

Continuously growing resources but discrete production units:

A probabilistic account of the development of
early utterance length

Qihui Xu

Martin Chodorow

Virginia Valian



Xu, 2020



1



Qihui Xu



Martin Chodorow



Virginia Valian



qxu@gradcenter.cuny.edu

But gradually lengthen as they grow older

(Bloom, 1973; Bowerman, 1973; Braine, 1976; Goldin-Meadow, 2003)

16 months

“Push.”

“Bottle rolling.”

“Key open door.”

26 months

“It fell.”

“Let me in.”

“But the crack in the track.”

(Providence. Evans & Demuth, 2012)



Xu, 2020

Importance of studying early utterance length development

- **Syntax** (Ambridge, Kidd, Rowland, Theakston, 2015; Goldberg, 2013)
- **Language learnability** (Valian, 1986; Seidenberg, 1997; Yang, 2013)
- **Cognitive vs. linguistic constraints** (Berk & Lillo-Martin, 2012; Bloom, Lightbown, & Hood, 1975)

the utterance length development itself still remains largely unclear

- ‘Single-word period’ -> ‘Two-word combinations’,... (Bloom, 1973; Braine, 1976; Brown, 1973)
 - Small sample size
 - Diary studies
 - Vague and inconsistent descriptions

Stage vs. continuous development

Research questions

- How do children's utterances change in length – continuously or in stages?
- What kind of underlying process could produce such behavioral change?

Research questions

- **Study 1:** How do children's utterances change in length – continuously or in stages?
 - Smooth change?
 - Stage transition

Study 1 - Methods

Data

English, longitudinal, CHILDES (MacWhinney, 2000)

25 children

Age: 14 – 43 months

Study 1 - Methods

- **Analyses**

The distribution of **utterance length 1-5 words** as a function of **age**.

Utterance length

‘xxx’ and ‘yyy’

‘pick up up up’

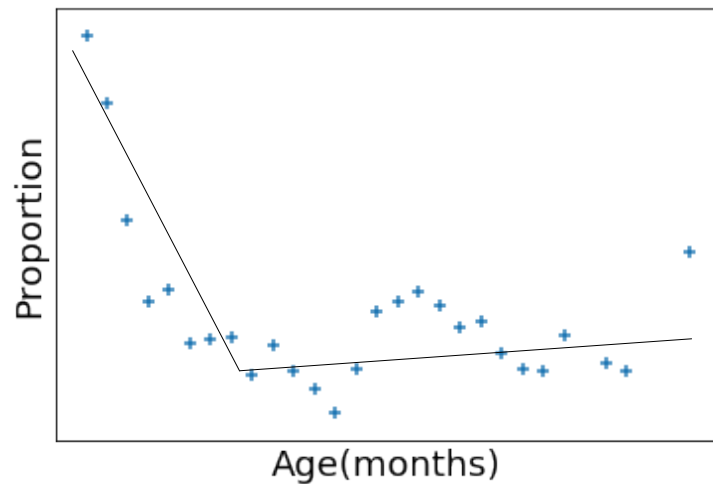
babbling and counting

immediate imitations

Study 1 - Methods

- **Analyses**

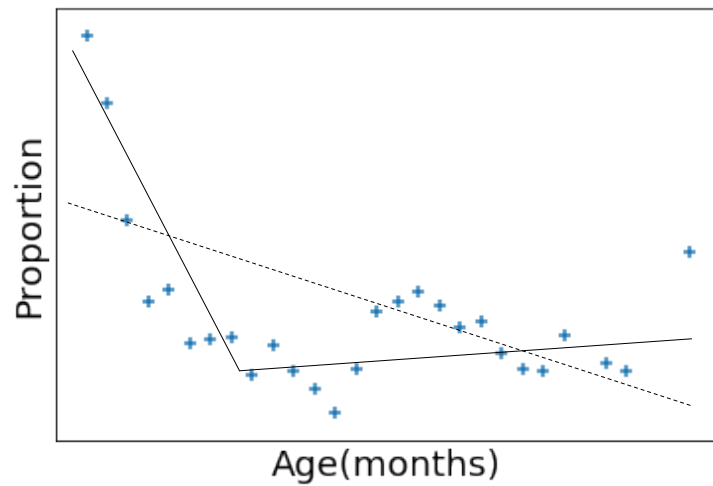
Track **stage** transitions – mixed-effect segmented regression



Study 1 - Methods

- **Analyses**

Track **stage** transitions – mixed-effect segmented regression

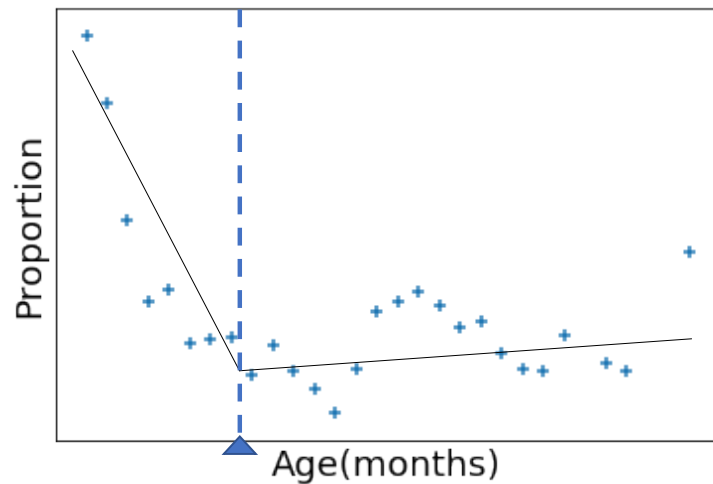


Significant transition

Study 1 - Methods

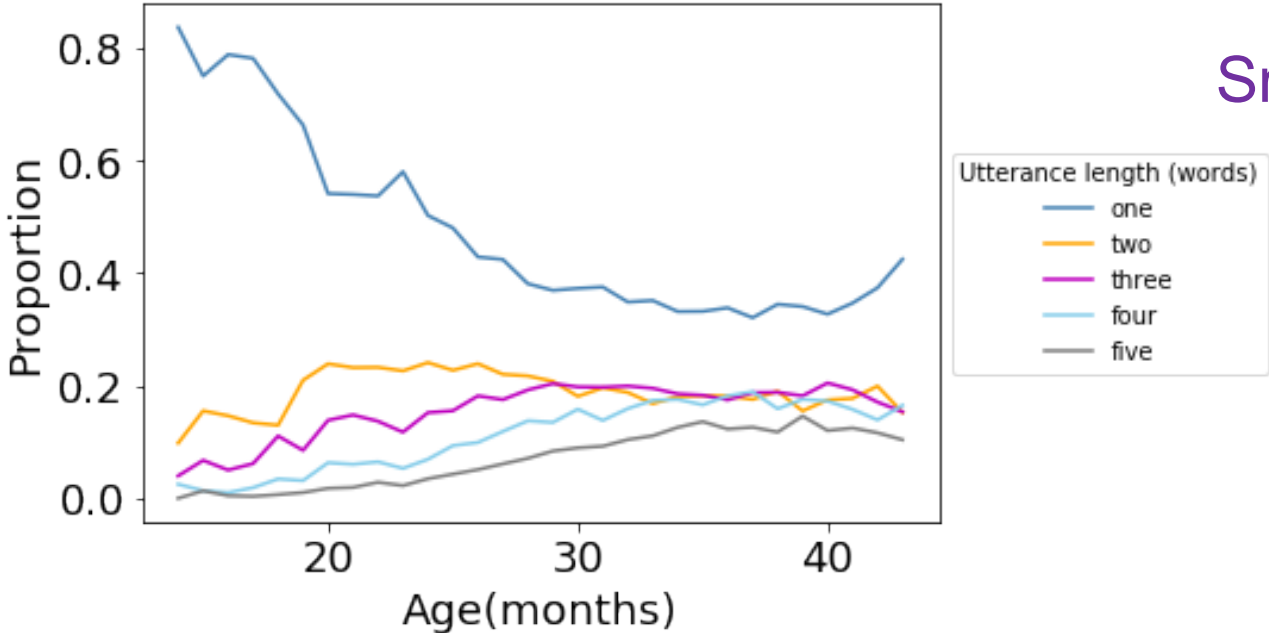
- **Analyses**

Track **stage** transitions – mixed-effect segmented regression



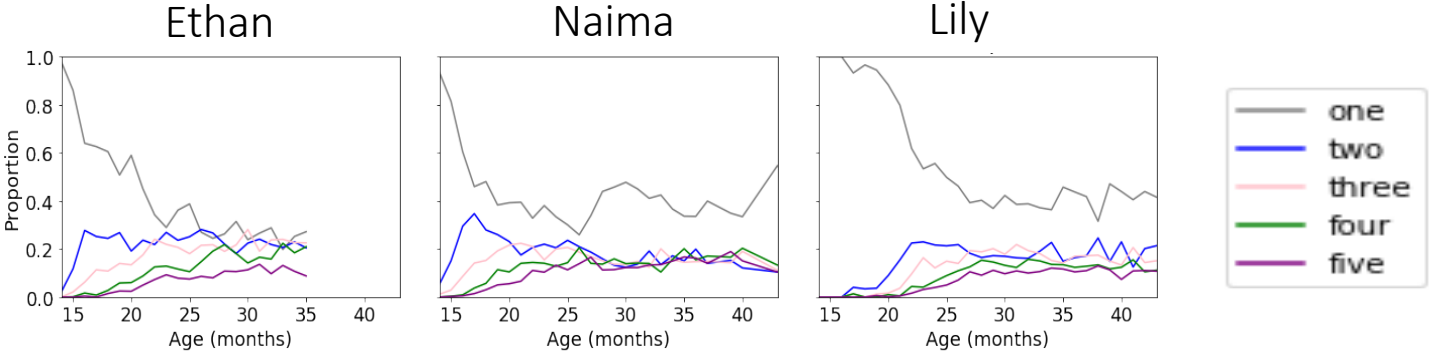
Breakpoint

Study 1 - Results



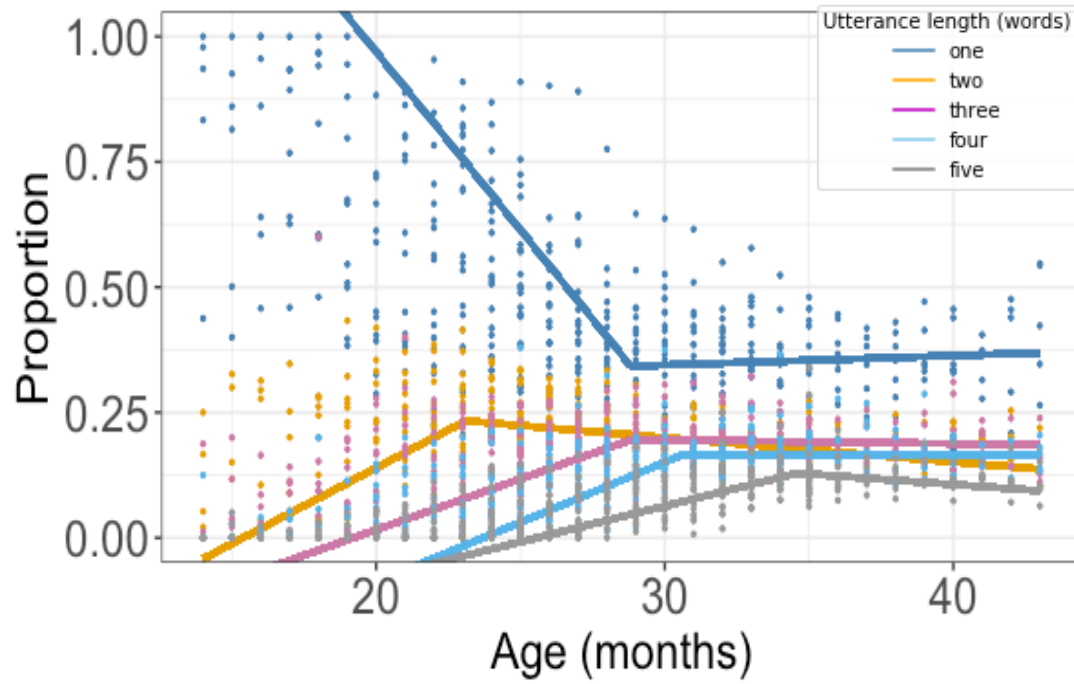
Smooth changes

Study 1 - Results



Smooth changes

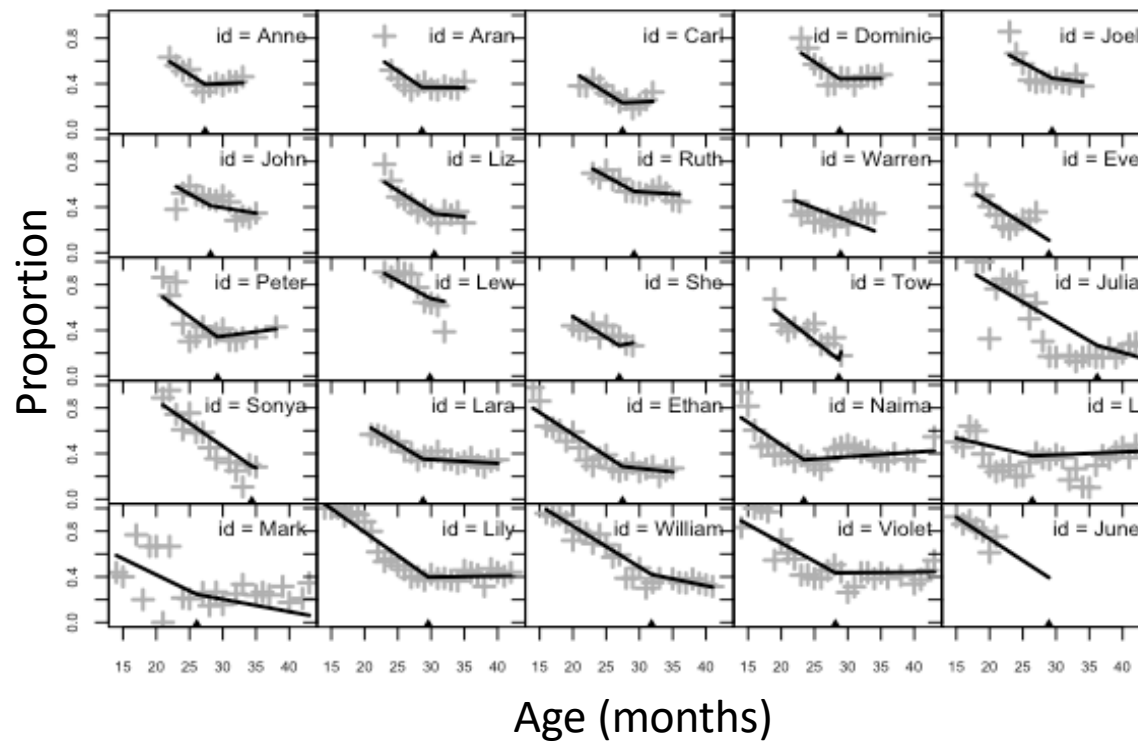
Study 1 - Results



Stage transition

$$\Delta_{\text{AIC}} \sim [30, 1279]$$

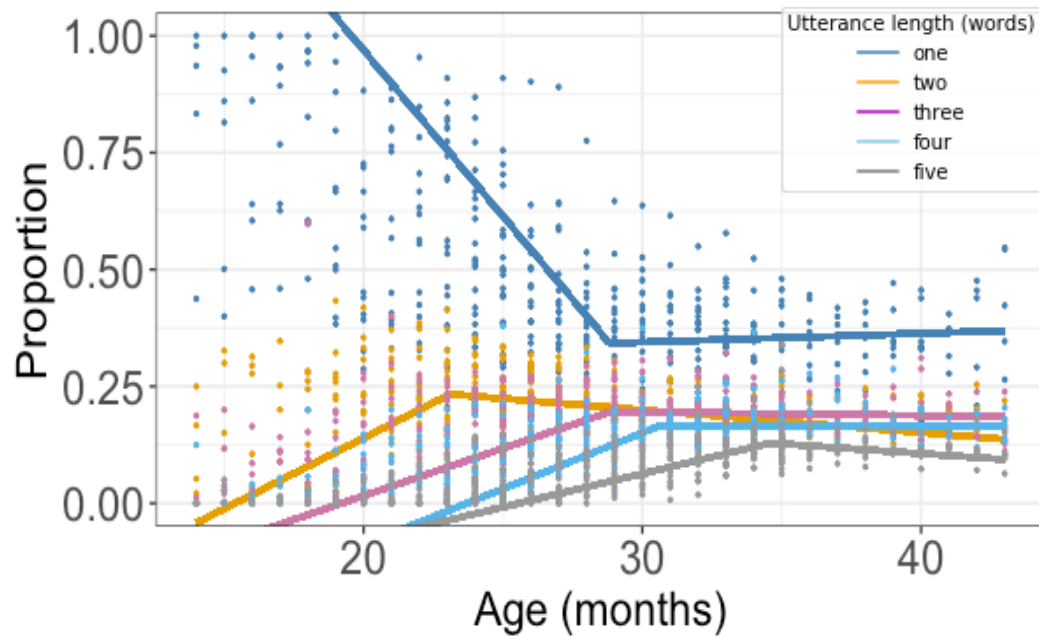
Study 1 - Results



Stage transition

$$\Delta_{\text{AIC}} \sim [30, 1279]$$

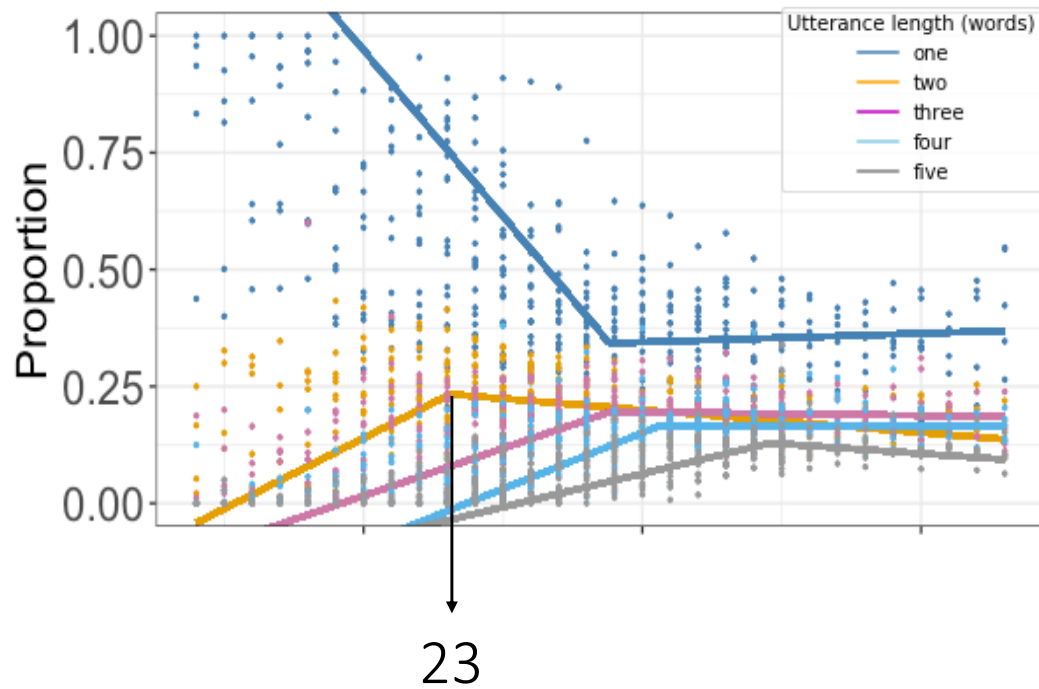
Study 1 - Results



Breakpoint by length

$$F(3, 72) = 147.27, p < .001$$

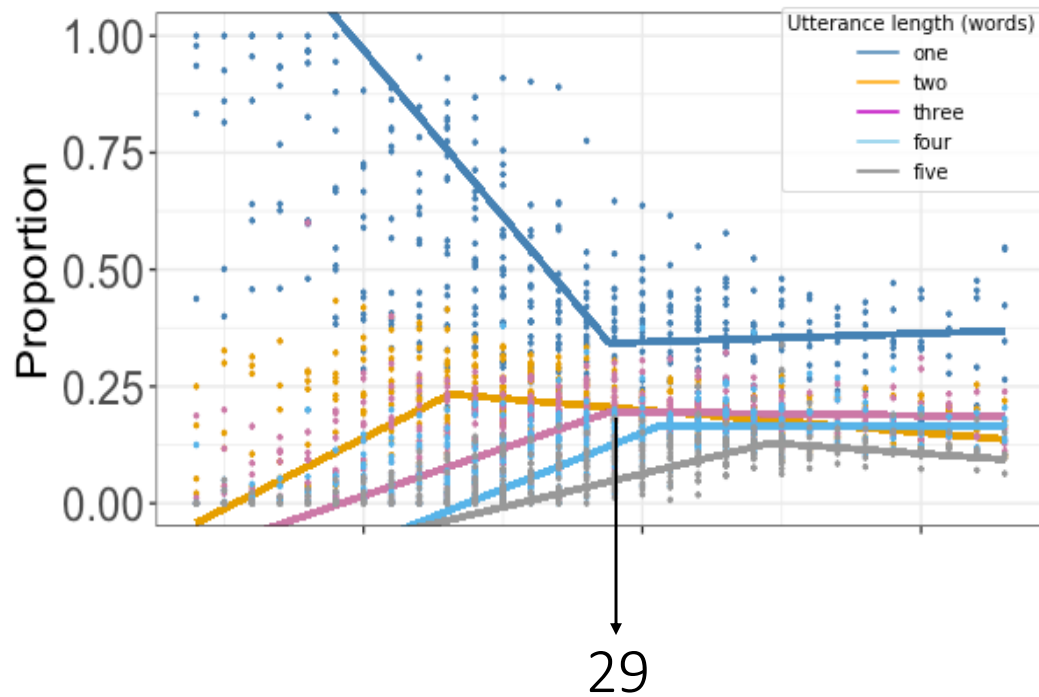
Study 1 - Results



Breakpoint by length

$$F(3, 72) = 147.27, p < .001$$

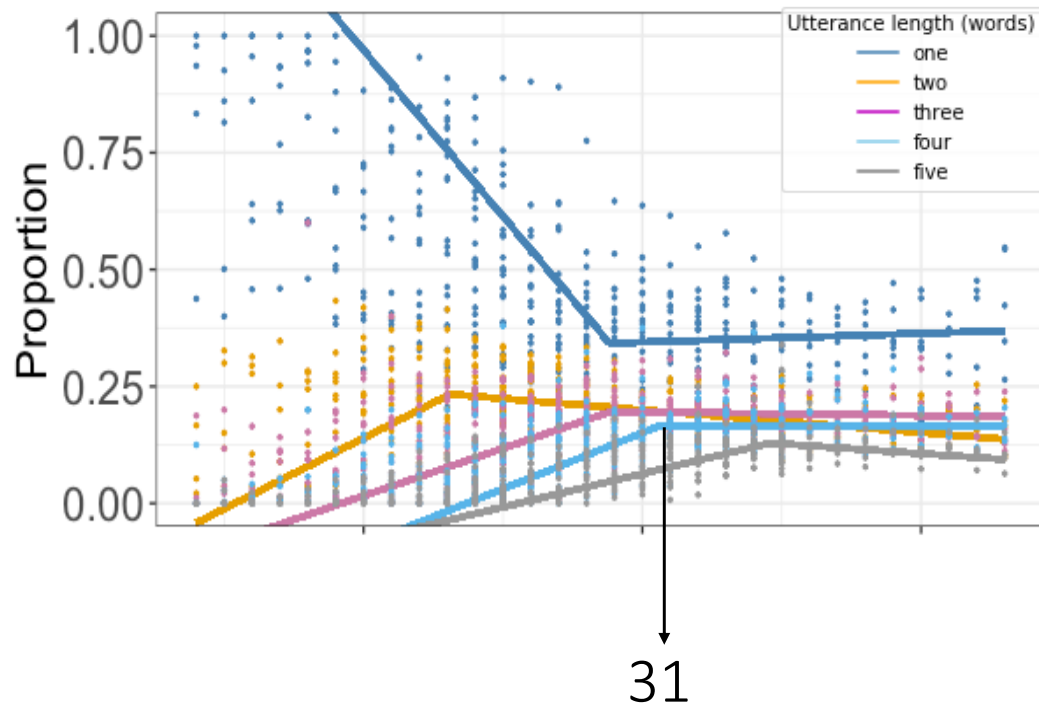
Study 1 - Results



Breakpoint by length

$$F(3, 72) = 147.27, p < .001$$

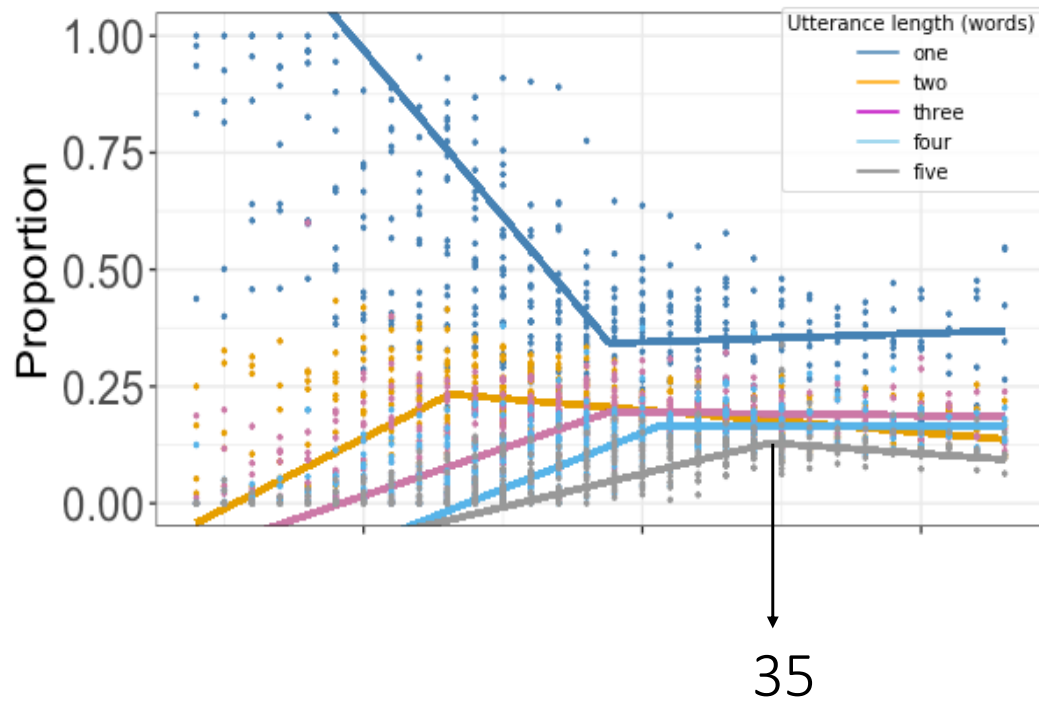
Study 1 - Results



Breakpoint by length

$$F(3, 72) = 147.27, p < .001$$

Study 1 - Results



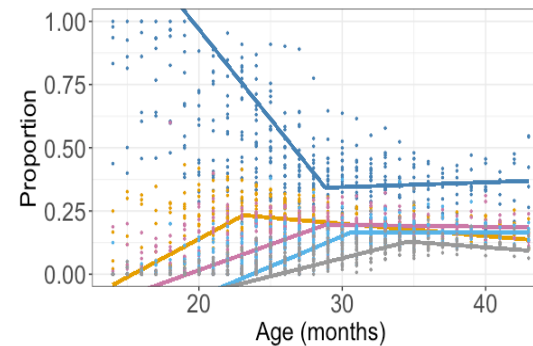
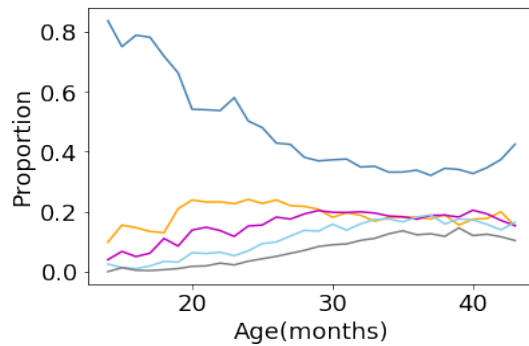
Breakpoint by length

$$F(3, 72) = 147.27, p < .001$$

Research questions

- **Study 1:** How do children's utterances change in length – continuously or in stages?

Smooth changes & Stage transitions



Research questions

- **Study 1:** How do children's utterances change in length – continuously or in stages?

Smooth changes & Stage transitions

The longer the utterances are, the later the breakpoint is

Research questions

- **Study 1:** How do children's utterances change in length – continuously or in stages?

Stage transition & Smooth changes

The longer the utterances are, the later the breakpoint is

- **Study 2:** What kind of underlying process could produce such behavioral change?

Study 2 – Probabilistic computational model

16 months

“Push.”

“Bottle rolling.”

“Key open door.”

26 months

“It fell.”

“Let me in.”

“But the crack in the track.”

(Providence. Evans & Demuth, 2012)



Xu, 2020

27

16 months

“Push.”

“Bottle rolling.”

“Key open door.”

26 months

“It fell.”

“Let me in.”

“But the crack in the track.”

(Providence. Evans & Demuth, 2012)



Xu, 2020

28

16 months

“Push.”

“Bottle rolling.”

“Key open door.”

26 months

“It fell.”

“Let me in.”

“But the crack in the track.”

(Providence. Evans & Demuth, 2012)



Xu, 2020

29



Two influences

Two scenarios

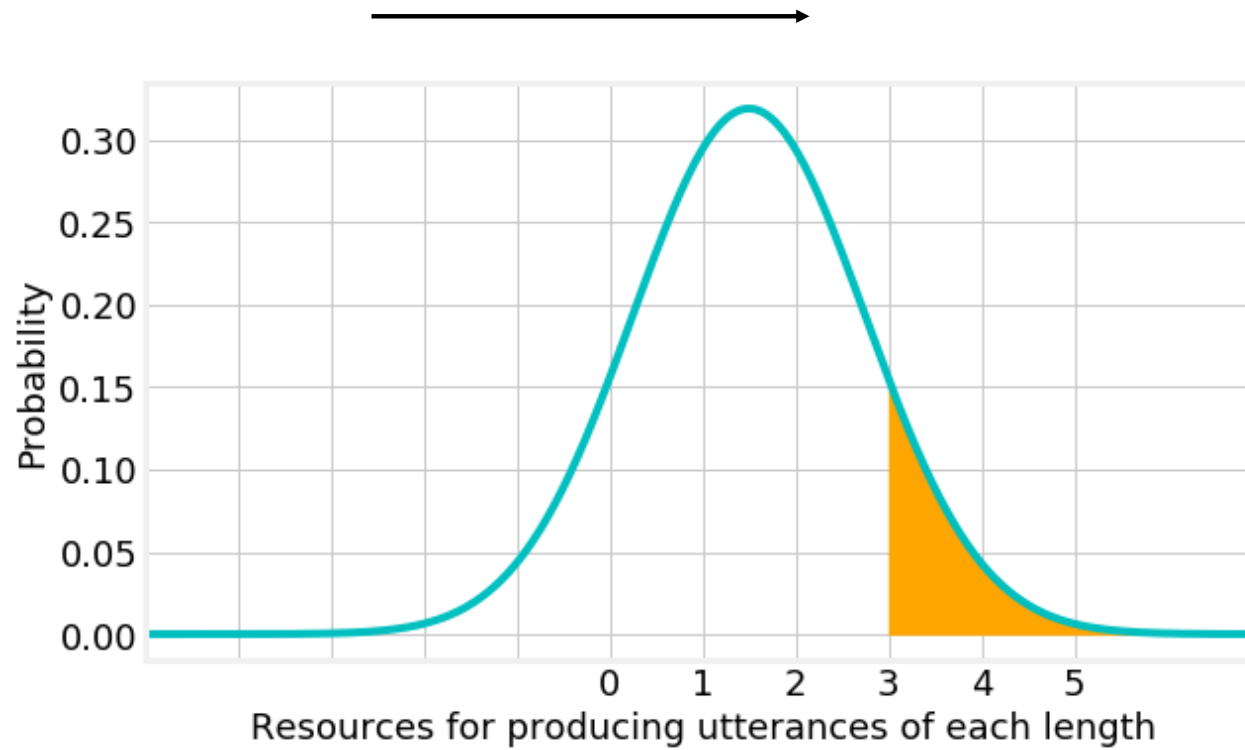
Influence 1



developmental resources

Working memory capacity, lexical knowledge...

Developmental resources

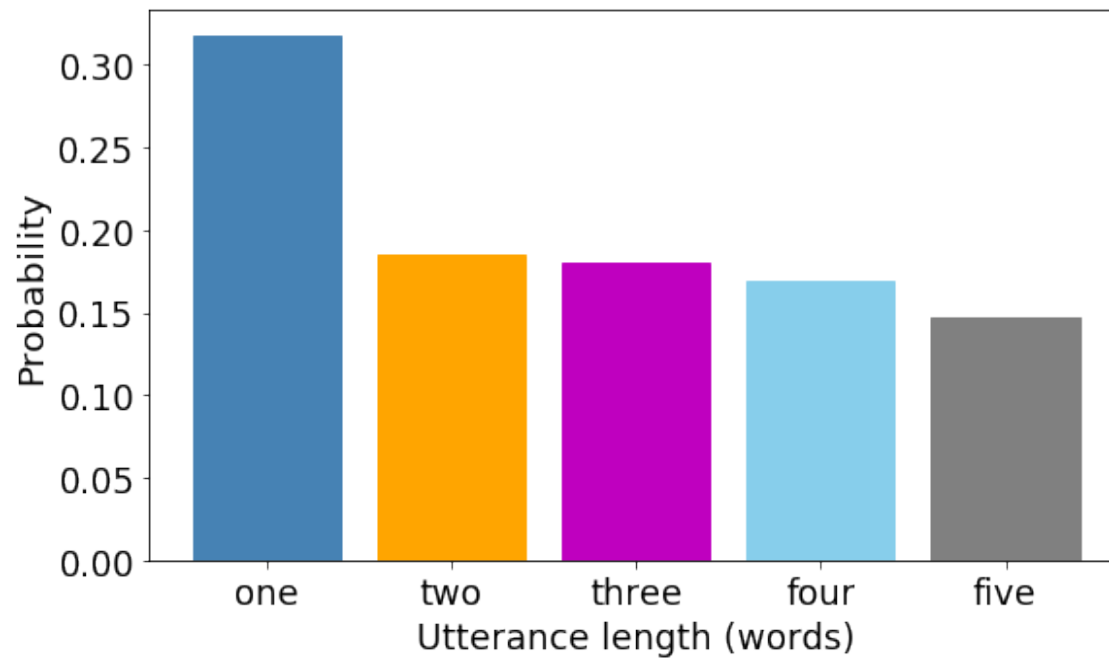


Influence 2

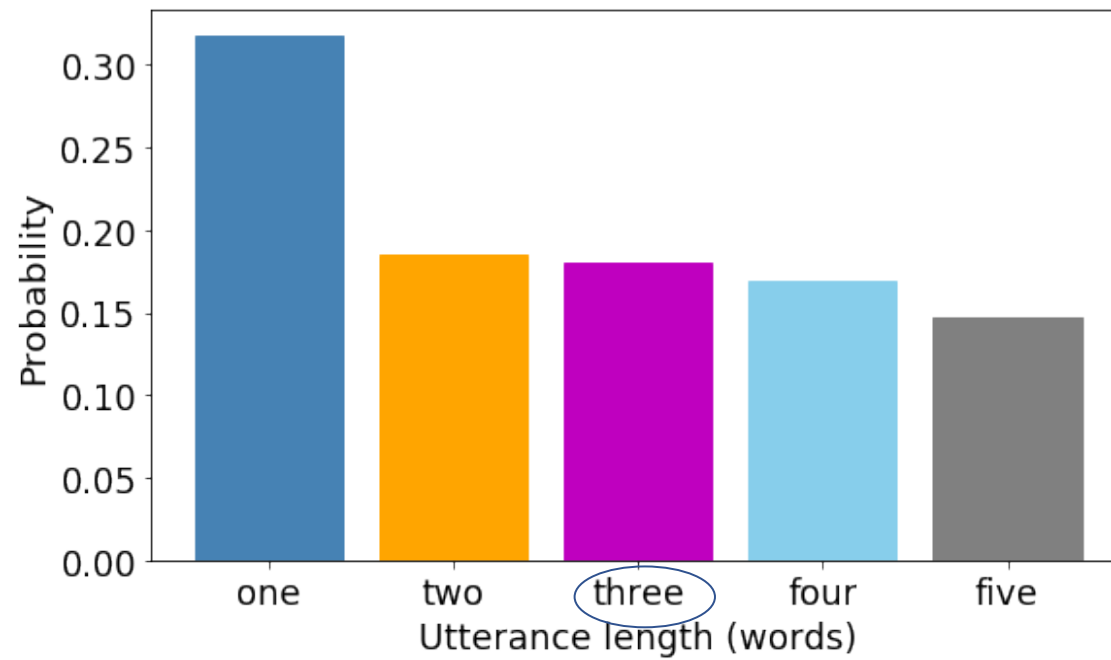


Intrinsic distribution

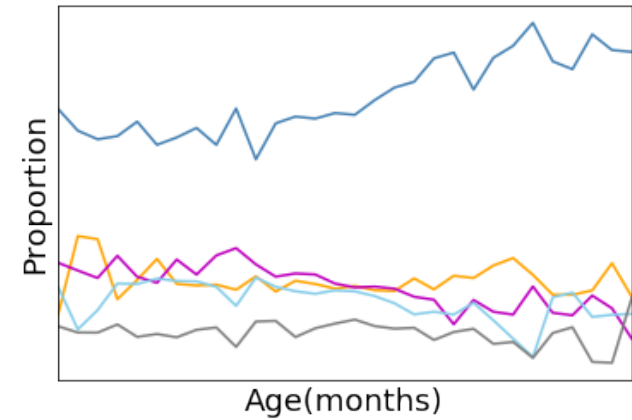
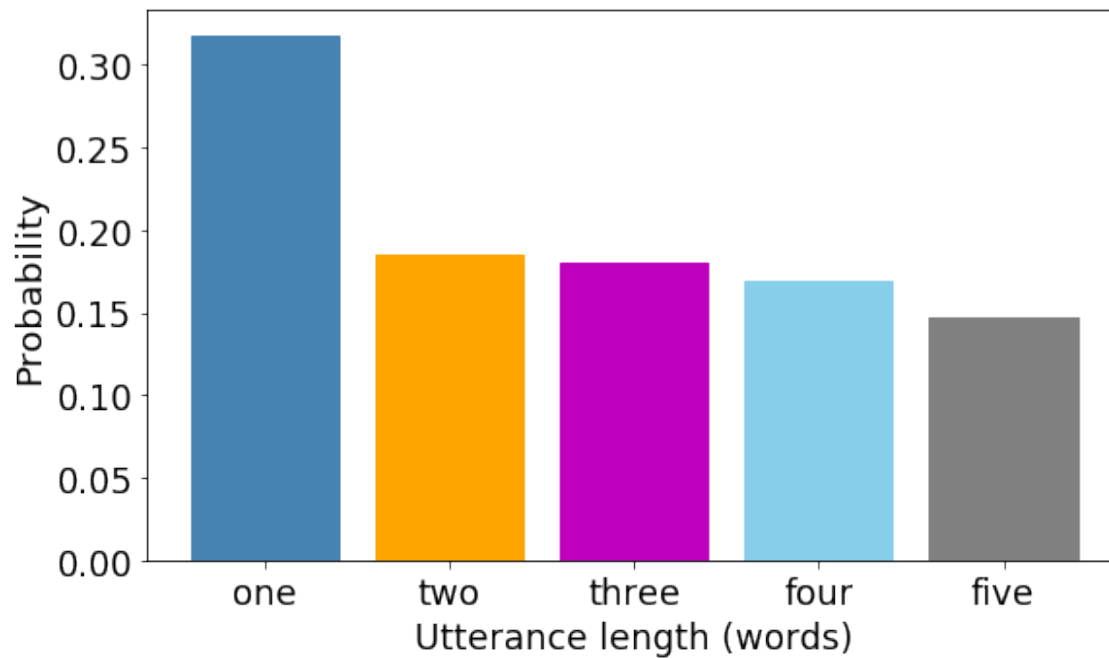
Intrinsic distribution – from parents' data



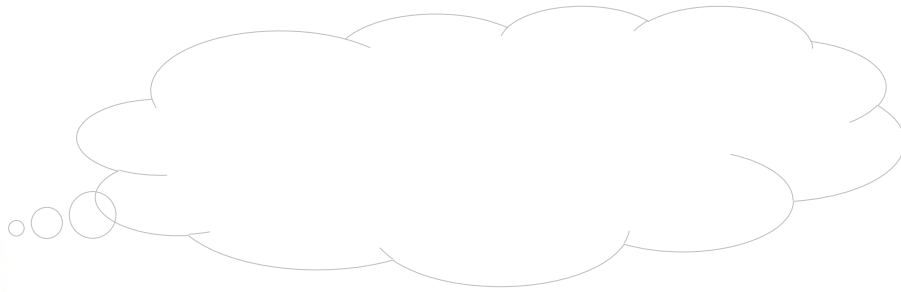
Intrinsic distribution – from parents' data



Intrinsic distribution – from parents' data

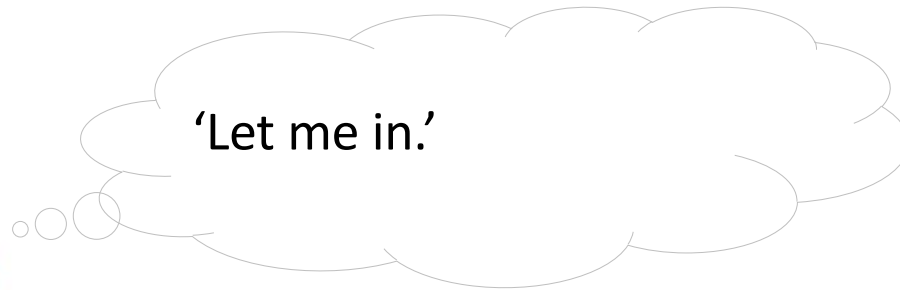


Two scenarios



Probability of producing N-word utterances

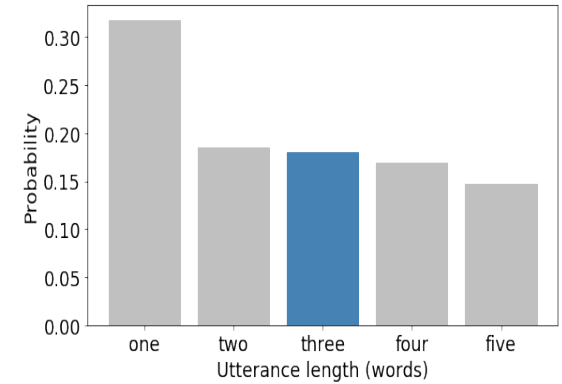
Scenario 1



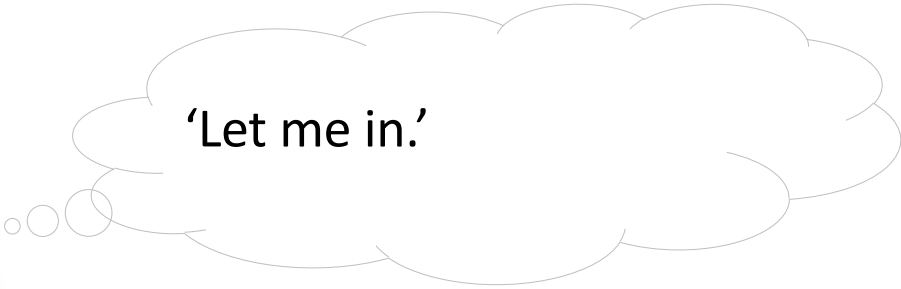
Adequate resources? YES



'Let me in.'



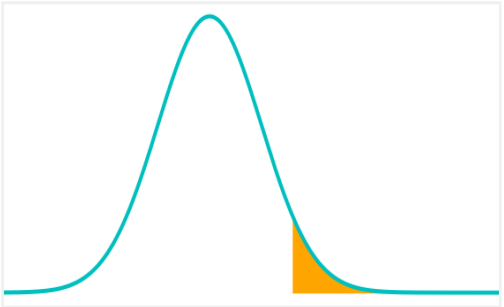
Scenario 1



Adequate resources? YES



'Let me in.'



Scenario 2

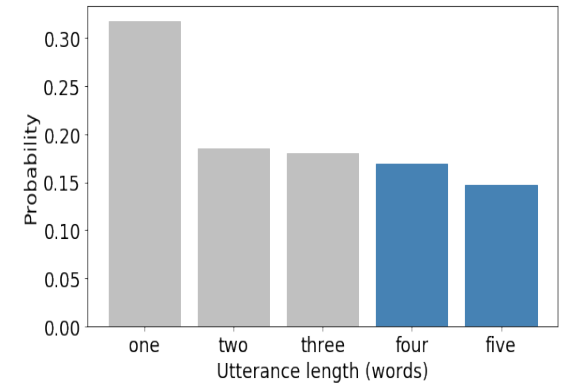


'The key opens the door.'

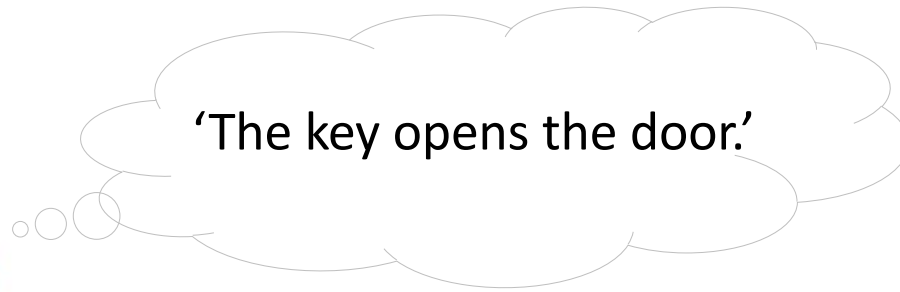
Adequate resources? **NO**



'Key open door.'



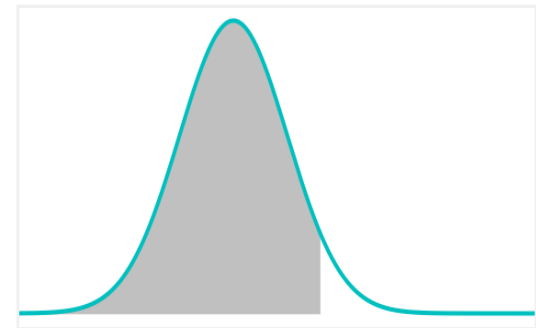
Scenario 2



Adequate resources? **NO**

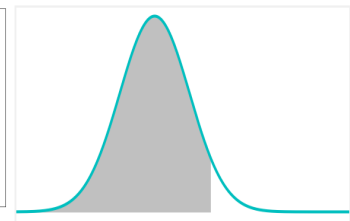
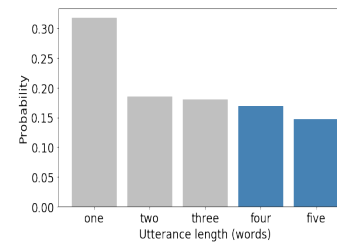
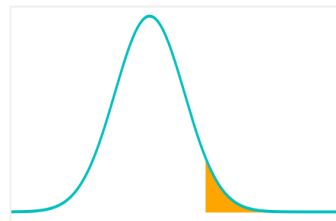
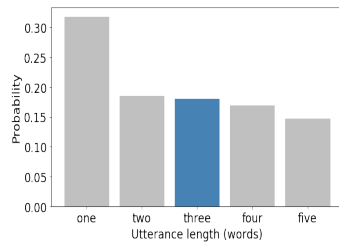


'Key open door.'



Two scenarios with two influences

Probability of producing N-word utterances



Study 2 - Probabilistic computational model

$$P(N)$$

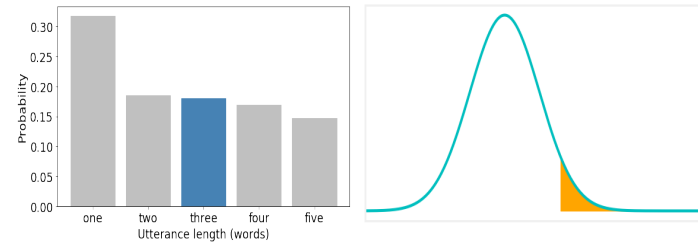
$$= ([P(T = N) * P(R \geq N)]$$

$$+ [P(T > N) * P(RMAX = N)]) * \left(\frac{1}{P(R \geq 1)} \right)$$

Study 2 - Probabilistic computational model

$$P(N) = ([P(T = N) * P(R \geq N)]$$

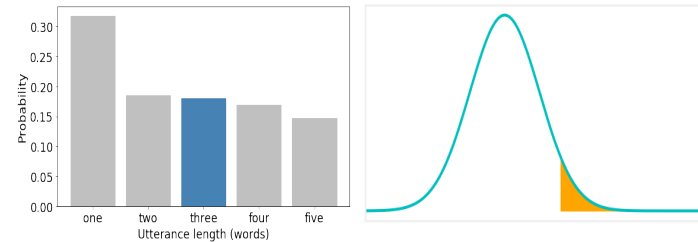
$$+ [P(T > N) * P(RMAX = N)]) * \left(\frac{1}{P(R \geq 1)} \right)$$



Study 2 - Probabilistic computational model

$$P(N) = ([P(T = N) * P(R \geq N)]$$

$$+ [P(T > N) * P(RMAX = N)]) * \left(\frac{1}{P(R \geq 1)} \right)$$

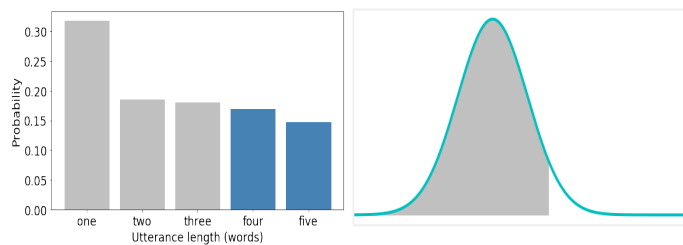


Study 2 - Probabilistic computational model

$$P(N)$$

$$= ([P(T = N) * P(R \geq N)]$$

$$+ [P(T > N) * P(RMAX = N)]) * \left(\frac{1}{P(R \geq 1)} \right)$$

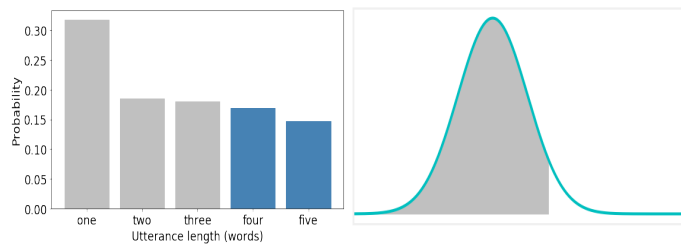


Study 2 - Probabilistic computational model

$$P(N)$$

$$= ([P(T = N) * P(R \geq N)]$$

$$+ [P(T > N) * P(RMAX = N)]) * \left(\frac{1}{P(R \geq 1)} \right)$$



Study 2 - Probabilistic computational model

$$\begin{aligned} P(N) &= ([P(T = N) * P(R \geq N)] \\ &+ [P(T > N) * P(RMAX = N)]) * \left(\frac{1}{P(R \geq 1)} \right) \end{aligned}$$

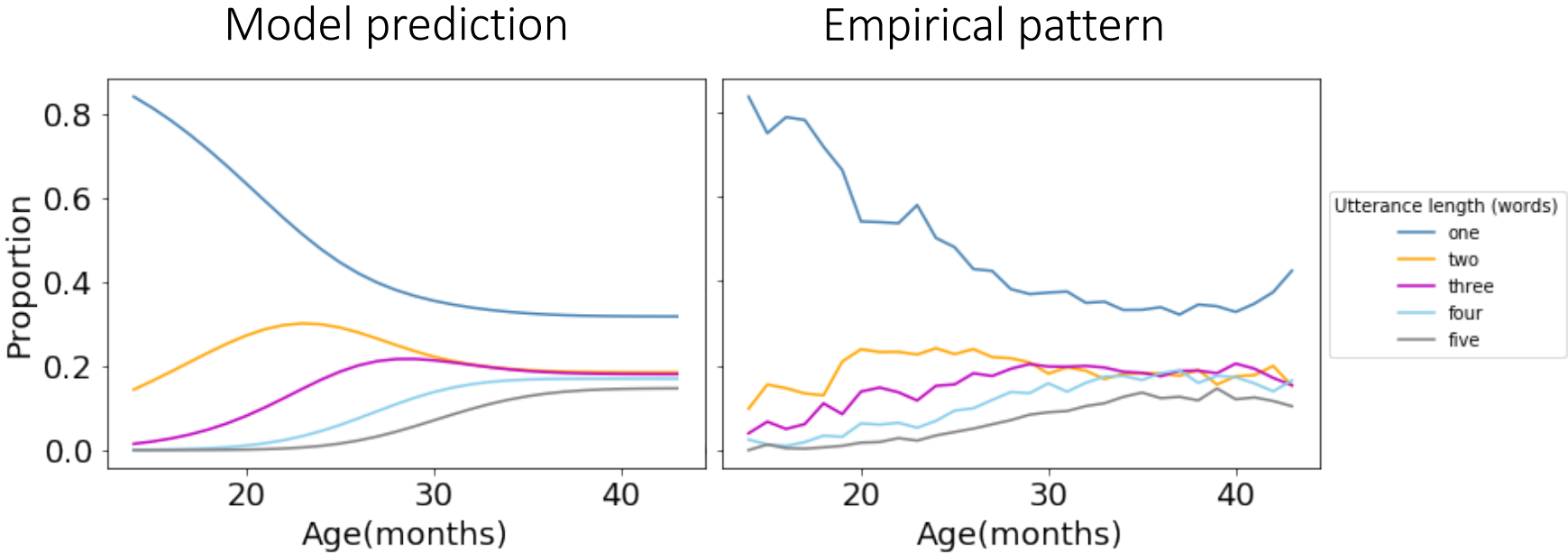
Study 2 - Methods

Analyses

The probabilities of utterance length 1-5 words as a function of age.

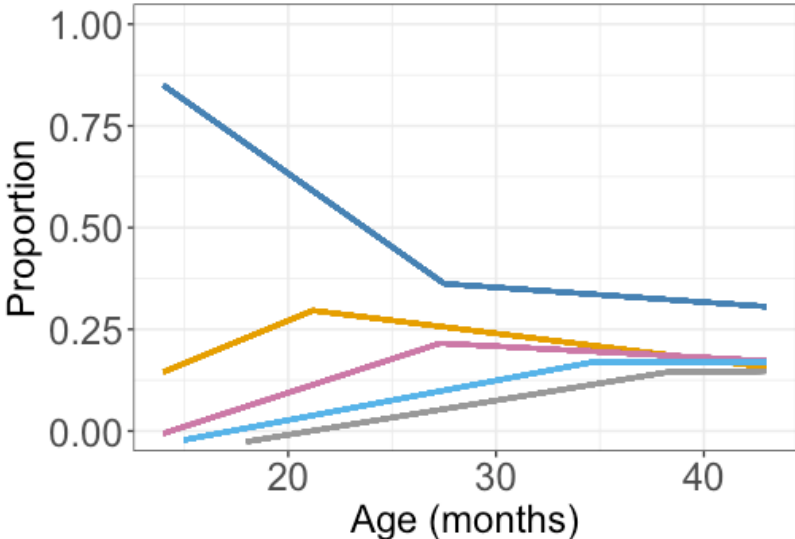
Segmented regression

Study 2 - Results

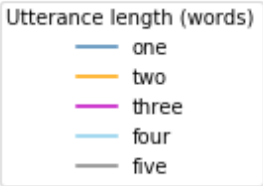
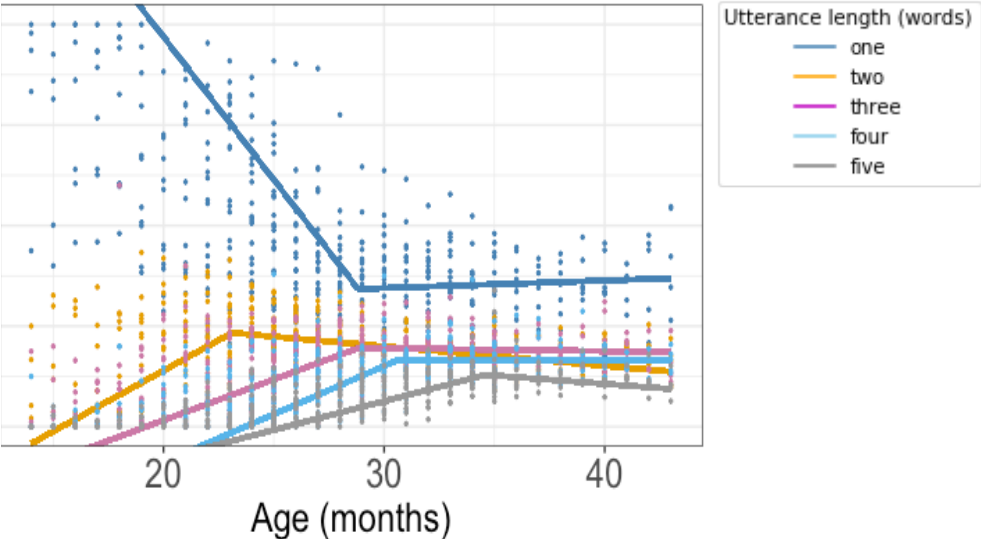


Study 2 - Results

Model prediction

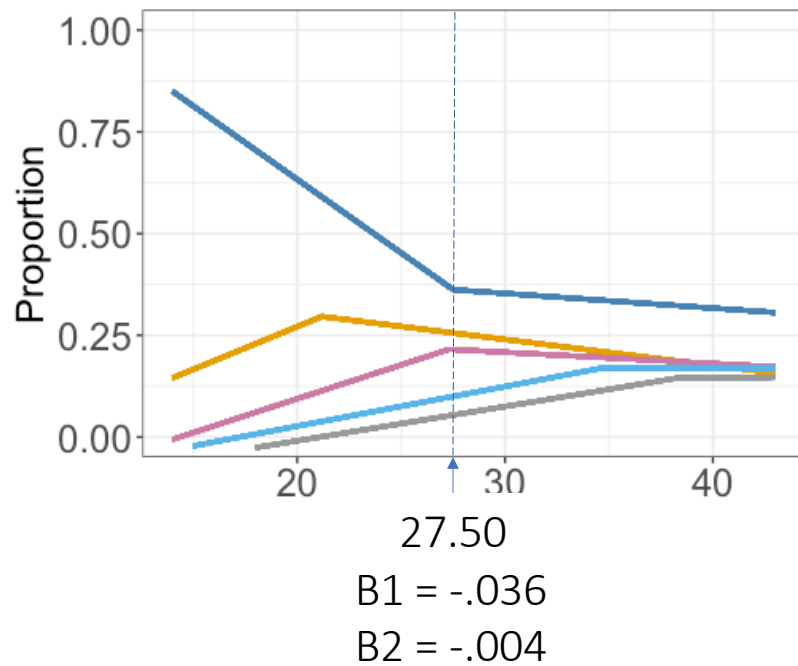


Empirical pattern

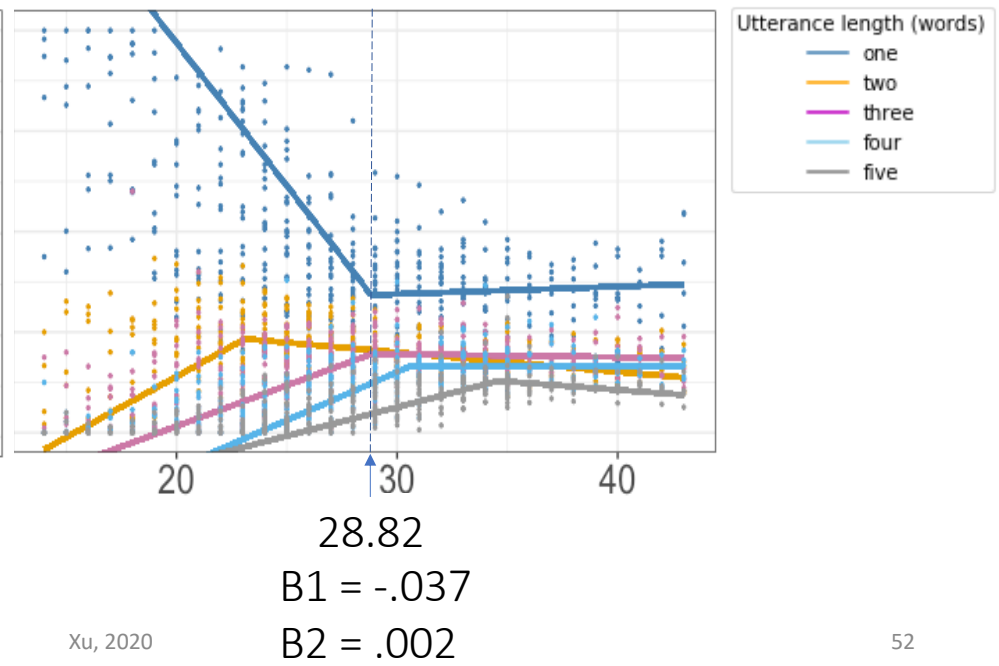


Study 2 - Results

Model prediction



Empirical pattern



Research questions

Study 2: What kind of underlying process could produce such developmental change?

- Continuously increasing resources x discrete production units
- Resources Adequate -> N
 Inadequate -> N-

Conclusion

- Smooth change & stage transitions
- Continuous development \leftrightarrow discrete behavior
- Inadequate resources \rightarrow drop elements

Conclusion

- Smooth change & stage transitions
- Continuous development x discrete behavior
- Inadequate resources -> drop elements **Ungrammaticality**

