



# Explaining differences between phonotactic learning biases in the lab & typological trends

Brandon Prickett ♦ bprickett@umass.edu ♦ brandon-prickett.com

Society for Computation in Linguistics 2025

## 1. Introduction

- Often, biases in artificial language learning experiments line up with typology:

**Goal:** Easier in lab = More frequent

- Here I show that an independently motivated model can help explain a mismatch observed by Moreton & Pertsova (2014) where **lab** ≠ **typology**

## 2. Background

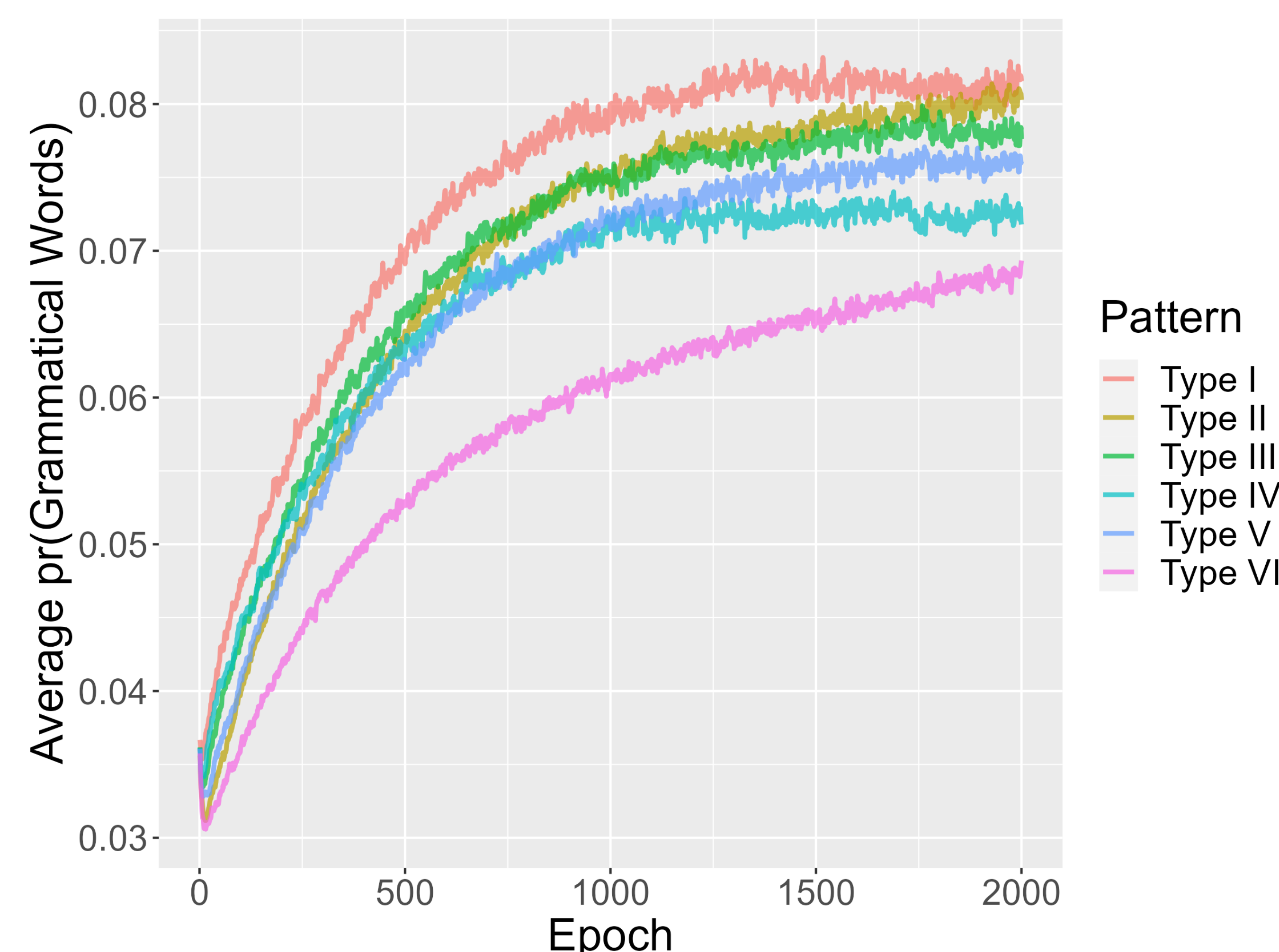
- Shepard et al. (1961) created 8 stimuli described by 3 features, with 6 ways to halve the space
  - When learning to categorize, learners had consistent biases for certain patterns
  - They numbered the patterns from easiest to hardest: **I > II > III > IV > V > VI**
- Moreton & Pertsova (2014) found a different ordering in phonotactics: **I > IV > III > V > II > VI**
  - But when they looked at typological frequency, it followed the original ordering
- Probabilistic Feature Attention (PFA) is a mechanism that assumes ambiguity in the learning process (Prickett 2023)
  - Features are sampled at each learning update to see which are attended to
  - This causes ambiguity (e.g., if [voice] isn't attended to, [t] and [d] look the same)

## 3. Methods

- MaxEnt phonotactic learner (Hayes & Wilson 2008)
  - Constraints represent all combinations of features (Moreton et al. 2017)
  - E.g., \*[+voice], \*[+voice, +labial], etc...
- The 6 Shepard et al. (1961) groupings were implemented as phonotactic patterns
  - Had to learn which words were grammatical
  - Words were 1 segment long: [p, b, t, d, f, v, s, z]
  - Features: [±continuant], [±voice], & [±labial]

## 4. Results

- The model matches the biases observed in the lab early in learning
- Later in learning, the model's biases match typology



## 5. Why does this Work?

- The initial biases match what would be expected for any MaxEnt learner like this (Moreton et al. 2017)
- But why does PFA cause this ordering to change?
  - In PFA, the more features that are relevant to a pattern, the more opportunities there are for ambiguity to interfere with learning
- The model's biases later in learning reflected this:
  - Type II had higher accuracy than III, IV, or V
  - And it also has fewer relevant features (2 vs. 3)

## 6. Conclusions

- PFA offers a potential explanation for the mismatch observed by Moreton and Pertsova (2014)
  - Early learning could reflect lab-like conditions
  - Later learning could better mirror 1<sup>st</sup> language
- Future work should test whether other models of phonotactic learning can capture this effect

## References

- Bruce Hayes and Colin Wilson (2008). A maximum entropy model of phonotactics and phonotactic learning.
- Elliott Moreton, Joe Pater, and Katya Pertsova. (2017) Phonological Concept Learning.
- Elliott Moreton and Katya Pertsova (2014). Pastry phonotactics: Is phonological learning special.
- Brandon Prickett (2023). Probabilistic feature attention as an alternative to variables in phonotactic learning.
- Roger N. Shepard, Carl I. Hovland, and Herbert M. Jenkins (1961). Learning and memorization of classifications.