



English word knowledge development of very young English learners in China: the best predictors

He Sun

Promotor: Kees de Bot

Supervisor: Rasmus Steinkrauss

University of Groningen



Outline

1. Background
2. Literature Review
3. Methodology
4. Results
5. Conclusion
6. Implication



Early English Education in China





Private Language Institutes :

- **14 billion Yuans in 2010**
(about 2.3 billion US dollars)
- At an annual growing rate of **12%**
- A **bigger** market in the future

Education and Training Industry Report
released by Deloitte in 2010



Relevant studies

- **Very little is** known about the very young learners
- **Hardly** any studies
- **Misconception** in the field of early language learning

(e.g. Mihaljević Djigunović, 2009)





Input and Usage in early SLA

Input quantity factors	Input quality factors	Language usage/output
<ul style="list-style-type: none">• Amount of exposure• Home / school• Length of exposure <p>(Chondrogianni & Marinis 2011; Unsworth 2013)</p>	<ul style="list-style-type: none">• Richness (TV, friends, etc.)• Parents' proficiency• Native vs. non-native• Variety of speakers <p>(Jia & Fuse 2007; Oller & Eilers 2002)</p>	<p>Child's output/use</p> <p>signf. predictor of children's developing language abilities</p> <p>(e.g. Hammer et al. 2012; Paradis 2011)</p>



Vocabulary Knowledge

Width

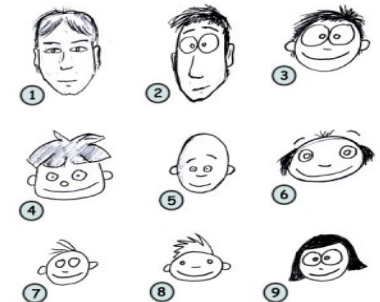
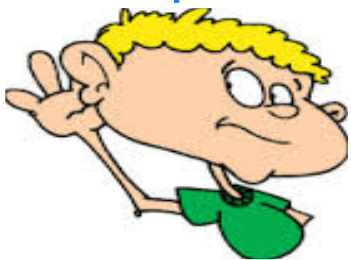
Depth

Receptive Vocabulary Size

Productive Vocabulary Size

Paradigmatic Knowledge

Syntagmatic Knowledge





Questions

1. To what extent do children develop their vocabulary skills 7 months during their onset period of instruction?
2. To the extent that such development exists, is it related to differences of English input and usage when children's internal factors and SES are controlled for?



Methodology

Participants

50 3-5 year-old children got involved
33 children participated two times

Tests

Time 1 (Apr. 2013): 45 children

- English Lexical knowledge tests
- Parental report: e.g. home English input quantity
- Language aptitude tests: e.g. memory

Time 2 (Nov. 2013): 41 children

- Tests as in Time 1



Predictors

Internal factors	External factors
Time	Length of exposure (Time 1)
Age	Weekly school input (in minutes)
<i>Age of onset</i>	Weekly media input (in minutes)
Gender	Media input richness (source and frequency)
Chinese proficiency (four aspects)	Native input ratio (native/total)
Nonverbal Intelligence (Raven's)	English usage (number of settings)
Short-term memory (CTOPP)	Mother's and father's English level (1-5)
	Mother's and father's educational level (0-5)



Outcome variables

Variables

English productive vocabulary
(EOWPV2)



English receptive vocabulary
(PPVT4)



English paradigmatic knowledge
(Word description; Verbal fluency)



English syntagmatic knowledge
(word description)





Data analysis

Why Linear mixed-effects models (lmer)?

- Can use all available data points
 - not affected by randomly missing data
 - can model time effects
- Can handle both covariates that change over time, and static covariates

NB: t-values given. We assume $t > 2$ means there is an effect



Analysis: precedure

1

- Data preperation (test type)

2

- Significant predictors

3

- Interaction between predictors and test type

4

- Testing by the importance of the significant predictors

5

- Other nonsignificant predictors

6

- Model criticism



Final model: Fixed effects

	Estimate	Std. Error	t value
Intercept	-2.53	0.44	-5.65
Time	0.49	0.10	4.99
Chinese Proficiency	0.33	0.08	3.85
Output	0.11	0.06	2.0
Length Of Exposure	0.06	0.02	2.66
Age	0.02	0.01	2.39



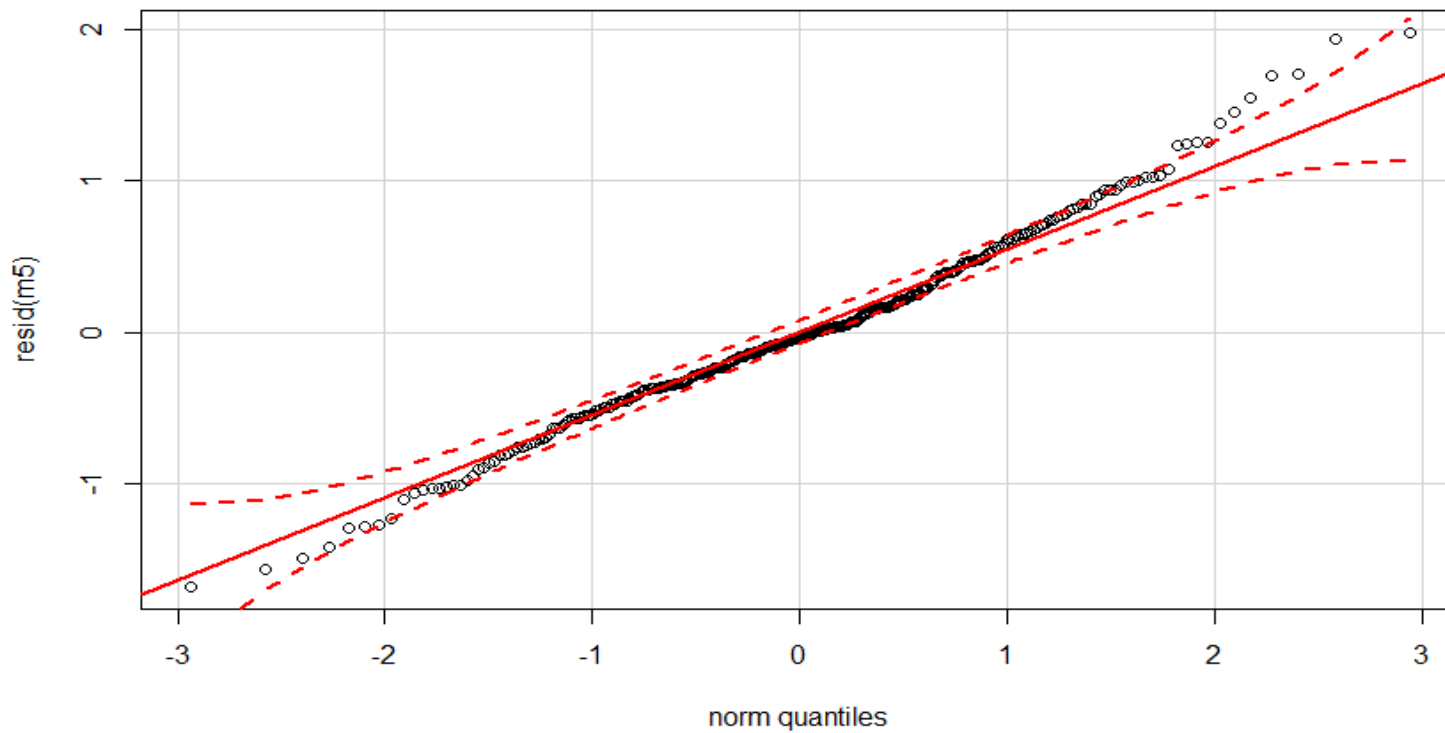
Final model: Random effects

Groups	Name	Variance	Std.Dev.
Subject	(Intercept)	0.2147	0.4634
Residual		0.4116	0.6415

Number of obs: 304, groups: Subject, 44



QQ plot





Preliminary conclusion

- › A significant growth in English lexical knowledge over 7 months
- › More English input and usage help English lexical acquisition
- › Older children learn faster
- › Better Chinese lexical knowledge facilitates the growth
- › Similar development pattern for different aspects of lexical knowledge



Acknowledgement





Questions on reporting the results

1. estimates or t value is more important?
2. r square? Generalized results?
3. small dataset and big dataset, how to report the two results?