

The Early Lexical Bias Revisited:
A Longitudinal Study of L1
Acquisition of Mandarin Chinese

LingLang Lunch Lite
Feb. 8th
Youtao Lu

Intro

- Lexical Bias:
 - for children at early stages of language acquisition, some lexical category, usually noun or verb, appears earlier and outnumbers other categories in children's production.
- Universal Noun Bias
 - Earlier and more
 - A variety of languages
- Language Specific
 - Frequency, saliency and transparency
 - Mainly verb bias in Korean and Mandarin

Doubts

- Age is not well controlled both within and between studies
 - e.g. For Mandarin, Gentner's 1;5 vs. Tardif's 1;10
 - Bassano (1998): a transition at 1;8
- The bias of language and the bias of early production is tangled

The Study

- Mandarin
 - Conflicting Empirical Evidence
 - Issues and doubts remain unsolved
- To solve the problems
 - Age is not well controlled → Longitudinal approach
 - Language is biased → Set a baseline to eliminate it
- Research Questions
 - What bias exists in Mandarin children's output (longitudinally)?
 - What is the relationship between adults' input and children's output?
 - Does the tentative removal of the bias of language change the answer to Q1?

Corpus & Coding

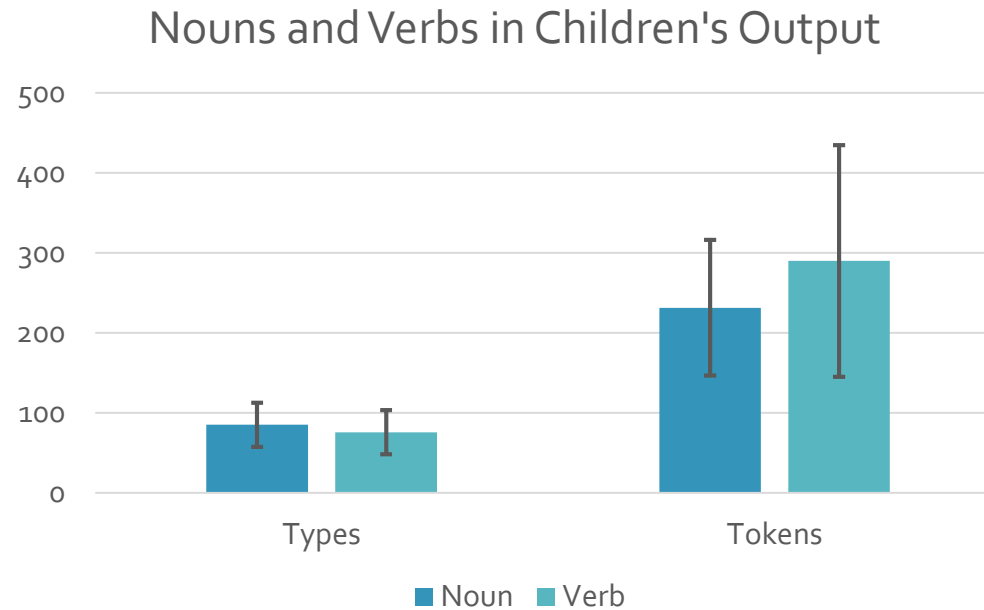
- Beijing Chinese Early Acquisition (BJCELA) corpus

Name (Initial letters)	CY	ZTX
Gender	female	male
Age range	1;6-2;4	1;6-2;4
Selection of age points	every two months	every two months
Sessions at each age point	two	two
Total sessions analyzed	12	12

- Only count common nouns and main verbs (Tardif, 1996)
- Deleted: Repetition, utterances under request ("say ..."), quoted speech (nursery rhymes, songs, poems).

Result I

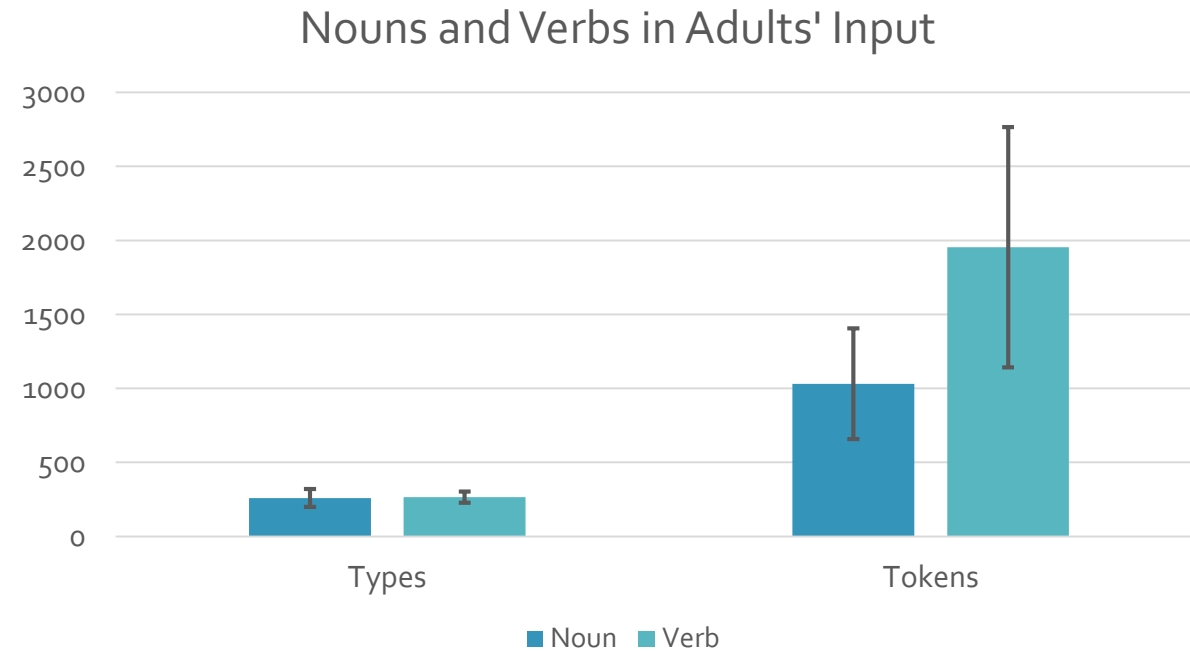
- I. Cross-sectional Analysis
 - Taken as independent observations
 - Comparable to existing studies



Type: pairwise $t(11) = 2.316, p = 0.041 (*)$. (Noun)

Token: pairwise $t(11) = -2.585, p = 0.025 (*)$. (Verb)

Result I



Type: pairwise $t(11) = -1.259$, $p = .234$. (ns)

Token: pairwise $t(11) = -7.251$, $p = .000$ (***) (Verb)

Result II

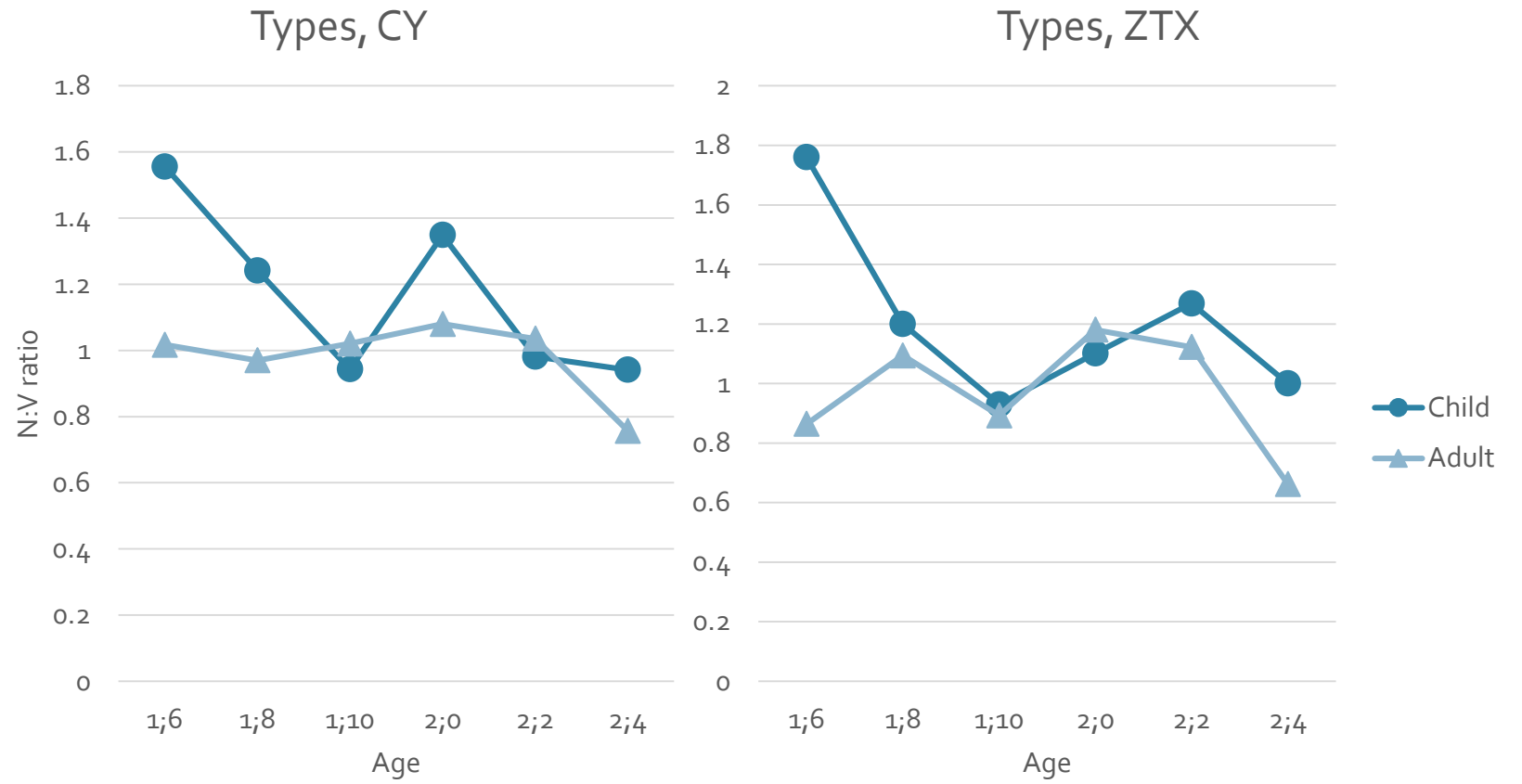
- II. Longitudinal: Noun/Verb Ratio
 - Use the number of nouns to divide the number of verbs
 - Ratio > 1: Noun bias

Name: CY							
Age		1;6	1;8	1;10	2;0	2;2	2;4
Child							
Type		1.56	1.24	0.94	1.35	0.98	0.94
Token		1.37	1.11	0.68	0.85	0.65	0.72
Adults							
Type		1.02	0.97	1.02	1.08	1.04	0.76
Token		0.49	0.51	0.44	0.44	0.45	0.38

Name: ZTX							
Age		1;6	1;8	1;10	2;0	2;2	2;4
Child							
Type		1.76	1.20	0.93	1.10	1.27	1.00
Token		1.91	1.12	0.74	0.71	0.69	0.74
Adults							
Type		0.86	1.09	0.89	1.18	1.12	0.66
Token		2.95	0.54	0.42	0.57	0.49	0.61

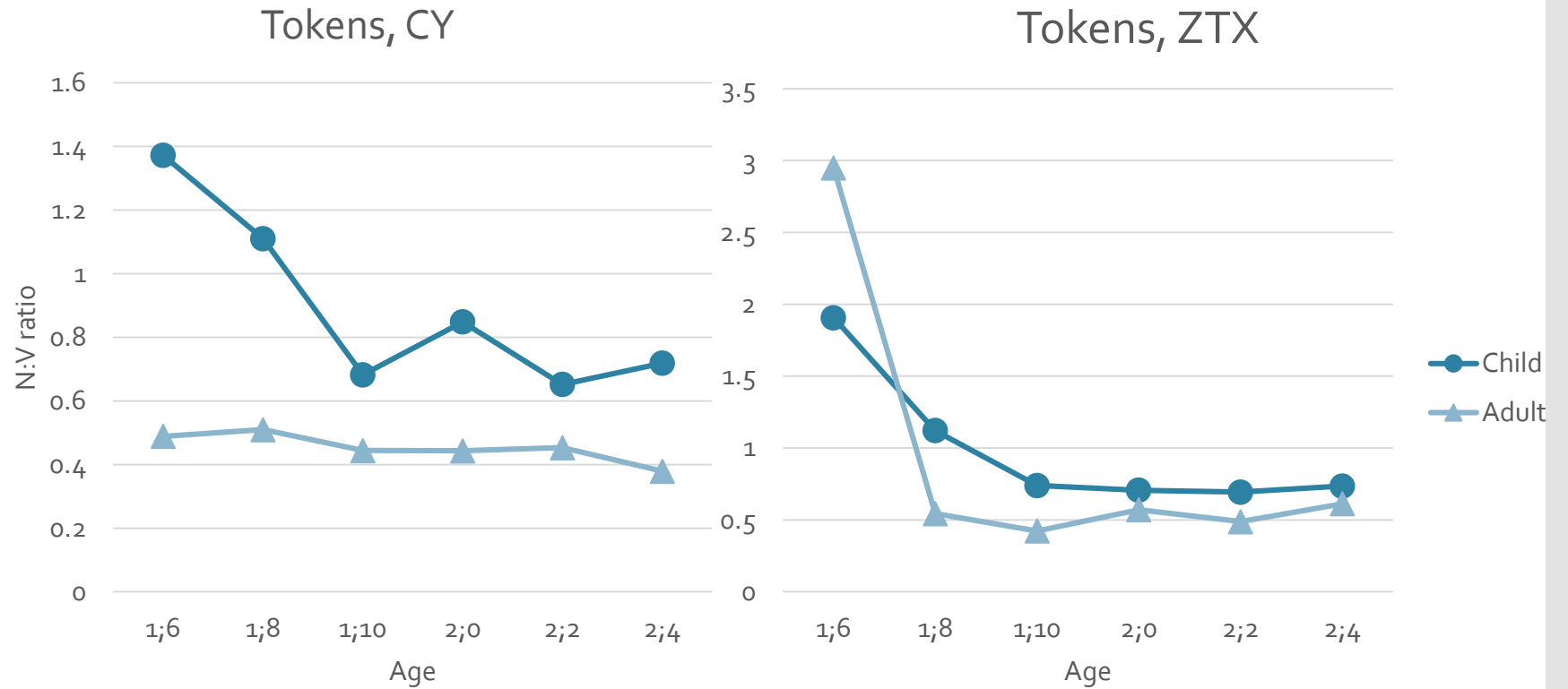
Result III

- III. Developmental Trend of N/V Ratio



A Trend of Approximation

Result III



Approximation
Mutual Influence?

Result IV

- IV. N/V Ratio Division
 - eliminate the lexical bias in language itself

$$\frac{N/V \text{ ratio in children's output}}{N/V \text{ ratio in adults' input}}$$

- Interpretation:
 - What if... a balanced distribution in language itself?

Age	1;6	1;8	1;10	2;0	2;2	2;4
CY						
Type	1.53	1.28	0.92	1.25	0.95	1.24
Token	2.81	2.18	1.53	1.91	1.44	1.90
ZTX						
Type	2.04	1.10	1.04	0.93	1.13	1.51
Token	0.65	2.06	1.75	1.24	1.42	1.20

Discussion & Conclusion

- Answers to the Questions:
 - Verb bias while age is not controlled, Noun bias while it is and while early enough
 - Children seem take adult's input as the paradigm of language using
 - To unravel the tangled bias: Mandarin is verb biased, but children's production is noun biased
- Evidence for:
 - A Universal Noun Bias
 - Maybe nouns are easier

Discussion & Conclusion

- Longitudinal Finding: 1;8, is it a coincidence?
 - Also Bassano (1998)
 - Between Gentner (1982) and Tardif et al. (1996, 1997, 1999)

What's Now/New

- Ambicategorical words (Conwell & Morgan, 2012)
- Polysemy/Homophony
 - Cross-linguistic differences?
- Homophones and phoneme inventory
 - larger → shorter (Nettle, 1995; 1998) → more (Piantadosi et al., 2012)
 - larger → fewer (?)

Thank You !