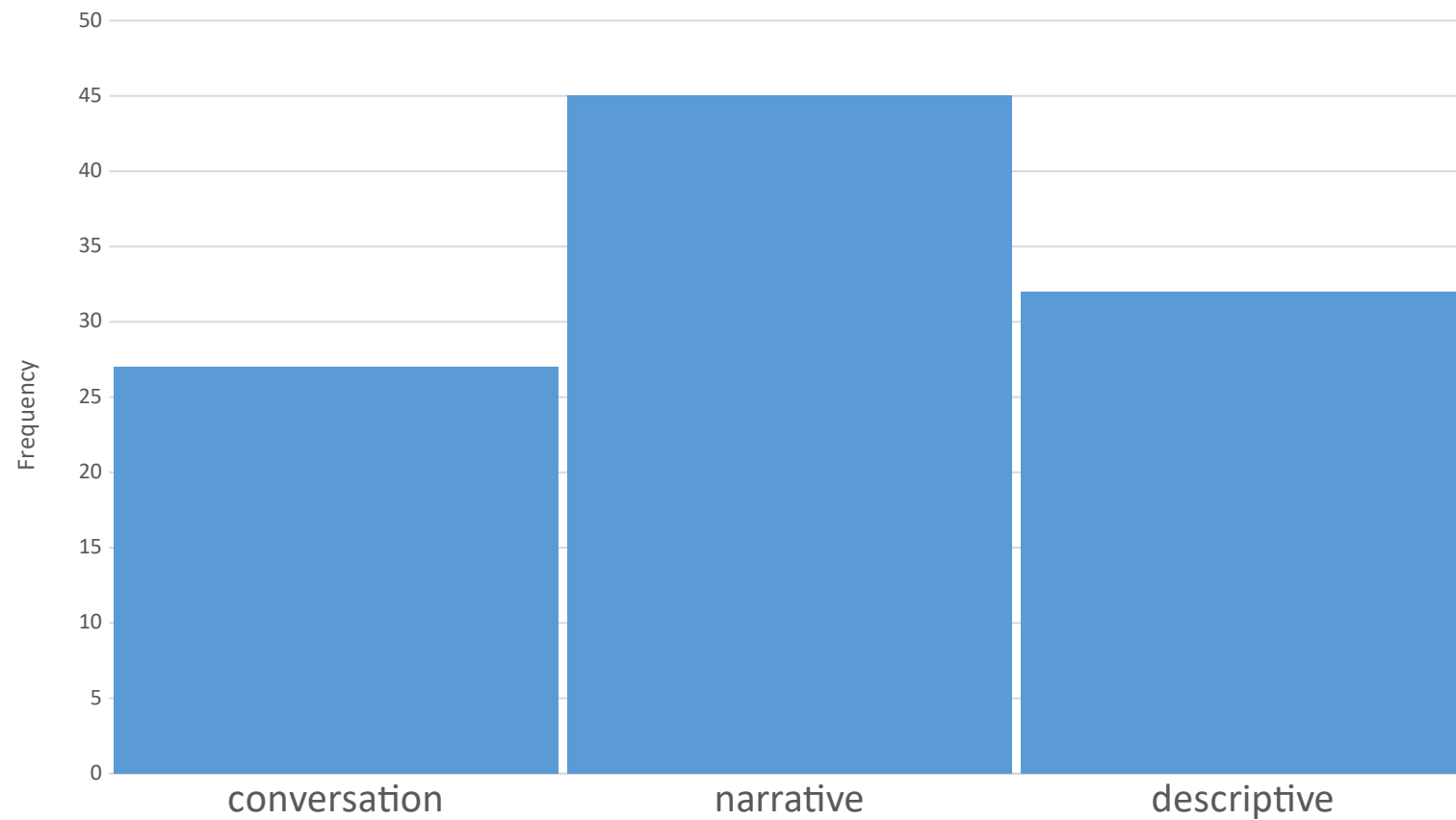
year of collection



clarity



type of text



Pichi

- Kofi Yakpo



amount of used examples

- 34 texts
- 46.000 words (avg 1350 per text)
- 1650 examples

examples per 1000 words

- range: 0 – 76 (one outlier)
- average 40

Pite Saami

- Joshua Wilbur



amount of used examples

- 23 texts
- 17.000 words (avg 730 per text)
- 236 examples

examples per 1000 words

- range: 0 – 43 (one outlier)
- average 19

Rapa Nui

- Paulus Kieviet



amount of used examples

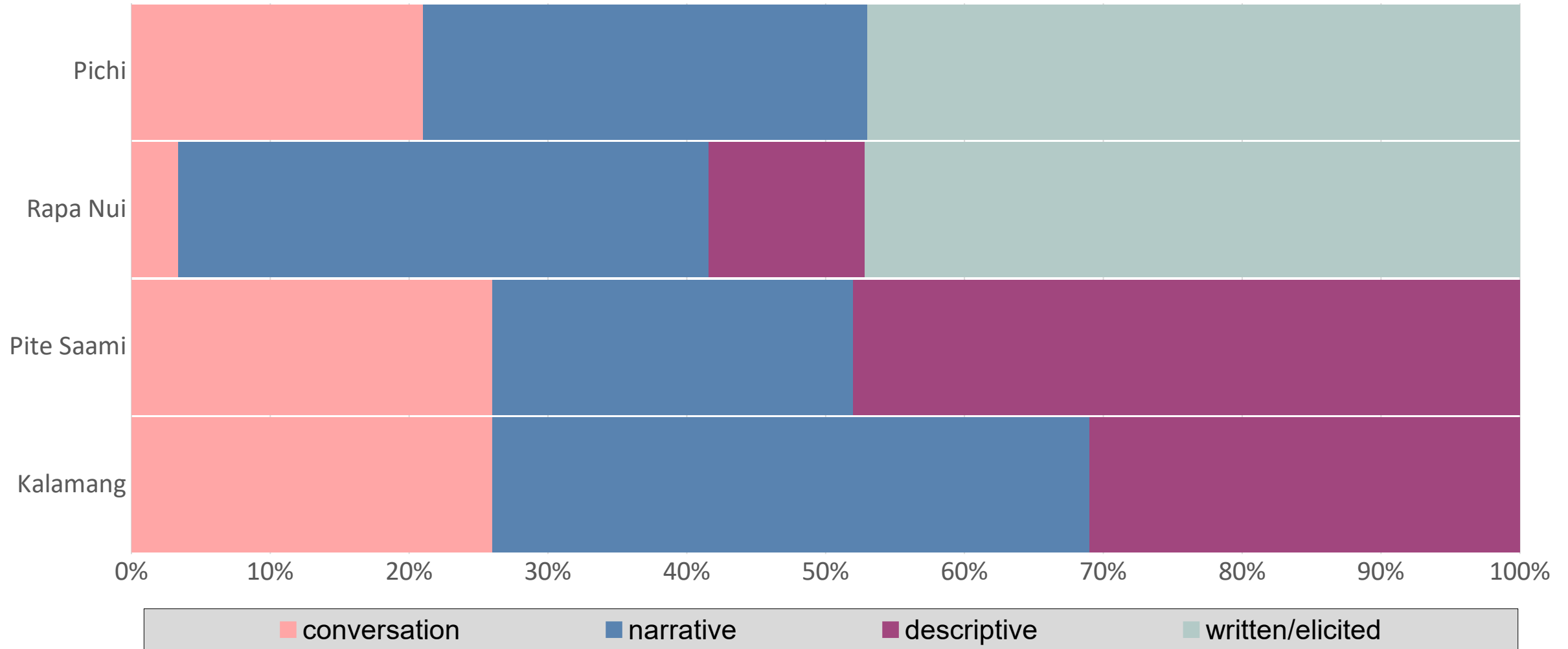
- 91 texts
- 990.000 words (avg 11.000 per text)
- 1750 examples

examples per 1000 words

- range: 0 – 24
- average 4

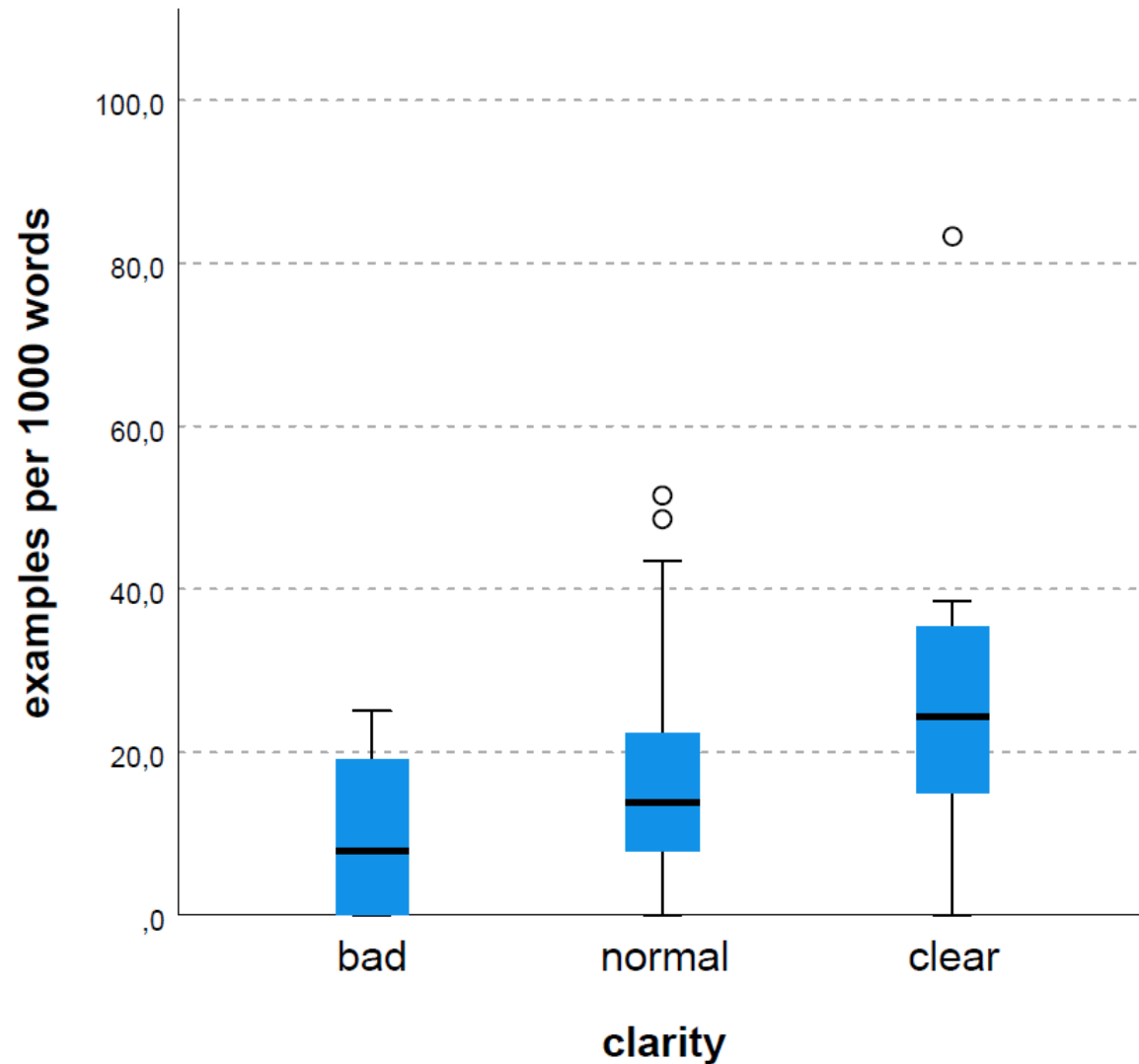
	no. of texts	words	words per text	no. of examples	examples per 1000 words	average examples per 1000 words	recording years
Kalamang	107	70000	670	1050	0-50	16	2010s
Pichi	34	46000	1350	1650	0-76	40	2000s
Pite Saami	23	17000	730	240	0-43	19	2000s, 2010s
Rapa Nui	91	990000	11000	1750	0-24	4	1940-2000
total	255	1123000		4690			

type of text



Results

clarity (Kalamang only)

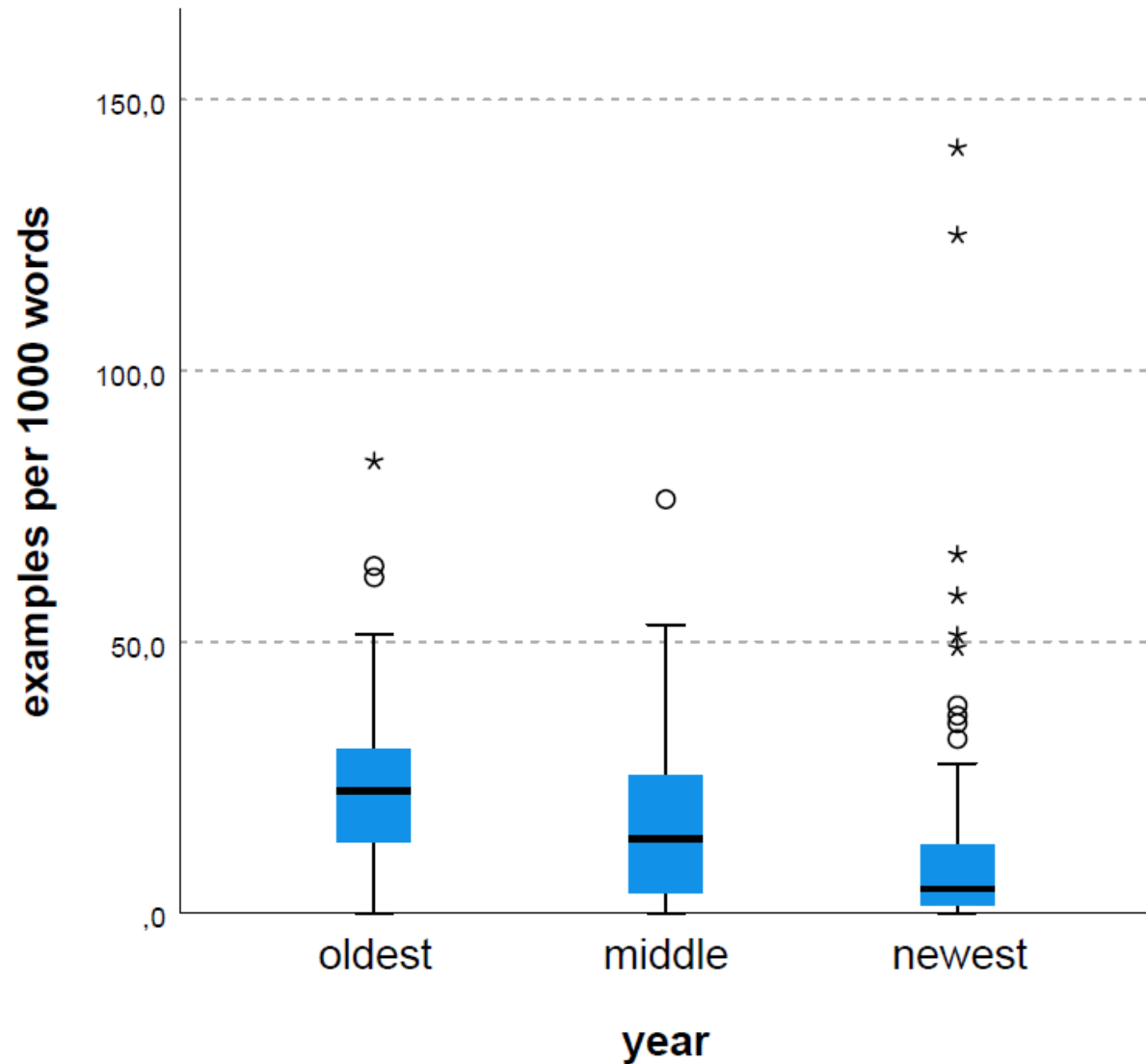


all differences significant

$\rho = .338$ (moderate)
P = .000

H(2) = 11.8
P = .003

year (all languages)

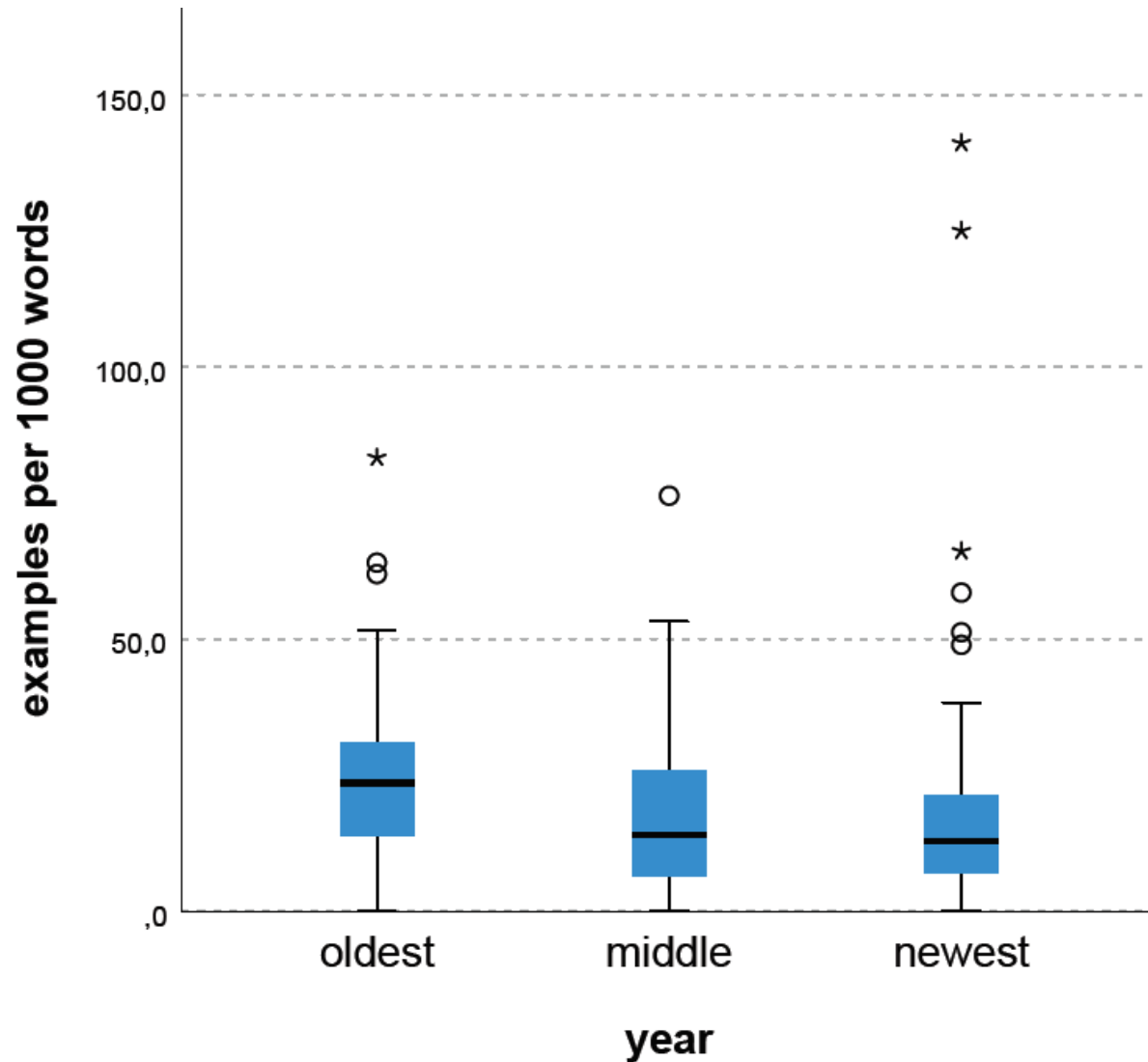


all differences significant

$H(2) = 52.0$

$P = .000$

year (all languages except Rapa Nui)

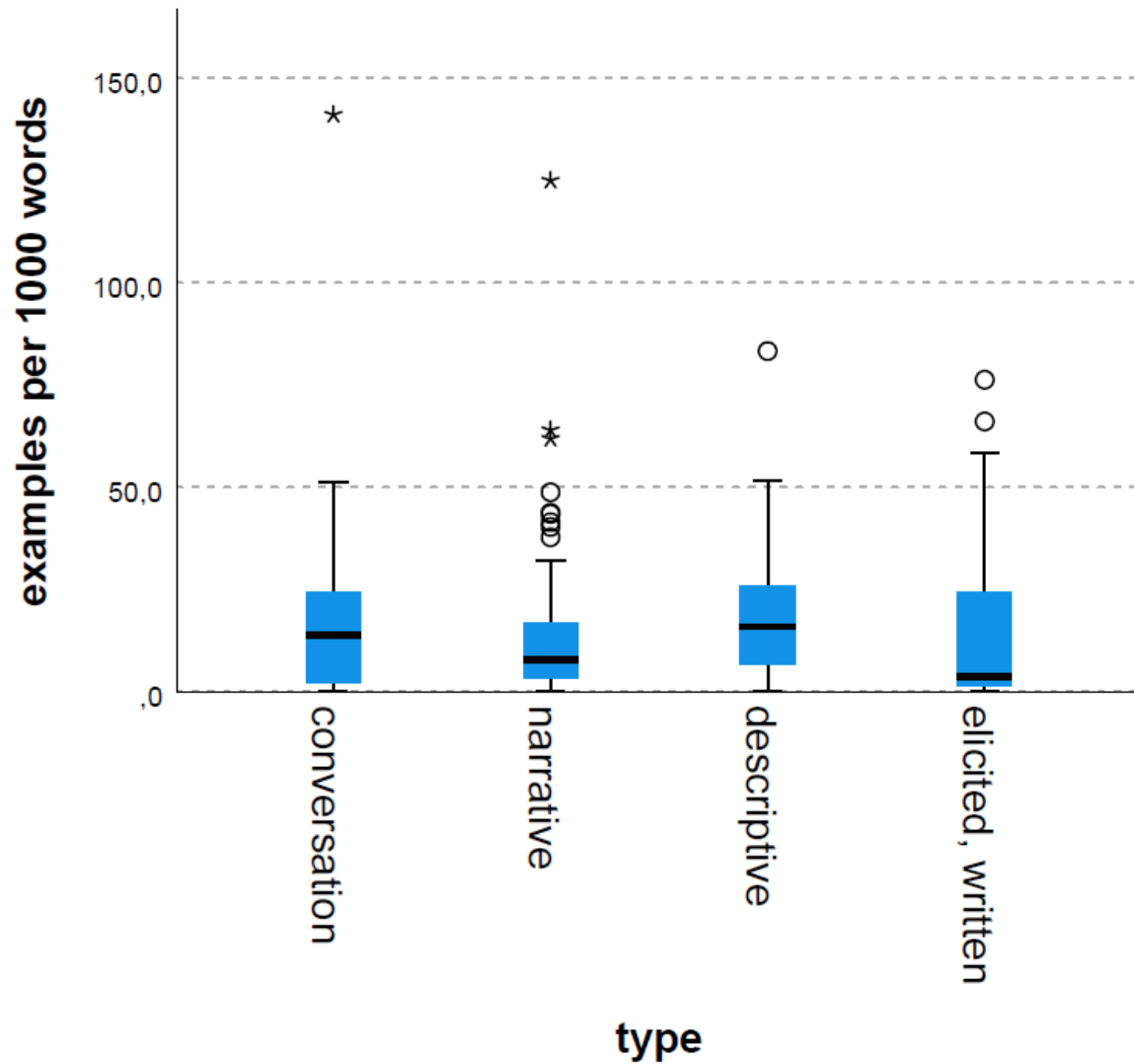


oldest > middle = newest

$H(2) = 12.5$

$P = .002$

type of text (all languages)



elicited < descriptive
narrative < descriptive

$H(3) = 9.335$
 $P = .025$

Conclusion

- no recording is useless
- annotate well!
- focus on good speakers and audio quality
- descriptive recordings are most useful

How do others use my grammar?

How do ~~others~~ use ~~my~~ grammar?

How do typologists use grammars?

Not considered

- language/areal specialists
- language users
- language revitalisers/planners

What is a good example?

questionnaire to LingTyp

20 responses

- short or long?
- genre?
- deciding factor when two equally good examples?
- find yourself in corpus?

Five principles

- 1) clarity
- 2) brevity
- 3) context
- 4) accessibility
- 5) good translation

1) clarity

- use consistent glosses
- follow conventions

ma=bon kiun=bon se bo-t
3SG=COM wife.3SG.POSS=COM IAM go-T
'He and his wife have gone.'

2) brevity

- start with short (“vanilla”) example
 - no room for confusion
- don’t shun elicitation

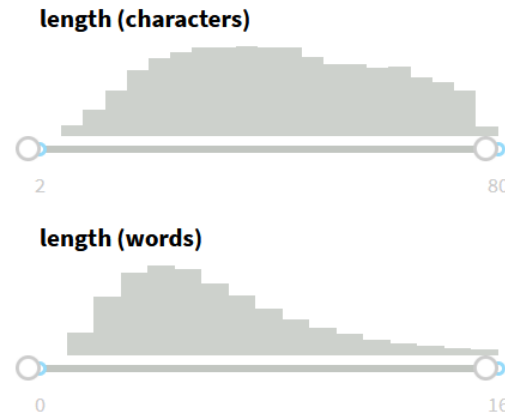
opa [som karuok] me mambara-n kon ladan kerkap
earlier person three TOP stand-N one shirt red

‘Earlier, there were three people standing, one [had] a red shirt.’

2) brevity: practice

IMT vault: all LangSci book examples (<https://imtvault.xitio.de/>)

- 40 characters
- 5 words



«This can probably be interpreted as the result of negotiating the competing motivations of naturalness, exhaustiveness and succinctness.» - Sebastian Nordhoff

3) context

- provide extralinguistic context
- facilitates interpretation

Sudah **diangkat** *barang* sama orang.

PFCT **di-**lift thing accompany person

[Landing at airport, arriving late at conveyor belt, passenger is worried]

‘The things may have already be taken by someone.’

David Gil – The prefixes di- and N- in Malay

Lena Terhart – A grammar of Paunake

Tamsin Donaldson – A grammar of Ngiyampaa

Anthony C. Woodbury (2007). On thick translation in linguistic documentation. In Peter K. Austin (ed.) Language Documentation and Description, vol 4. London: SOAS. pp. 120-135

4) accessibility

- add a tag
- make tag informative

(2) *bal se sor=at koraru*
dog IAM fish=OBJ bite
'The dog has bitten the fish.'

[stim2_3:45]

(3) *bal na*
dog eat
'The dog eats.'

[stim1_0:52]

4) accessibility

- make archive easy to navigate
 - informative titles
 - metadata
- update your email
- choose searchable archive
- think about language! (regional/international)

5) good translations

- use idiomatic sparsely
- add literal at phrase/sentence level

b. *məpanə* *məčáda* *layrik pahəllí*
mə -pa -nə mə -čá -tə layrik pa -hən -í
3P -male -AGN 3P -small -LOC book read -CAUS -NHYP
father to son book cause to read
'The father had the book read through his son (to someone).'

c. *Ramnə* *Shyamda* *líhəlləmmí*
Ram -nə Shyam -tə lí -hən -ləm -í
Ram -AGN Shyam -LOC narrate -CAUS -EVD -NHYP
Ram to Shyam caused to narrate
'Ram had (a story) narrated through Shyam.'

discussion point: replicability

- be forthcoming about alternative analyses
- code for certainty?
- «thick analysis»
- gloss ≠ truth

Language	Family	Voice marker		Person marker		Prob.
		Gloss	Form	Gloss	Form	
Soninke	Mande	DETR	-i	3PL	i=	possible
Mandinka	Mande	ANTIP, REFL	í	3PL	i=	possible
Mandinka	Mande	ANTIP, REFL	í	2SG	í=	unlikely
Krongo	Kadugli-Krongo	ANTIP	-tì	3.INAN.OBL	-tí	probable
Krongo	Kadugli-Krongo	ANTIP	-tì	1SG.NOM	-tí	unlikely
Krongo	Kadugli-Krongo	ANTIP	-Àkú	3SG.F	àakù	unlikely
Koyraboro Senni	Songhay	ANTIP	-a	3SG	a	unlikely

Table 3: Antipassive-person overlaps in Africa

discussion point: replicability

- be forthcoming about alternative analyses
- code for certainty?
- «thick analysis»
- gloss ≠ truth

P.QI 1 10.A.i.6–8 ● ἀγγελος εφοδῆ μαρτυροςγο[υκα] τεν
ματτλᾱ αυλοσιδολ̄ τα[ρου] πετρῶσιδαλ̄ Δουᾶρᾱ.

aggelos ηod-in marturos-gou-ka ten mat-t-ila
angel Lord-GEN martyr-PL-ACC 3PL.GEN afflict-NMLZ-DAT
aul-os-ij-ol-∅ tarou petiros-idal
save-PFV-PLACT-PST1.DET-NOM 3SG Peter-COM
dou-ar-a
exist-PST1-PRED

‘The angel of the Lord who saved the martyrs from their affliction stayed himself with Peter.’

