Why do late talkers not say all the words that they understand?

Elissa Schiff, Mary Alt, Heidi Mettler, Allison Staib, Nora Evans-Reitz
Department of Speech, Language, and Hearing Sciences

THE UNIVERSITY OF ARIZONA Undergraduate Biology Research Program

VAULT

Introduction

- During language acquisition, children develop two vocabularies: a receptive vocabulary (the words they understand) and an expressive vocabulary (the words they say)
- Late talkers are children who do not say 200 words or use two-word phrases by the time they are two years old; they have small expressive vocabularies
- Children may learn words through focusing on words' phonological (sound) characteristics
  - Previous research into word learning has shown that for typically developing word learners, words that sound less like other words may be easier to learn because words with common sounds may shift attention to similar-sounding, already known words; words with unique sounds tend to stand out more and require less time for recognition
  - Late talkers have been found to use sounds as a word-learning cue for longer than their typically-developing peers (Stokes et al, 2011)
- So: late talkers do not say all the words they understand. The purpose of our study is to determine if this may be because of words' phonological characteristics
- Understanding why late talkers learn certain words may be helpful in selecting words to use for treatment

Our question: Is there a difference between the phonological characteristics of the words in late talkers' expressive and receptive vocabularies?

Our hypothesis: The expressive vocabularies will be of higher PND, PND frequency, PS average, and BP average than the receptive vocabularies.

Methods

1. We collected data on 4 late talkers' expressive and receptive vocabularies through parent report with the MacArthur-Bates Communicative Development Inventory. 677 commonly used words were included in the study

2. We used online neighborhood density (Vitevitch & Luce "English (Child Corpus)") and phonotactic probability calculators (Vitevitch & Luce, 2004) to analyze the children's vocabularies through four phonological characteristics:
   - **Phonological neighborhood density**
   - **Phonological neighborhood density frequency**
   - **Positional segment average**
   - **Biphone average**

3. Using these values, we compared the 4 late talker's expressive and receptive vocabularies using Bayesian t-tests

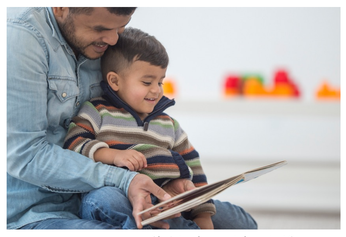
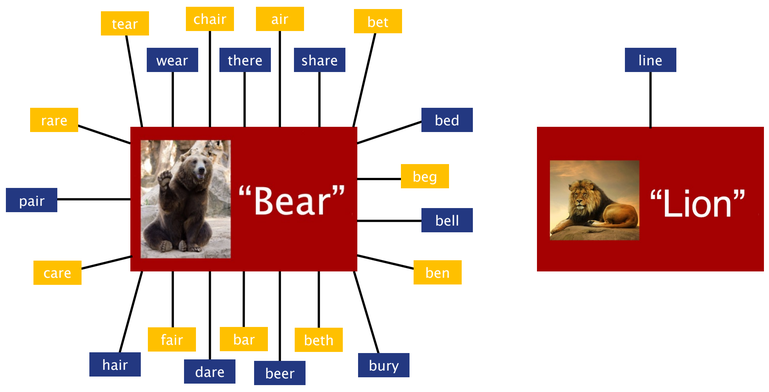


Image: Christopher Futcher, *FatCamera*

Image: Yan Lev, *Shutterstock*

Terms to know

- **Phonological neighborhood density**: The number of other words that differ from the word by the addition, deletion, or substitution of one phoneme



- **Neighborhood density frequency**: The average of the frequencies of occurrence of a word's phonological neighbors

- **Phonotactic probability**: How frequently a certain phoneme or phoneme sequence occurs in a certain position in a word
  - **Positional segment average**: Measures phonotactic probability with the frequency of a word's single phonemes occurring in the given position

PS = 1.22
PS avg. = 0.41

>

PS = 1.11
PS avg. = 0.28

  - **Biphone average**: Measures phonotactic probability with the frequency of a word's two-phoneme sequences occurring in the given position

BP = 1.02
BP avg. = 0.51

>

BP = 1.01
BP avg. = 0.36

Images: *Shutterstock*

Results

Phonological neighborhood density and frequency

	NEIGHBORHOOD DENSITY #	
	Receptive	Expressive
Valid	1249	143
Missing	0	0
Mean	7.548	10.126
Std. Deviation	7.382	8.810
Minimum	0.000	0.000
Maximum	33.000	29.000

$BF_{10} = 141$

	NEIGHBORHOOD DENSITY Frequency	
	Receptive	Expressive
Valid	1249	143
Missing	0	0
Mean	4330.861	10117.994
Std. Deviation	11524.838	17713.051
Minimum	0.000	0.000
Maximum	212837.000	122724.889

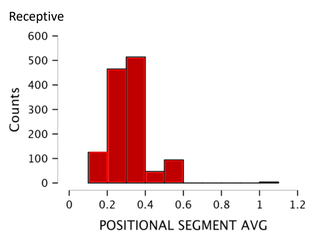
$BF_{10} = 89000$

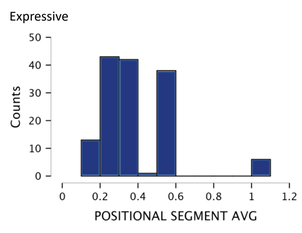
Bayesian t-tests show extreme evidence for differences in PND and PND frequency

Positional segment average

	POSITIONAL SEGMENT AVG	
	Receptive	Expressive
Valid	1249	143
Missing	0	0
Mean	0.324	0.392
Std. Deviation	0.101	0.176
Minimum	0.116	0.154
Maximum	1.019	1.019

$BF_{10} = 1.31 \times 10^{+9}$



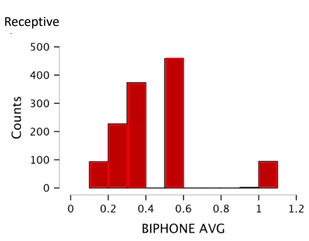


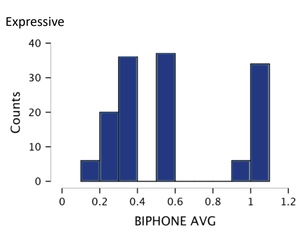
Bayesian t-tests show extreme evidence for differences in positional segment average

Biphone average

	BIPHONE AVG	
	Receptive	Expressive
Valid	1247	139
Missing	2	4
Mean	0.418	0.549
Std. Deviation	0.206	0.306
Minimum	0.101	0.144
Maximum	1.016	1.004

$BF_{10} = 2.45 \times 10^{8}$





Bayesian t-tests show extreme evidence for differences in biphone average

Discussion

- The results aligned with our hypothesis; the average values of all four of the phonological characteristics tested were higher for the late talkers' expressive vocabularies then their receptive vocabularies .
- This suggests **that out of all the words that late talkers understand, they learn to say the words with more common sounds before those with rarer sounds.**
  - This may be because late talkers need strong phonological representations of speech sounds (i.e., in many different words and in different positions within words) before they are able to produce these speech sounds in words
  - Late talkers also have limited words and speech sounds in their receptive vocabularies; this may explain why words of less common sounds do not "pop" more (and may in fact stand out less) than words of more common sounds during early word learning
- Our next steps will be to expand our sample size and apply these same analyses to more children.
- Our results may be used by clinicians in therapy settings and families at home to decide what words to use in speech therapy treatment for late talkers to facilitate their word learning and communication skills.

Acknowledgements

This project was funded by the Undergraduate Biological Research Project through the BIO5 Institute and NIDCD grant #R01 DC015642-06.

Thank you to all the L4 lab members who have contributed to MCDI data entry for VAULT and to all our wonderful VAULT toddlers and families for their participation.

Thank you to Mrs. Andrea Jeffers Wellington and Mr. Doug Wellington for funding my UBRP position.

References

Stokes, S. F., Kern, S., & Dos Santos, C. (2011). Extended statistical learning as an account for slow vocabulary growth*. Journal of Child Language, 39 (1), 105-129. https://doi.org.10.1017/s0305000911000031

Vitevitch, M. S., and Luce P. A. A Web-Based Interface to Calculate Phonotactic Probability for Words and Nonwords in English, Behavior Research Methods, Instruments, & Computers, 36 (3), 481-487. https://doi.org/10.3758/bf03195594.

Vitevitch, M. S., and Luce, P. A. English (Child Corpus). Similarity Neighborhood Calculator, https://calculator.ku.edu/density/English/words.