The effect of population size on intergenerational language convergence:
An artificial language learning paradigm

Jayden Ziegler (ziegler@g.harvard.edu), Annemarie Kocab, Jesse Snedeker

SnedLab
Lab for Developmental Studies
Departmenet of Psychology • FAS

1. Introduction

- When input is variable, children regularize while adults frequency-match, except in cases of highly variable and unpredictable input[1]
- Artificial language learning paradigms using transmission chains allow us to probe the effect of iterated learning on linguistic structure[2]
- Typological work and artificial language learning studies demonstrate a bias for subject-initial word orders[3-5]
- How do learning biases interact with features of the input such as the number of speakers and tokens per speaker?

Question 1

When one word order is dominant within and across multiple speakers, will adults regularize or frequency-match?

Question 2

Is regularization of SOV driven by the number of speakers or the number of tokens per speaker?

Question 3

Will we see the same pattern of regularization vs. frequency-matching in a more ecologically valid context?

2. Methods & Results

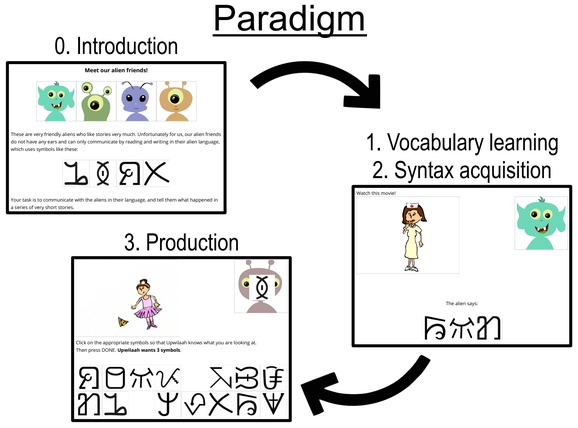
Iterated artificial grammar learning
Output of one generation input to next generation

Version A:
- ✧ 16 hieroglyph symbols
- ✧ Mean output of GenN becomes each alien's production in GenN+1

Version B:
- ✧ 24 nonce words
- ✧ Output of one learner in GenN becomes production for one alien in GenN+1

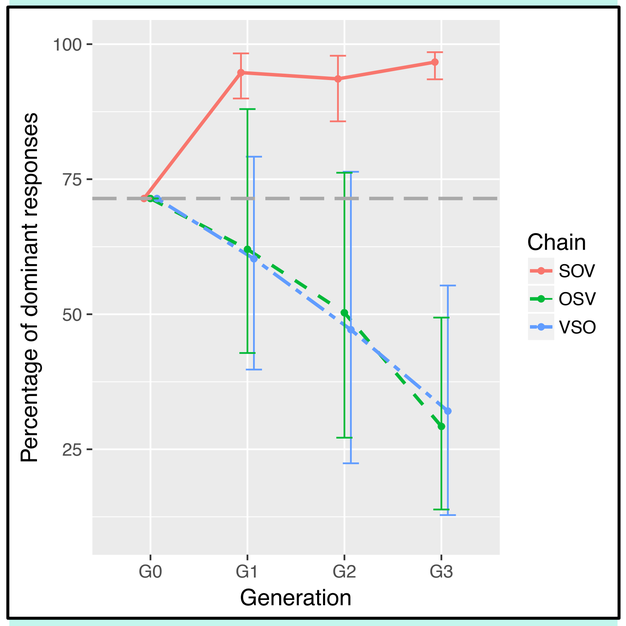
*Checks confirm attending to alien identity

Paradigm



Version A

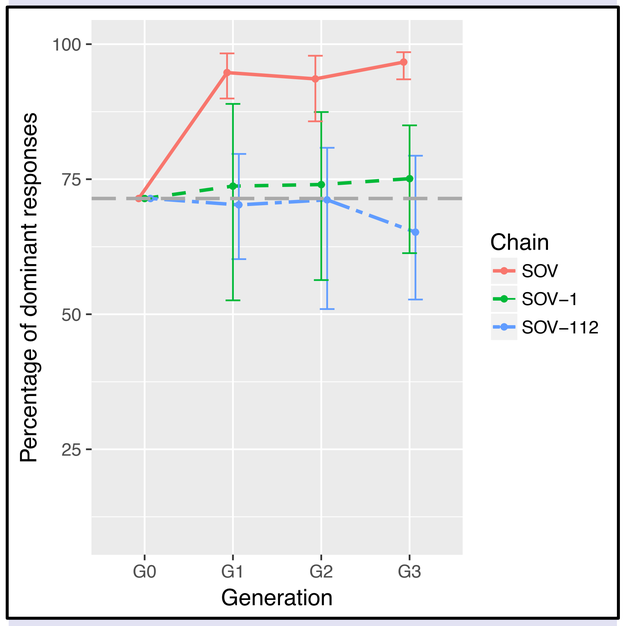
Aliens	Sentences	Dominant	Non-dom.
4	7*4=28	71% SOV	29% OSV
4	7*4=28	71% OSV	29% SOV
4	7*4=28	71% VSO	29% SOV



N = ~10 per generation

Version A

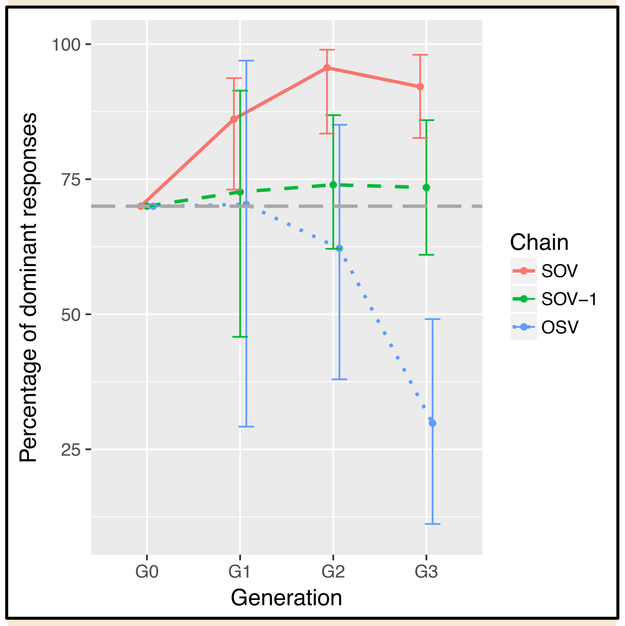
Aliens	Sentences	Dominant	Non-dom.
4	7*4=28	71% SOV	29% OSV
1	28	71% SOV	29% OSV
4	28*4=112	71% SOV	29% OSV



N = ~10 per generation

Version B

Aliens	Sentences	Dominant	Non-dom.
6	10*6=60	70% SOV	30% OSV
1	60	70% SOV	30% OSV
6	10*6=60	70% OSV	30% SOV



N = 6 per generation

4. Discussion

- Subject-initial word orders are privileged in non-gestural artificial language tasks with native English speakers
  - Even when the alternate word order shares as many features with participants' native language and is more common in the input, SOV is preferred to both OSV and VSO
- Whether adults frequency-match or regularize depends on both the amount of data and its distribution across speakers
  - Regularization occurs when there is insufficient information about each speaker

References: [1] Hudson Kam, C. L., & Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology, 59*, 30-66. [2] Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, *105*, 10681-10686. [3] Greenberg, J. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. Greenberg (Ed.), *Universals of Language* (pp. 73–113). Cambridge, MA: MIT Press. [4] Dryer, M. S. (2005). Order of subject, object, and verb. In M. Haspelmath, M. S. Dryer, D. Gil, & B. Comrie (Eds.), *The World Atlas of Language Structures* (pp. 330–333). Oxford, UK: Oxford University Press. [5] Goldin-Meadow, S., So, W. C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences*, *105*, 9163–9168.
Acknowledgments: This work was supported by a Pershing Square Fund grant from Harvard's Foundations of Human Behavior initiative. Special thanks to Melissa Kline for the codebase, Hannah Lam for stim creation, and SnedLab for critical discussion of this work.