Sound change, lexical frequency, and ambiguous input: a computational implementation

MANCHESTER 1824 The University of Manchester

School of Arts, Languages and Cultures

Stefano Coretta

The University of Manchester

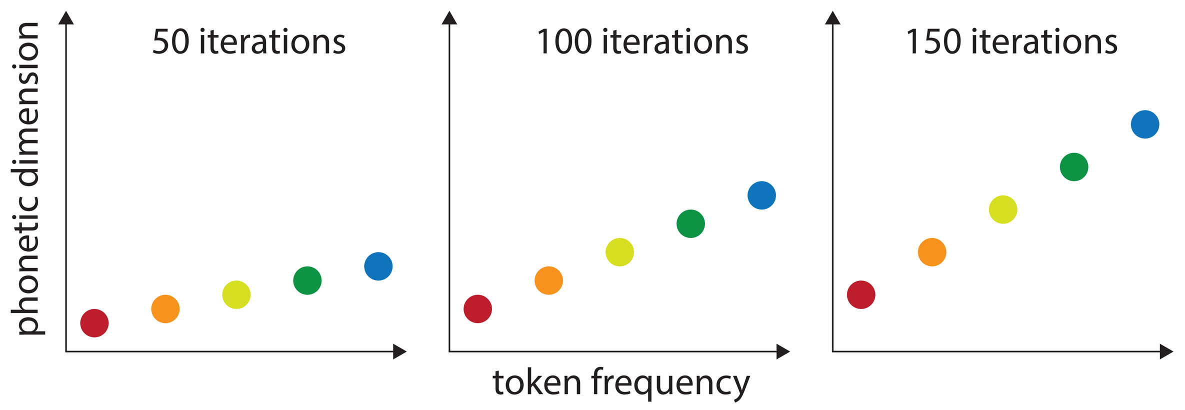
stefano.coretta@manchester.ac.uk
StefanoCoretta
stefanocoretta.altervista.org

1. Background

1.1 Sound change

- **exemplar theory**: each token of experience is stored in memory (Hintzman 1986, Nosofsky 1988)
- **sound change**: lexical (token) frequency plays a role (Bybee 2002)

~ high frequency words should lead ~



Hay *et al.* 2015 found that *low frequency* words were changing at a higher rate!

1.2 NZE vowel shift

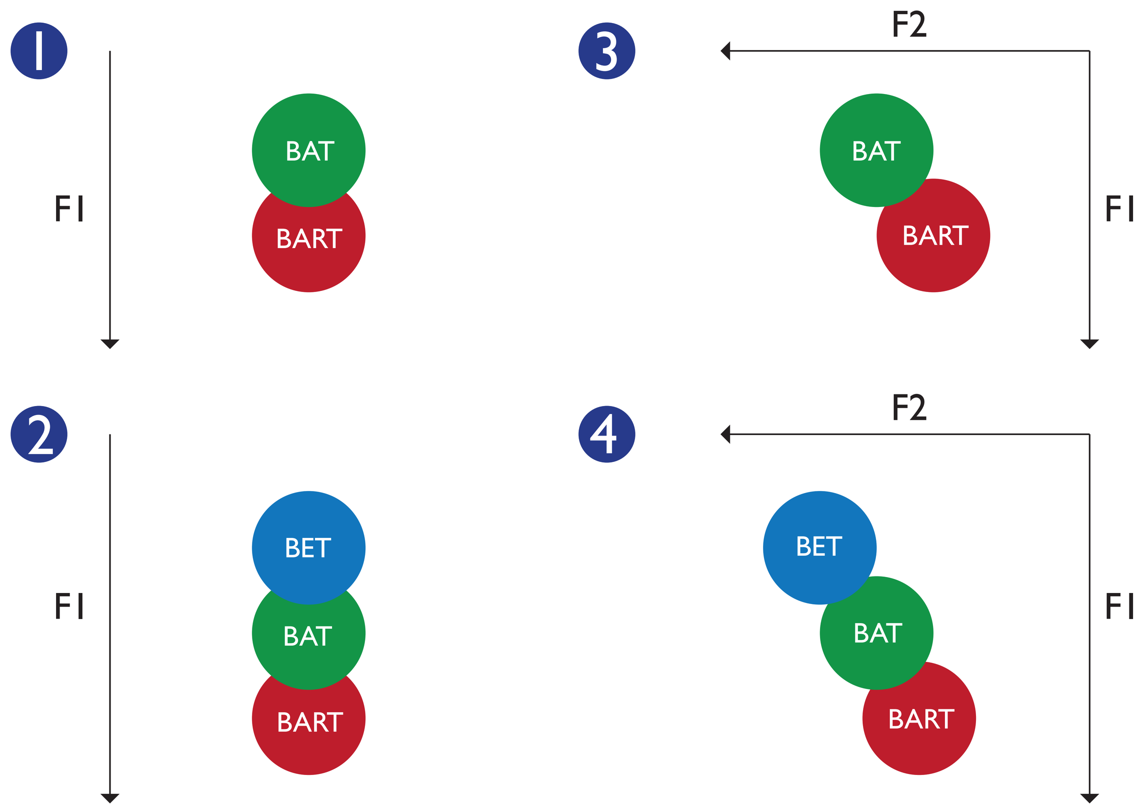
- **New Zealand English**: historical change whereby the BAT and BET vowel raise and the BIT vowel is backer and lower
- **push chain shift**, perhaps initiated by BART fronting
- *low frequency words were ahead and changed at a higher rate*

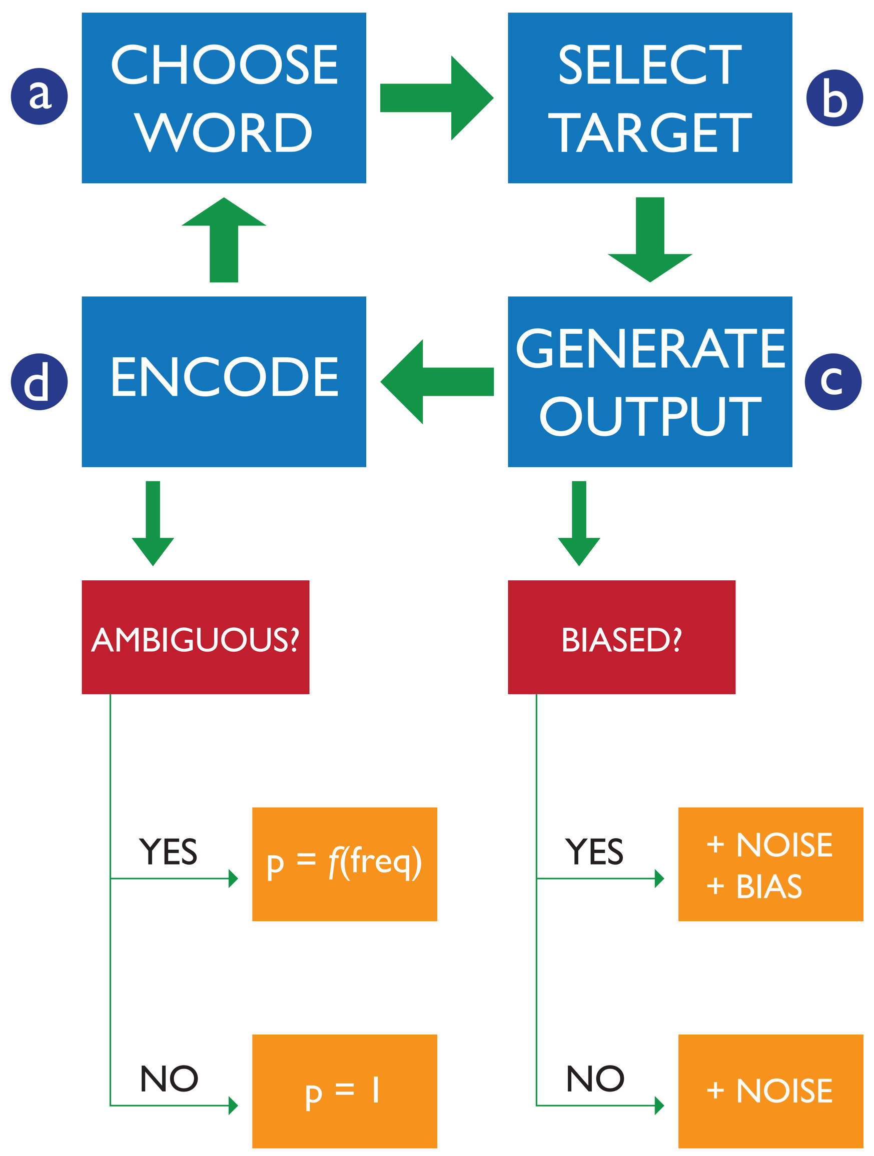
1.3 Encoding of ambiguous input

- Hay *et al.* 2015 propose an *encoding algorithm* that discards ambiguous input based on its **lexical frequency**
- **ambiguous input** has less chance to be encoded in memory

3. Methodology

- single-agent simulations (feedback loop)
- 4 simulations of the NZE vowel shift in R (R Core Team 2017)
- only BART vowel is subject to bias
- simplified scenarios to keep computational effort low





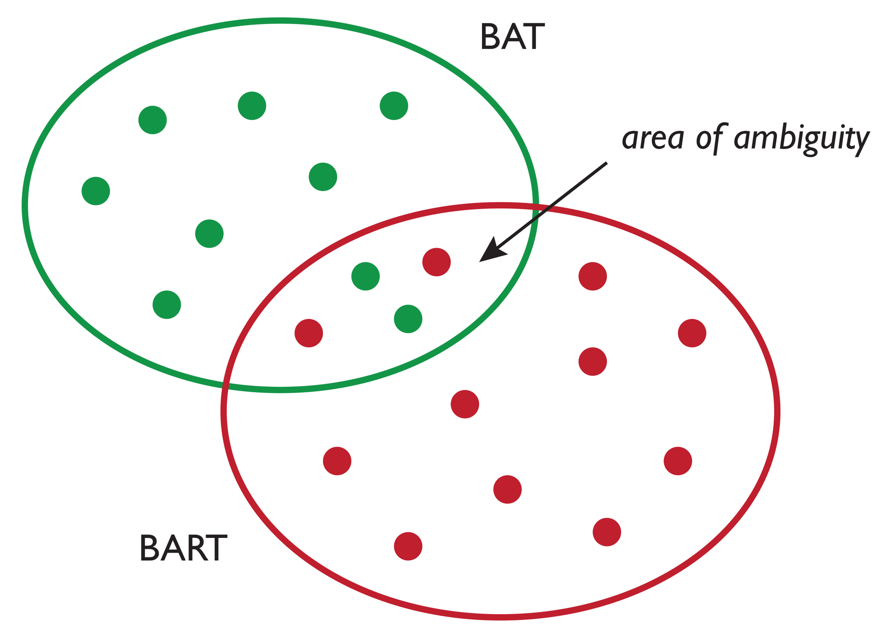
Encoding Probability Function

if x is ambiguous

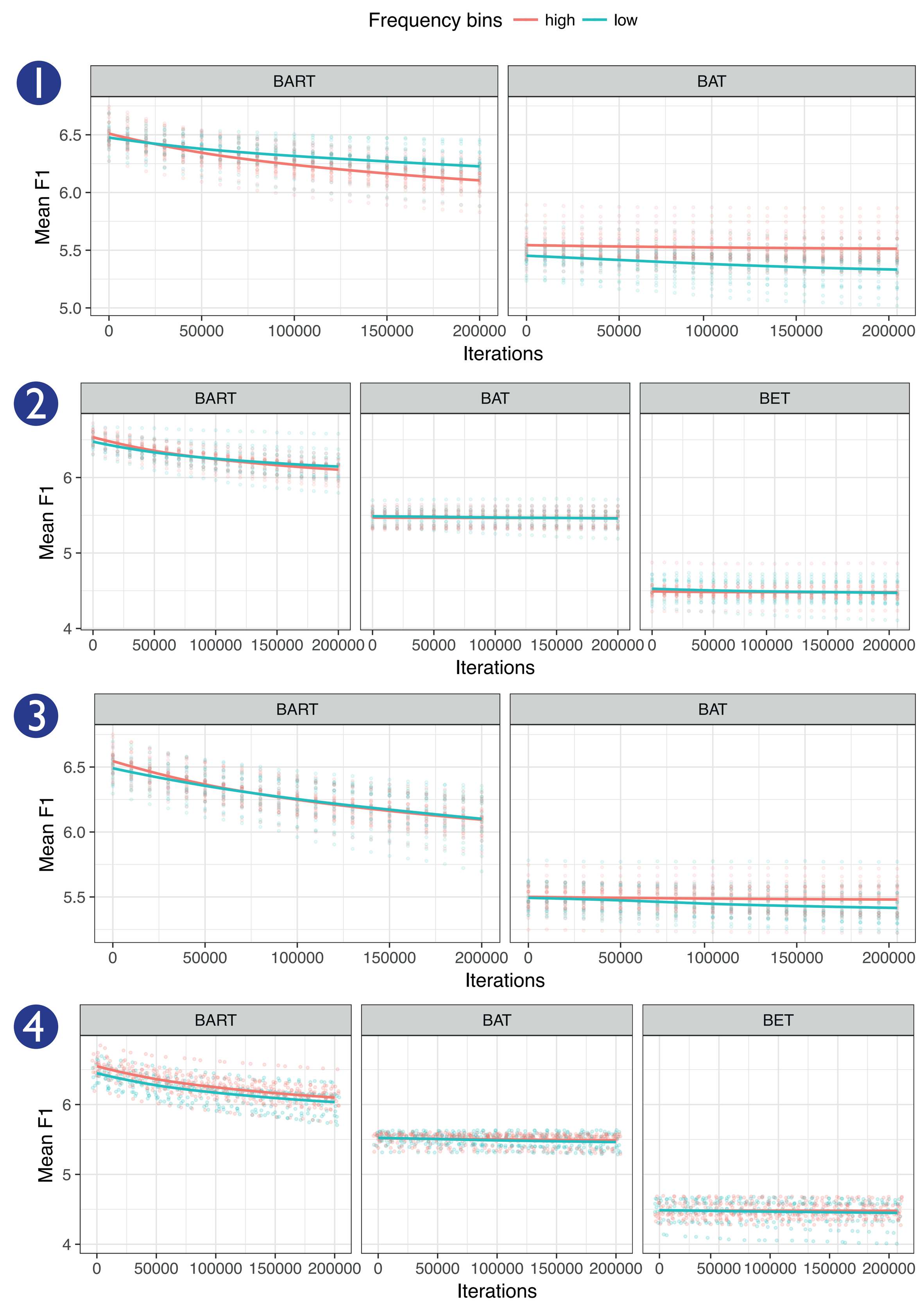
$$p(x|M) = freq(x) / maximum(freq)$$

if x is not ambiguous

$$p(x|M) = 1$$

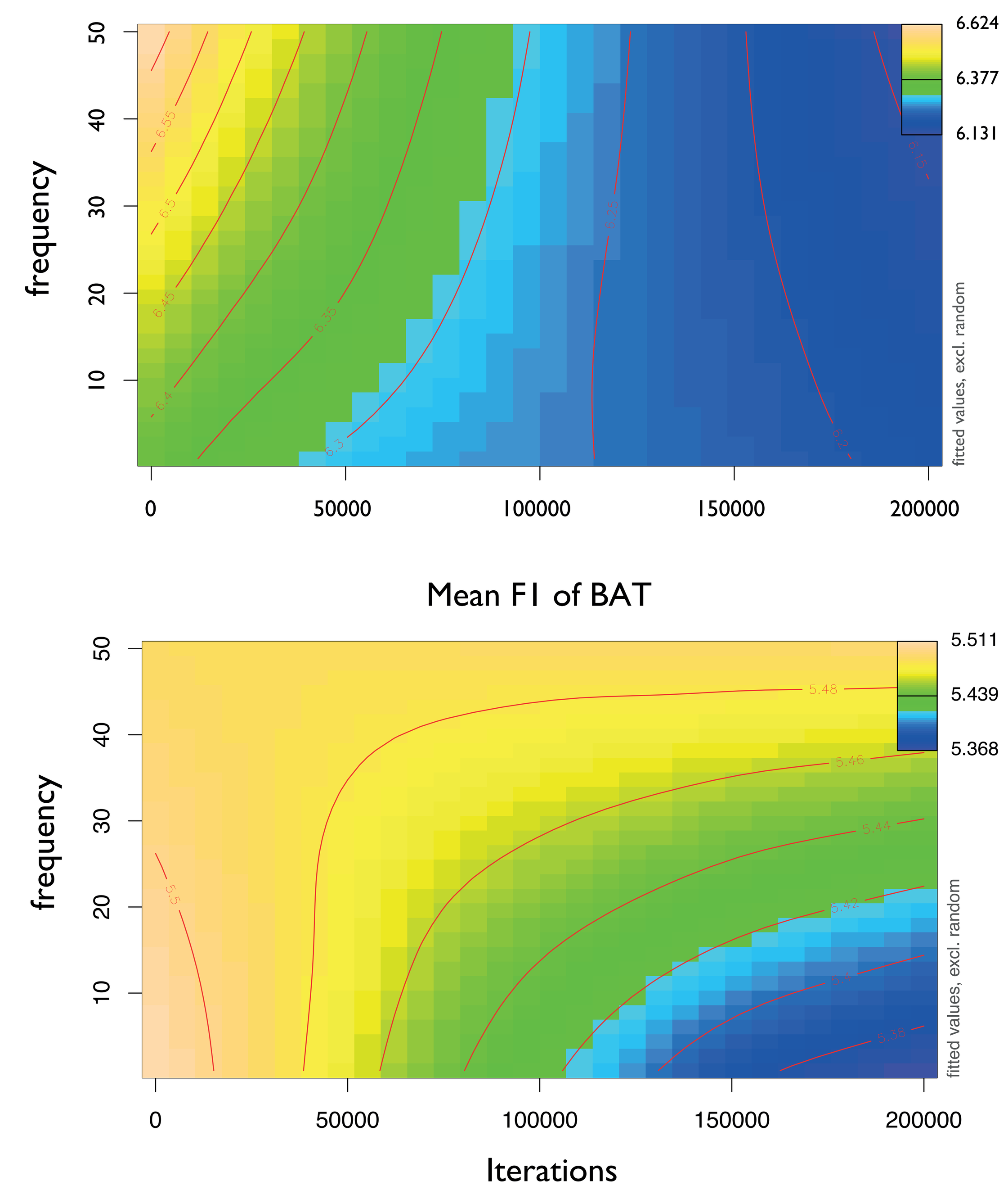


4. Results



4.1 Statistical analysis

- generalised additive mixed models (GAMM) fitted to the output of the simulations (Wood 2006, Zuur 2012)



5. Conclusions

- **low frequency** words change at a higher rate in the *pushed* categories (BAT, BET)
- **high frequency** words change at a higher rate in the *pushing* category (BART); not discussed in Hay *et al.* 2015
- *future directions*: increase complexity; include memory decay; try non-chain sound changes

The code is available at http://github.com/stefanocoretta/2017-eshiph

References: Bybee, Joan. 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. Language variation and change 14(03). 261–290. • Hay, Jennifer B., Janet B. Pierrehumbert, Abby J. Walker & Patrick LaShell. 2015. Tracking word frequency effects through 130 years of sound change. Cognition 139. 83–91. • Hintzman, Douglas L. 1986. "Schema abstraction" in a Multiple-Trace Memory model. Psychological Review 93(4). 411–428. • Nosofsky, Robert M. 1988. Similarity, frequency, and category representations. Journal of Experimental Psychology: Learning, Memory, and Cognition 14(1). 54–65. • Wood, Simon. 2006. Generalized additive models: An introduction with R. CRC Press. • Zuur, Alain F. 2012. A beginner's guide to generalized additive models with R. Highland Statistics Limited: Newburgh. • **Acknowledgements**: I wish to thank Dr Márton Sóskuthy for having introduced me to the world of computational modelling, and the members of the Linguistics and English Language division of the University of Manchester for the invaluable input which shaped this work. All errors are mine.