Child language documentation

The sketch acquisition project

Birgit Hellwig (bhellwig@uni-koeln.de)

Linguistic diversity

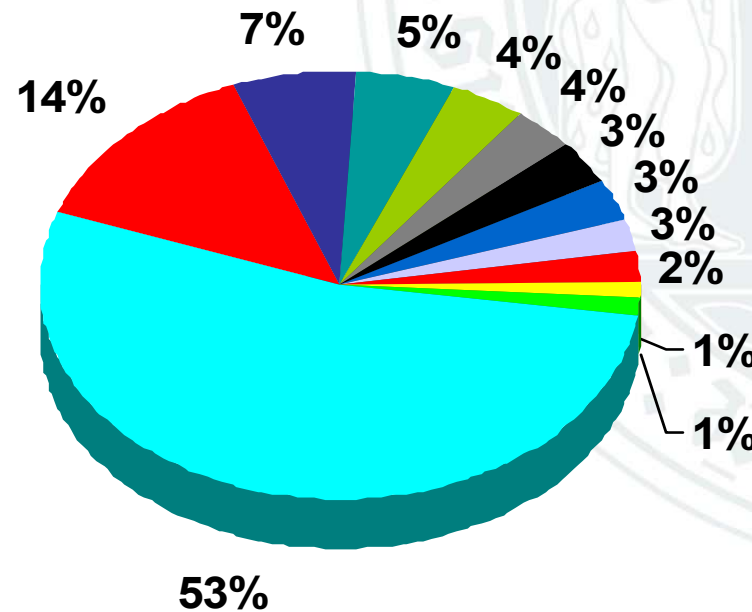47% speaks 11 languages

96% speaks 250 languages

4% speaks the remaining ~ 7000 languages

Legend:
- 🟥 Mandarin
- 🟦 Spanisch
- 🟦 Englisch
- 🟩 Hindi
- ⬛ Arabisch
- ⬛ Portugiesisch
- 🟦 Bengali
- 🟪 Russisch
- 🟥 Japanisch
- 🟨 Javanisch
- 🟩 Deutsch

Pie chart values: 7%, 5%, 4%, 4%, 14%, 3%, 3%, 3%, 2%, 1%, 1%, 53%

Challenge

"to show how the child's mind can learn and the adult's mind can use, with approximately equal ease, any one of this vast range of alternative systems." (447)

Nick Evans & Steve Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32: 429–492.

LINGUISTIC DIVERSITY in LANGUAGE EXPERIMENTS and CORPORA
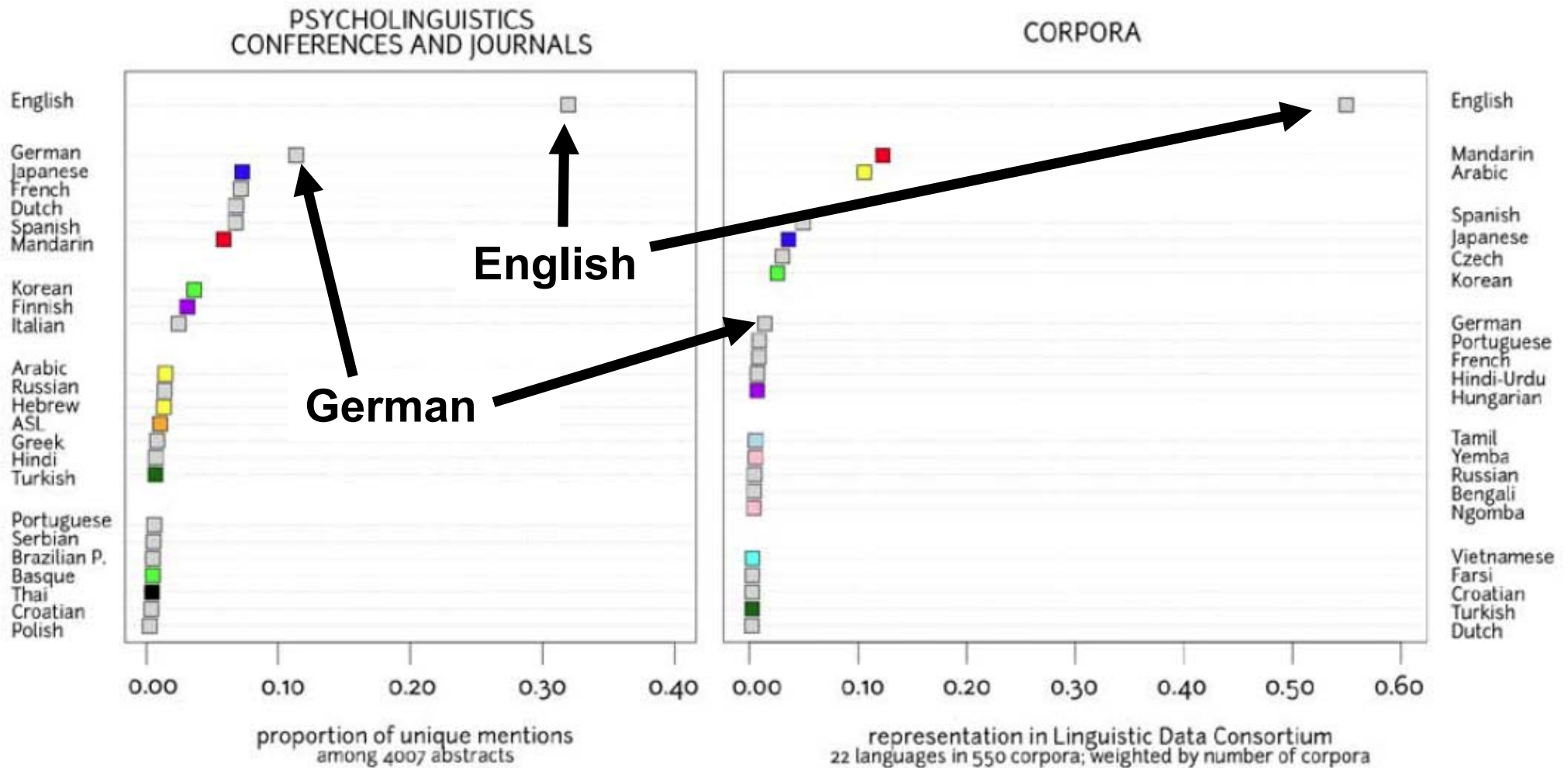
(Anand, Chung & Wagers 2011: 3)
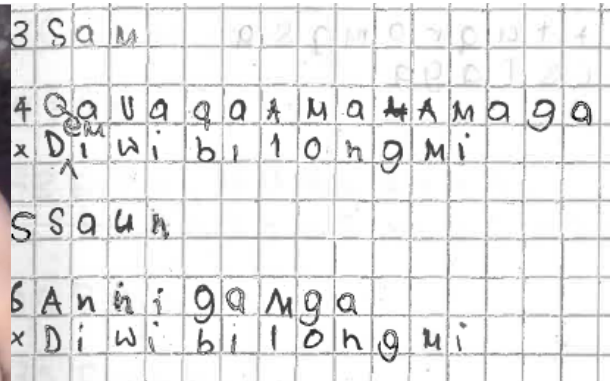
Language Acquisition

Elena Lieven & Sabine Stoll. 2010. Language. In Marc H. Bornstein (ed.). *The Handbook of Cultural Developmental Science*. New York & London: Psychology Press. 143–160.

"If we take all the acquisition studies together (experiments and longitudinal studies), we know something about the acquisition of approximately 70 to 80 languages (i.e., approximately 1% of all the languages spoken today).This 1% of languages also includes languages for which only one acquisition study of a single feature exists […]." (p. 144)

Psycholinguistics in the field?

- "These conditions often make it difficult to follow the best-practice approaches to data collection which are commonly assumed in lab-based FLA research." (Kelly et al. 2015: 287)

- "Still, specialized studies are perhaps best done in larger, less endangered language communities, especially given that many larger, unendangered language communities are also understudied." (Whalen & McDonough 2015: 3)

Sketch Acquisition Project

- **Language documentation**
  - plus language acquisition & socialisation
- **Project:**
  - child corpus of manageable size
  - plus acquisition sketch
- **Core group:**
**Rebecca Defina, Birgit Hellwig,
Shanley Allen, Lucy Davidson, Barb Kelly, Evan Kidd**
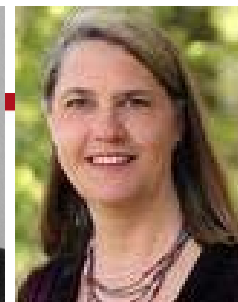
Sketch Acquisition Project

Workshop series

Manual

Workshop on the acquisition of lesser-documented languages
Cologne 25-26 January 2019
Programme

Friday 25th January

Presentation
9.00-9.10
9.10-9.40

Workshop on the Acquisition of Lesser-studied Languages

Melbourne 9 August 2019

Participants: Lucy Davidson, Rebecca Defina, Maria Graziano, Birgit Hellwig, Caroline Jone
Dagmar Jung, Barb Kelly, Carmel O'Shannessy, Penelope Schmidt, Gianna Urbanczik, Gillia
Wigglesworth, Wanyima Wighton

Archiving & Publishing

LAC Language Archive Cologne

LANGUAGE DOCUMENTATION & CONSERVATION

Piloting under way: Qaqet, Totoli, Pitjantjatjara,
Inuktitut, Dëne Sųłıné, Eegimaa, German, …

Acquisition Corpus

- **Severe constraints on:**
  - **selection of participants**
  - **sampling intervals**
  - **amounts of data**

  (e.g. Behrens 2008; Demuth 1996; 2008; Eisenbeiß 2006; 2010; Parisse 2019; Tomasello & Stahl 2004)

≈ 1 hour/week recording
  x 2 years (2;0-4;0)
  x 2 children   = 208 hours

"[...] the majority of existing child speech samples [...] represent only a very small proportion of all the language the child produces and hears – on average around 1%. [...] and in some cases 1% sampling is not adequate to answer the question at hand." (Tomasello & Stahl 2004: 118)

Sketch Corpus

Sketch corpus (ideal scenario): Longitudinal

Age (±2 months)	2;0	2;6	3;0	3;6	4;0
Child A	30(60)	30(60)	30(60)	30(60)	30(60)
Child B	30(60)	30(60)	30(60)	30(60)	30(60)
Total	60(120)	60(120)	60(120)	60(120)	60(120)

- **5 hrs analyzed (10 hrs recorded)**
- **Scenarious**
  - **longitudinal (2 children at 5 time points)**
  - **cross-lagged (2+ children at 2-3 time points each)**
  - **cross-sectional (10 children at 1 time point each)**

Universität zu Köln

Sketch Corpus

Sketch corpus (ideal scenario): Longitudinal

Age (±2 months)	2;0	2;6	3;0	3;6	4;0
Child A	30(60)	30(60)	30(60)	30(60)	30(60)
Child B	30(60)	30(60)	30(60)	30(60)	30(60)
Total	60(120)	60(120)	60(120)	60(120)	60(120)

More feasible example (Qaqet): Cross-lagged

Age (±2 months)	2;0	2;6	3;0	3;6	4;0
ZDL	2;0	2;8	–	–	–
YJL	–	–	3,1	3;7	4;0
YDS	2;0	2;6	–	–	–
YRA	–	–	3;2	3;7	4;0
Total	60(120)	60(120)	60(120)	60(120)	60(120)

Example

Qaqet

Papuan
(Baining)

PNG, East
New Britain

~ 15.000
speakers

- **2014-2022 (VolkswagenStiftung)**
- **Longitudinal study**

4 families
ages 2-4
natural setting
1hr/week
• video

Child	Age range	Recorded	Annotated	Partly annotated
ZJS	4;3-7;10	112 hrs	37 hrs	38 hrs
YJL	2;8-6;3			
ZDL	0;7-4;3			
YRA	3;2-4;7	120 hrs	37 hrs	45 hrs
YDS	1;11-4;10			
Total		**232 hrs**	**74 hrs**	**83 hrs**

+ 5 families dropped out (59 hrs); + 2 families from multilingual subcorpus (151 hrs)

Sketch corpus

More feasible example (Qaqet): Cross-lagged

ZDL	2;0	2;8	–	–	–
YDS	2;0	2;6	–	–	–
YJL	–	–	3,1	3;7	4;0
YRA	–	–	3;2	3;7	4;0
Minutes	60	60	60	60	60
IUs (Total)	1704	1454	2405	1145	1462
IUs (Adults)	569	228	585	136	139
IUs (Children)	1135	1226	1820	1009	1323
IUs (Focal child)	452	510	952	331	711

Settings

Missing setting
- children alone in the bush

Village
- in/around the house
- + many different interlocutors (adults, children)

Garden
- in garden or garden hut
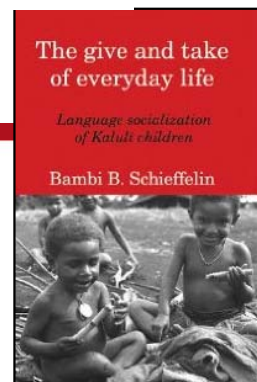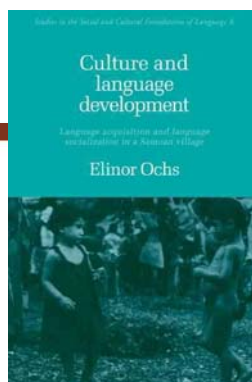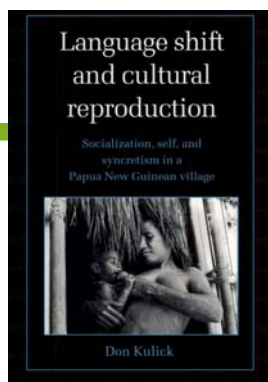- + few interlocutors (a parent, a sibling)

Case study: Child-directed speech

- ## CDS register:
    - short, correct & complete; few hesitations & errors
    - exaggerated pitch, high F0, long duration, more pauses
    - restricted vocabulary, here & now
    - nursery vocabulary
    - many questions & imperatives
    - repetitions & variations
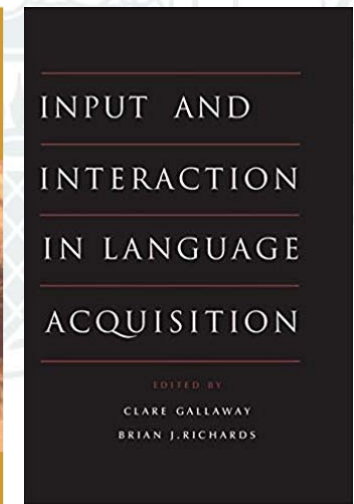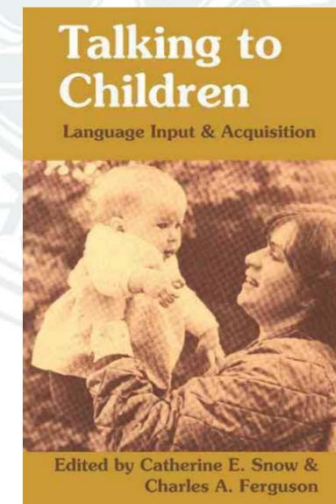
Universality?

Father: They will have to find money for a chicken, a small chicken.
        Boy [tries to scare his baby brother away]
        Boy [to uncle]: Look over there, at the baby's head.
Father: Like this.
Uncle: The chickens used to be here.
Mother: Those ones there.
        Boy [to uncle]: Uncle, he fell down.
Father: Is that our custom, or what?
Uncle: It isn't.
Father: It's really not our custom.
Uncle: The custom won't kill you. Last time I simply gave away a chicken.
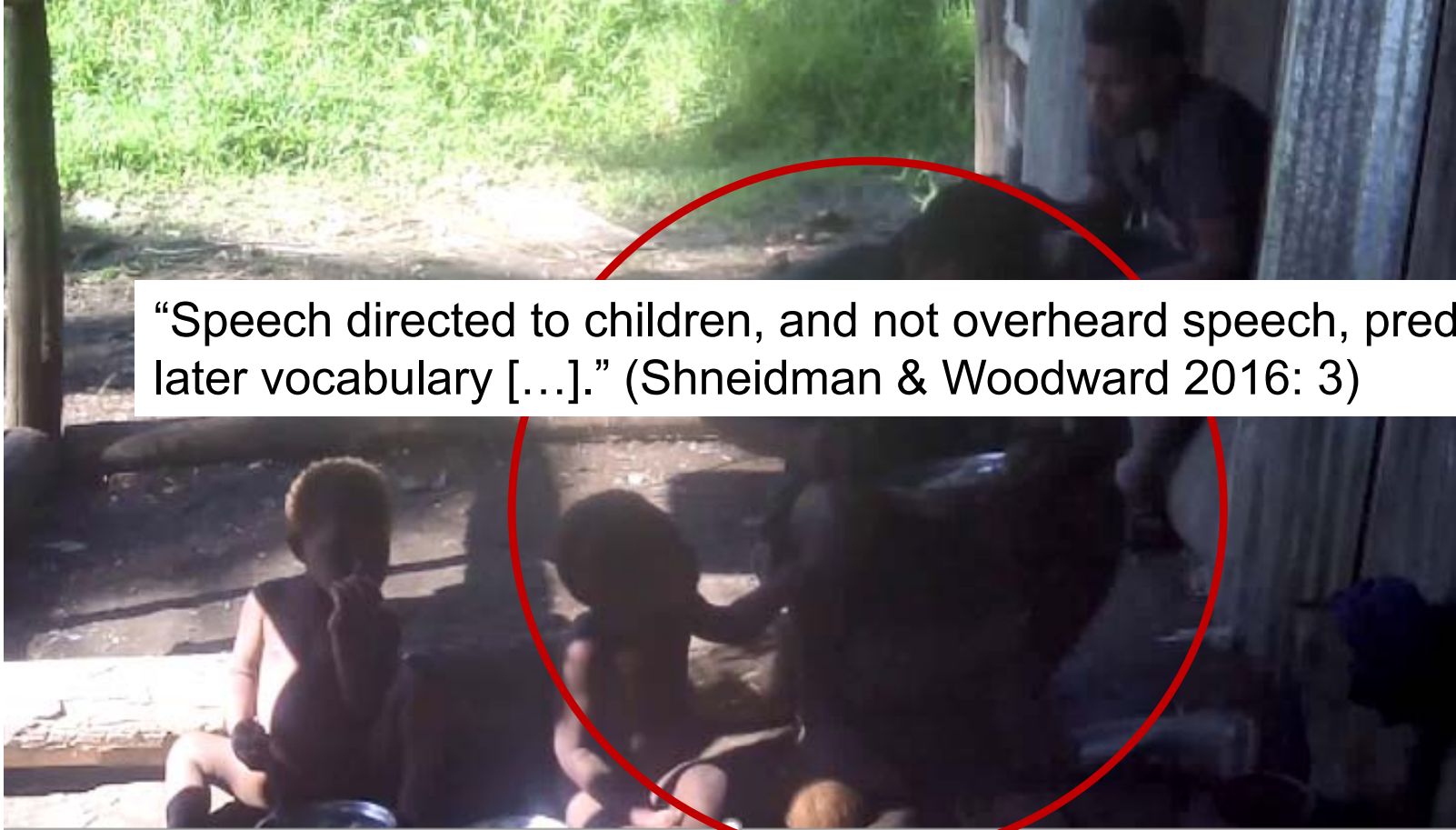
"Speech directed to children, and not overheard speech, predicts both children's later vocabulary […]." (Shneidman & Woodward 2016: 3)

But (Shneidman et al. 2009: 276):

A child is exposed to overheard speech in her everyday life	→	The child forms attention strategies that are adaptive to her role as an observer	→	The child employs these attention strategies when overhearing conversations	→	Child is able to learn novel words through overhearing
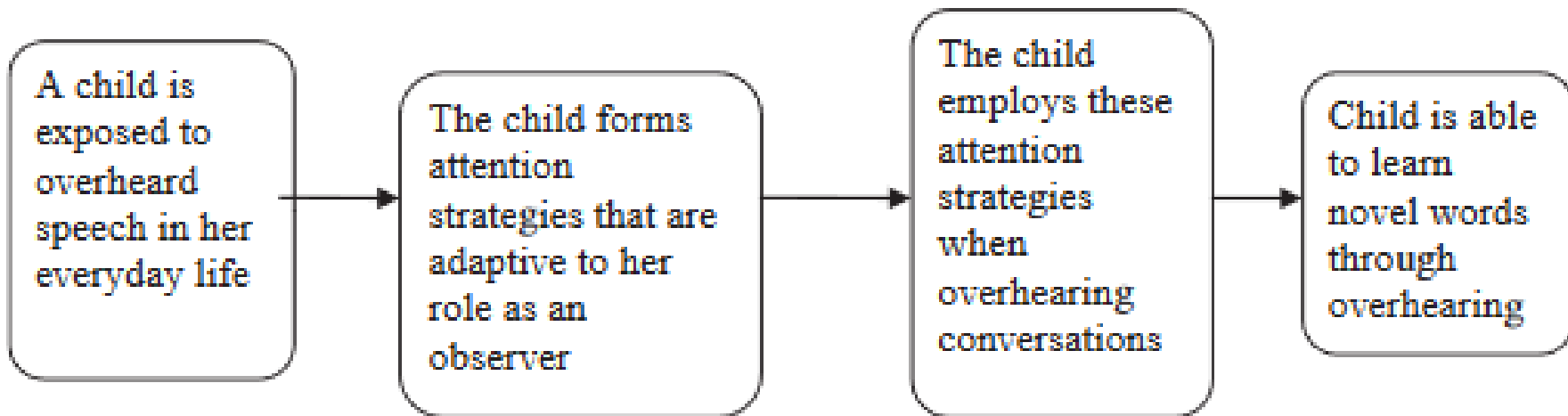
CDS in Qaqet

	Sketch corpus (5 hrs)	
Child-directed IUs	3.662	
Speaker: Adult	1.323	36%
Speaker: Older child	2.339	64%

- **Controlled study (The Qaqet Pear Story Corpus)**
  - **longer pauses, higher pitch & greater frequency range**
  - **fewer disfluencies & hesitations**
  - **short, less complex, lower MLU**
  - **(mostly) correct and complete**
  - **more imperatives & questions**

Frye, Henrike. 2019. *Child-directed speech in Qaqet*. University of Cologne, PhD Thesis.

Universitä

Sketch corpus: Prosody

ip ngulidresiit nanget nana? paqani nana?
'when do they tell stories? at what time?'
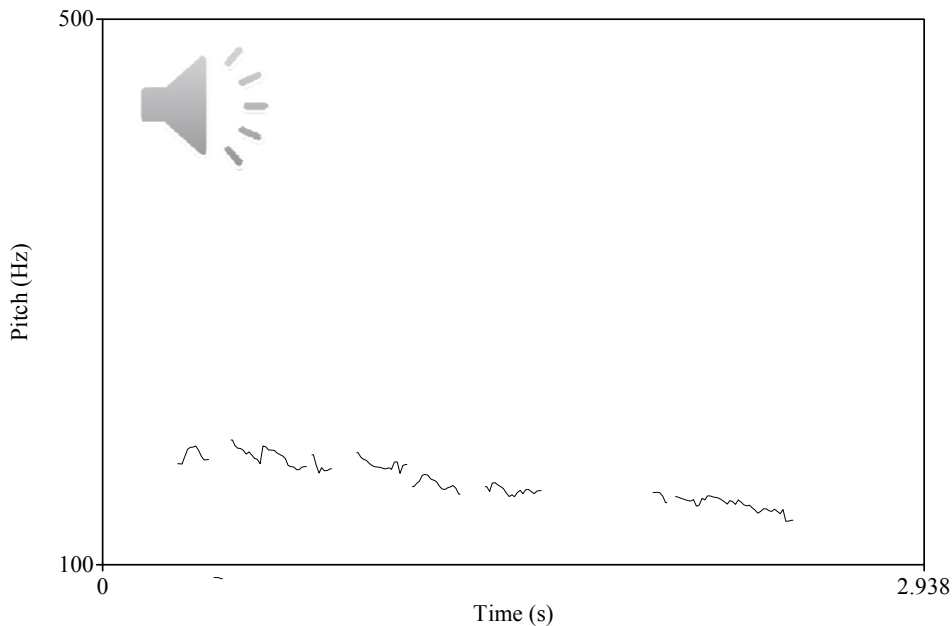
kua nyitlamaivip? amaivivim?
'do you see the snakes? the two snakes?'
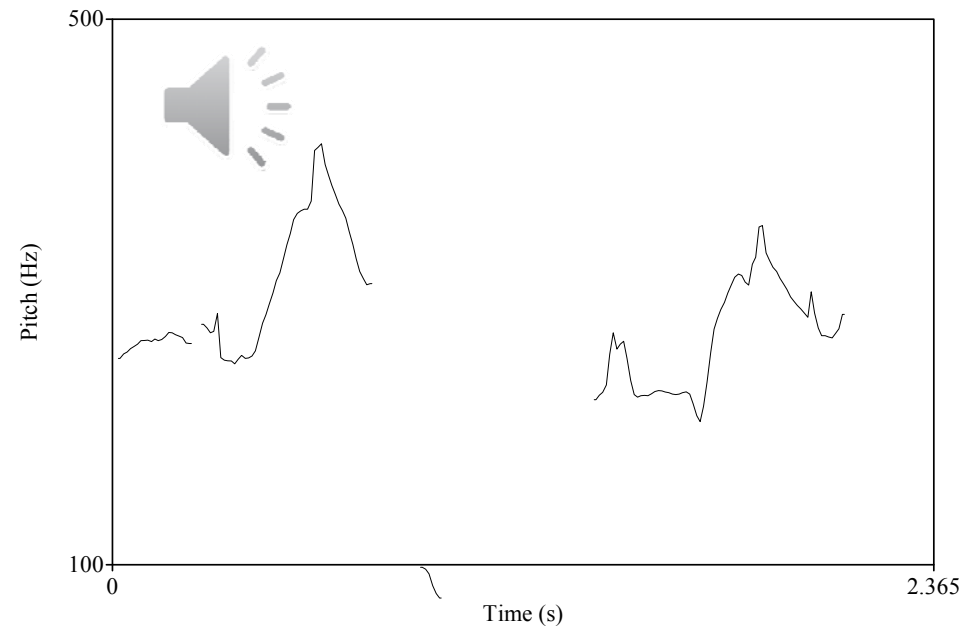
Sketch corpus: Morphology

ADS

ART=noun
SBJ=verb

CDS

ART=noun ~ noun
SBJ=verb ~ verb
+ morphophonology

aɣuukuka 'sweet potato'

a=kuukuk-ka 'NM=sweet.potato-SG.M'

YDS	*kaakak	sweet potato
AMT (mother)	ee, *kuukuka	yes, sweet potato
YDS	*kaakak	sweet potato
YRA	*akuukuka	sweet potato
AMT	ee	yes
YDS	*kaakak	sweet potato
YRA	ussh!	ussh!

22

Feedback

Expansion		
YDS	*tit	go
YDS	*gaka	my friend
AMT (mother)	mh?	yes?
YDS	*tit	go
AMT	undit	we go

Correction, Imitation, Laughter

23

YRA (3;2) to YDS (2;0)	
nyikut	you dig
unekut	we two dig
nyikut	you dig
nyikut	you dig
YDS, nyikut	YDS, you dig
nyikut tamasinep	you dig for spiders
sinep	spiders
nyikut iara	you dig here
nyikura	you dig now

Repetition & Variation

"partial repetitions […], with changes in lexical items, grammatical morphology, and/or word order, maintaining a constant communicative intent" (Küntay & Slobin 1996: 267).

Pronoun	Verb	Object	Adverb	
nyi	*kut*			you dig
une	*kut*			we two dig
nyi	*kut*			you dig
nyi	*kut*			you dig
nyi	*kut*			you dig
nyi	*kut*	*tamasinep*		you dig for spiders
		sinep		spiders
nyi	*kut*		*iara*	you dig here
nyi	*kur*		*a*	you dig now

Repetition & Variation

Sketch Corpus (5 hrs)

	Total	Speaker: Adult		Speaker: Older child	
CDS	3.662	1.323		2.339	
Varied repetitions	895	503	39%	392	17%
Exact repetitions	352	98	7%	254	11%

Longitudinal Corpus (23 hrs)

	Total	Speaker: Adult		Speaker: Older child	
CDS	12.968	4.142		8.826	
Varied repetitions	2.724	1.291	31%	1.433	16%
Exact repetitions	1.315	227	5%	1.088	12%

Literature: ages 2-4: ~25-30% variations, <1% exact repetitions

Repetition & Variation

AMT (mother) to YDS & YJS

utit	we go
utiravit	we go up
utiravit, **serama akun**	we go up, into the corn
YDS, utit maavit	YDS, we go up here
utit	we go
YJS, utiravit	YJS, we go up

YRA (3;3) to YDS & YJS

|	
guaki, nyan	my friend, come
nyan	come
YJS, nyan	YJS, come
nyan	come
nyan	come
guaki, nyan	my friend, come

+ many more topics

- **Conversational routines**
- **Lexicon: early words, word classes (nouns/verbs)**
- **Phonology: first sounds, realization of target sounds, phonological processes**
- **Morphology: noun/verb morphology, non-target-like forms, MLU**
- **Combinations: early combinations of words, word order, argument realization**
- **...**

A note of caution

- **Limited database – impacts on the kinds of statements we can make:**
  - descriptions of what the children hear & do & say – not of what they know
  - clear focus on qualitative analyses – not quantitative
  - raise topics/questions, to be explored and tested in a larger data set

- **Major contribution: to broaden our understanding of the problem space**

Not new …
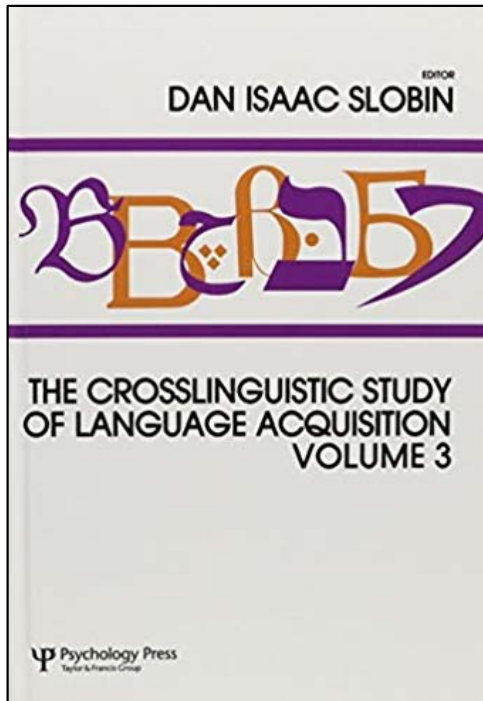
A FIELD MANUAL

for cross-cultural study of the acquisition of communicative competence

(Second draft -- July 1967)

Edited by

Dan I. Slobin

And written by

Susan M. Ervin-Tripp
John J. Gumperz
Dan I. Slobin
Jan Brukman
Keith Kernan
Claudia Mitchell
Brian Stross

"to guide investigators in the collection of comparable cross-linguistic and cross-cultural data on the acquisition of communicative competence" (Slobin et al. 1967: ix)

DAN ISAAC SLOBIN — EDITOR

THE CROSSLINGUISTIC STUDY OF LANGUAGE ACQUISITION VOLUME 3

Psychology Press
Taylor & Francis Group

- "However, despite the promise of this earlier typologically diverse work, when we fast-forward 30 years we do not see a boom in the study of FLA in typologically diverse languages and culturally different communities [...]." (Kelly & Nordlinger 2014: 180)

Danke – Amatlungena – Tenkyu tru