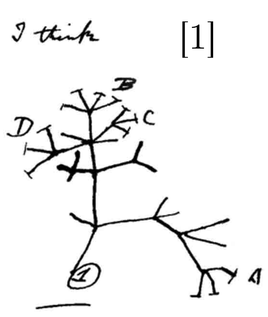
Cognitive factors influence rate and type of linguistic change in the vocabulary

James Brand[1] and Padraic Monaghan[1,2]
[1]Lancaster University, [2]Max Planck Institute for Psycholinguistics
j.brand@lancaster.ac.uk

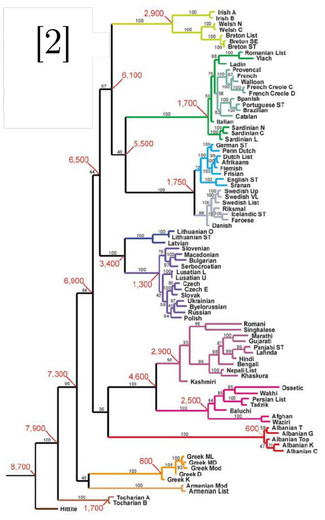
Linguistic Change

- Languages, like living organisms, change, evolve and diverge over time through processes of cultural transmission
- Methodological advances have enabled such processes to be studied e.g. cladistics [3], iterated learning [4]



[1]

[2]



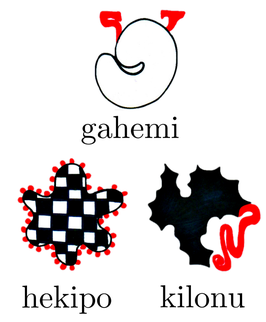
- These approaches have provided strong support for the role of cognitive factors predicting rates of linguistic change e.g. short, early acquired, low frequency words change most rapidly [5,6]
- Words can be subjected to large rapid replacements, but they can also remain stable, changing through small adjustments over time
- Distinguishing between what is a replacement and what is an adjustment is a considerable challenge

Research Outline

- ➢ Provide an experimental approach to studying language change in the lab
- ➢ Provide an empirical approach for distinguishing between different types of linguistic change in the vocabulary
- ➢ Examine the influence of frequency effects during linguistic change – rate and types of change
- ➢ Explore where changes are most likely to occur in a word

Method

- Artificial language learning paradigm with 12 form-meaning mappings [7]
- Exp 1: participants learnt low, med, high frequency words (6 letters long), then were tested during a production task
- Exp 2: same design was then iterated over 8 generations of learners across 4 chains



Quantifying Error

- Compare training word to testing output word for error by calculating Normalised Levenshtein Distance (*NLD*)
- Number of deletions, insertions or substitutions required to transform the output string into the input

'gahemi' → *'gahumo'*

$NLD = \text{error}(3)/\text{length}(6) = 0.5$

Quantifying Change

$$CT = \frac{\mu_{\text{LD}} - \alpha\sigma}{len}$$

CT = critical threshold for adjustment/ replacement classification, if error > CT, output is replacement, else adjustment
μ_{LD} = mean Levenshtein distance between the training word and Monte Carlo sample of 1,000,000 permutations
α = value at which a Z-score is significant at the 95% confidence level
σ = standard deviation of μ_{LD}
len = length of the training word

Results

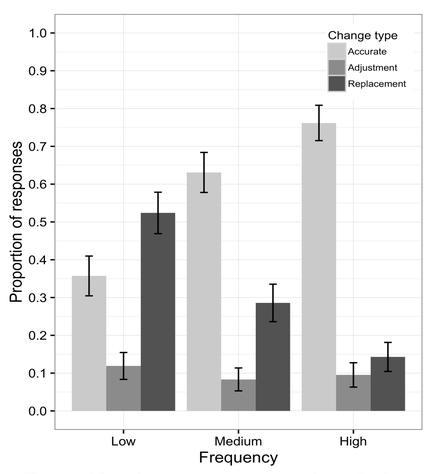


Fig 1. Number of responses classified as accurate, adjustment or replacements

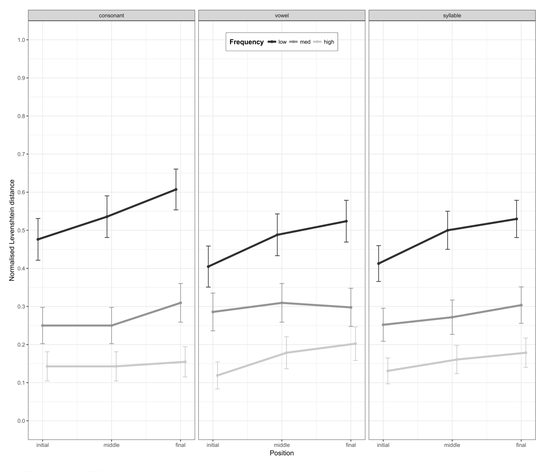


Fig 2. Error based on part of word (initial, middle, final) for consonants, vowels and syllables

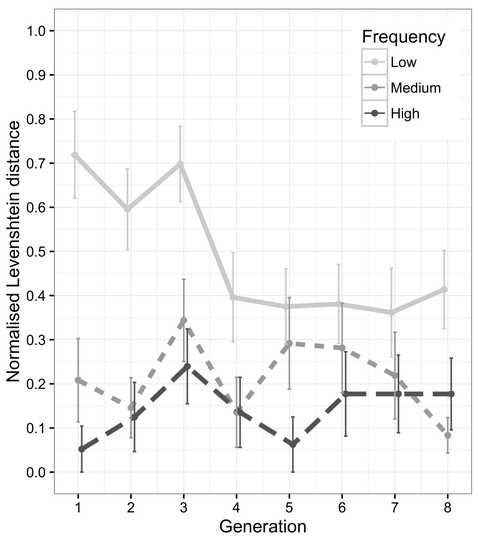


Fig 3. Error calculated at each generation

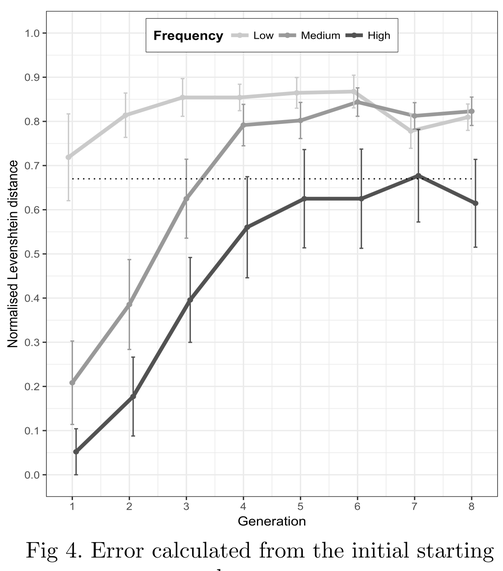


Fig 4. Error calculated from the initial starting language

Exp 1

- ➢ Low frequency words undergo more replacements
- ➢ High frequency words more accurate with very little error
- ➢ Low frequency words experience most change in the final part of the word

Exp 2

- ➢ Low frequency words become more learnable over the course of cultural transmission
- ➢ Low frequency words are replaced rapidly, losing traces of their original forms early on
- ➢ High frequency words are much slower to diverge away from their original word forms

Conclusions

- ➢ We present here an empirical approach to studying different types of linguistic change under controlled experimental conditions in the lab
- ➢ We generate previously reported findings from corpus analyses
- ➢ Lexical replacements occur when words are lower in frequency
- ➢ The final part of the word is most susceptible to change
- ➢ Words that are high in frequency exhibit the least change in their forms
- ➢ High frequency words are conserved in languages, even when undergoing cultural transmission

References

[1] Darwin (1837) *Notebook B: (Transmutation of species).*
[2] Gray & Atkinson (2003) *Nature.*
[3] Pagel & Meade (2006) *Phylogenetic methods and the prehistory of languages.*
[4] Kirby, Cornish & Smith (2008) *PNAS.*
[5] Pagel, Atkinson & Meade (2007) *Nature.*
[6] Monaghan (2014) *Cognition.*
[7] Kirby et al (2015) *Cognition.*

Lancaster University
MAX PLANCK INSTITUTE FOR PSYCHOLINGUISTICS
LuCiD The ESRC International Centre for Language and Communicative Development