Coleman Haley coleman.c.haley@gmail.com

Colin Wilson colin.wilson@jhu.edu

Prior work has demonstrated that neural networks can induce both concatenative and non-concatenative morphological patterns, including infixation and reduplication (e.g., Kann & Schütze, 2017; Nelson et al. 2020)

Research questions

1. Can the networks learn ***unattested*** non-concatenative patterns that are formally simple but unlike those found in natural languages?
2. Do the networks have ***inductive biases*** that favor natural (or unnatural!) non-concatenative morphology?

We studied learning and generalization by LSTM and GRU encoder-decoder networks for a variety of natural and unnatural infixation and reduplication patterns under various low-resource conditions

Read the full paper here.

This research was supported by NSF grant BCS-1844780 to Colin Wilson.

Deep neural networks easily learn unnatural infixation and reduplication patterns.

Society for Computation in Linguistics, 2021

Which patterns were learned most easily?

Unattested counting patterns (e.g., place the infix before the third segment), were learned reliably in even low-data contexts. **Overall, unattested patterns were learned at least as robustly as attested patterns.**

Which patterns were most difficult to learn?

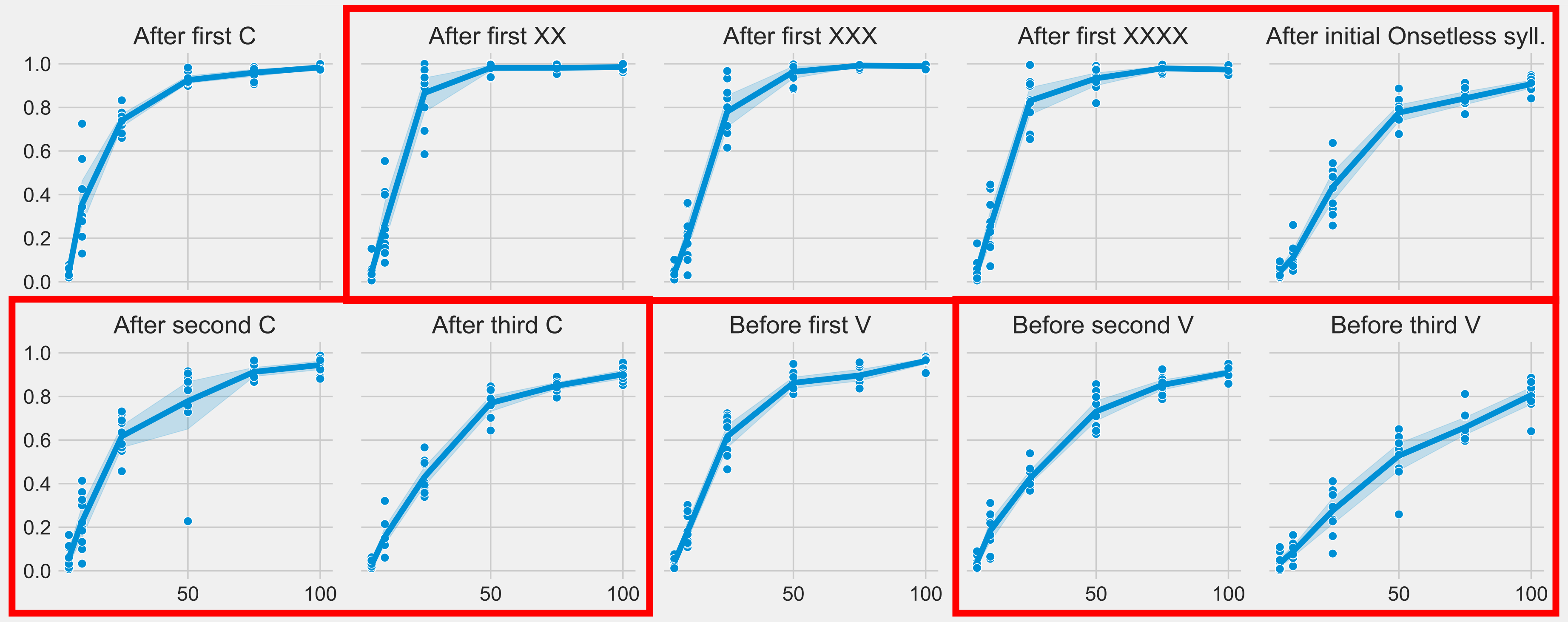
Unattested patterns that involve both counting and distinguishing consonants and vowels (e.g., place infix before the third vowel) required the most training data.

Do LSTM and GRU networks differ?

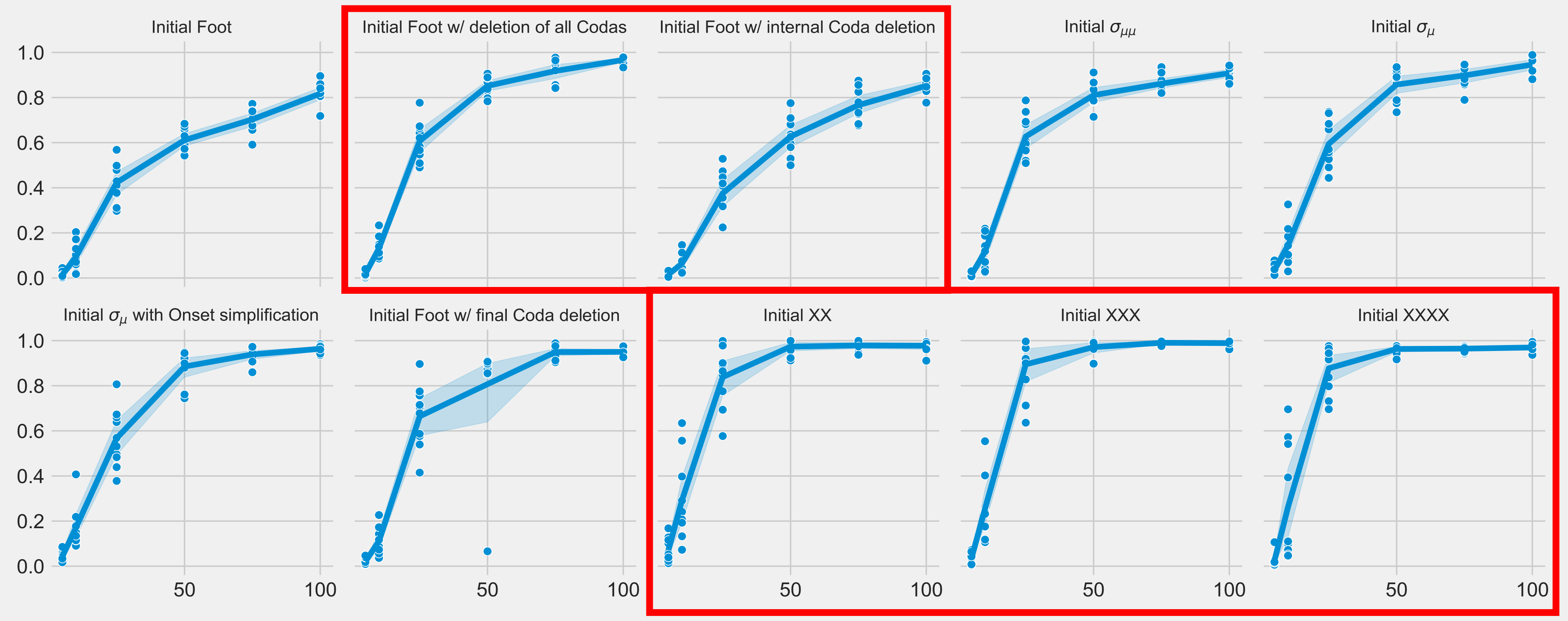
Despite the limited counting ability of the GRU unit, it showed the same patterns as LSTM models, including easily learning unattested counting patterns.

Infixation pattern		LSTM	GRU
Initial 2 segments	$\underline{X_1X_2}X_1X_2...$	.99	1.0
Initial 3 segments	$\underline{X_1X_2X_3}X_1X_2X_3...$	1.0	1.0
Initial 4 segments	$\underline{X_1X_2X_3X_4}X_1X_2X_3X_4...$	.98	.97

Table: Average test set performance for LSTM and GRU networks for 50 training examples.



(a) ***Infixation*** test accuracy as a function of training size (red = unnatural pattern)



(b) ***Reduplication*** test accuracy as a function of training size (red = unnatural pattern)

How many examples are required to learn the patterns?

Unattested segment-counting patterns were learned from as few as 25 examples, other patterns required at least 50-100+ training examples