



University of  
Zurich<sup>UZH</sup>

Psycholinguistics Research Unit

---

# The role of complexity in the acquisition of morphology

Sabine Stoll

# Overview of the talk

---

- main question in acquisition research
- what role does complexity play in the acquisition of grammatical categories?
- 2 cases studies on the acquisition of nouns and verbs
  - Russian
  - Chintang
- conclusions

# Main questions in language acquisition research?

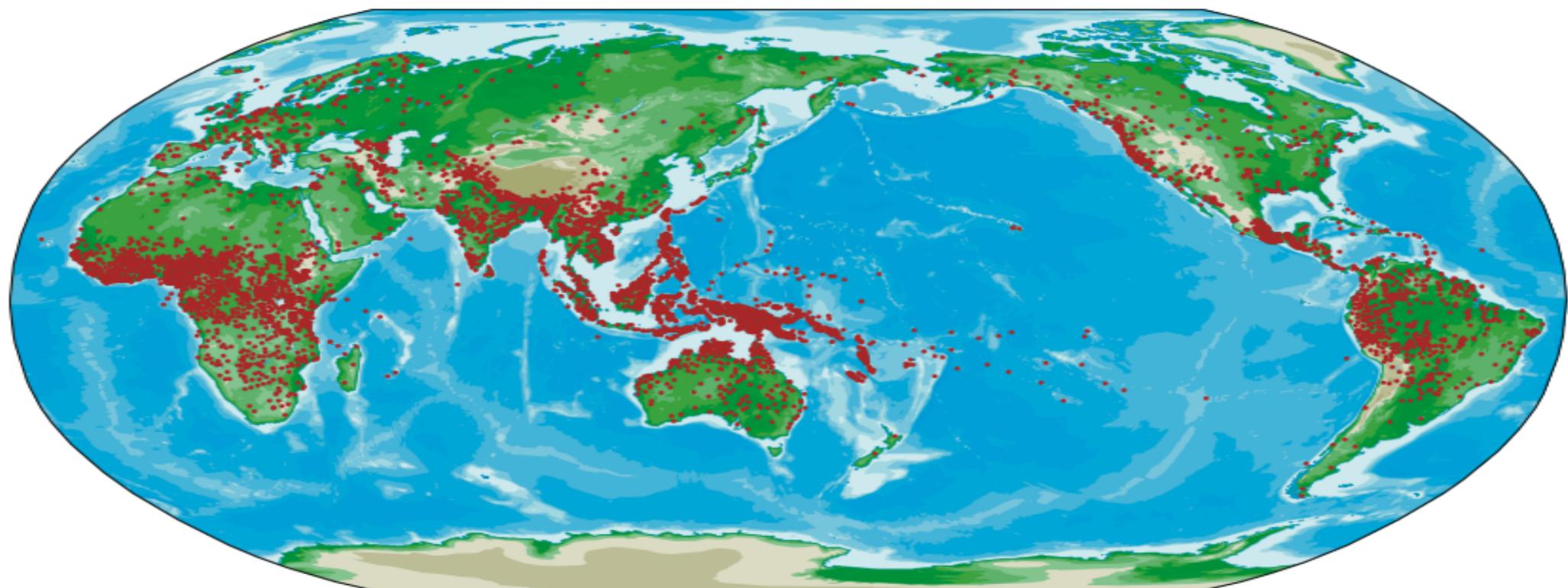
---

- Are there universal strategies in language acquisition?
- If there are different preferences, do they depend on the structure of the language?
- What are the factors responsible for the order of acquisition (frequency, saliency, transparency etc.)?
- Challenge: typological variation

# Challenges for the learner: typological variation

---

- Children are able to learn any of the approx. 7000 spoken languages or hundreds (??) of sign languages used today, provided they grow up in the appropriate social context.



# Challenges for the learner: typological variation

- Extreme variation in inflectional morphology in the languages of the world ranging from isolating/analytic (e.g. English, Vietnamese) to polysynthetic (e.g. Inuktitut, Chintang)

khi tôi đến nhà bạn tôi, chúng tôi bắt đầu làm bài.

when I come house friend I Plural I begin do lesson

*When I came to my friend's house, we began to do lessons.*

[Comrie 1989](#) 43

- *tuntussuqatarniksaitengqiggtuq*
- tuntu -ssur -qatar -ni -ksaite -ngqiggte -uq
- reindeer -hunt -FUT -say -NEG -again -3SG:IND
- 'He had not yet said again that he was going to hunt reindeer.'

# Challenges for the learner: analytic languages

---

Present tense

I meet

you meet

he/she meets

we meet

you meet

they meet

Past tense

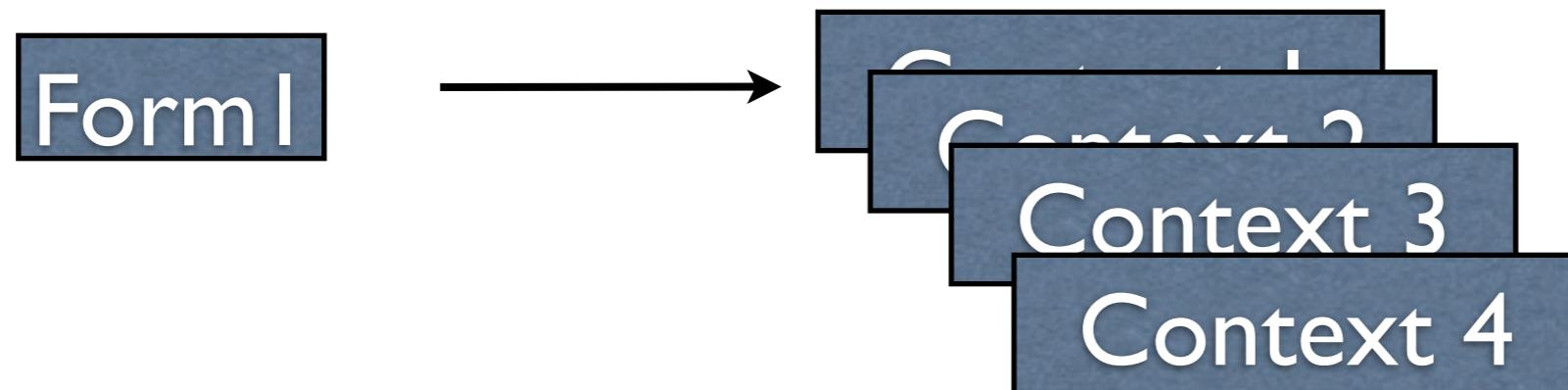
met

==> 3 synthetic verb forms (meet, meets, met)

# Challenges for the learner: analytic languages

---

- Generalisation procedure in analytic languages



# Challenges for the learner: synthetic languages

## Verb paradigm of the verb *tupma* ‘to meet’

	1s	1di	1pi	1de	1pe	2s	2d	2p	3s	3ns	intransitive
1s						<i>tupna?</i> <i>tupna?</i> <i>tupna?</i> <i>tupnehē</i> <i>matupyoknehē</i>	<i>tupna?</i> <i>tupna?</i> <i>tupna?</i> <i>tupnace</i> <i>matupyoknace</i>	<i>tupna?</i> <i>tupna?</i> <i>tupna?</i> <i>tupnaihē</i> <i>matupyoknaihē</i>	<i>tubukuj</i> <i>tubukuj</i> <i>tubuhē</i> <i>matupyoktuhē</i>	<i>tubukujciŋ</i> <i>tubukujciŋ</i> <i>tubuhēcihē</i> <i>matupyokturjcihē</i>	<i>tupma?</i> <i>tupma?</i> <i>tubehē</i> <i>matupyoktehē</i>
1di									<i>tupcoko</i> <i>tupcoko</i> <i>tubace</i> <i>matupyoktace</i>	<i>tubumcum</i> <i>tubumcum</i> <i>tubumcumhē</i> <i>matupyoktumcumhē</i>	<i>tupceke</i> <i>tupcekeniŋ</i> <i>tubace</i> <i>matupyoktace</i>
1pi									<i>tubukum</i> <i>tubukum</i> <i>tubumhē</i> <i>matupyoktumhē</i>	<i>tubiki</i> <i>tubikiniŋ</i> <i>tubihē</i> <i>matupyoktihē</i>	
1de									<i>tupcokoŋa</i> <i>tupcokoŋanij</i> <i>tubacehē</i> <i>matupyoktacehē</i>	<i>tubumcumma</i> <i>tubumcummanij</i> <i>tubumcummehē</i> <i>matupyoktumcummehē</i>	<i>tupcekeŋa</i> <i>tupcekejanij</i> <i>tubacehē</i> <i>matupyoktacehē</i>
1pe									<i>tubukumma</i> <i>tubukummanij</i> <i>tubummehē</i> <i>matupyoktummehē</i>	<i>tubikiŋa</i> <i>tubikitjanij</i> <i>tubiehē</i> <i>matupyoktiehē</i>	
2s	<i>atupma?</i> <i>atupma?</i> <i>atubehē</i> <i>amatupyoktehē</i>					<i>amatupceke</i> <i>amatupcekeniŋ</i> <i>amatubace</i> <i>amamatupyoktace</i>	<i>amatupno</i> <i>amatupnikniŋ</i> <i>amatube</i> <i>amamatupyokte</i>	<i>tupna?</i> <i>tupna?</i> <i>tupnacyehē</i> <i>matupyoknacyehē</i>	<i>atuboko</i> <i>atuboko</i> <i>atube</i> <i>amatupyokte</i>	<i>atubukuce</i> <i>atubukucenij</i> <i>atubuce</i> <i>amatupyoktuce</i>	<i>atupno</i> <i>atupnikniŋ</i> <i>atube</i> <i>amatupyokte</i>
2d	<i>atupma?</i> <i>atupma?</i> <i>atubacihē</i> <i>amatupyoktaŋcihē</i>								<i>atupcoko</i> <i>atupcoko</i> <i>atubace</i> <i>amatupyoktace</i>	<i>atubumcum</i> <i>atubumcum</i> <i>atubumcumhē</i> <i>amatupyoktumcumhē</i>	<i>atupceke</i> <i>atupcekeniŋ</i> <i>atubace</i> <i>amatupyoktace</i>
2p	<i>atupma?</i> <i>atupma?</i> <i>atubajnihē</i> <i>amatupyoktaŋnihē</i>								<i>atubukum</i> <i>atubukum</i> <i>atubumhē</i> <i>amatupyoktumhē</i>	<i>atubiki</i> <i>atubikiniŋ</i> <i>atubihē</i> <i>amatupyoktihē</i>	
3s	<i>utupma?</i> <i>utupma?</i> <i>utubehē</i> <i>umatupyoktehē</i>					<i>maitupceke</i> <i>maitupcekeniŋ</i> <i>maitubace</i> <i>maimatupyoktace</i>	<i>maitupno</i> <i>maitupnikniŋ</i> <i>maitube</i> <i>maimatupyokte</i>	<i>natupno</i> <i>natupnikniŋ</i> <i>natube</i> <i>namatupyokte</i>	<i>tuboko</i> <i>tuboko</i> <i>tube</i> <i>matupyokte</i>	<i>tubukuce</i> <i>tubukucenij</i> <i>tubuce</i> <i>matupyoktuce</i>	<i>tupno</i> <i>tupnikniŋ</i> <i>tube</i> <i>matupyokte</i>
3d	<i>utupma?</i> <i>utupma?</i> <i>utubacihē</i> <i>amatupyoktaŋcihē</i>								<i>utupcoko</i> <i>utupcoko</i> <i>utubace</i> <i>amatupyoktace</i>	<i>utubukuce</i> <i>utubukucenij</i> <i>utubace</i> <i>amatupyoktace</i>	<i>utupceke</i> <i>utupcekeniŋ</i> <i>utubace</i> <i>amatupyoktace</i>
3p	<i>utupma?</i> <i>utupma?</i> <i>utubajnihē</i> <i>amatupyoktaŋnihē</i>								<i>utuboko</i> <i>utuboko</i> <i>utube</i> <i>umatupyokte</i>	<i>utupno</i> <i>utupnikniŋ</i> <i>utube</i> <i>umatupyokte</i>	

Table 1: Chintang agreement paradigm of the verb *tupma* ‘to meet’, with stem *tup* (identical in all forms) (Vertical axis: subject agreement; horizontal axis: object agreement. Within each cell, the forms denote (in vertical order) nonpast affirmative, nonpast negative, past affirmative, and past negative tenses, respectively. Abbreviations: s ‘singular’, d ‘dual’, p ‘plural’, ns ‘nonsingular (dual or plural)’, i ‘inclusive of addressee’, e ‘exclusive of addressee’, 1-3 denote persons.)

# Variables determining complexity

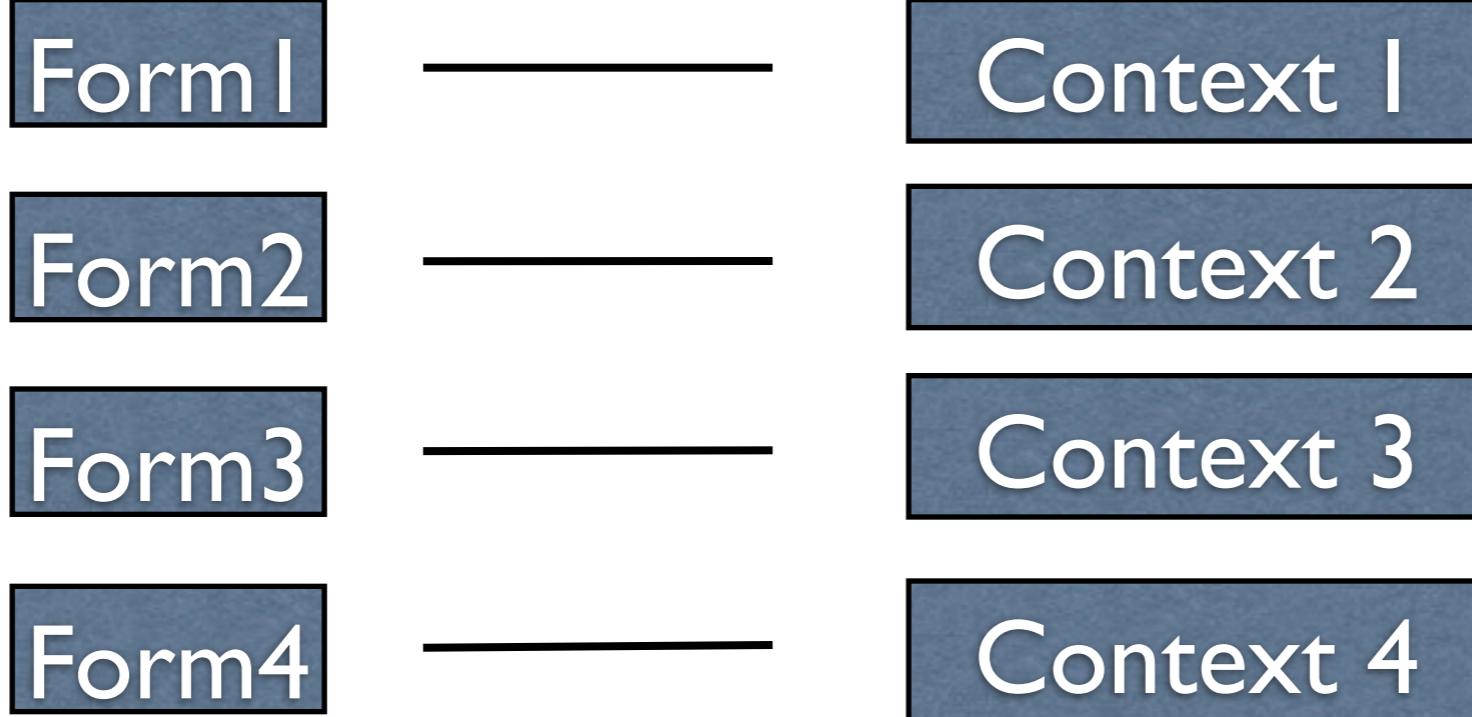
---

- number of morphological markers
- type of morphological markers
- location of marking
- regularity of marking
- phonological fusion
- presence or absence of conjugation/declension classes
- degree of homonymy
- presence of a default and number of exceptions
- transparency of morpheme boundaries

# Challenges for the learner

---

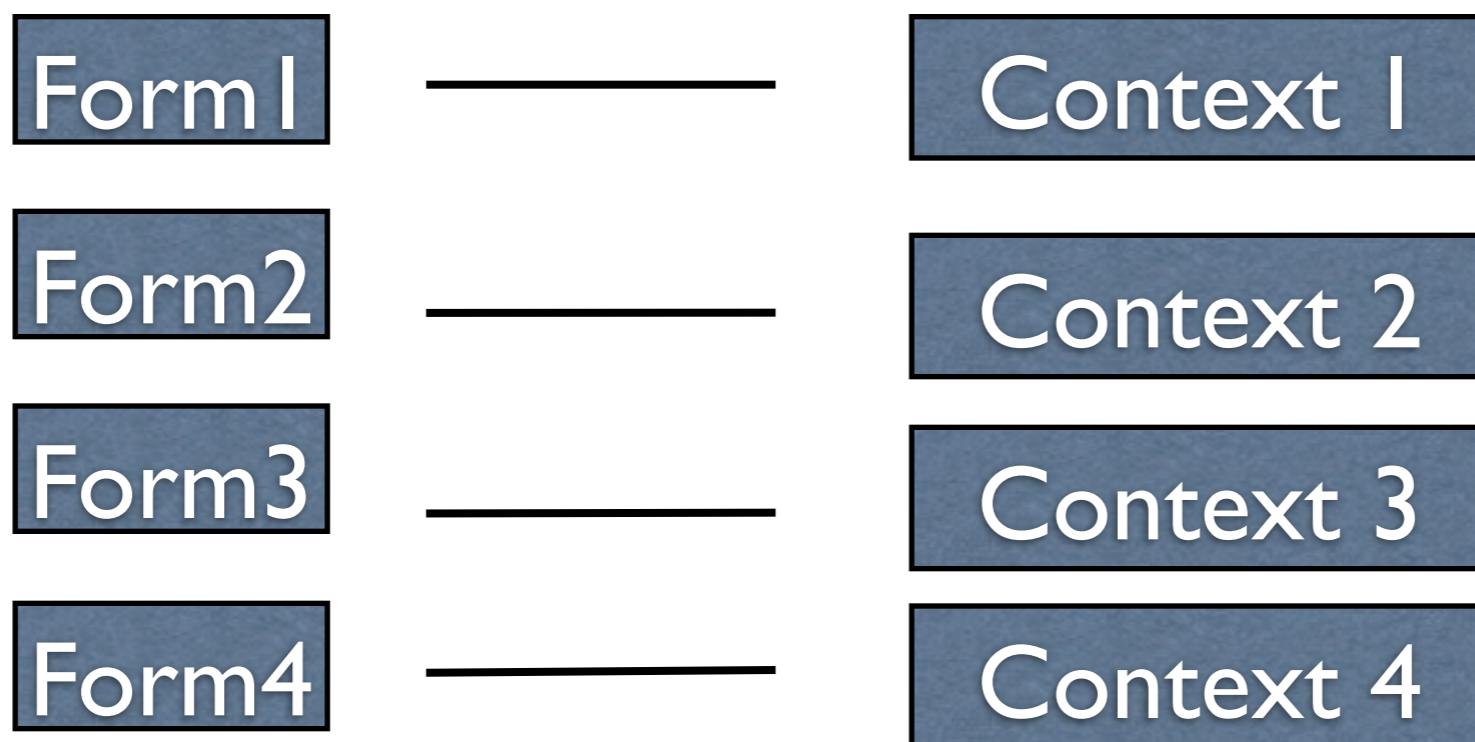
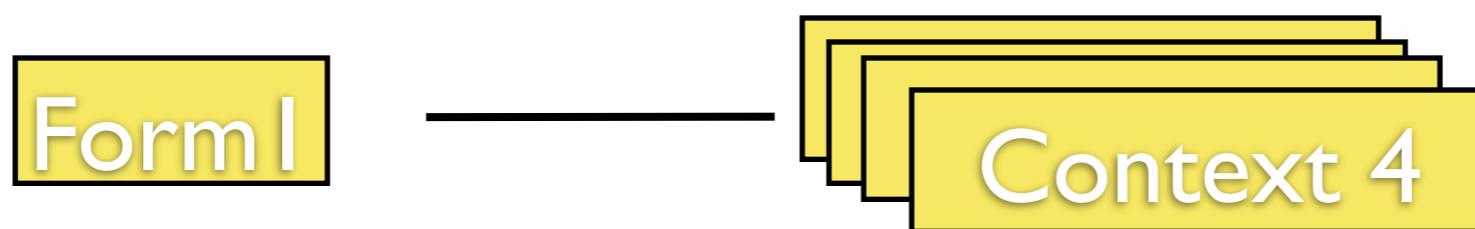
- Generalisation procedure in synthetic languages



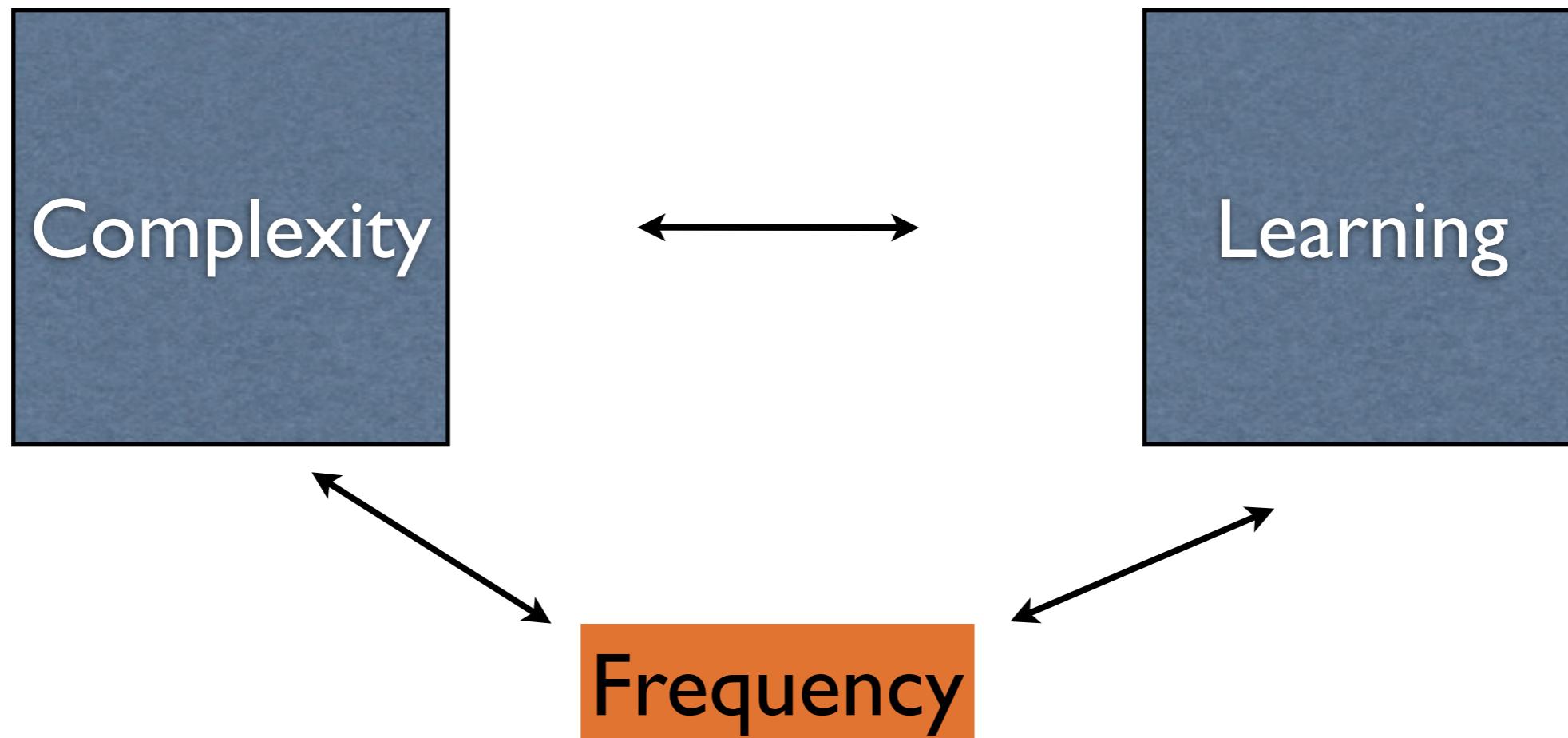
# Challenges for the learner

---

- Generalisation procedure in analytic vs. synthetic languages



# Challenges for the learner: typological variation



# Research Questions

---

- Do children first learn nouns or verbs?
- Can the meaning and the form of these words tell us something about acquisition strategies and cognitive development?
- Is there a common development in how children learn words or what kind of variation (qualitative and/or quantitative) do we find?
- Are there differences across languages? If yes, what are the relevant factors for these differences?

Collaboration with:  
Balthasar Bickel, University of Zurich,  
Elena Lieven, University of Manchester

# Early vocabulary composition

---

- Cognitive universals or language specific processes?
- Hypothesis about a cognitive universal: ‘Natural Partition Hypothesis’ (Gentner, 1982):
  - Distinction between nouns and verbs and preference for nouns in early vocabulary is due to:  
‘a preexisting perceptual-conceptual distinction between concrete concepts such as persons or things and predicative concepts of activity, change-of-state, or causal relations.’ (Gentner, 1982: 301).

# Early vocabulary composition

## Results from the literature

	Questionnaire (mostly MCDI)	Naturalistic Data
English	N>V	N>V
Italian	N>V	N=V
Korean	N>V	N=V
Mandarin	N>V	V>N
Ngas	V>N	not tested
Tzeltal	not tested	N=V

# Early vocabulary composition

---

- Result: In some languages nouns are learned first in others verbs
- Factors relevant for these differences:
  - distributions in the input
  - ‘teaching’ strategies
  - morphological complexity

# Early vocabulary composition

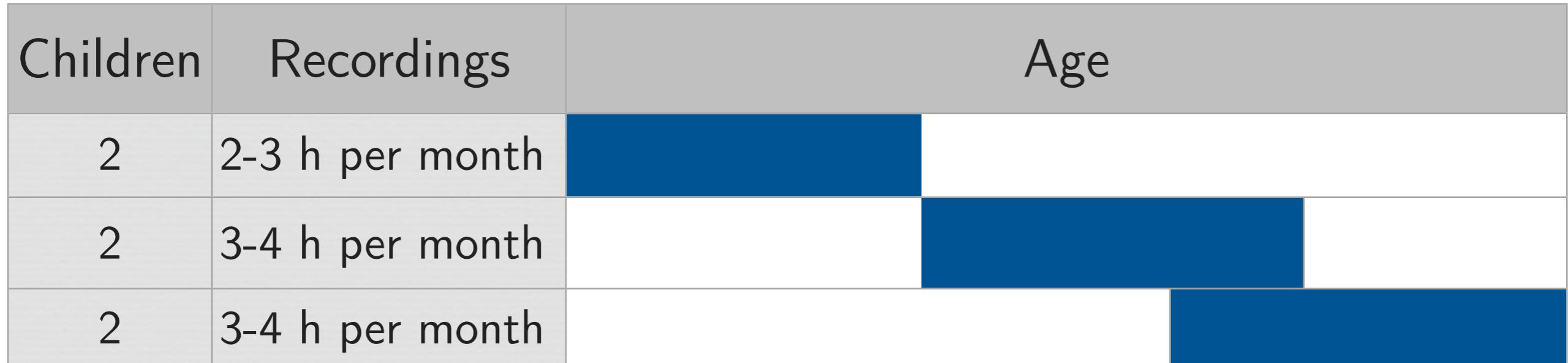
---

- Role of morphology for the early preference of either nouns and verbs?
- Repetition factor of one and the same form. The more forms a word has, the less each individual form is repeated. Thus, we expect that morphology in combination with frequency is an important factor.
- To check the role of morphology: 2 case studies with languages with complex morphology: Chintang and Russian.

# Chintang corpus (with E. Lieven)



# Chintang language acquisition corpus (with E. Lieven)



Age of the target children      0;6    1;0    1;6    2;0    2;6    3;0    3;6    4;0    4;6

- 220 h of recording
- close to 1 mio words transcribed
- most translated and annotated



महान शहीदका  
र्थ आए रहने !

जातिगत भ्रतोग आत्म  
स्वाक्षर्त युवक  
सिवायक

प्रति लागि सविचान सत्ताका  
धनकुटा

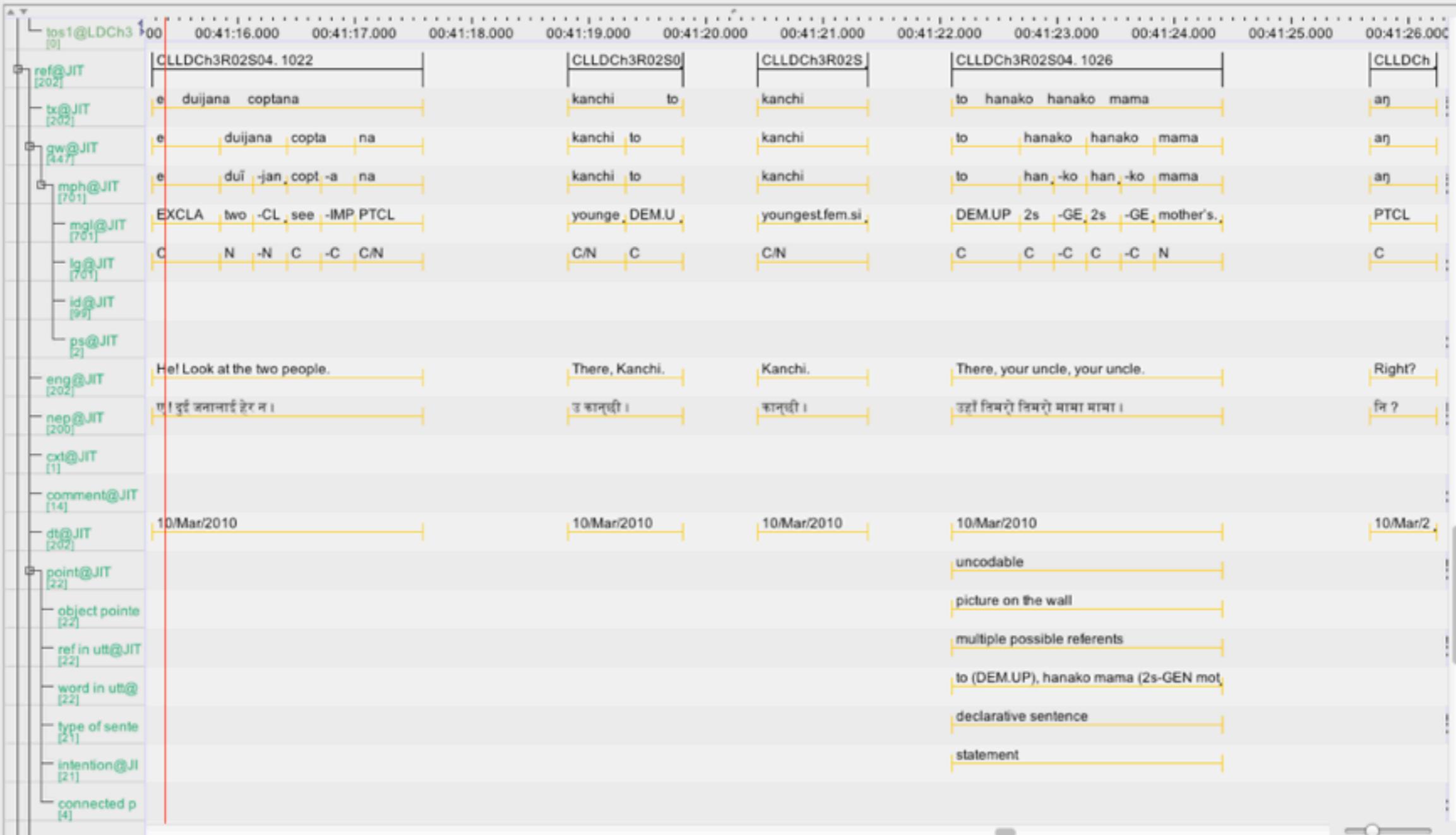




tx@JIT		Begin Time	End Time	Duration
>	112 kristina	00:41:14.010	00:41:14.800	00:00:00.790
▶	113 e duijana coptana	00:41:15.275	00:41:17.580	00:00:02.305
	114 kanchi to	00:41:18.824	00:41:19.800	00:00:00.976
	115 kanchi	00:41:20.446	00:41:21.380	00:00:00.934
	116 to hanako hanako mama	00:41:22.109	00:41:24.410	00:00:02.301
	117 arj	00:41:25.670	00:41:26.240	00:00:00.570
	118 bago salo kha	00:41:27.540	00:41:28.720	00:00:01.180
	119 topattigo	00:41:29.905	00:41:30.580	00:00:00.675
	120 ena bago salo kha	00:41:30.660	00:41:31.870	00:00:01.210
	121 hale cekta	00:41:32.888	00:41:33.860	00:00:00.972

00:41:15.380

Selection: 00:41:27.540 - 00:41:28.720 1180

 DS  S  ←  →  ↓  ↑  Selection Mode  Loop Mode 


# Russian longitudinal corpus

	Age	Recordings (1h each)
Child 1	1;8 - 5;4	138
Child 2	1;11 - 4;0	67
Child 3	1;4 - 5;6	131
Child 4	1;3 - 4;11	125
Child 5	3;1 - 6;7	95
<b>Sum</b>		<b>556</b>

more than 2.1 Mio words

# Nouns and verbs in Chintang and Russian

	noun forms	verb forms
Russian	12	ca. 100
Chintang	ca. 30	ca. 1900

# Nouns vs. verbs in Chintang and Russian

---

- Noun to verb ratio in Chintang and Russian
- **Expectation about child surrounding speech:** Adults will produce more verbs than nouns, i.e. the noun/verb ratio will be small.
- **Expectation about child speech:** Children will start out with more nouns than we would expect from the input.
- **Reason:** extremely complex verb morphology
- **Hypothesis:** First, in early development, children will prefer nouns because of the complex verb morphology. As soon as they get more productive with verb morphology they will adapt to the adult pattern and use more verbs than nouns.

# Nouns vs. verbs in Chintang

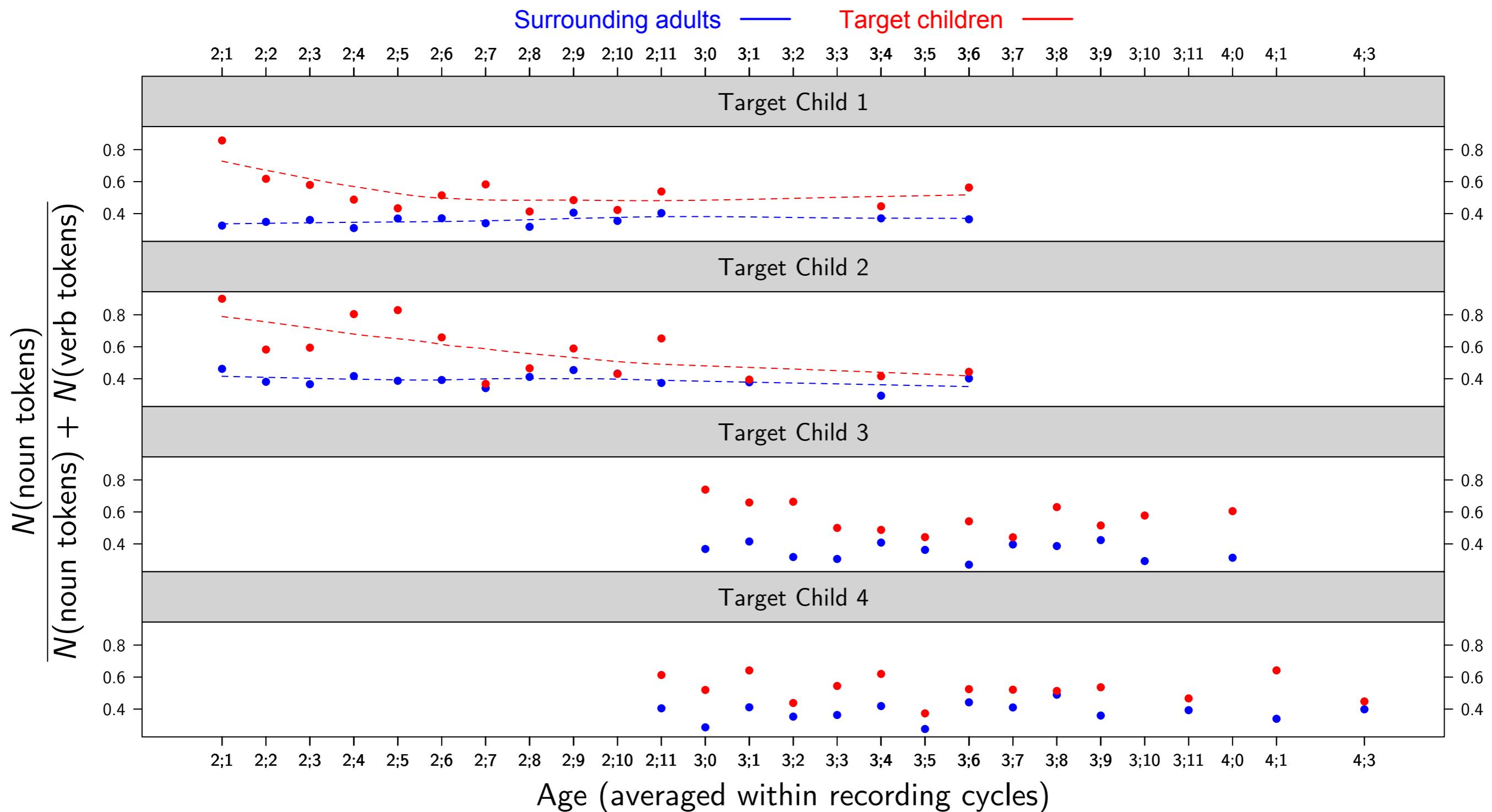
---

- Noun-to-verb ratio both for types and tokens:

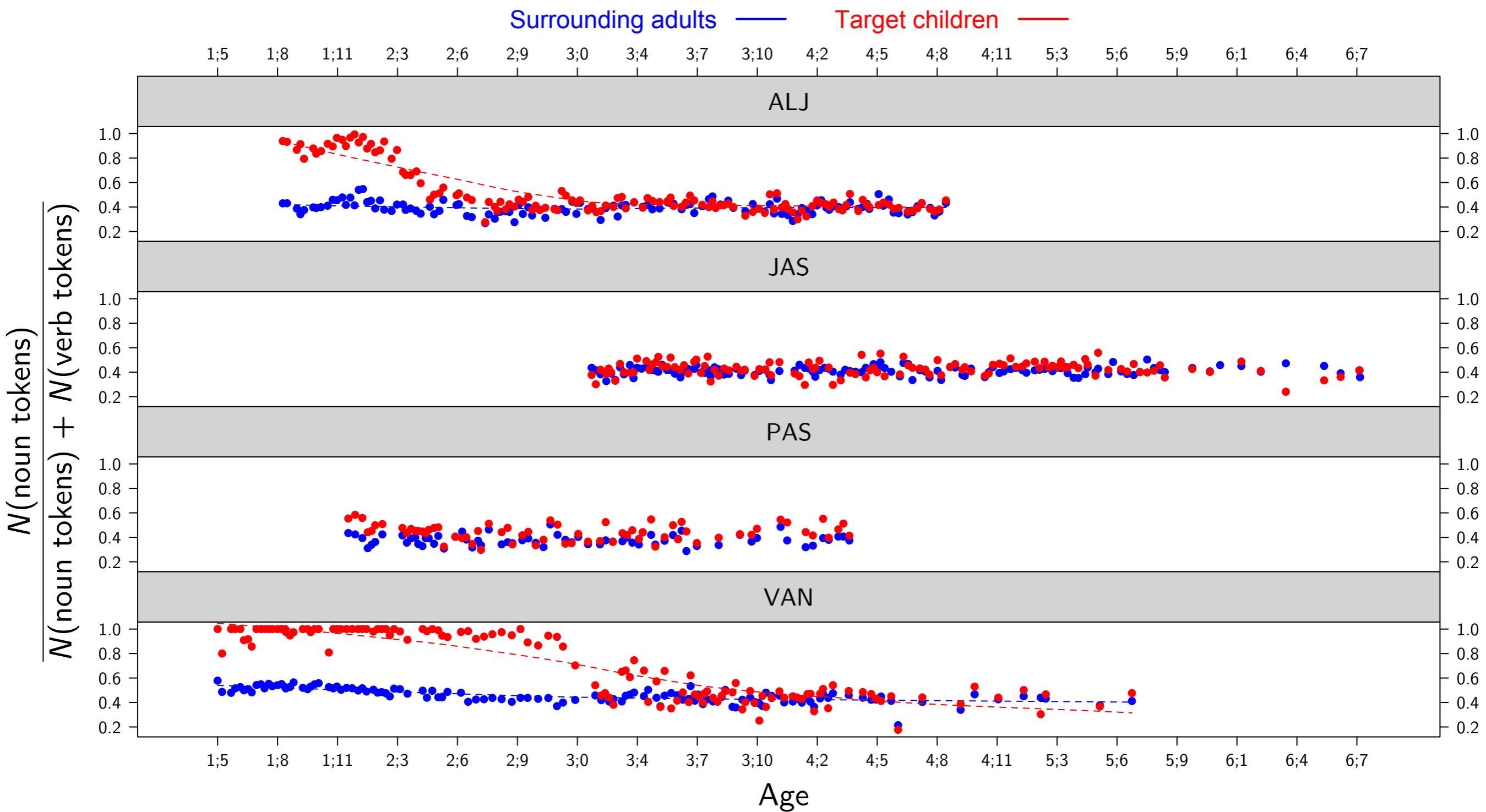
$$R_{N/V} = \frac{N(\text{nouns})}{N(\text{nouns}) + N(\text{verbs})}$$

- Noun-to-to verb ratios were measured both for types and tokens for each child per recording cycle and for the surrounding adults (pooled).

# Noun to verb ratio in Chintang



# Noun to verb ratio in Russian

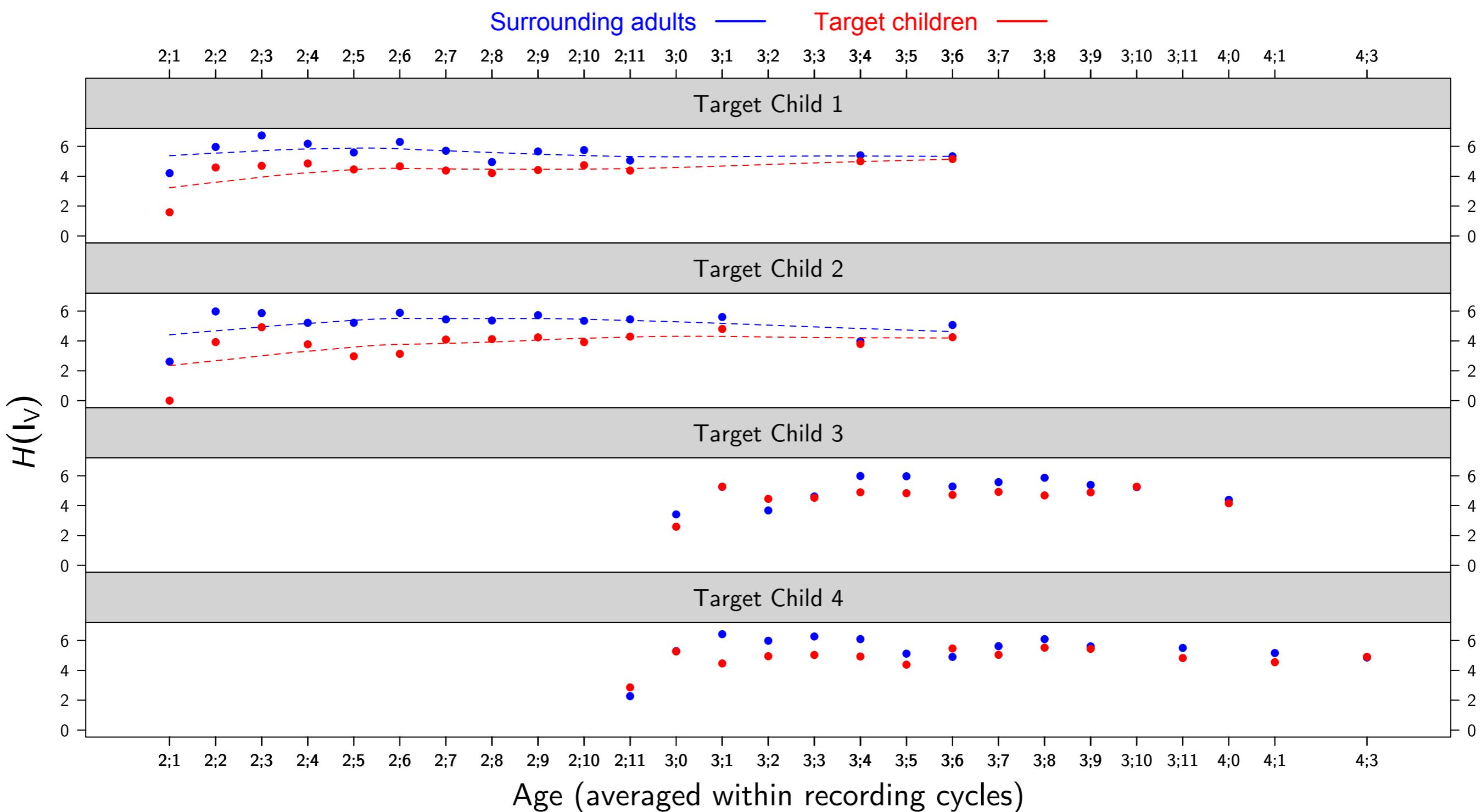


# Nouns vs. verbs in Chintang: morphological complexity

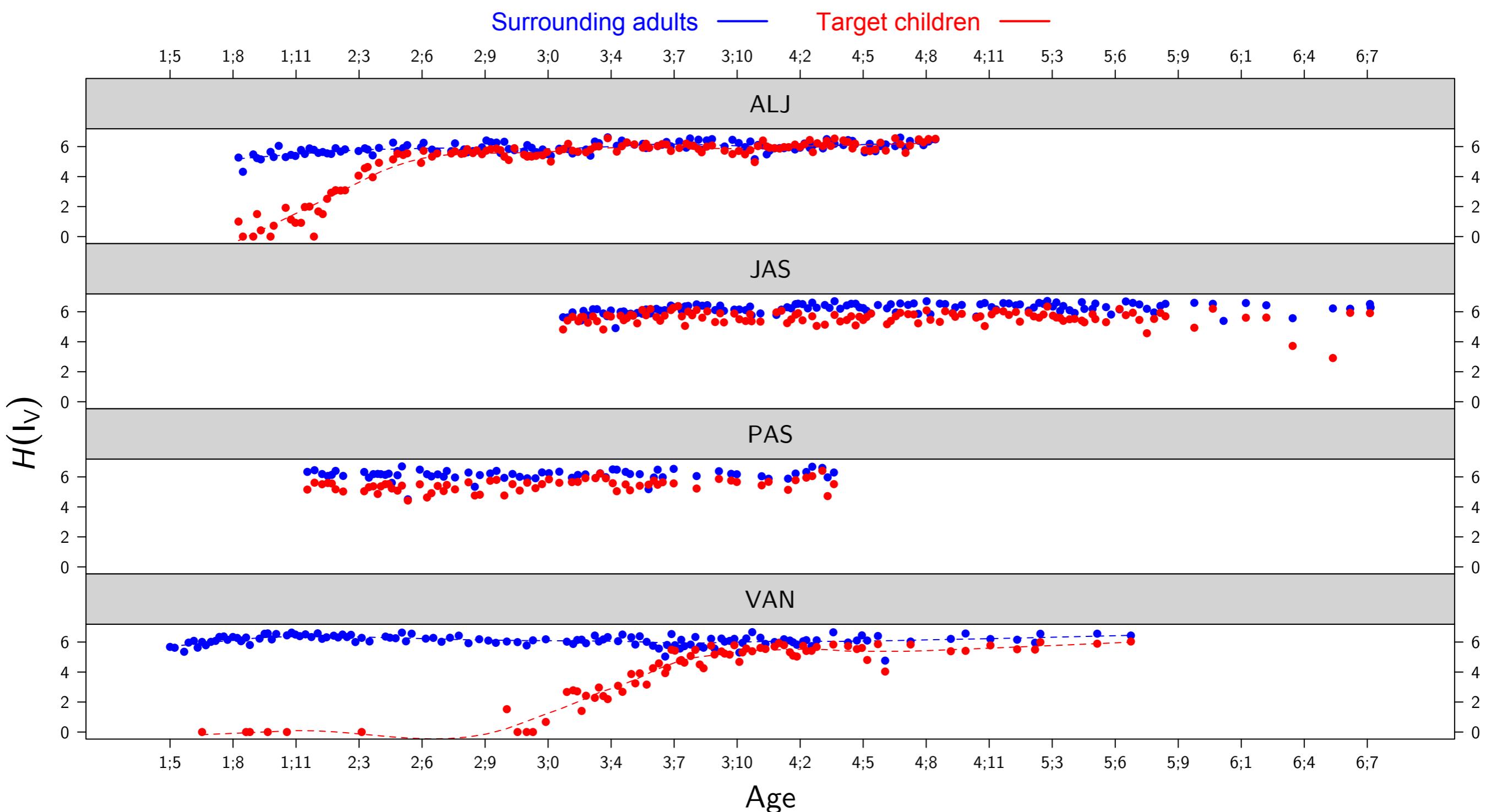
---

- Measure morphological complexity in the use of nouns and verbs
- Entropy captures the intuition that a child masters a paradigm the better the more distinct forms she uses and the less she prefers some over others.
- A measure of uncertainty in a system. The more various forms used the higher the entropy.

# Development of verbal morphology in Chintang



# Development of verbal morphology in Russian

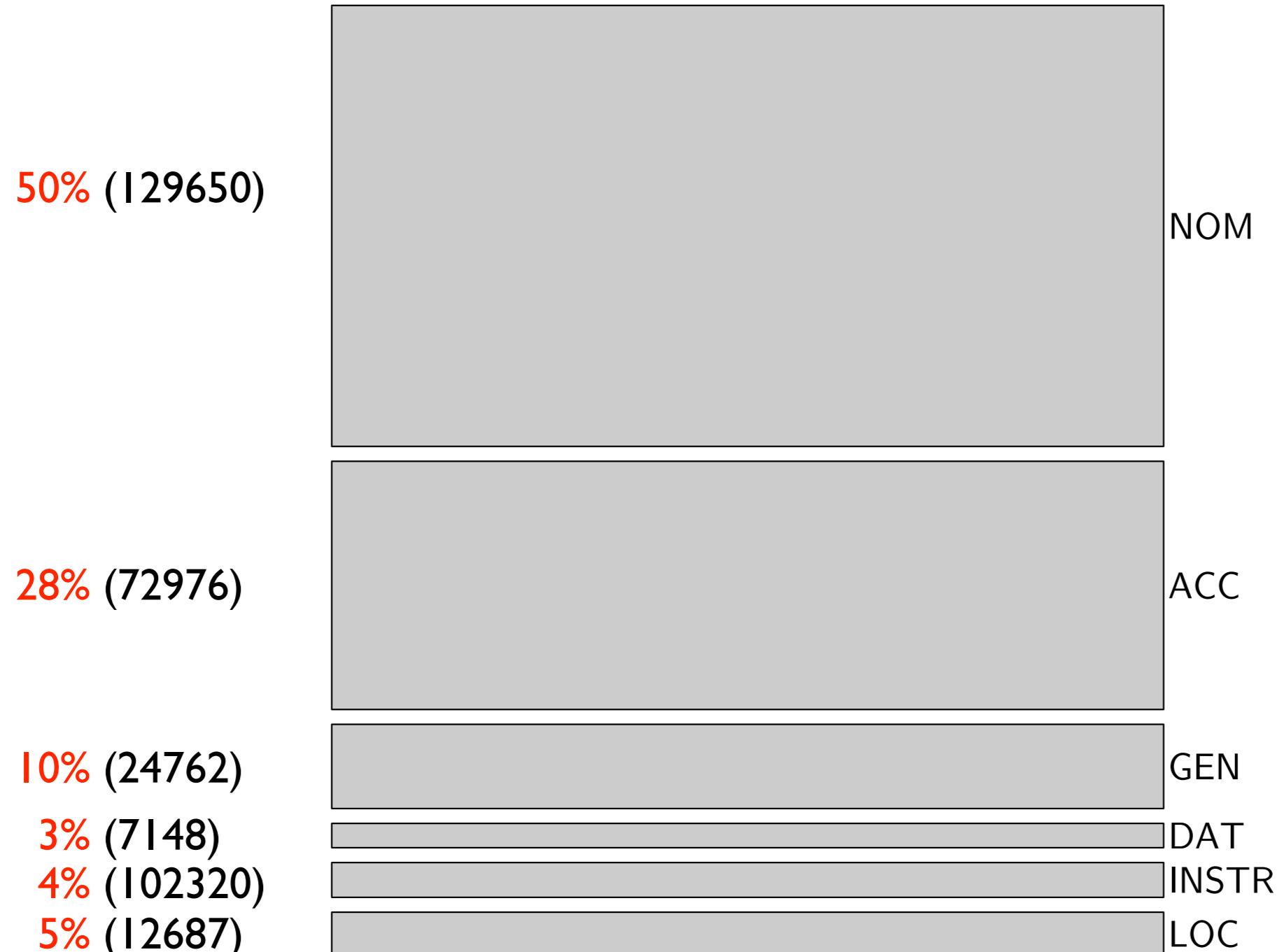


# Role of morphological complexity

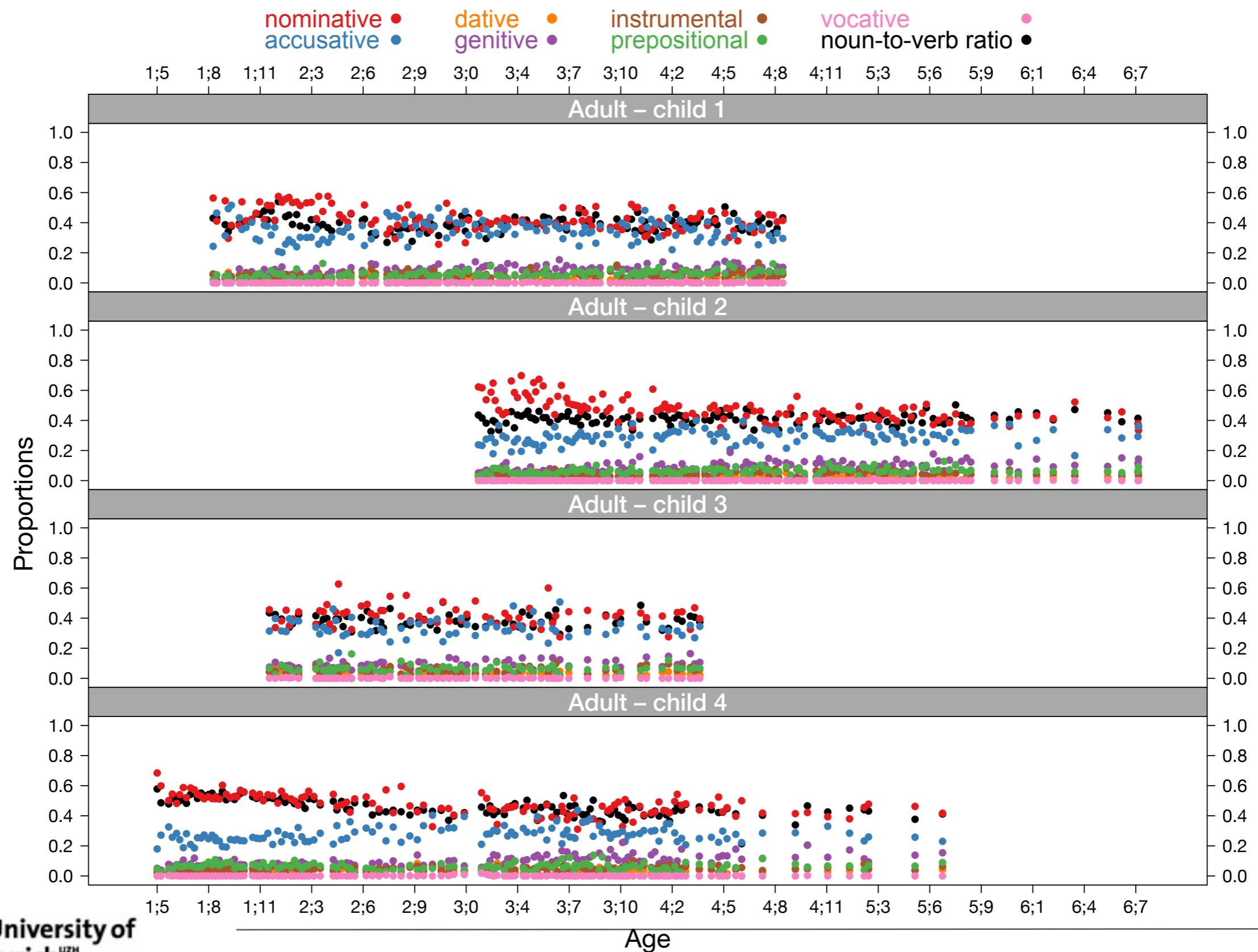
---

- Complexity in verbal morphology (at least for Chintang and Russian) seems to be an important factor in the development of the noun-to-verb distributions.
- Only when the children are similar in their productivity of verbal morphology to the adults they also show a similar noun-to-verb ratio.
- Now: closer look at the development of noun morphology (here in Russian only)

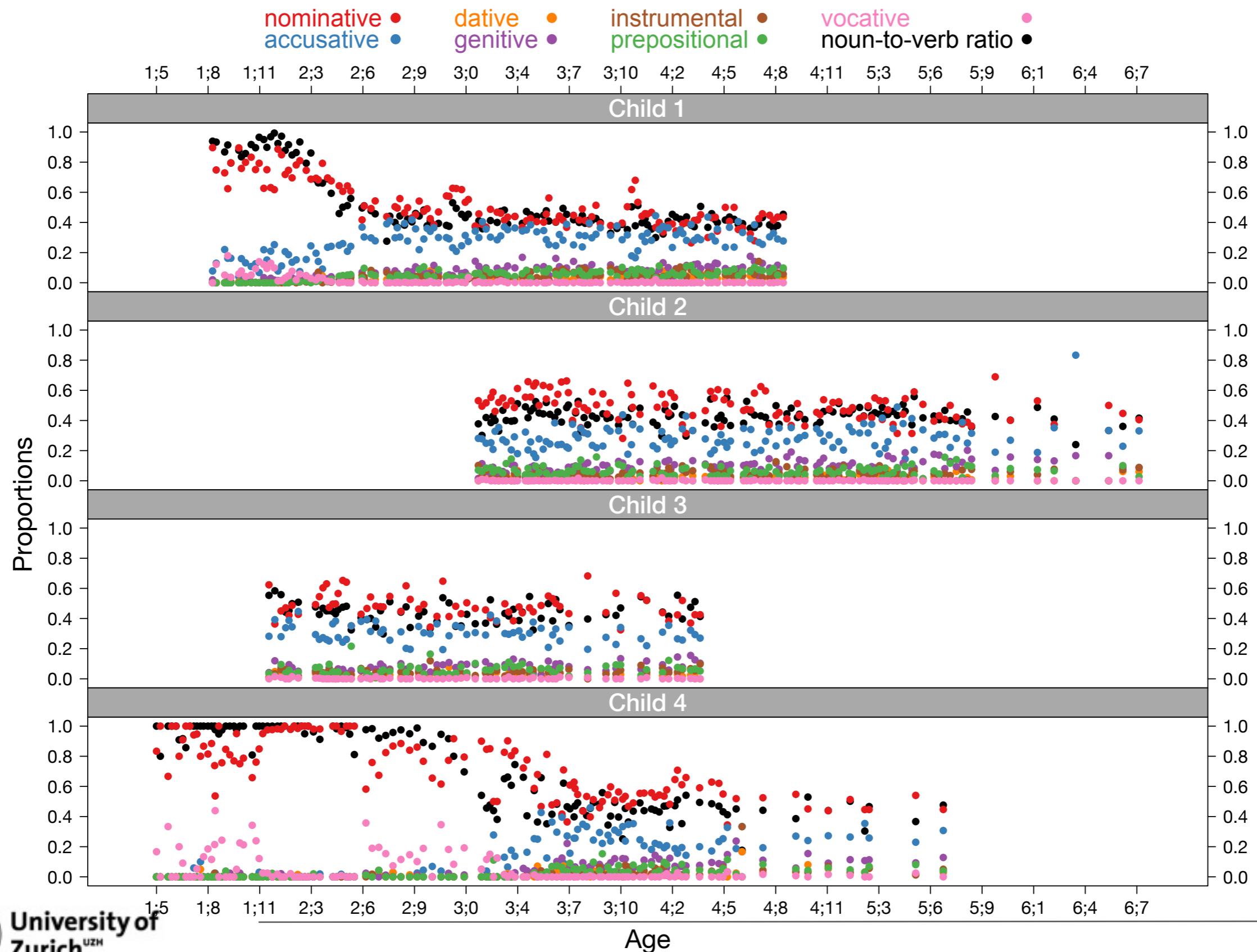
# Distribution of cases in the input of Russian nouns



# Distribution of cases in the surrounding adults



# Distribution of cases in children



# Conclusions

---

- The complexity of a morphological subsystems influences the acquisition process.
- Nouns are preferred in early acquisition both in Chintang and Russian because they are morphologically less complex.
- In Russian, also in noun morphology children clearly focus on the least complex and most frequent case, i.e. nominative in early acquisition.
- As soon as other cases and verbal morphology gets more varied the noun to verb ratio adapts to the distributions of the surrounding adults.
- Finegrained analyses of correlations to describe development necessary

Thank you very much for your attention!

# Verbal morphology of Chintang

	1s	1di	1pi	1de	1pe	2s	2d	2p	3s	3ns	intransitive
1s						tupna?ā tupna?āniŋ tupnehē matupyoknehē	tupna?āce tupna?āceniŋ tupnace matupyoknace	tupna?āni tupna?āninij tupnane matupyoknanihē	tubukuj tubukunij tubuhē matupyoktuhē	tubukujcuŋ tubukucuŋniŋ tubuŋcihē matupyoktuŋcihē	tupma?ā tupma?āniŋ tubehē matupyoktehē
1di									tupcoko tupcokonij tubace matupyoktace	tubumcum tubumcumniŋ	tupceke tupcekeniŋ tubace matupyoktace
1pi									tubukum tubukumnim tubumhē matupyoktumhē	tubumcumhē matupyoktumcumhē	tubiki tubikiniŋ tubihē matupyoktihē
1de						tupna?ānciyā tupna?ānciyāniŋ tupnanciyehē matupyoknanciyehē	tupcokorja tupcokorjanij tubacehē matupyoktacehē	tubumcumma tubumcummanij	tupcekeja tupcekejanij tubacehē matupyoktacehē		
1pe							tubukumma tubukummanij tubummehē matupyoktummehē	tubumcummehē matupyoktumcummehē	tubikija tubikijanij tubiehē matupyoktihē		
2s	atupma?ā atupma?āniŋ atubehē amatupyoktehē								atuboko atubokonij atube amatupyokte	atubukuce atubukucenij atubuce amatupyoktuce	atupno atupnikniŋ atube amatupyokte
2d	atupma?anciy atupma?anciŋniŋ atubacihē amatupyoktacyihē					amatupceke amatupcekenij amatubace amamatupyoktace	amatupno amatupnikniŋ amatube amamatupyokte	atupcoko atupcokonij atubace amatupyoktace	atubumcum atubumcumniŋ	atupceke atupcekenij atubace amatupyoktace	
2p	atupma?anij atupma?aninij atubajnihē amatupyoktajnihē								atubukum atubukumnim atubumhē amatupyoktumhē	atubukumhē amatupyoktumcumhē	atubiki atubikiniŋ atubihē amatupyoktihē
3s	utupma?ā utupma?āniŋ utubehē umatupyoktehē								tuboko tubokonij tube matupyokte	tubukuce tubukucenij tubuce matupyoktuce	tupno tupnikniŋ tube matupyokte
3d	utupma?anciy utupma?anciŋniŋ utubacihē amatupyoktacyihē	maitupceke maitupcekenij maitubace maimatupyoktace	maitupno maitupnikniŋ maitube maimatupyokte	matupceke matupcekenij matubace mamatupyoktace	matupno matupnikniŋ matube mamatupyokte	natupno natupnikniŋ natube namatupyokte	natupceke natupcekenij natubace namatupyoktace	natubiki natubikiniŋ natubihē namatupyoktihē	utupcoko utupcokonij utubace umatupyoktace	utubukuce utubukucenij utubuce umatupyoktace	utupceke utupcekenij utubace umatupyoktace
3p	utupma?anij utupma?aninij utubajnihē amatupyoktajnihē								utuboko utubokonij utube umatupyokte	utubuce umatupyoktuce	utupno utupnikniŋ utube umatupyokte

Table 1: Chintang agreement paradigm of the verb *tupma* ‘to meet’, with stem *tup* (identical in all forms) (Vertical axis: subject agreement; horizontal axis: object agreement. Within each cell, the forms denote (in vertical order) nonpast affirmative, nonpast negative, past affirmative, and past negative tenses, respectively. Abbreviations: s ‘singular’, d ‘dual’, p ‘plural’, ns ‘nonsingular (dual or plural)’, i ‘inclusive of addressee’, e ‘exclusive of addressee’, 1-3 denote persons.)

# Challenges for the learner: synthetic languages

## Verb paradigm of the verb *tupma* ‘to meet’

	1s	1di	1pi	1de	1pe	2s	2d	2p	3s	3ns	intransitive
1s						<i>tupna?</i> <i>tupna?</i> <i>tupna?</i> <i>tupnehē</i> <i>matupyoknehē</i>	<i>tupna?</i> <i>tupna?</i> <i>tupna?</i> <i>tupnace</i> <i>matupyoknace</i>	<i>tupna?</i> <i>tupna?</i> <i>tupna?</i> <i>tupnaihē</i> <i>matupyoknaihē</i>	<i>tubukuj</i> <i>tubukuj</i> <i>tubuhē</i> <i>matupyoktuhē</i>	<i>tubukujciŋ</i> <i>tubukujciŋ</i> <i>tubuhēcihē</i> <i>matupyokturjcihē</i>	<i>tupma?</i> <i>tupma?</i> <i>tubehē</i> <i>matupyoktehē</i>
1di									<i>tupcoko</i> <i>tupcoko</i> <i>tubace</i> <i>matupyoktace</i>	<i>tubumcum</i> <i>tubumcum</i> <i>tubumcumhē</i> <i>matupyoktumcumhē</i>	<i>tupceke</i> <i>tupcekeniŋ</i> <i>tubace</i> <i>matupyoktace</i>
1pi									<i>tubukum</i> <i>tubukum</i> <i>tubumhē</i> <i>matupyoktumhē</i>	<i>tubiki</i> <i>tubikiniŋ</i> <i>tubihē</i> <i>matupyoktihē</i>	
1de									<i>tupcokoŋa</i> <i>tupcokoŋanij</i> <i>tubacehē</i> <i>matupyoktacehē</i>	<i>tubumcumma</i> <i>tubumcummanij</i> <i>tubumcummehē</i> <i>matupyoktumcummehē</i>	<i>tupcekeŋa</i> <i>tupcekejanij</i> <i>tubacehē</i> <i>matupyoktacehē</i>
1pe									<i>tubukumma</i> <i>tubukummanij</i> <i>tubummehē</i> <i>matupyoktummehē</i>	<i>tubikiŋa</i> <i>tubikitjanij</i> <i>tubiehē</i> <i>matupyoktiehē</i>	
2s	<i>atupma?</i> <i>atupma?</i> <i>atubehē</i> <i>amatupyoktehē</i>					<i>amatupceke</i> <i>amatupcekeniŋ</i> <i>amatubace</i> <i>amamatupyoktace</i>	<i>amatupno</i> <i>amatupnikniŋ</i> <i>amatube</i> <i>amamatupyokte</i>	<i>tupna?</i> <i>tupna?</i> <i>tupnacyehē</i> <i>matupyoknacyehē</i>	<i>atuboko</i> <i>atuboko</i> <i>atube</i> <i>amatupyokte</i>	<i>atubukuce</i> <i>atubukucenij</i> <i>atubuce</i> <i>amatupyoktuce</i>	<i>atupno</i> <i>atupnikniŋ</i> <i>atube</i> <i>amatupyokte</i>
2d	<i>atupma?</i> <i>atupma?</i> <i>atubacihē</i> <i>amatupyoktaŋcihē</i>								<i>atupcoko</i> <i>atupcoko</i> <i>atubace</i> <i>amatupyoktace</i>	<i>atubumcum</i> <i>atubumcum</i> <i>atubumcumhē</i> <i>amatupyoktumcumhē</i>	<i>atupceke</i> <i>atupcekeniŋ</i> <i>atubace</i> <i>amatupyoktace</i>
2p	<i>atupma?</i> <i>atupma?</i> <i>atubajnihē</i> <i>amatupyoktaŋnihē</i>								<i>atubukum</i> <i>atubukum</i> <i>atubumhē</i> <i>amatupyoktumhē</i>	<i>atubiki</i> <i>atubikiniŋ</i> <i>atubihē</i> <i>amatupyoktihē</i>	
3s	<i>utupma?</i> <i>utupma?</i> <i>utubehē</i> <i>umatupyoktehē</i>					<i>maitupceke</i> <i>maitupcekeniŋ</i> <i>maitubace</i> <i>maimatupyoktace</i>	<i>maitupno</i> <i>maitupnikniŋ</i> <i>maitube</i> <i>maimatupyokte</i>	<i>matupno</i> <i>matupnikniŋ</i> <i>matube</i> <i>mamatupyokte</i>	<i>tuboko</i> <i>tuboko</i> <i>tube</i> <i>matupyokte</i>	<i>tubukuce</i> <i>tubukucenij</i> <i>tubuce</i> <i>matupyoktuce</i>	<i>tupno</i> <i>tupnikniŋ</i> <i>tube</i> <i>matupyokte</i>
3d	<i>utupma?</i> <i>utupma?</i> <i>utubacihē</i> <i>amatupyoktaŋcihē</i>								<i>utupcoko</i> <i>utupcoko</i> <i>utubace</i> <i>amatupyoktace</i>	<i>utubumcum</i> <i>utubumcum</i> <i>utubumcumhē</i> <i>amatupyoktumcumhē</i>	<i>utupceke</i> <i>utupcekeniŋ</i> <i>utubace</i> <i>amatupyoktace</i>
3p	<i>utupma?</i> <i>utupma?</i> <i>utubajnihē</i> <i>amatupyoktaŋnihē</i>								<i>utuboko</i> <i>utuboko</i> <i>utube</i> <i>umatupyokte</i>	<i>utubukuce</i> <i>utubukucenij</i> <i>utubuce</i> <i>umatupyoktuce</i>	<i>utupno</i> <i>utupnikniŋ</i> <i>utube</i> <i>umatupyokte</i>

Table 1: Chintang agreement paradigm of the verb *tupma* ‘to meet’, with stem *tup* (identical in all forms) (Vertical axis: subject agreement; horizontal axis: object agreement. Within each cell, the forms denote (in vertical order) nonpast affirmative, nonpast negative, past affirmative, and past negative tenses, respectively. Abbreviations: s ‘singular’, d ‘dual’, p ‘plural’, ns ‘nonsingular (dual or plural)’, i ‘inclusive of addressee’, e ‘exclusive of addressee’, 1-3 denote persons.)