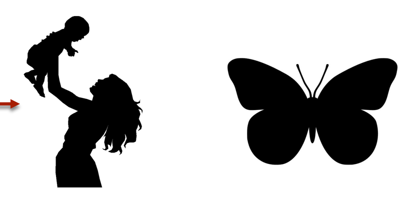
Predictors of lexical stability in an artificial language

James Brand & Padraic Monaghan
Lancaster University

Introduction

- Lexical items in the vocabulary are susceptible to change, but what predicts these changes? and what types of change are they subject to?
- Some words undergo changes much more rapidly than others, leading to some word forms being conserved across different languages, whilst others appear to be unrelated



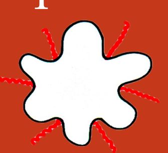
- It has been shown that certain psycholinguistic properties of these words can be used to reliably predict the rate at which a word form is replaced by a new, unrelated form
- These are *FREQUENCY, LENGTH* and *AGE OF ACQUISITION (AoA)* (*Pagel et al, 2008; Monaghan, 2014*)
- Additionally, word forms can also be subjected to more minor changes where they are adjusted to provide optimal communicative efficiency, reducing effort in speech production (*Zipf, 1949*)
- Here, we aim to generate these findings using a novel artificial language learning paradigm in the lab

Hypotheses

- ➢ Low frequency, long, late acquired words undergo more lexical replacements
- ➢ High frequency, short, early acquired words undergo more adjustments
- ➢ High frequency, short, early acquired words are conserved in the language during processes of cultural transmission

Methods

- Artificial language with 12 images, each paired with an unfamiliar word (adapted from *Kirby et al, 2015*)



- Training phase where participants saw the image-word pairings over several blocks (total 120 trials), then tested in a production recall task
- 3 different experimental conditions:

1. Frequency: Weight number of exposures during training (4 words per frequency condition)

training block	low	medium	high
1	1	3	6
2	1	3	6
3	1	3	6
total	3	9	18

2. Length: Vary the number of characters in the words (either 4, 6, or 8 with 4 words per length condition)

training block	short	medium	long
1	5	5	5
2	5	5	5
total	10	10	10

3. AoA: Present words early or late during training (6 words per acquisition condition)

training block	early	late
1	6	0
2	1	3
3	1	3
4	1	3
5	1	1
total	10	10

Analysis

Each participant learned an artificial language and produced one new testing output language

1. Compare words presented in training language to testing output
2. Calculate error as *normalised Levenshtein edit distance (NLED)* by quantifying the number of character insertions, substitutions or replacements, then dividing that value by the longest word length

e.g. '*hello*' → '*helicopter*' = 7/10 = 0.7

3. Classify errors based on the mean *NLED* between all words in the initial training language, producing a threshold value of 0.67:

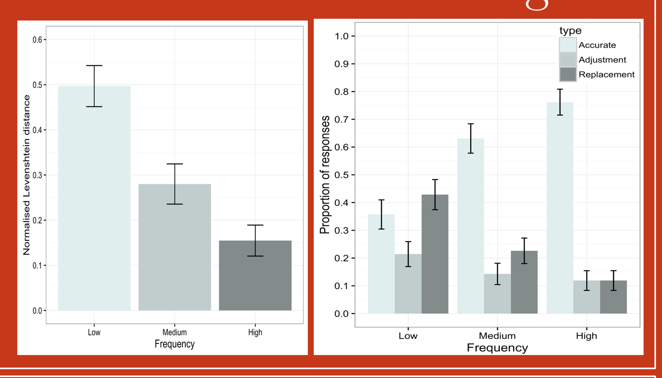
accurate: 0, ***adjustment***: 0 < 0.67, ***replacement***: > 0.67

Experiment 1 – *One generation learning*

- 21 participants in each experimental condition (total n = 63)
- Analysis using three mixed-effects models for each *psycholinguistic property* (fixed effect), with *error, number of adjustments* and *number of replacements*, as dependent variables.

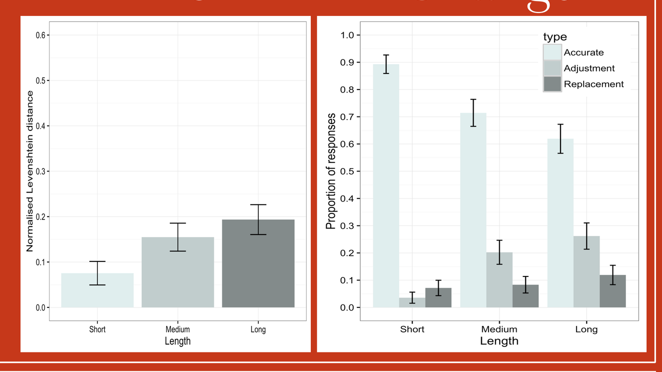
Frequency

Error: decreases as frequency increases $\chi^2(2) = 46.2$, $p<.001$
Adjustments: no significant difference.
Replacements: decreases as frequency increases $\chi^2(2) = 29.5$, $p<.001$



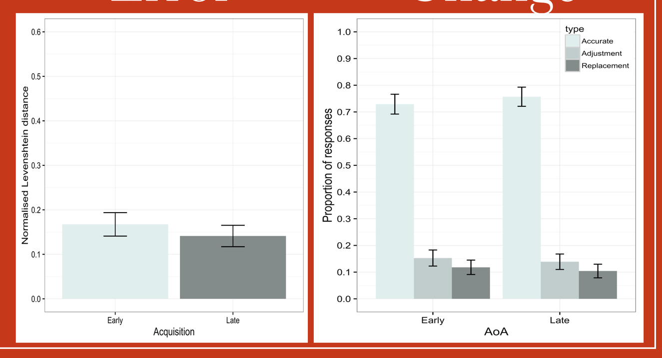
Length

Error: increases as length increases $\chi^2(2) = 10.8$, $p = .005$
Adjustments: increases as length increases $\chi^2(2) = 20.1$, $p<.001$
Replacements: no significant difference.



AoA

Error: no significant difference.
Adjustments: no significant difference.
Replacements: no significant difference.

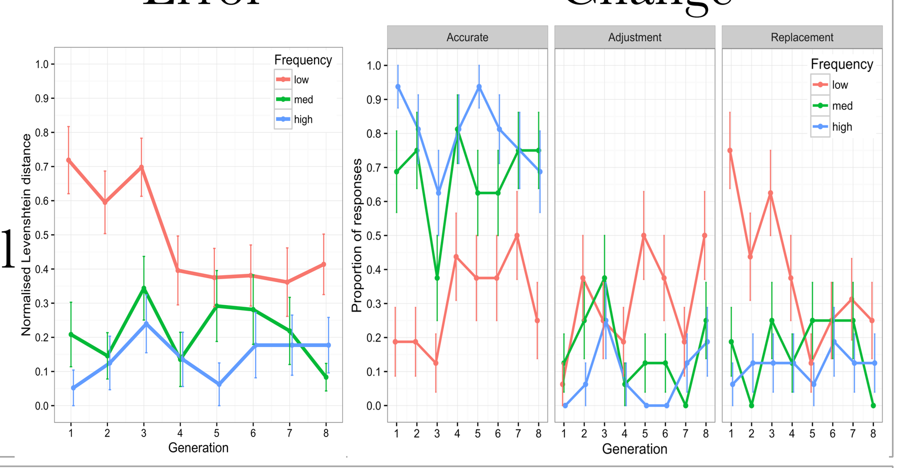


Experiment 2 - *Iterated learning*

- Participant gets trained on previous participant's testing output, with 4 chains of 8 participants in each experimental condition (see *Kirby et al, 2008*) each participant representing a generation in the learning chain (total n = 96)
- Analysis using mixed-effects modelling, with *generation* as predictor variable

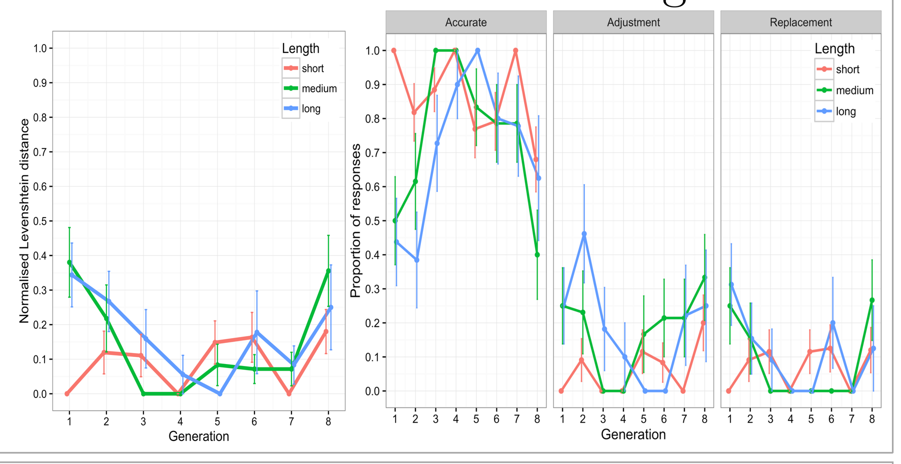
Frequency

Low: error decreases: $\chi^2(1)=13.2$, $p<.001$
adjustments marginal increase: $\chi^2(1) = 3.6$, $p = .058$
replacements decrease: $\chi^2(1)=13.6$, $p<.001$
Medium: no significant changes.
High: no significant changes.



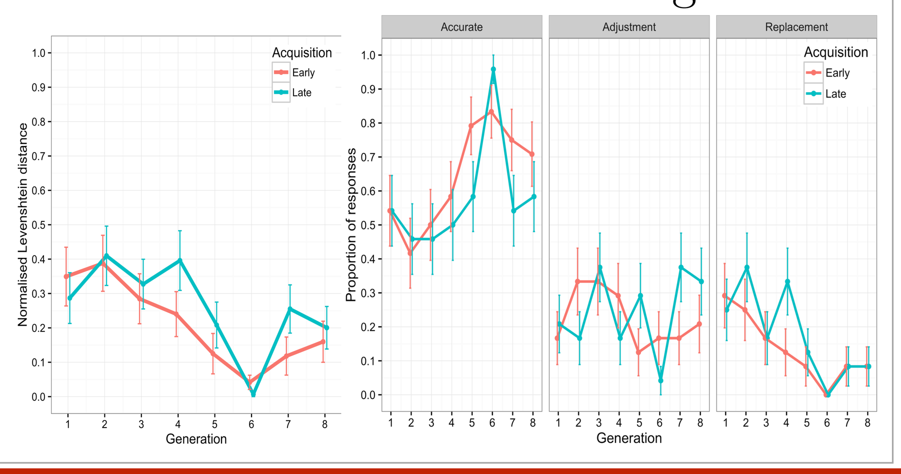
Length

Short: no significant changes in error or replacements.
adjustments increase: $\chi^2(1) = 4.1$, $p = .04$
Medium: no significant changes.
Long: no significant changes.



AoA

Early: error decreases: $\chi^2(1) = 16.9$, $p<.001$
no significant change in adjustments.
replacements decrease: $\chi^2(1) = 10.2$, $p = .001$
Late: error decreases: $\chi^2(1) = 7.4$, $p = .01$
no significant change in adjustments.
replacements decrease: $\chi^2(1) = 11.3$, $p<.001$



Conclusions

- ➢ Low frequency words undergo more error, but over time can become more learnable, whilst high frequency words remain reliably recalled and transmitted with little change over time.
- ➢ Longer words undergo more error than shorter words, but over time shorter words are becoming adjusted more.
- ➢ No difference in AoA, but both early and late acquired words become more learnable over time.

References:
Kirby et al (2015). Compression and communication. *Cognition.*
Kirby et al (2008). Cumulative cultural evolution in the laboratory. *PNAS.*
Monaghan (2014). AoA predicts rate of lexical evolution. *Cognition*
Pagel et al (2007). Frequency of word-use. *Nature.*
Zipf (1949). *Human behavior and the principle of least effort.*

j.brand@lancaster.ac.uk