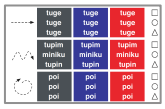
CENTRE for LANGUAGE EVOLUTION

Variation and selection in an iterated language learning experiment

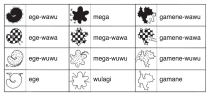
Monica Tamariz & Joleana Shurley

Darwinian dynamics = descent + modification + selection

We usually assume that this is what is going on when properties of language evolve to adapt to our learning biases and communicative usage, e.g.:



Kirby, Cornish & Smith (2008)



Kirby, Tamariz, Cornish & Smith (2015)

But is it?

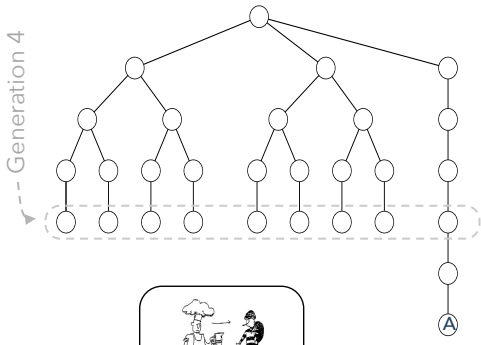
Question: Do iterated learning experimental languages actually follow Darwinian Dynamics?

Materials

We collected a **family of languages** originating in a a single language using an Iterated Learning design.

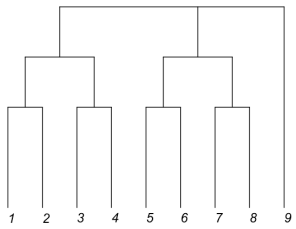
The **initial language** had 27 randomly constructed signals which referred to 27 graphical scenes:

All the elements

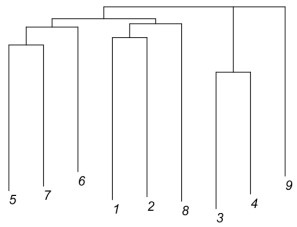


A training item. During testing, the participant was shown the picture only and had to type the signal.

1. Evidence of descent with modification



Tree 1 showing relatedness of Generation 4 languages



Tree 2 based on similarity between Gen. 4 languages

If there is descent with modification, similarity-based Tree 1 should be similar (correlated), but not identical to, relatedness Tree 2.

Monte Carlo analysis results:

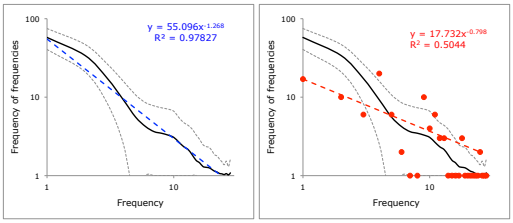
Trees based on languages in Gen. 4, *Correlation*=0.44 z=2.5, p<0.01

Trees based on all 29 languages, *Correlation* = 0.62, z=8.6, p<0.001

Highly significant correlation support descent with modification in the experiments!

2. Evidence of selection

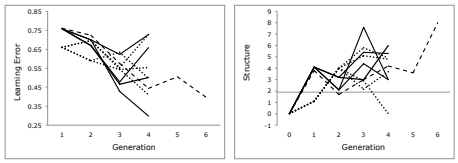
Random evolution (drift) results in power law frequency distributions *(e.g. Hahn & Bentley 2003)*. Significant deviations from power law in our languages indicate selection. We analyze n-gram frequency.



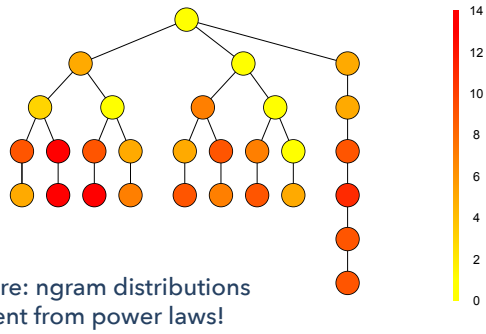
Simulated drift (in black) follows a power law (regression shown in blue).

N-gram frequencies from language A (in red), differ significantly from drift.

Like Kirby et al. (2008, 2015), we also find cumulative increases in structure and decreases in transmission error.



We plot how different from drift the frequency distributions are in all the languages: z-score of slopes (z>1.96 is significant (p<0.05)



Selection is a work here: ngram distributions are increasingly different from power laws!