Regularization and probability-matching as generalization

Christine Soh Yue, Charles Yang, and Katie Schuler

University of Pennsylvania

Introduction

How do learners learn when to **regularize** vs. **probability match**?

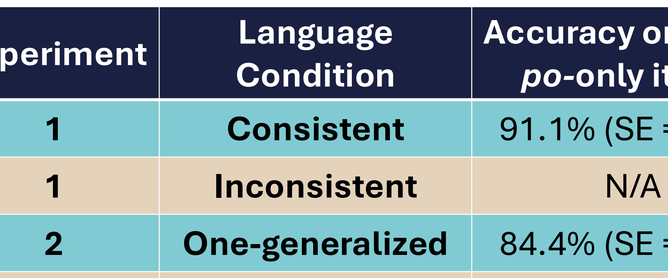
- Two separate learning mechanisms?
  - Productive generalizations (e.g. Pinker, 1999; Yang, 2016) vs. probabilistic mechanisms (e.g. Kroch, 1989)
- Regularization:
  - a necessary first step to learning variation? (e.g. Shin & Miller, 2022)
  - exceptional behavior, with probability matching as the default? (e.g. Hudson Kam & Newport, 2005; 2009)

The Proposal

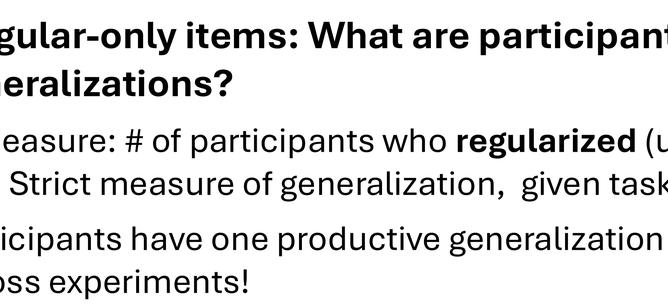
Productive generalization underlies regularization and probability-matching

Key first step: distinguish the two learning scenarios

1 generalization → categorical rule (regularization)



Multiple generalizations → systematic variation



Experimental Design

Artificial language: **plural markers** with no structural conditioning (e.g. Schuler et. al., 2016)

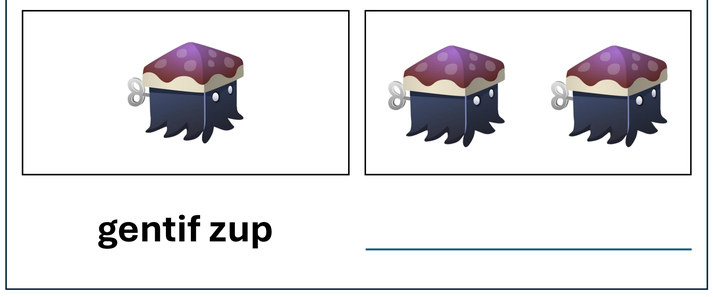
- Singular: "*gentif* NOUN" Plural: "*gentif* NOUN *ka/po*"
- Plural markers: fixed token frequencies (2:1), varied type frequencies (application to nouns) by condition

Experimental Task

- Exposure followed by a wug test (Berko, 1958): **singular-only items** and **observed-plural items** (seen in both singular and plural forms) mixed in exposure and at test, decreasing task effects for adults

Exposure: Active noun-learning

Test: Wug test



Experiment 1

- Perfectly consistent and inconsistent languages (e.g., Austin et al., 2022; Wonnacott et. al., 2008)
- 60 adult participants (n=30 in each condition) after exclusions recruited from Prolific

Experiment 2

- Partially consistent languages (e.g., Wonnacott et al., 2017)
- Inputs matched for mutual information & entropy
- 60 adult participants (n=30 in each condition)

Experiment	Language Condition	Singular (# nouns)	No plural	Plural				# Generalizations supported
				ka	ka / po	po	Total	
1	Consistent	14	6	6	0	2	8	1
1	Inconsistent	14	6	0	8	0	8	2
2	One-generalized	14	6	5	2	1	8	1
2	Both-generalized	14	6	3	2	3	8	2

Results

Observed-plural items: Did participants learn the input? Yes!

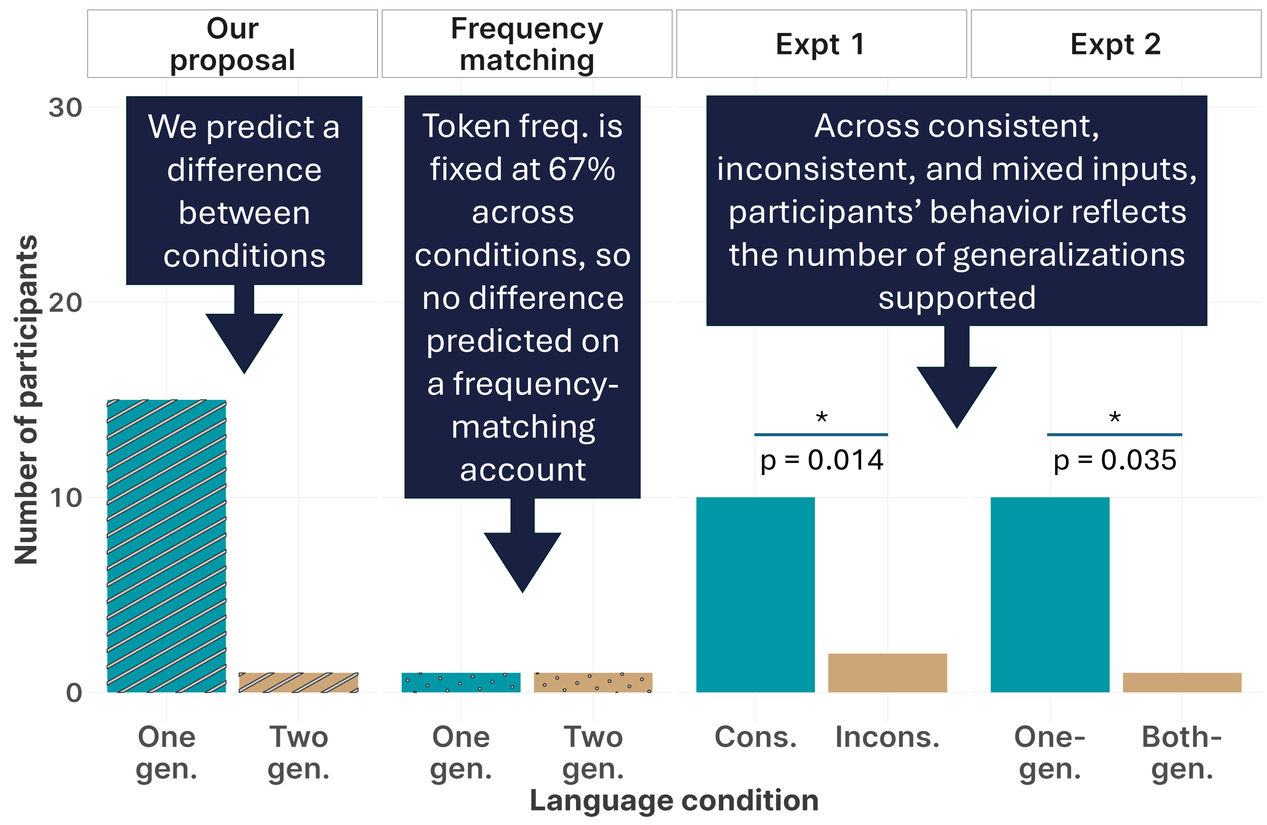
- High accuracy on *ka*-only and *po*-only items
- Approximately probability matching behavior on alternating items

Experiment	Language Condition	Accuracy on *ka*- or *po*-only items	% Dominant marker on alternating items
1	Consistent	91.1% (SE = 1.06)	N/A
1	Inconsistent	N/A	62.2% (SE = 1.32)
2	One-generalized	84.4% (SE = 1.51)	57.7% (SE = 2.70)
2	Both-generalized	78.4% (SE = 1.82)	54.5% (SE = 2.72)

Singular-only items: What are participants' productive generalizations?

- Measure: # of participants who **regularized** (used only the dominant marker)
  - Strict measure of generalization, given task effects (e.g. Perfors, 2016)

Participants have one productive generalization when only one is supported across experiments!



Contributions

- Adult learners generalize across consistent, inconsistent, and mixed patterns in the input
- Generalization is driven by type frequency: token frequency is fixed across conditions
- Methodological innovation of "hiding" singular-only items with observed-plural items ameliorates task effects for adults in regularization tasks

Future Directions

- What is the role of token frequency? What is probability matching? How do alternating items affect learning?
- What are the implications for developmental differences?

Selected References

Austin et. al., 2022. "Learning a language from inconsistent input: Regularization in child and adult learners" *Language Learning & Development*.

Berko, 1958. "The child's learning of English morphology" *Word*.

Hudson Kam & Newport, 2005. "Regularizing unpredictable variation: The roles of adult and child learners in language formation and change" *Language Learning & Development*.

Shin & Miller, 2022. "Children's acquisition of morphosyntactic variation" *Language Learning & Development*.

Acknowledgments: Thanks to the Penn Child Language Lab and the Penn Syntax Lab, as well as the audience of NELS 56 for discussion. Thanks also for our funding sources: NSF GRFP under Grant No. DGE-1845298 and the University of Pennsylvania.