University of Zurich[UZH]
Department of Comparative Linguistics

How to study the acquisition of verbal morphology in languages of different complexity?

Sabine Stoll
Jekaterina Mazara

GOAL:

find general, potentially universal mechanisms of how morphological patterns are learned

CHALLENGES:

1. how to sample languages

and

2. how to measure the learning process

How to sample languages?

- Key idea **Maximum-Diversity Approach**

Language features
word order
synthesis
exponence
case marking
inflectional compactness of categories
existence of inflectional classes
…

Satellites (sign language)
Kata Kolok
Tzeltal

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Turkish	Indonesian	Inuktitut	Sesotho	Dene
Japanese	Yucatec	Chintang	Russian	Cree

Stoll & Bickel 2013

How to sample languages?

- Key idea **Maximum-Diversity Approach**

http://www.acqdiv.uzh.ch

Language features
word order
synthesis
exponence
case marking
inflectional compactness of categories
existence of inflectional classes
…

Satellites (sign language)
Kata Kolok
Tzeltal

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Turkish	Indonesian	Inuktitut	Sesotho	Dene
Japanese	Yucatec	Chintang	Russian	Cree
			English	

Stoll & Bickel 2013

University of Zurich[UZH]

erc

Data

- longitudinal corpora of English and Chintang (Sino-Tibetan, Eastern Nepal)
  - Chintang: polysynthetic
  - English: very little morphology

University of Zurich[UZH]

erc

The Chintang corpus

Chintang Language Research Program
(http://www.clrp.uzh.ch)

University of Zurich[UZH]

Chintang Language Research Program
(http://www.clrp.uzh.ch)

erc

The Chintang corpus

Linguistics:
Grammar
Discourse
Lexicon

How is the language **structured**?

Ethnography:
Rituals and religion
Social structure
Everyday life

How is the language **used**?

Psycholinguistics:
Acquisition strategies
Multilingualism

How is the language **learned**?

Balthasar Bickel
Robert Schikowski
Novel K. Rai
Martin Gaenszle
Sabine Stoll
Elena Lieven

Chintang Language Research Program (http://www.clrp.uzh.ch)

University of Zurich[UZH]

erc

The Chintang corpus

Genre	Transcribed	Transcribed	Translated	Translated	Glossed duration	Glossed # words
conversation	**232:44:54**	1'064'109	232:39:18	1'045'254	207:23:27	**903'645**
description	3:09:10	20'934	3:06:24	20'617	1:52:56	14'433
narrative	6:15:13	46'044	6:12:12	45'826	5:54:52	42'922
experimental	4:11:31	43'780	3:57:37	33'273	2:48:08	24'110
ritual language	11:15:55	40'634	10:03:34	32'736	6:36:02	18'270
songs	00:57:27	3'264	0:31:00	1'588	0:21:48	1'231
total	**258:34:10**	**1'218'765**	256:30:05	1'179'294	225:02:07	**1'004'837**

Chintang Language Research Program (http://www.clrp.uzh.ch)

University of Zurich[UZH]

erc

The Chintang Language Acquisition Corpus

children	recordings (monthly)	age
2	2-3 h	
2	4 h	
2	4 h	

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

- 220 h acquisition corpus

Chintang Language Research Program (http://www.clrp.uzh.ch)

University of Zurich[UZH]

erc

The Manchester Language Acquisition Corpus

children	recordings (monthly)	age		
2	4h			
1	4 h			

Age of the target children 0;6 2;0 3;0 3;6 4;11

- high density acquisition corpus, matched to the same number of recordings per child as in the Chintang corpus

Morphological complexity: English vs.

	Number of affixes	Number of unique verb forms
English	3	3
Chintang	148	4745

Morphological complexity: English vs. Chintang

English

I go
you go
he goes
we go
you go
they go
went

Chintang

	1s	1di	1pi	1de	1pe	2s	2d	2p	3s	3ns	intransitive
1s						*tupnaʔã* *tupnaʔãniŋ* *tupnehẽ* *matupyoknehẽ*	*tupnaʔãce* *tupnaʔãceniŋ* *tupnace* *matupyoknace*	*tupnaʔãni* *tupnaʔãniniŋ* *tupnanihẽ* *matupyoknanihẽ*	*tubukuŋ* *tubukuŋniŋ* *tubuhẽ* *matupyoktuhẽ*	*tubukuŋcuŋ* *tubukuŋcuŋniŋ* *tubuŋcihẽ* *matupyoktuŋcihẽ*	*tupmaʔã* *tupmaʔãniŋ* *tubehẽ* *matupyoktehẽ*
1di									*tupcoko* *tupcokoniŋ* *tubace* *matupyoktace*	*tubumcum* *tubumcumnim* *tubumcumhẽ* *matupyoktumcumhẽ*	*tupceke* *tupcekeniŋ* *tubace* *matupyoktace*
1pi									*tubukum* *tubukumnim* *tubumhẽ* *matupyoktumhẽ*		*tubiki* *tubikiniŋ* *tubihẽ* *matupyoktihẽ*
1de						*tupnaʔãncĩyã* *tupnaʔãncĩyãniŋ* *tupnancĩyehẽ* *matupyoknancĩyehẽ*			*tupcokoŋa* *tupcokoŋaniŋ* *tubacehẽ* *matupyoktacehẽ*	*tubumcumma* *tubumcummaniŋ* *tubumcummehẽ* *matupyoktumcummehẽ*	*tupcekeŋa* *tupcekeŋaniŋ* *tubacehẽ* *matupyoktacehẽ*
1pe									*tubukumma* *tubukummaniŋ* *tubummehẽ* *matupyoktummehẽ*		*tubikiŋa* *tubikiŋaniŋ* *tubiehẽ* *matupyoktiehẽ*
2s	*atupmaʔã* *atupmaʔãniŋ* *atubehẽ* *{a-ma}tupyoktehẽ*			*{a-ma}tupceke* *{a-ma}tupcekeniŋ* *{a-ma}tubace* *{a-ma-ma}tupyoktace*	*{a-ma}tupno* *{a-ma}tupnikniŋ* *{a-ma}tube* *{a-ma-ma}tupyokte*				*atuboko* *atubokoniŋ* *atube* *amatupyokte*	*atubukuce* *atubukuceniŋ* *atubuce* *{a-ma}tupyoktuce*	*atupno* *atupnikniŋ* *atube* *{a-ma}tupyokte*
2d	*atupmaʔanciŋ* *atupmaʔanciŋniŋ* *atubaŋcihẽ* *{a-ma}tupyoktaŋcihẽ*								*atupcoko* *atupcokoniŋ* *atubace* *amatupyoktace*	*atubumcum* *atubumcumnim* *atubumcumhẽ* *{a-ma}tupyoktumcumhẽ*	*atupceke* *atupcekeniŋ* *atubace* *{a-ma}tupyoktace*
2p	*atupmaʔaniŋ* *atupmaʔaniniŋ* *atubaŋnihẽ* *{a-ma}tupyoktaŋnihẽ*								*atubukum* *atubukumnim* *atubumhẽ* *amatupyoktumhẽ*		*atubiki* *atubikiniŋ* *atubihẽ* *{a-ma}tupyoktihẽ*
3s	*utupmaʔã* *utupmaʔãniŋ* *utubehẽ* *{u-ma}tupyoktehẽ*	*maitupceke* *maitupcekeniŋ* *maitubace* *{mai-ma}tupyoktace*	*maitupno* *maitupnikniŋ* *maitube* *{mai-ma}tupyokte*	*matupceke* *matupcekeniŋ* *matubace* *{ma-ma}tupyoktace*	*matupno* *matupnikniŋ* *matube* *{ma-ma}tupyokte*	*natupno* *natupnikniŋ* *natube* *{na-ma}tupyokte*	*natupceke* *natupcekeniŋ* *natubace* *{na-ma}tupyoktace*	*natubiki* *natubikiniŋ* *natubihẽ* *{na-ma}tupyoktihẽ*	*tuboko* *tubokoniŋ* *tube* *matupyokte*	*tubukuce* *tubukuceniŋ* *tubuce* *matupyoktuce*	*tupno* *tupnikniŋ* *tube* *matupyokte*
3d	*utupmaʔanciŋ* *utupmaʔanciŋniŋ* *utubaŋcihẽ* *{u-ma}tupyoktaŋcihẽ*								*utupcoko* *utupcokoniŋ* *utubace* *{u-ma}tupyoktace*	*utubukuce* *utubukuceniŋ* *utubuce* *{u-ma}tupyoktuce*	*utupceke* *utupcekeniŋ* *utubace* *{u-ma}tupyoktace*
3p	*utupmaʔaniŋ* *utupmaʔaniniŋ* *utubaŋnihẽ* *{u-ma}tupyoktaŋnihẽ*								*utuboko* *utubokoniŋ* *utube* *{u-ma}tupyokte*		*utupno* *utupnikniŋ* *utube* *{u-ma}tupyokte*

jo-go-yaŋ *na-khutt-i-**ca**-i-**hatt**-i-**bir**-i.*
whatever-NMLZ-ADD 3[s]>2-steal-2pO-V2:eat-2pO-V2:move.away.TR-2pO-V2:do.for-[SBJV.]2pO

‘It (a cat) may steal everything from you and eat it all up!’ [story.cat.204]

University of Zurich[UZH]

Consequences of complexity for input distributions

Consequences of complexity for input distributions

University of Zurich[UZH]

but they learn it at the same time:

University of Zurich[UZH]

erc

How can such complex system be learned?

Possible solution:

general statistical learning mechanisms

HEAD
DEP
DEP
independent stress
onset requirement, prosodic subcat.
voicing
u ca ŋa ta hai? ya ʔã na kina
3sA eat 1sO FOC move.away.TR 1sO IND.NPST INSIST SEQ
[:V] [:] [:V] [:X] [:V2σ] [:V] [:V] [:VP] [:XP]
insertion and displacement potential
cross-slot dependencies
fixed ordering

Frequencies of heads and dependents: Chintang verbs

University of Zurich[UZH]

GAM dependents vs. heads

Predictability

Predictable:
head/dependent occurs with few dependents/heads

Not very predictable:
head/dependent occurs with many dependents/heads

Modelling AoA

Age of acquisition ~

smooth component of frequency

+random intercepts for children

+type of morpheme (head/dependent)

+predictability

+predictability by type of morpheme

University of Zurich[UZH]

erc

Conclusions

- despite extreme differences in acquisition challenges, children learn Chintang and English morphology around the same age.
- children use statistical learning strategies that allow them to learn such extremely different systems
- they adapt to the affordances of their language.
- Chintang children first learn elements with higher repetition rates within verbs, the heads. Dependent elements in verbs are learned at a slower pace, there are more of them and the individual instances are less frequent. This reduces the cognitive load an makes the system learnable.

The Chintang team and funding sources

Banjade, Goma
Bibiko, Hans-Jörg
Bickel, Balthasar
Berthold, Falko
Cannedo, Laura
Domberg, Andreas
Erlach, Michael
Ghimire, Sandesh
Gaenszle, Martin
Hegetschweiler, Shirin
Harbodt, Silke
Kantak, Nicole
Klein, Felix
Krämer, Tatjana
Labs, Kristina
Lieven, Elena
Lorenz, Nicole
Paudyal, Netra
Pettigrew, Judith
Polkau, Claudia
Poppitz, Sindy
Prokopchuk, Joel
Rai, Chandra
Rai, Anita
Rai, Daya
Rai, Durga B.
Rai, Durga K.
Rai, Ganesh
Rai, Janaki
Rai, Lash K.
Rai, Rikhi
Rai, Novel K.
Rai, Shanti M.
Rai, Manoj
Rai, Rabindra
Sanjuan, Marcel
Sauppe, Sebastian
Schackow, Diana
Schikowski, Robert
Schneider Sarah
Seeg, Jenny
Strassberg, Joel
Stoll, Sabine
Vissiennon, Kodjo
Washington, Kerstin
Weymuth, Rachel
Wienholz, Anna
Wohlgemuth, Jan
Wolters, Kristin
Wunder, Saskia
Zakharko, Taras
Zeiske, Mathias

MAX-PLANCK-GESELLSCHAFT

EUROPEAN SCIENCE FOUNDATION

DFG

FNSNF
FONDS NATIONAL SUISSE
SCHWEIZERISCHER NATIONALFONDS
FONDO NAZIONALE SVIZZERO
SWISS NATIONAL SCIENCE FOUNDATION

VolkswagenStiftung

University of Zurich[UZH]

erc