Language acquisition in crosslinguistic perspective

Elena Lieven
Sabine Stoll

Max Planck Institute for Evolutionary Anthropology

OVERVIEW

- Day 1: Background, questions, data, methods
- Day 2: Approaches to comparative language acquisition
- Day 3: Naturalistic, experimental and modelling studies across languages
- Day 4: The role of input and cultural context

QUESTIONS

- How do children learn language? How much is innate?
- Are there universals in language acquisition or at least universal strategies?
- Which factors are relevant for the order and time of acquisition (cognitive factors, language specific factors)?

LANGUAGE ACQUISITION IN CROSSLING. PERSPECTIVE

- Crosslinguistic language acquisition: study of the acquisition of individual languages other than English.
- Comparative language acquisition: study of how the acquisition of specific variables compares in different langugages

Ultimate goal --> typological language acquisition

Linguistics

Corpus linguistics

Typology

Comparative Language Acquisition Research

Psycholinguistics

Ethnolinguistics

Developmental Psychology

Psychology

Ethnography

LANGUAGE ACQUISITION THEORIES

NATIVIST THEORIES

- characteristics of grammar are due to innate principles
- universal grammar is innate, language specific principles are innate
- deductive theories
- criterion for theory: descriptive adequacy, simplicity

USAGE-BASED THEORIES

- characteristics of grammar are due to communicative principles
- only general cognitive abilities are innate no language specific structures
- inductive theories
- criterion: learnability

PRELINGUISTIC DEVELOPMENT

Birth - 6 months:

- Recognition of mothers voice (de Casper & Fifer 1980)
- Distinguish native language from other languages (Mehler et al. 1988, Moon et al. 1993)
- Categorical perception of speech sounds (Eimas et al. 1971)
- Recog. of identity of sounds across contexts (Kuhl 1980)

HIGH-AMPLITUDE SUCKING (BIRTH - 4 MONTHS)

Abbildung aus Atkinson et al.(2000). Copyright Harcourt Publishers

1. [ba], [ba] acquisition phase (child learns that she can alter the sounds).

2. Habituation phase

control group: no change in stimulus

test group: change in stimulus

Leipzig Spring School 2008, Lieven & Stoll

PRELINGUISTIC DEVELOPMENT

6 - 12 months:

- Discrimination of phonetic contrasts. Up to approx. 10 months discrimination of all contrasts. Then, only contrasts of native language/s are discriminated. (Werker & Tees 1984)
- By 7 1/2 months children listen longer to familiarized words (Jusczyk & Aslin 1995) within longer sentences.

PRELINGUISTIC DEVELOPMENT

Around 9 months (e.g. Tomasello 2003)

- recogniton of symbols
- pointing
- imitation
- joint attention
- intention reading

Leipzig Spring School 2008, Lieven & Stoll

DEVELOPMENT OF LANGUAGE

- 9 months to 1 year of age children start to use their first words

- strong variation when children start speaking and how they progress.

Table 4.3. Age of acquisition for 10- and 50-word vocabularies in six children

Child	Sex	Lexicon size	
		10 words	50 words
S1	M	1;0	1;5
S2	F	1;1	1;6
S3	M	1;2	1;7
S4	F	1;2	1;8
S5	M	1;4	1;10
S6	F	1;3	1;7

Source: Robb, Bauer, & Tyler 1994:40. Used with permission from Alpha Academic.

LEARNING WORDS

Very different task in different languages, e.g. polysynthetic languages vs. isolating languages

Example (Yup'ik Inuit)

- *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.'

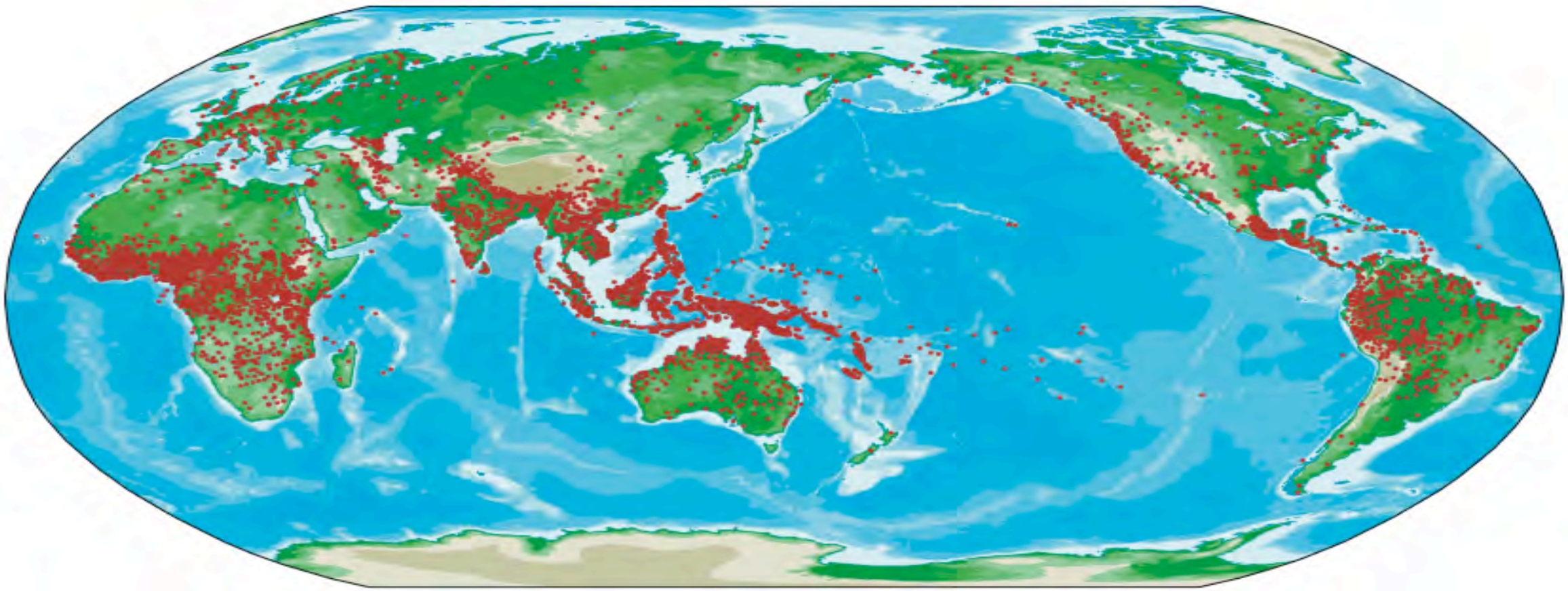
Source: Eliza Orr, cited by Payne, T. 1997a 27–28

COMPARATIVE LANGUAGE ACQUISITION

Basic Idea

If we want to understand the ability of children to learn any language and not only how a specific language is learned (like English or German) we need to look at the acquisition of a wide range of languages from different families and cultures.

LANGUAGES TODAY



VARIATION

Phoneme Inventories (Maddieson, 2005):

1. Consonant inventories
   6 (Rotokas, Papua New Guinea) - 122 (!Xóõ, Southern Khoisan) out of a sample of 562 languages

2. Vowel inventories
   - 2 (Yimas (Papua New Guinea) -14 (German)

WORD FORMATION

Inflectional synthesis of the verb (Bickel & Nichols, 2005)

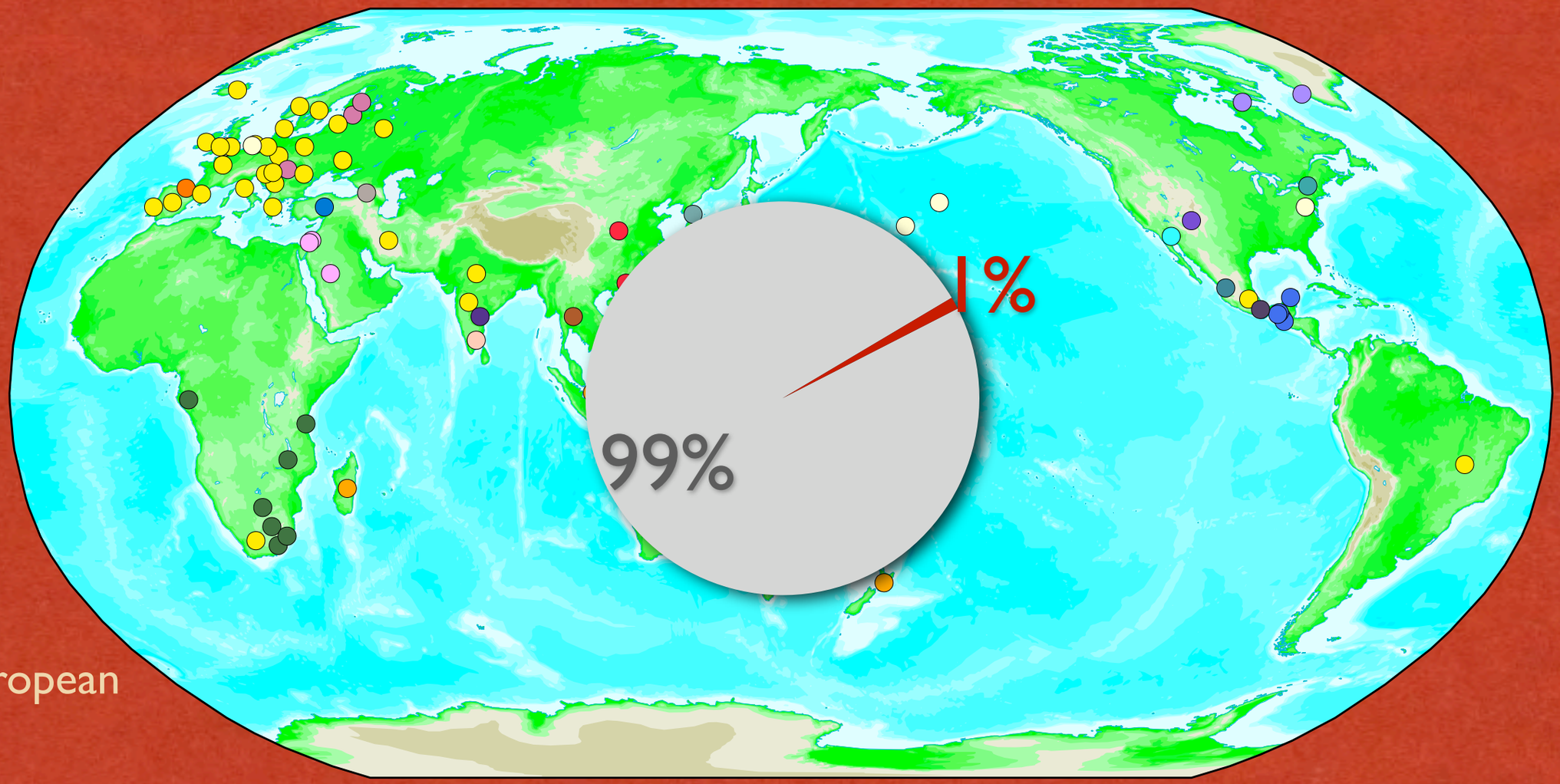
- Degree of synthesis as defined by the number of elements that make up a synthetic verb form
- Large variation form 0 categories per verb form (Vietnamese) to 13 (Koasati)

VARIATION IN WORD ORDER

Word order (Subject, Verb, Object) (Dryer, 2005)

- All 6 logically possible orders are attested
- SOV (Japanese)
- SVO (Mandarin)
- VSO (Irish)
- VOS (Nias, Austronesian)
- OVS (Hixkaryana, Carib, Brazil)
- OSV (Nadëb, Brazil)
- No predominant order

LANGUAGES WITH ACQUISITION STUDIES



- Bantu
- Indo-European
- Semitic
- Maya
- Sinitic

Leipzig Spring School 2008, Lieven & Stoll

DATA WE WORK WITH

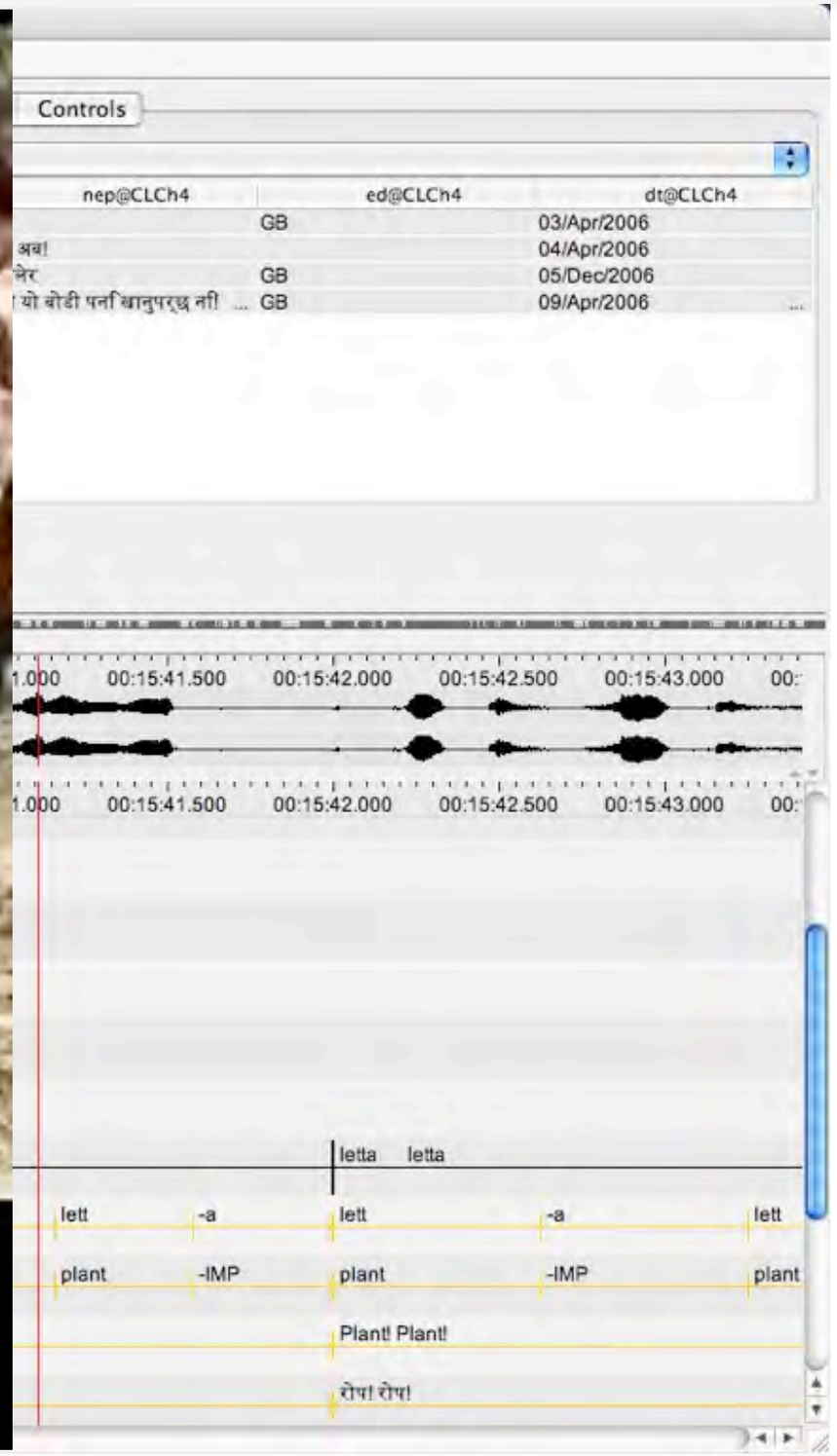
- Diaries
- Experiments
- Longitudinal studies (corpora)

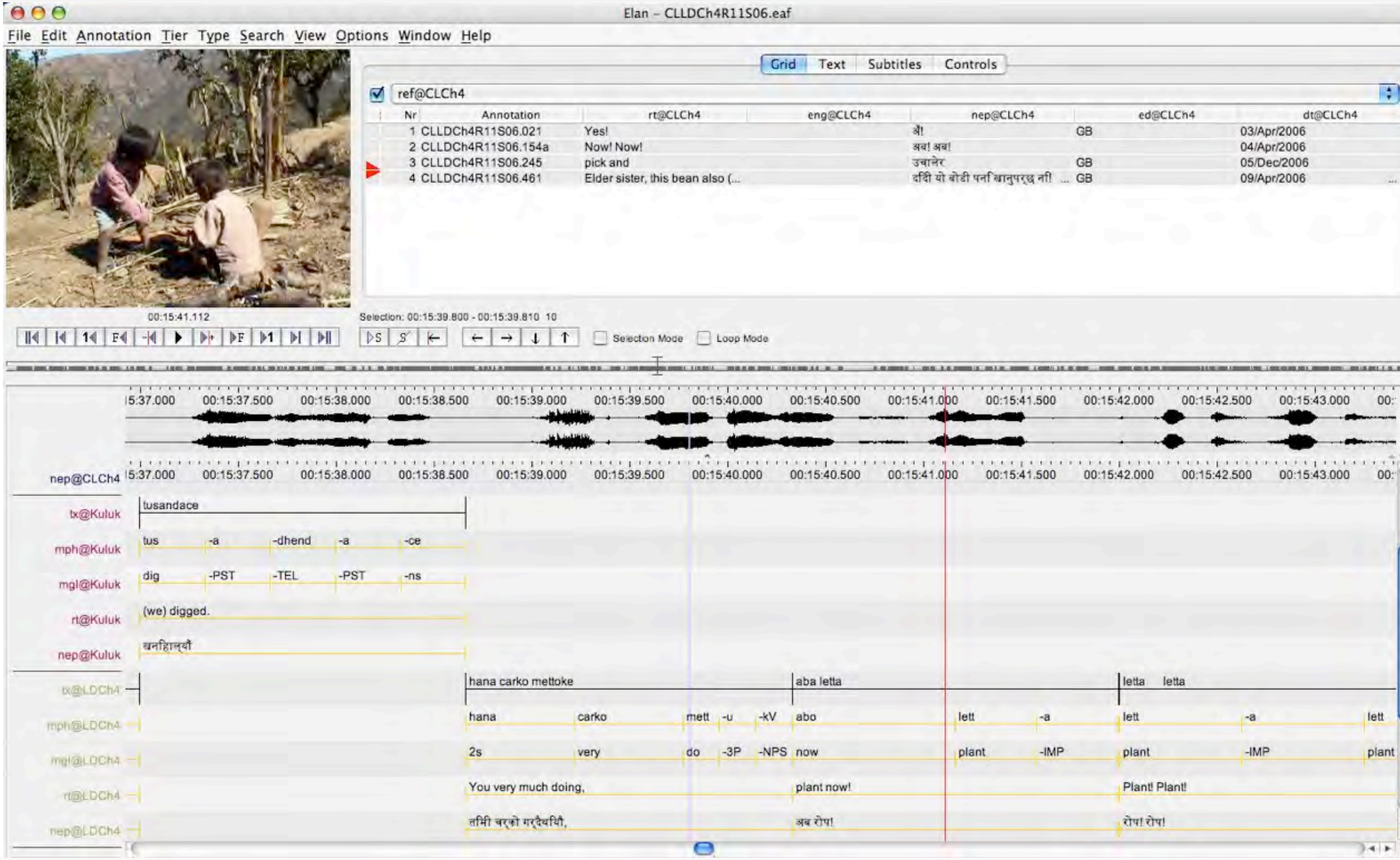
LONGITUDINAL STUDIES

1. Audio or video record the child and her caretakers Transcribe the data.
2. (Translate the data).
3. Tagging of the data (morphological glossing, and parts of speech glossing).
4. Further annotations, if necessary.
5. (Link to video or audio).
6. Analyze

Daten


Ia Kuluke tusande
Look, Kuluke, digged!


Controls
nep@CLCh4
ed@CLCh4
dt@CLCh4
GB
03/Apr/2006
अब!
04/Apr/2006
GB
05/Dec/2006
GB
09/Apr/2006
00:15:41.500
00:15:42.000
00:15:42.500
00:15:43.000
letta
letta
lett
-a
lett
-a
lett
plant
-IMP
plant
-IMP
plant
Plant! Plant!
रोप! रोप!

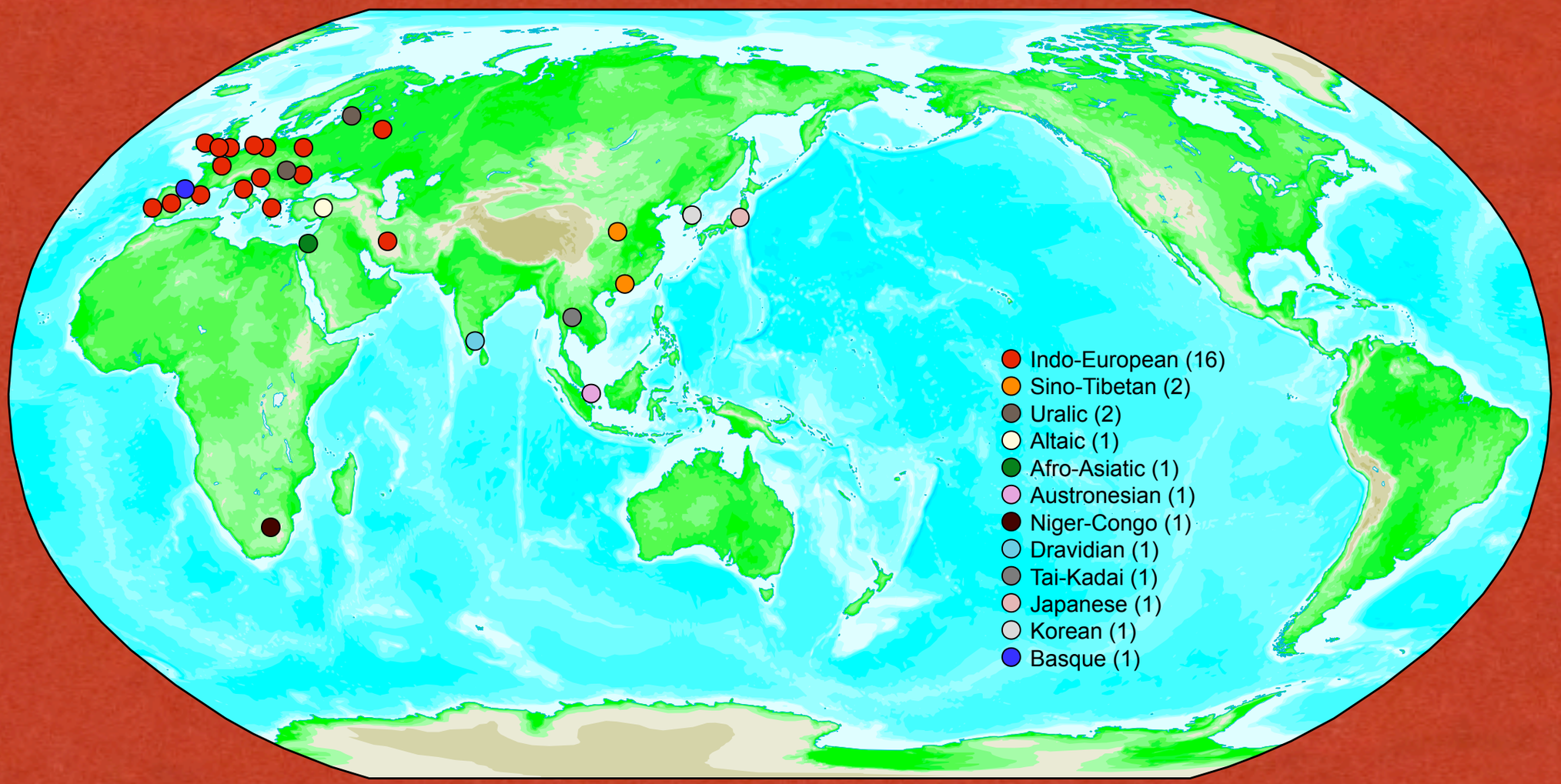

Elan - CLLDCh4R11S06.eaf
File Edit Annotation Tier Type Search View Options Window Help
Grid Text Subtitles Controls
ref@CLCh4
Nr Annotation rt@CLCh4 eng@CLCh4 nep@CLCh4 ed@CLCh4 dt@CLCh4
1 CLLDCh4R11S06.021 Yes! GB 03/Apr/2006
2 CLLDCh4R11S06.154a Now! Now! 04/Apr/2006
3 CLLDCh4R11S06.245 pick and GB 05/Dec/2006
4 CLLDCh4R11S06.461 Elder sister, this bean also (... GB 09/Apr/2006
00:15:41.112
Selection: 00:15:39.800 - 00:15:39.810 10
Selection Mode
Loop Mode
nep@CLCh4
tx@Kuluk
mph@Kuluk
mgl@Kuluk
rt@Kuluk
nep@Kuluk
tx@LDCh4
mph@LDCh4
mgl@LDCh4
rt@LDCh4
nep@LDCh4
tusandace
tus -a -dhend -a -ce
dig -PST -TEL -PST -ns
(we) digged.
hana carko mettoke
hana carko mett -u -kV
2s very do -3P -NPS
You very much doing,
aba letta
abo lett -a
now plant -IMP
plant now!
letta letta
lett -a lett
plant -IMP plant
Plant! Plant!

CHILD LANGUAGE DATA EXCHANGE SYSTEM

http://childes.psy.cmu.edu/

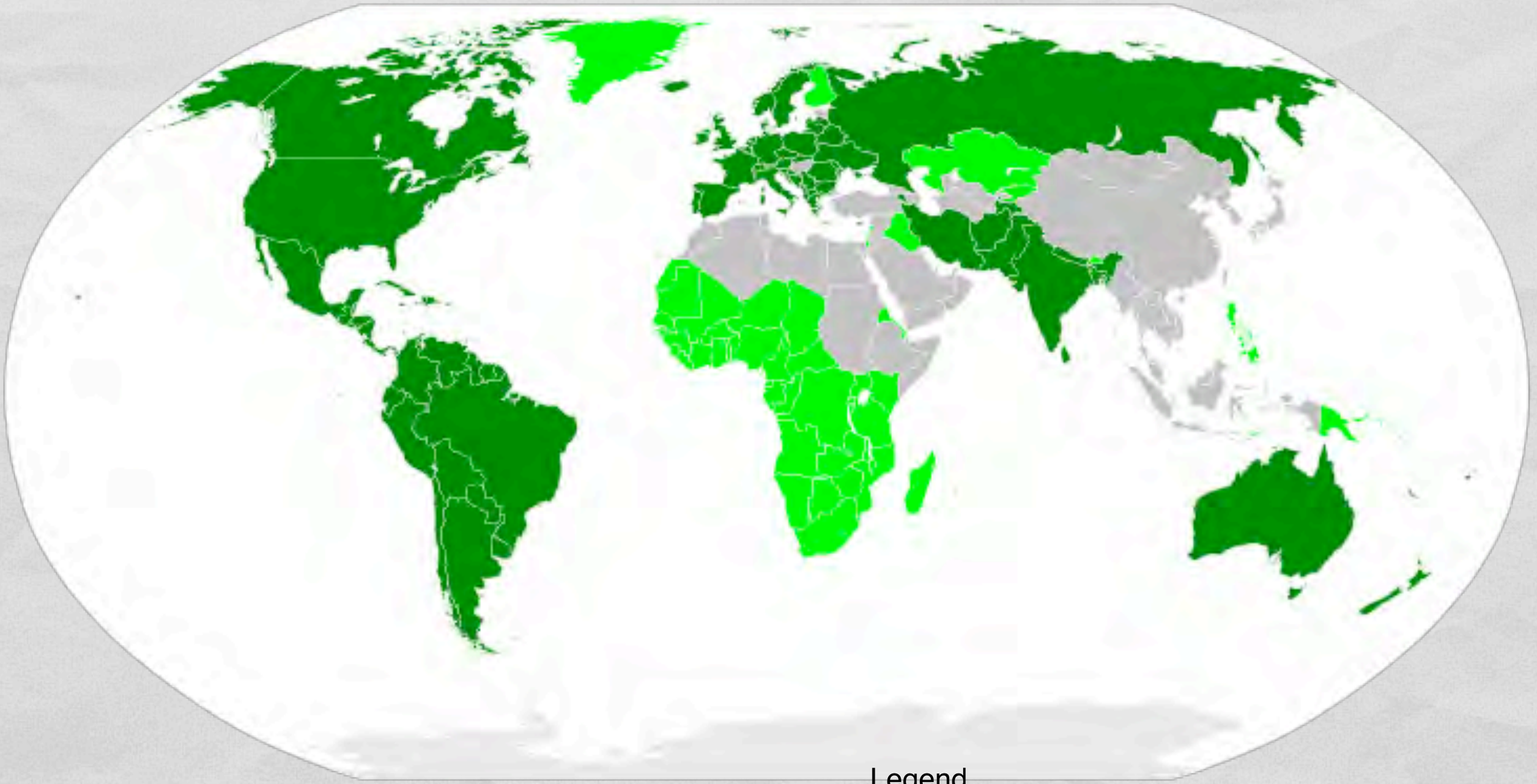
- 130 corpora of 30 different languages
- Tools include:
  - Methods for linguistic coding
  - Systems for linking transcripts to digitized audio and video
  - Programs for computer analysis of transcripts

CORPORA CHILDES



Leipzig Spring School 2008, Lieven & Stoll

IndoEuropean Languages today



Legend

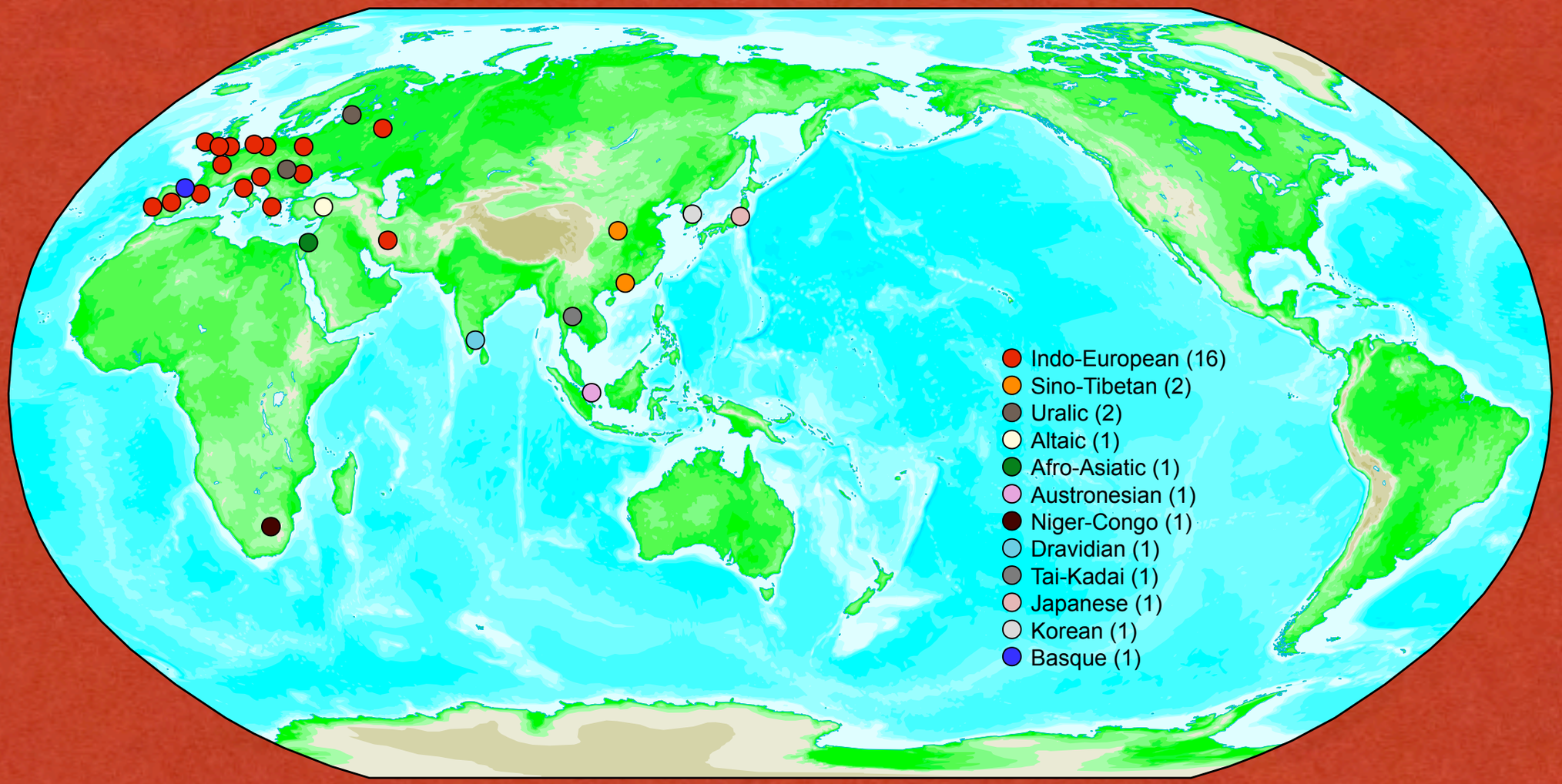
- Dark Green: countries with a majority of speakers of IE languages.
- Light Green: countries with an IE minority language with official status.
- Grey: other countries

Source: Wikipedia

AREA WITH MOST RARE FEATURES

(Cysouw, 2005)

CORPORA IN CHILDES



Leipzig Spring School 2008, Lieven & Stoll

SAMPLING PROBLEMS

```
@Begin
@Languages:    kor
@Participants:  CHI Target_Child, MOT Mother
@ID:    kor|Jiwon|CHI|2;0.20|female|||Target_Child||
@ID:    kor|Jiwon|MOT|||||Mother||
@Comment:       collected on 27-JUL-1992
@Comment:       coded on 5-April-2004
*MOT:  누구니, 이사람?
*CHI:  엄마 .
*MOT:  엄마가 어떡하고 있어?
*CHI:  코자 .
*MOT:  코자?
*MOT:  하하 야 엄마가 지원이 안고 있잖아 눈 감았다고
       코 잔다고?
*MOT:  아니야, 엄마 안 잤어 .
*CHI:  아냐
*MOT:  지원이 하도 이뻐서 웃다 보니까, 눈이 감긴거야
*CHI:  아니야
*MOT:  코 자는 거야?
*MOT:  응 그래 코 자
*MOT:  엄마 코 자고 누구냐?
*CHI:  아빠 코 자
*MOT:  아빠가 또 코 자구?
*MOT:  응 .
U1303[E][CHAT]  1
```

Leipzig Spring School 2008, Lieven & Stoll

CANTONESE (CHILDES)

*MOT: aa3_2 zeon3zeon3 gong2 bei2_2 ze4ze1_1 teng1 jau5 mou5 lok6 jyu5 ?

%ort: 阿 俊俊 講 比 姐姐 聽 有 冇 落 雨 ?

%mor: adj|aa3_2=the n:prop|zeon3zeon3=propernoun vt|gong2=say prep|bei2_2=to
n|ze4ze1_1=sister vt|teng1=hear vf|jau5=have neg|mou5=not_have
vdir|lok6=down n|jyu5=rain ?

*CHI: lok6 jyu5 .

%ort: 落 雨 .

%mor: vdir|lok6=down n|jyu5=rain .

DATA

- To make general claims about language acquisition and not the acquisition of a particular language we need to diversify sampling. It is not enough to restrict studies of language acquisition to one family (or even subfamily).

Leipzig Spring School 2008, Lieven & Stoll

POSSIBLE WORDS

German

ge-troff-en
an-ge-troff-en

but:

**ge-an-troff-en*
**ge-troff-an-en*

Chintang

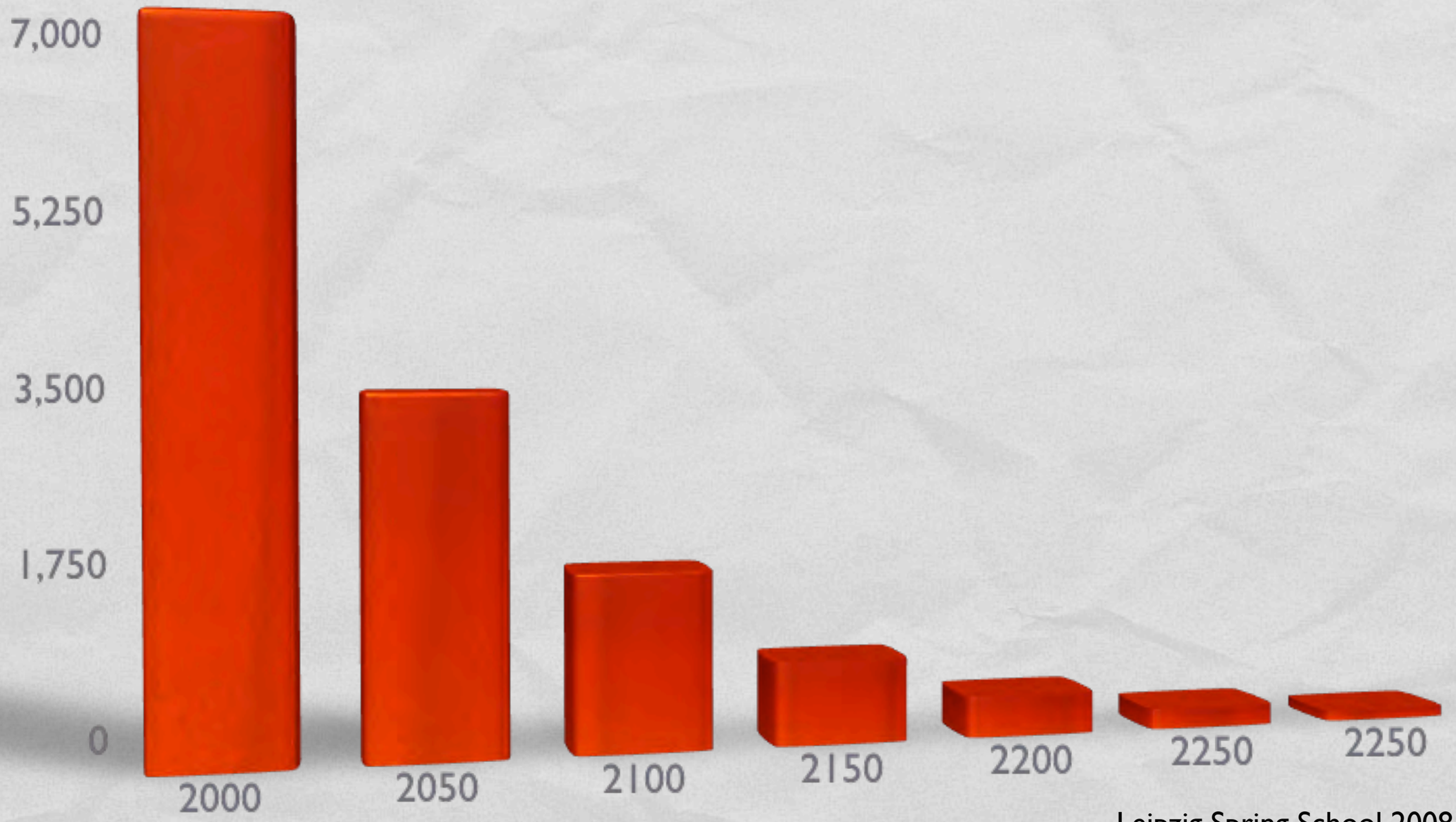
u-ma-tup-yokt-e-hẽ
3-NEG-meet-NEG-PST-1
'He didn't me'

or:

ma-u-top-yokt-e-hẽ
ma-top-u-yokt-e-hẽ

All theories about the word have assumed that there is a fixed order of morphemes in words. (But cf. Bickel et al. 2007)

ESTIMATION OF LANGUAGE DEATH (50% EVERY 50 YEARS)



Leipzig Spring School 2008, Lieven & Stoll

LANGUAGES AND THEIR NUMBER OF SPEAKERS

Leipzig Spring School 2008, Lieven & Stoll

COMPARATIVE LANGUAGE ACQUISITION (SLOBIN)

Crosslinguistic method to reveal developmental universals and language-specific patterns:

H0: Language development is everywhere the same

H1: Hypothesis of specific language effects

Leipzig Spring School 2008, Lieven & Stoll

LANGUAGES IN THE SAMPLE

Indo-European:

(a) Germanic: English, German

Romance: French, Italian, Portuguese, Romanian, Spanish

Slavic: Polish

- Semitic: Hebrew

Finno-Ugric: Hungarian

Ural-Altaic: Turkish

Japanese

Trans-New Guinea Non-Austronesian: Kaluli

Polynesian: Samoan

Leipzig Spring School 2008, Lieven & Stoll

STRUCTURE OF LANGUAGE CHAPTERS IN SLOBIN

1. Introductory materials about the language
2. Language acquisition data (errors, error-free acquisition, time of acquisition)
3. Data on the setting of language acquisition (cognitive pacesetting, linguistic pacesetting, input and adult-child interaction

Leipzig Spring School 2008, Lieven & Stoll

PROBLEMS TO COMPARE ACQUISITION PATTERNS

- Huge variation within a language
- Small sample longitudinal studies
- Variation in sampling and methods across languages
- Different age ranges
- How and what exactly to compare?

Leipzig Spring School 2008, Lieven & Stoll

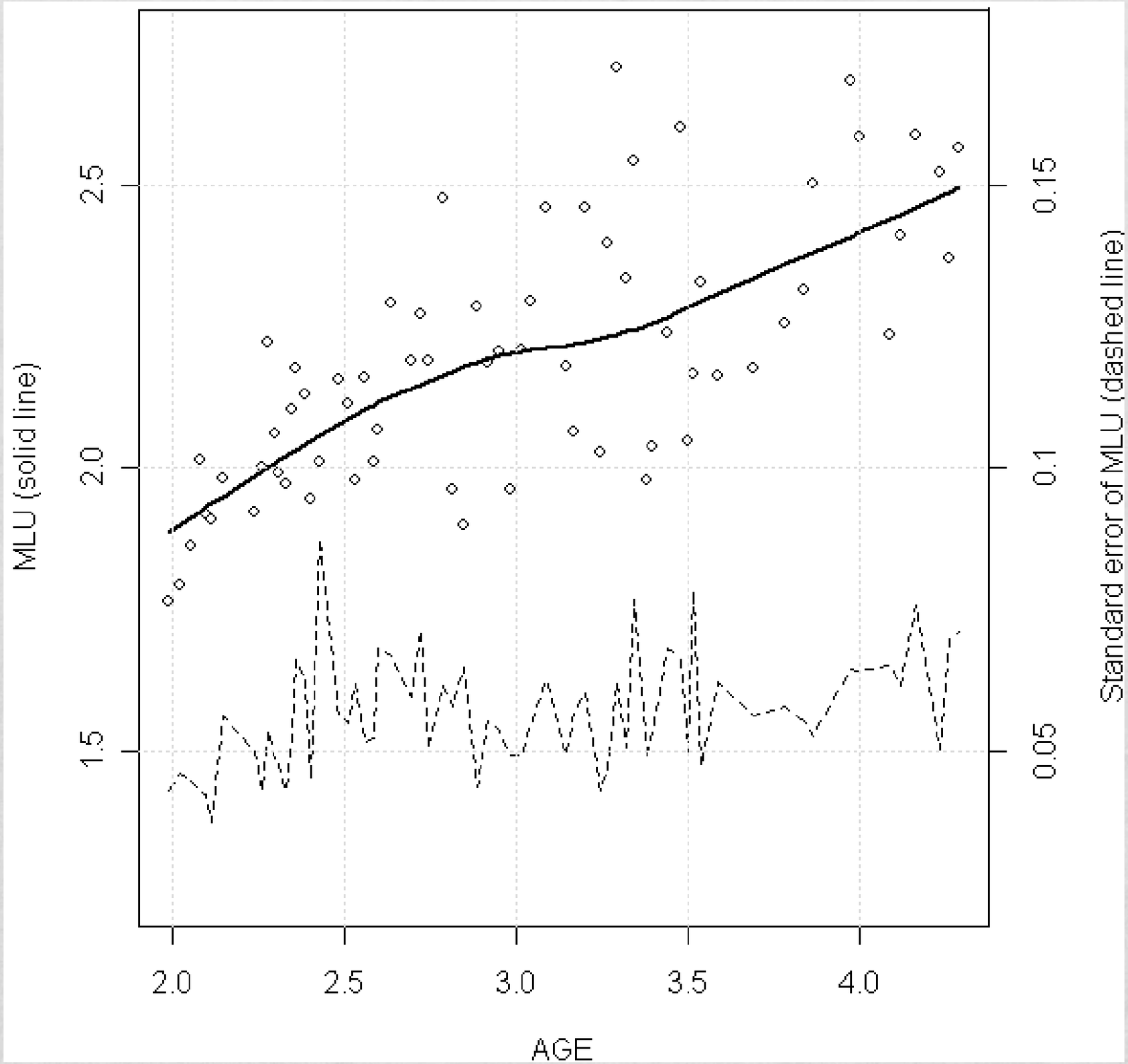
PROBLEMS TO COMPARE ACQUISITION PATTERNS

- Criteria to make comparisons:
  - Age
  - Mean length of utterance (MLU)

Table 3.4 *A summary of Brown's (1973) five stages of grammatical development*

Stage	Range of MLU (morphemes)	Upper Bound	Midpoint	Stage name and description
	1.00			*The period of single-word utterances* The use of single words without any grammatical knowledge
I	1.00–1.99			*Semantic roles and syntactic relations*
early	1.00–1.49			The onset and acquisition of the basic semantic
late	1.50–1.99	5	1.75	relations used in language like Agent, Patient. Word order is first syntactic device acquired.
II	2.00–2.49	7	2.25	*Modulation of meaning* The child begins to acquire inflections and grammatical morphemes. Most are actually acquired in subsequent stages.
III	2.50–2.99	9	2.75	*Modalities of the simple sentence* The active acquisition of the English auxiliary as it appears in yes–no questions, *wh*-questions, imperatives, and negative questions.
IV	3.00–3.99	11	3.50	*Embedding of one sentence within another* Complex sentences appear with object noun phrase complements, embedded *wh*-questions, and relative clauses.
V	4,00 and up	13	4.00	*Coordination of simple sentences and propositional relations* The active development of sentence, noun phrase, and verb phrase coordination with the use of conjunctions.

Ingram (1989: 50)



(Gries & Stoll, in press)

PROBLEMS TO COMPARE ACQUISITION PATTERNS

- Relevance problem
- Variability problem
- Data-sparsity problem
- Arbritrariness problem

(Gries & Stoll, in press)

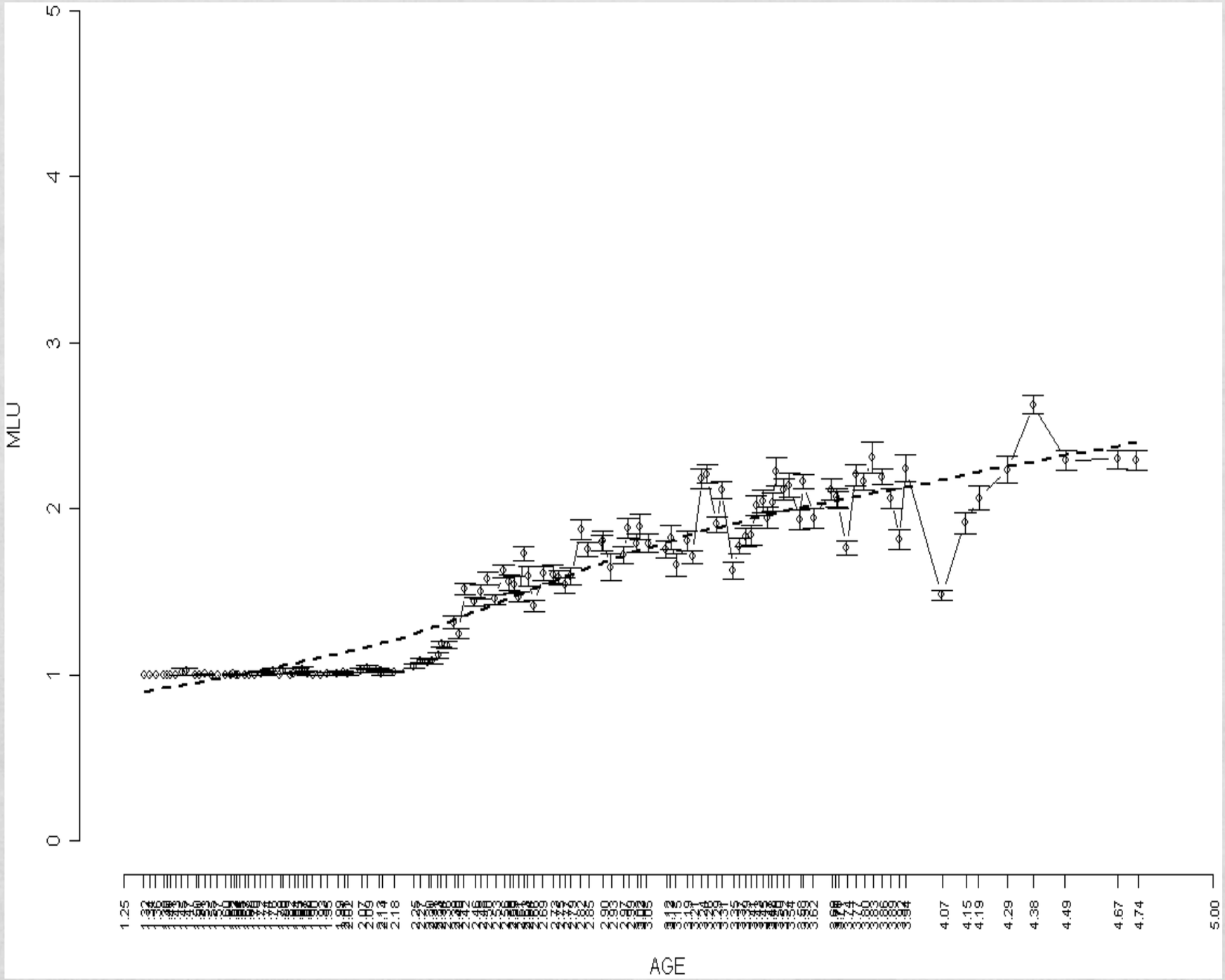
DEVELOPEMTAL STAGES IN CORPORA

- Suggestion: Variability based-neighbor clustering
    - all neighboring elements are compared on the basis of some distance measure
    - the two neighboring elements that are most similar to each other get merged
    - compute the distance measure on the basis of this merger

(Gries & Stoll, in press)

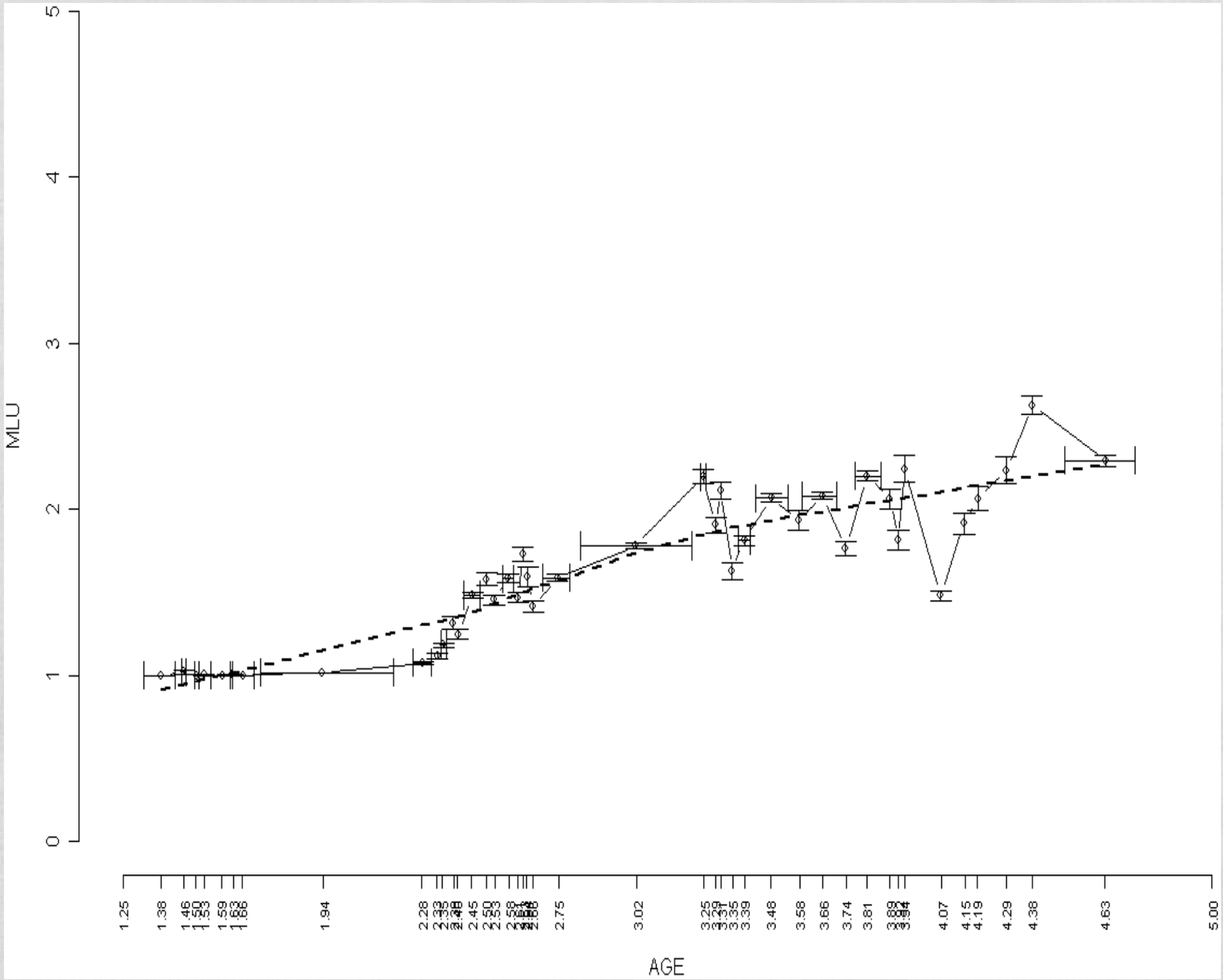
```
1 compute a distance or a similarity matrix,[3] which provides the (dis-)
  similarity of all elements to each other on the basis of some distance
  measure
2 repeat
3     identify the two elements that are most similar to each other
      (in the case of ties, choose one pair randomly);
4     merge the two elements that are most similar to each other and compu
      new distances on the basis of this merger
5 until the number of elements is one
6 draw a dendrogram that summarizes the groupings arrived at in steps 1 to
```

(Gries & Stoll, in press)



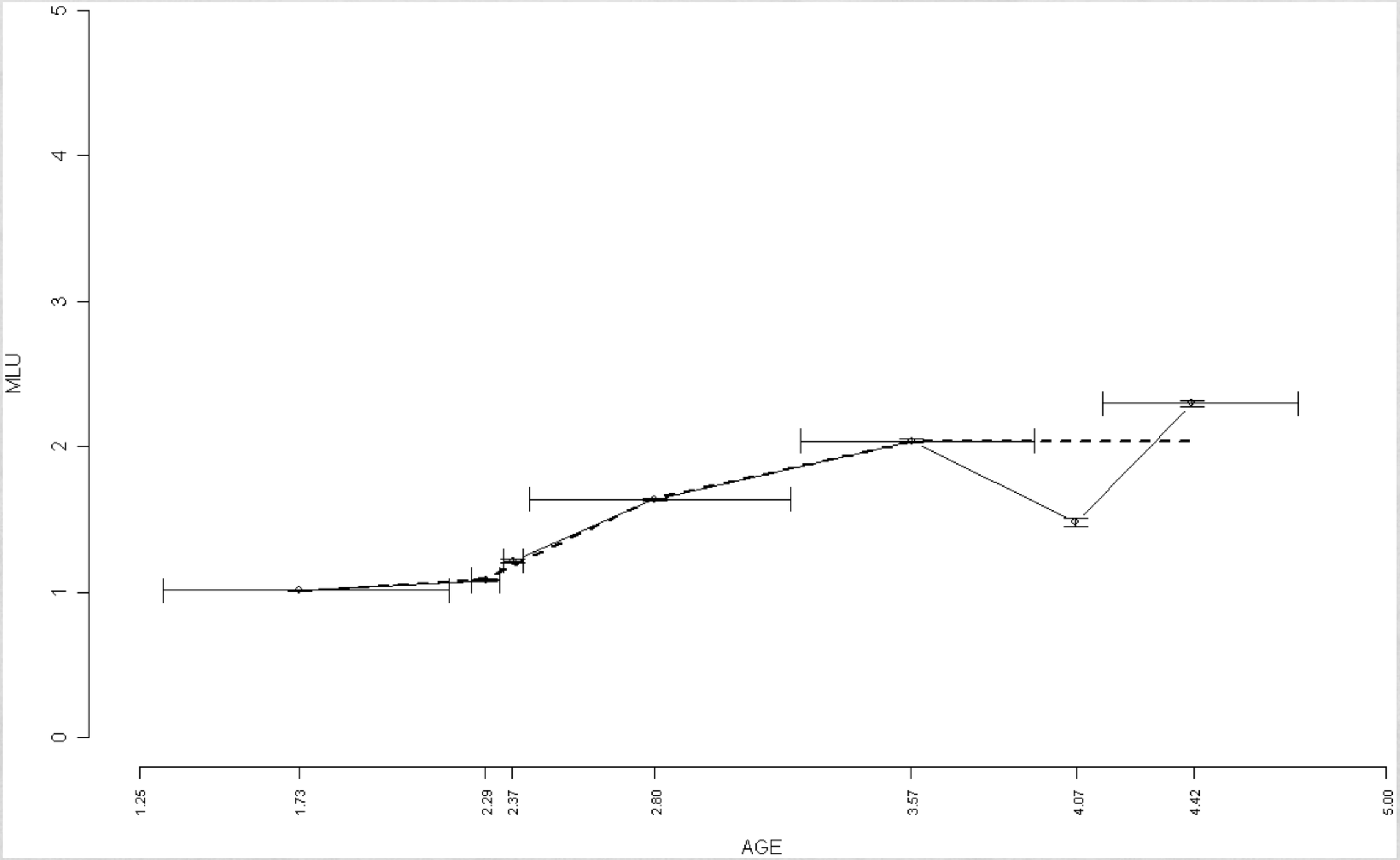
(Gries & Stoll, in press)

MLUs of 123 recordings of Child 3 between 1;03.26 and 4;09.30: step 40

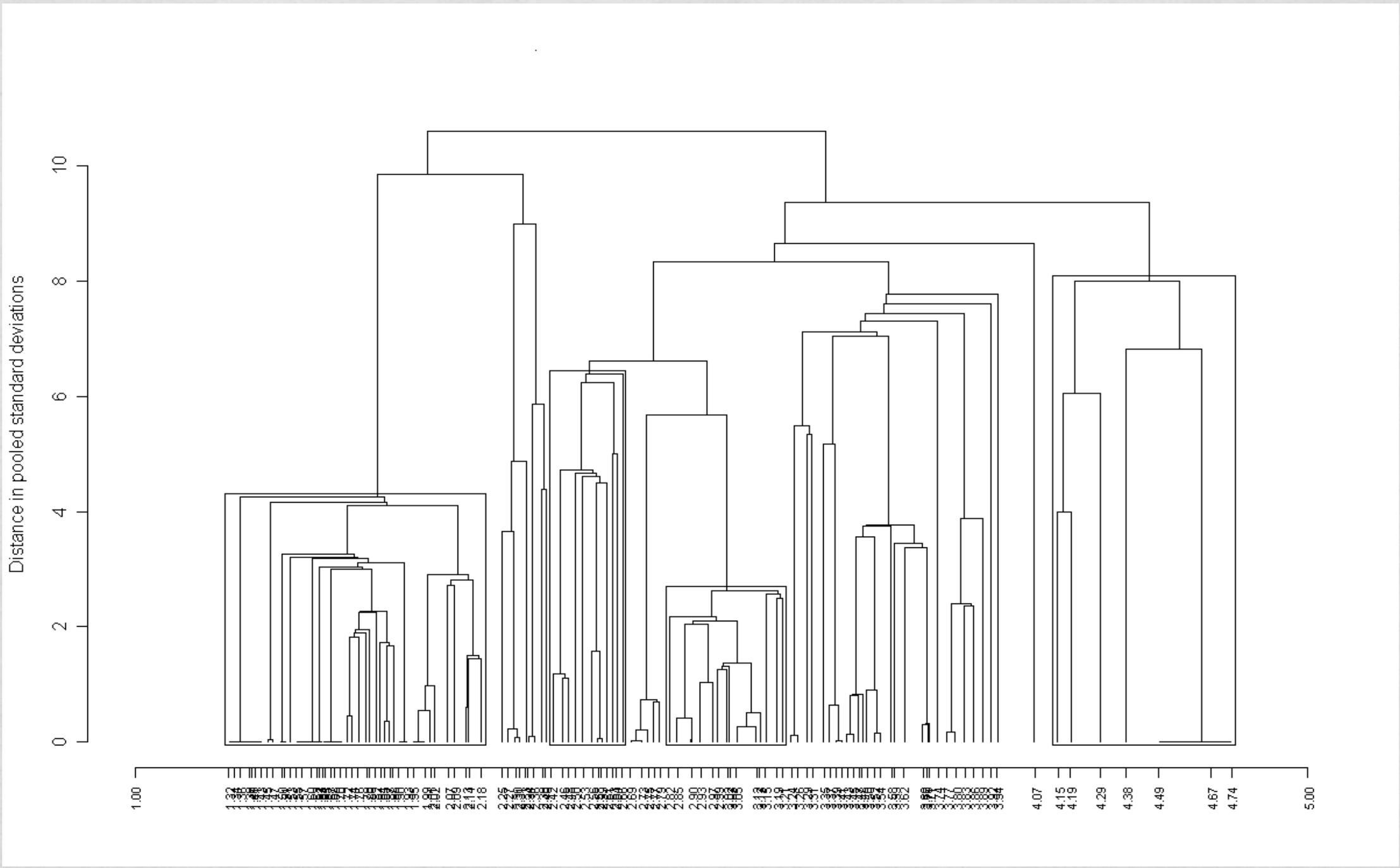


(Gries & Stoll, in press)

MLUs of 123 recordings of Child 3 between 1;03.26 and 4;09.30: step 115



(Gries & Stoll, in press)



(Gries & Stoll, in press)

SUGGESTION HOW TO PROCEED

- Identify pattern in individual children
- Compare each pattern to respective caretaker
- Compare caretakers across sessions
- Compare children across sessions
- Compare across languages.

(Stoll in prep.)

EXPERIMENTS

- Experiments are used to test a specific question with a clearly stated hypothesis.

- Experiments in general:
  - large enough number of participants for the variables tested.
  - control of variables possible.
  - good knowledge about the language/languages.

Leipzig Spring School 2008, Lieven & Stoll

EXPERIMENTS

- Potential interfering variables
- Participants
- Stimuli
- Procedure
- Practical questions

Leipzig Spring School 2008, Lieven & Stoll

Thank you!