More Than Wordplay: An Analysis of Word-form Variability in Speech to Infants

Charlotte Moore & Elika Bergelson
Duke University

Introduction

Children hear many surface forms of the same lemma.
E.g. lemma = dog
Surface forms = dogs, doggy, puppy dog, etc.
Some added morphology has no systemic meaning, i.e. diminutives*: "doggy" and "dog" can refer to same object.
Semantic-syntactically void morphology = **wordplay.**
Exploratory analysis of:
• How words actually appear in infants' daily lives
• What the implications of that variability are

Research Questions:
1. **What does word form variability in an infant's input look like, by word and by family?**
2. **Which words undergo wordplay and why?**
3. **What does surface form variability mean for early word learning?**

*Diminutives primarily denote familiarity in English (Savickiene & Dressler, 2007).

Methods

Corpus analysis of SEEDLingS. 44 infants from 6-17 mo
Data collection:
• Monthly day-long audio recordings
• Monthly hour-long video recordings
Annotation of object words:
Coded as said (e.g. birdies) and in lemma form (e.g. bird)
Also coded for speaker, utterance type, and object presence

Variability analysis:
spoken words and lemma differed were coded for up to 3 alternations.
E.g. "birdy-birdies": plural, reduplication, y-epenthesis

Patterns

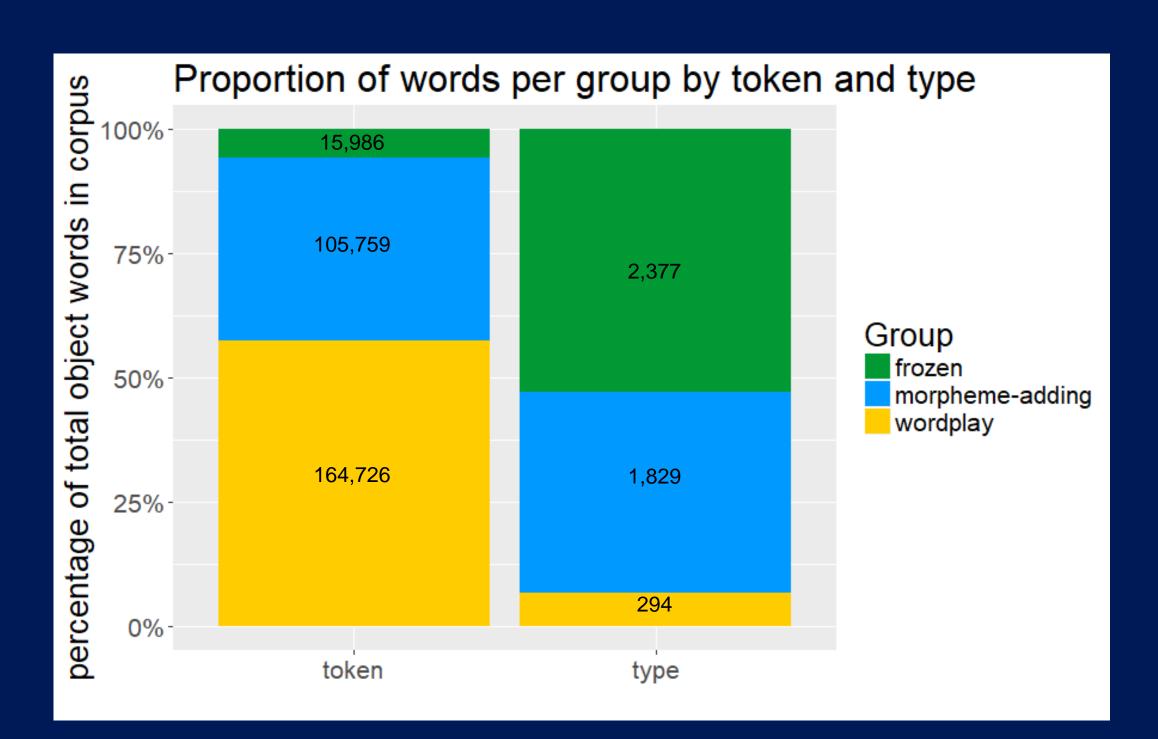**Frozen words** – the lemma and the surface form always match.
• Examples: stomach, stairs, comb
• Letters & numbers: one, A (crystalized, rote-learned)

Morpheme adding words – the lemma and surface form differ, but through typical, meaningful alternations like pluralization or modifier addition.
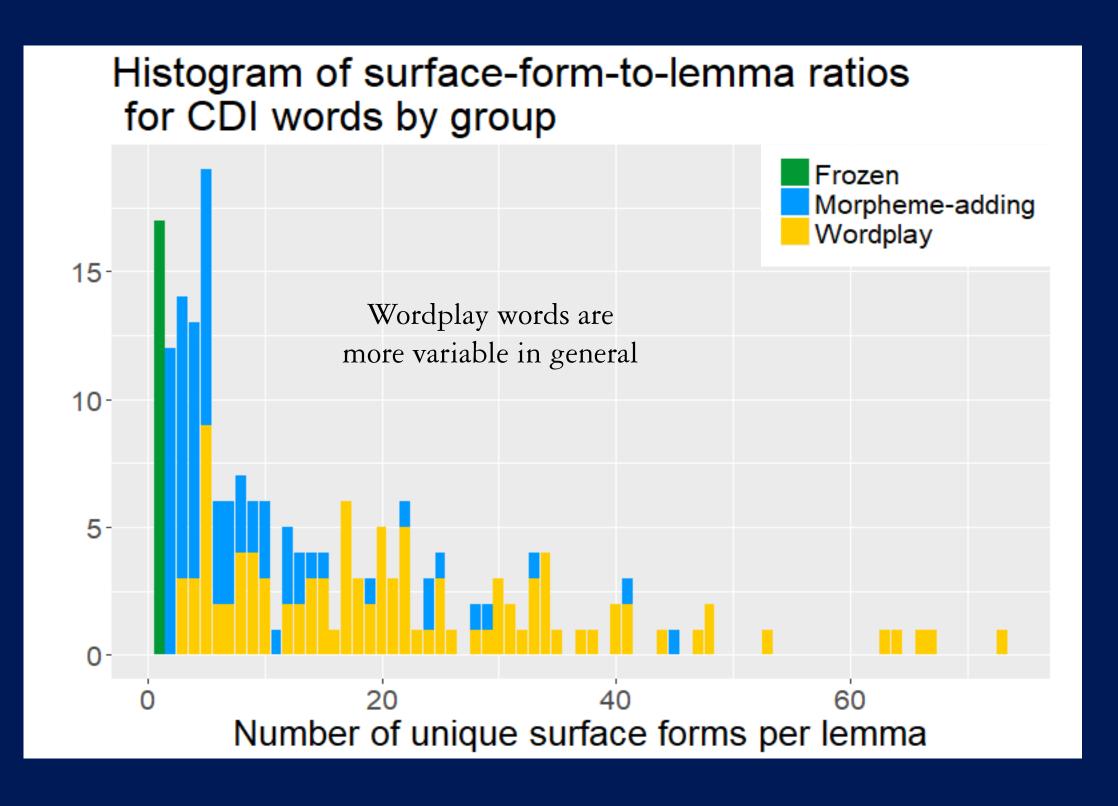• Examples: knee, apple, tree
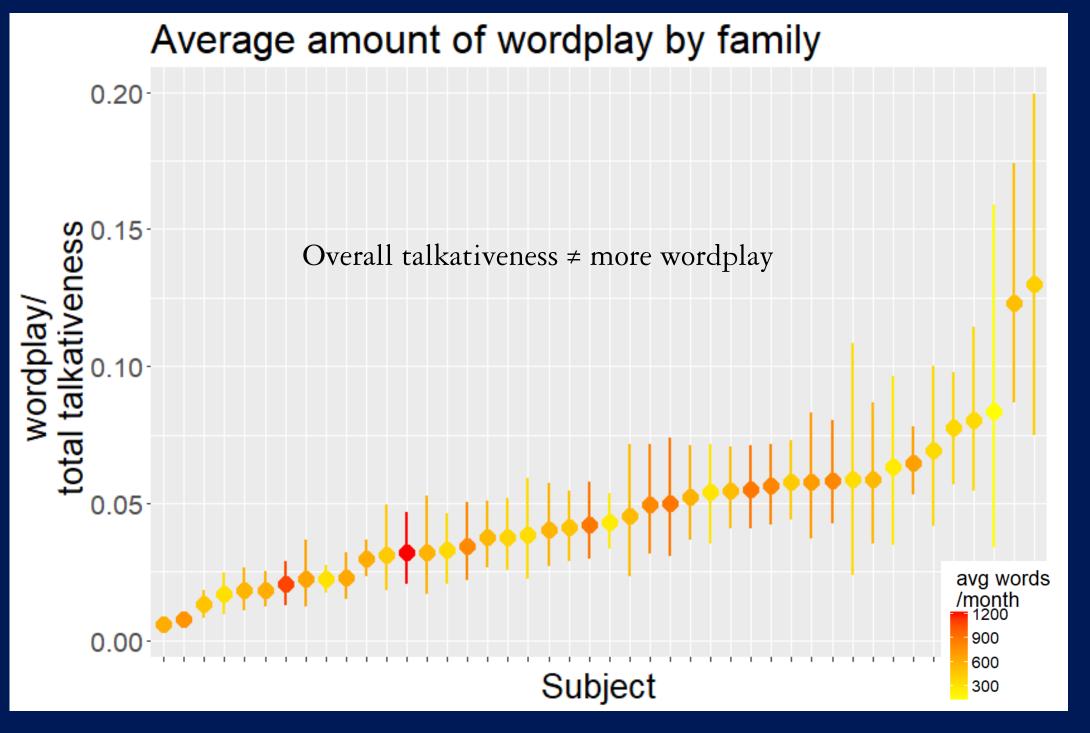• 90% of all types of change were morpheme-adding changes

Wordplay words – lemma and surface forms can differ in either meaningful ways like morpheme adding, but <u>also</u> include meaningless alternations.
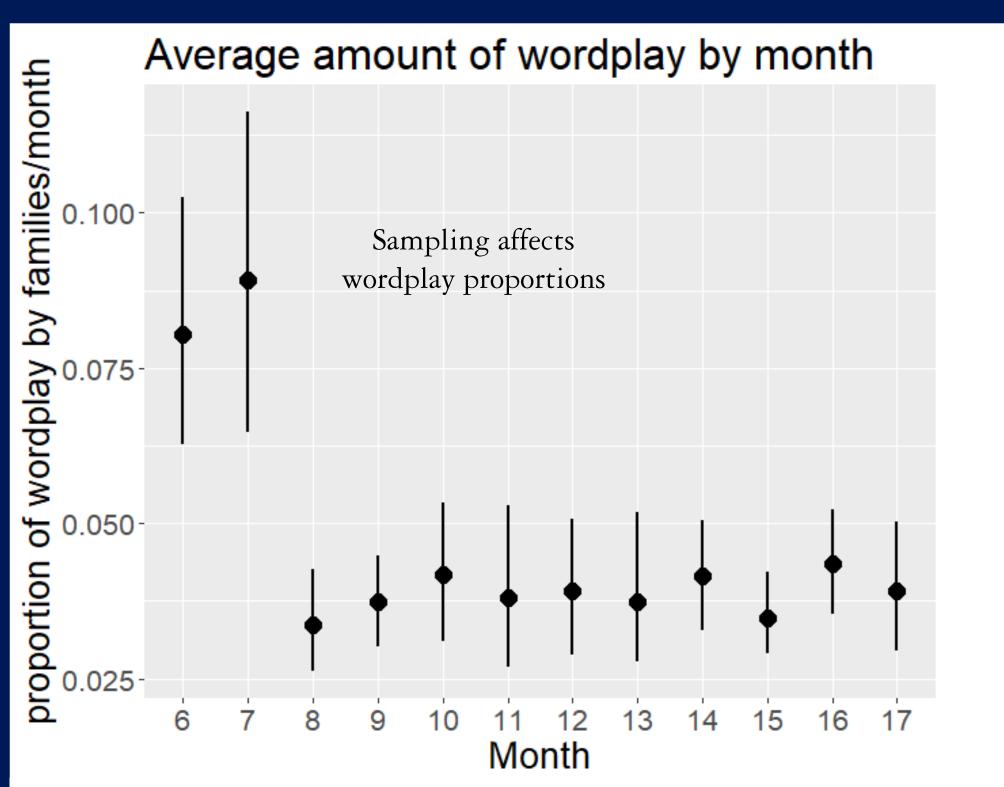• Examples: dog, baby, foot

Proportion of words per group by token and type

Word	Example alternations	# of forms
Tooth	Toofer, tootheroo, teethies	34
Baby	Bebe, robo-baby, baby-baby-baby	48
Diaper	Diapey, didey bidey, diaperdoo	66
Apple	Mini apple, apple slices, apples	22
Coat	Raincoat, fur coat, coats	13
Wallet	Wallet (no alternation)	1

Histogram of surface-form-to-lemma ratios for CDI words by group

Wordplay words are more variable in general

Average amount of wordplay by family

Overall talkativeness ≠ more wordplay

Average amount of wordplay by month

Sampling affects wordplay proportions

Only mos 6&7 were annotated from morning to bedtime, vs 8-17mos, where only most talkative hrs were annotated. Qualitative difference throughout the day!
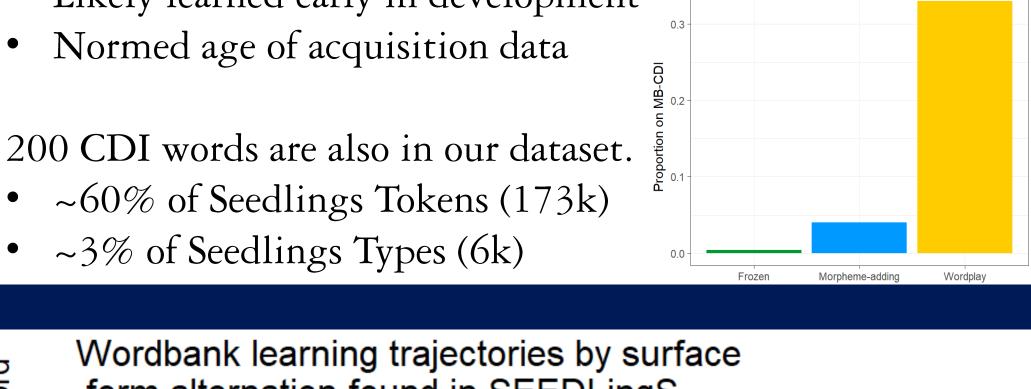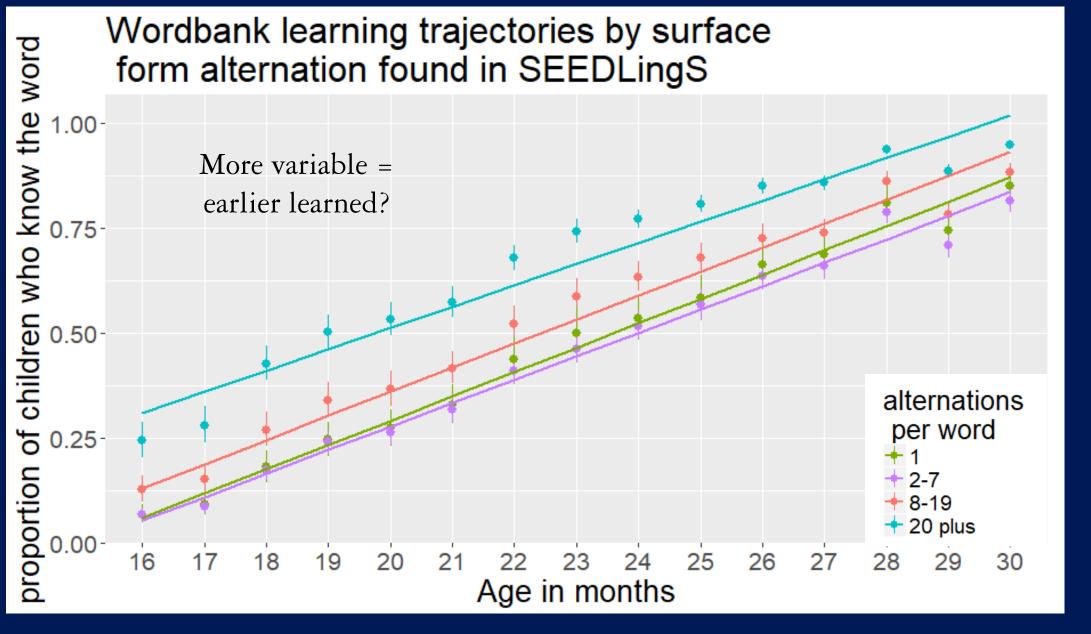
Early Learning & Alternations

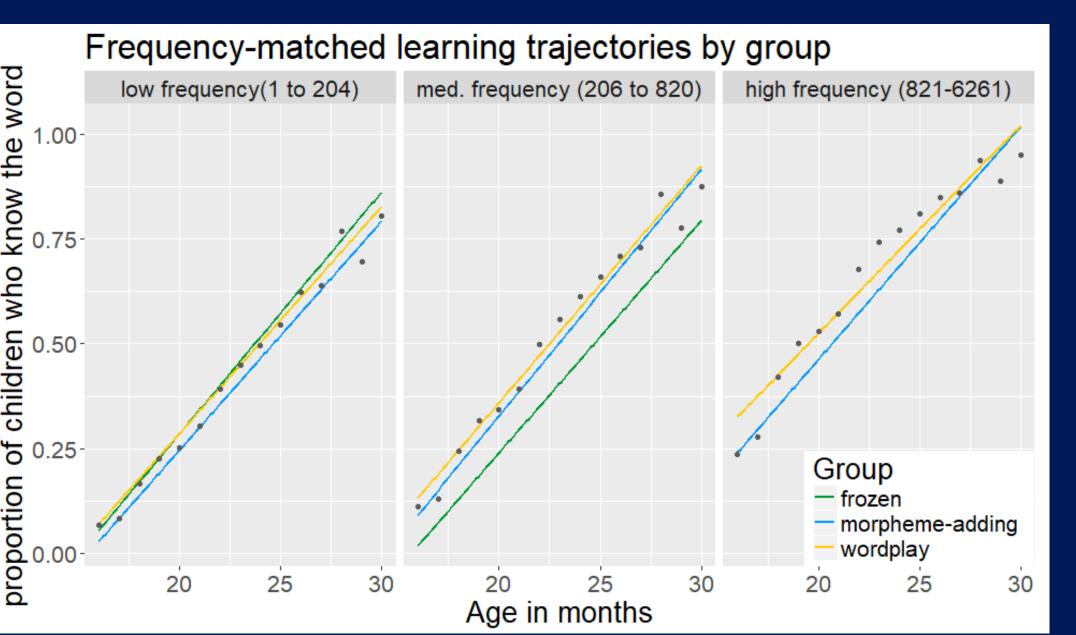Wordbank CDIs (n=~14k) as proxy for learning trajectories (Frank et al., 2016)
Why focus on CDI words?
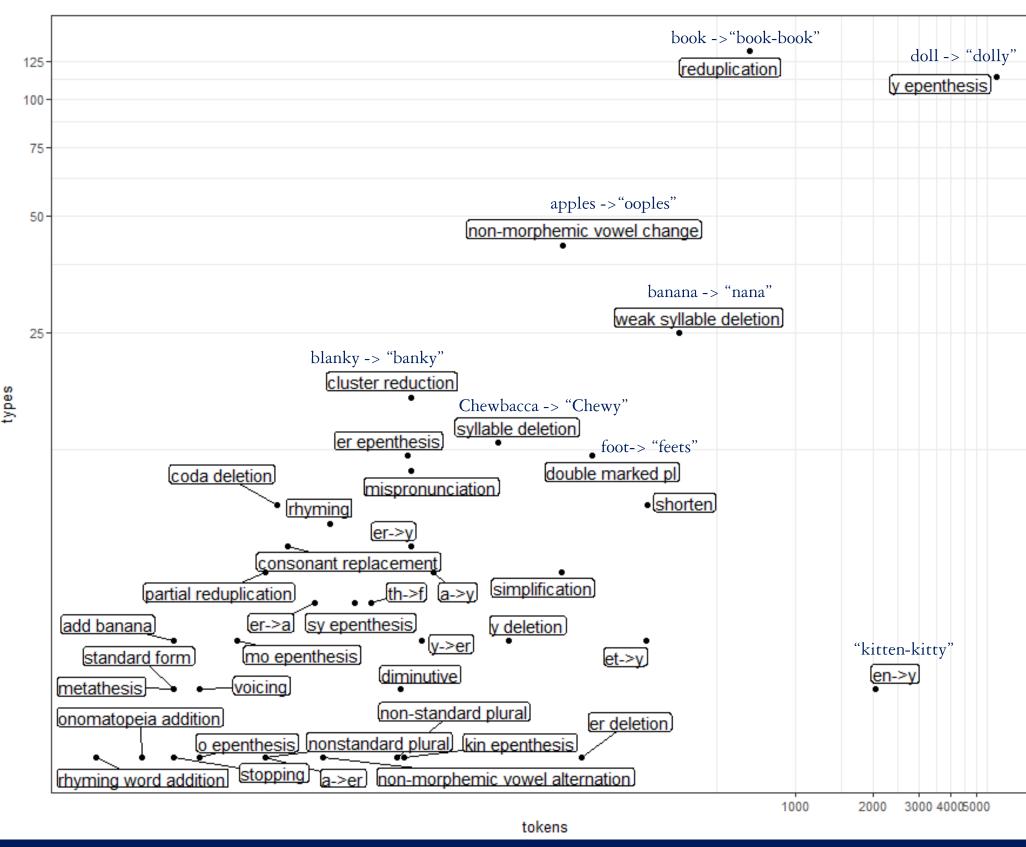• Likely learned early in development
• Normed age of acquisition data

200 CDI words are also in our dataset.
• ~60% of Seedlings Tokens (173k)
• ~3% of Seedlings Types (6k)

Words in each group represented on the MB-CDI

Wordbank learning trajectories by surface form alternation found in SEEDLingS

More variable = earlier learned?

Frequency-matched learning trajectories by group

Wordplay types

Discussion & Conclusions

1. **What does variability look like by word? by family?**
• Tokens usually vary from the dictionary form.
• Wordplay is idiosyncratic by family.

2. **Which words undergo wordplay?**
• More frequent = more likely to have wordplay
• On MB-CDI = more likely to have wordplay
• Less talkative time of day = more wordplay

3. **Surface form variability and early word learning?**
• Variability & frequency are correlated (r = 0.82).
• Both play some role in learning trajectory.

Wordplay could also:
• Aid word segmentation (Kempe et al., 2005; Brooks et al., 2011)
• Increase word saliency (Gervain & Werker, 2008)
• Act as pragmatic cue for infant (Savickiene & Dressler, 2007)

Future directions:
• Semantic variation (count vs mass distinctions)
• Look at effects of different kinds of wordplay
• Increase scope beyond CDI

References

Brooks, P. J., Kempe, V., & Donachie, A. (2011). Second Language Learning Benefits from Similarity in Word Endings: Evidence from Russian: Diminutives Benefit L2 Learning. *Language Learning*, *61*(4), 1142–1172.

Gervain, J., & Werker, J. F. (2008). How Infant Speech Perception Contributes to Language Acquisition. *Language and Linguistics Compass*, *2*(6), 1149–1170.

Kempe, V., Brooks, P. J., & Gillis, S. (2005). Diminutives in child-directed speech supplement metric with distributional word segmentation cues. *Psychonomic Bulletin & Review*, *12*(1), 145–151.

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2016). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*.

Savickiene, I., & Dressler, W. U. (2007). *The Acquisition of Diminutives: A cross-linguistic perspective*. John Benjamins Publishing.