Explaining lexical frequency effects: a critique and an alternative account

Márton Sóskuthy
University of York

29 May 2014

Outline

Introduction

- lexical frequency effects (e.g. Schuchardt 1972; Bybee 2001; Pierrehumbert 2001):
  lexical frequency (partly) determines the speed at which lexical items undergo sound changes
- in this talk: only looking at phonetically gradient sound change
- rapidly growing literature → many different explanations & findings

Introduction

- different sources for frequency effects:
  - exposure (Pierrehumbert, 2001)
    high-frequency items 'erode' faster
  - independence (Bybee, 2001)
    high-frequency items are more independent $\rightarrow$ *more resistant to change*

    ...

Introduction

- problem:
  many different explanations $\rightarrow$ predictions not clear and not sufficiently distinct
- main goal:
  clarify the predictions of the exposure-based and independence-based accounts
- solution:
  computational and mathematical modelling

Outline

Exposure

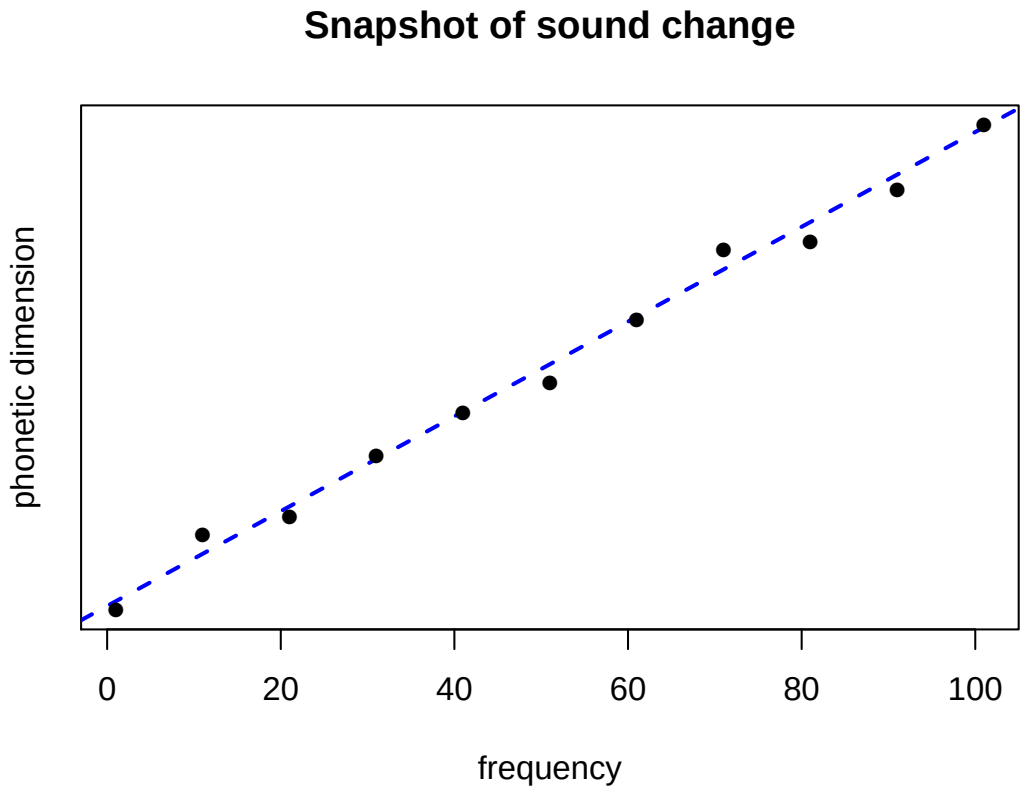
Prediction

- high-frequency forms are exposed to phonetic biases more often $\rightarrow$ they change faster
- Pierrehumbert (2002):
  '[...H]igh frequency words are affected more because they are produced more often and so more memories of them in their lenited form accrue, once the lenition gets underway.'

Exposure

Prediction

- apparent prediction: positive linear frequency effect (e.g. Pierrehumbert 2001)



Exposure

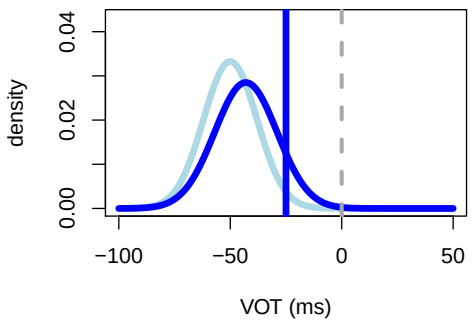
Simulation architecture

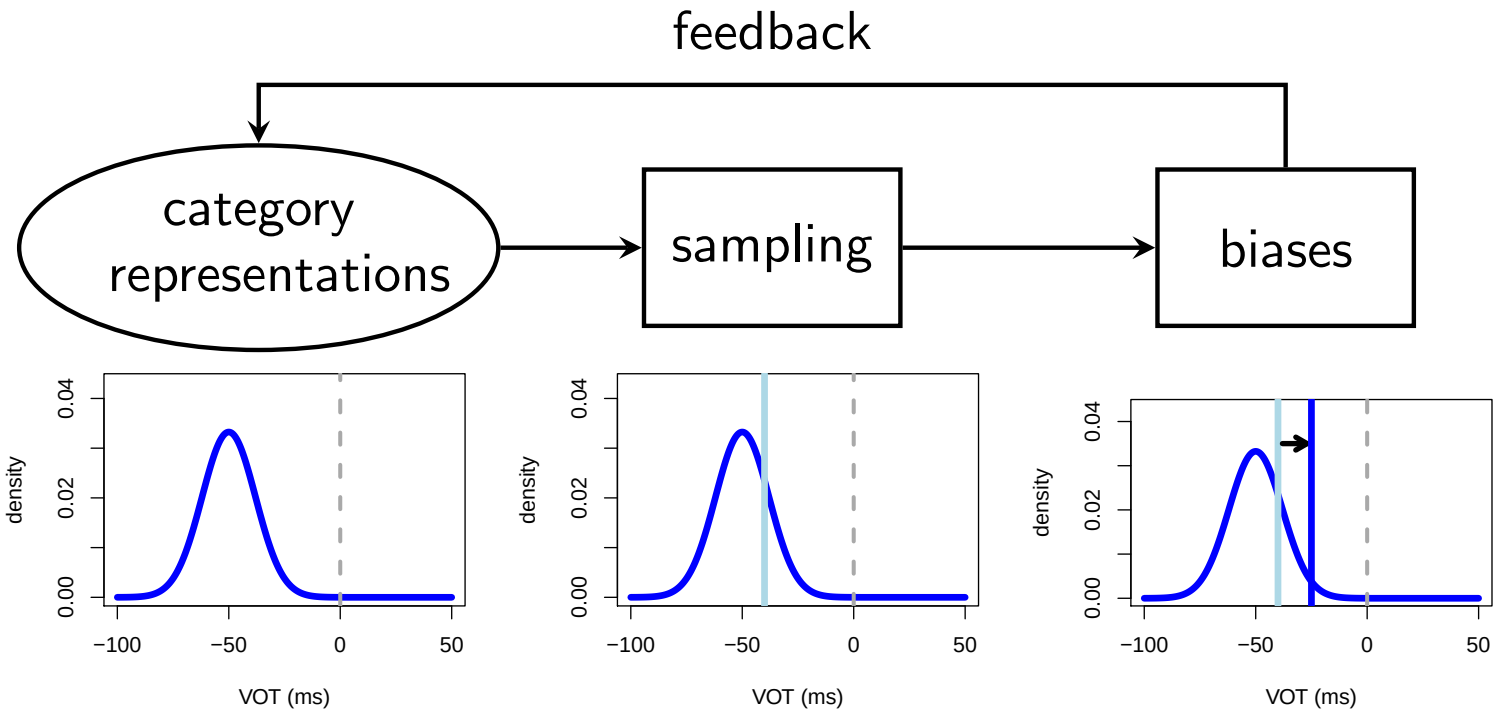
- modelling framework: simplified version of Pierrehumbert (2001)
- each word has a separate representation
- looking at how those representations evolve as a function of frequency
- parametric (prototype-based) representations

$\rightarrow$ the results hold for the original model as well

Exposure

Simulation architecture





Exposure

Simulation architecture

- in exemplar models:
  - memory activation of exemplars decreases with time
  - memory activations summed (MASS) $\propto$ frequency
- current model also includes MASS
- words with higher MASS (i.e. high frequency categories) are more resistant to incoming stimuli

Exposure

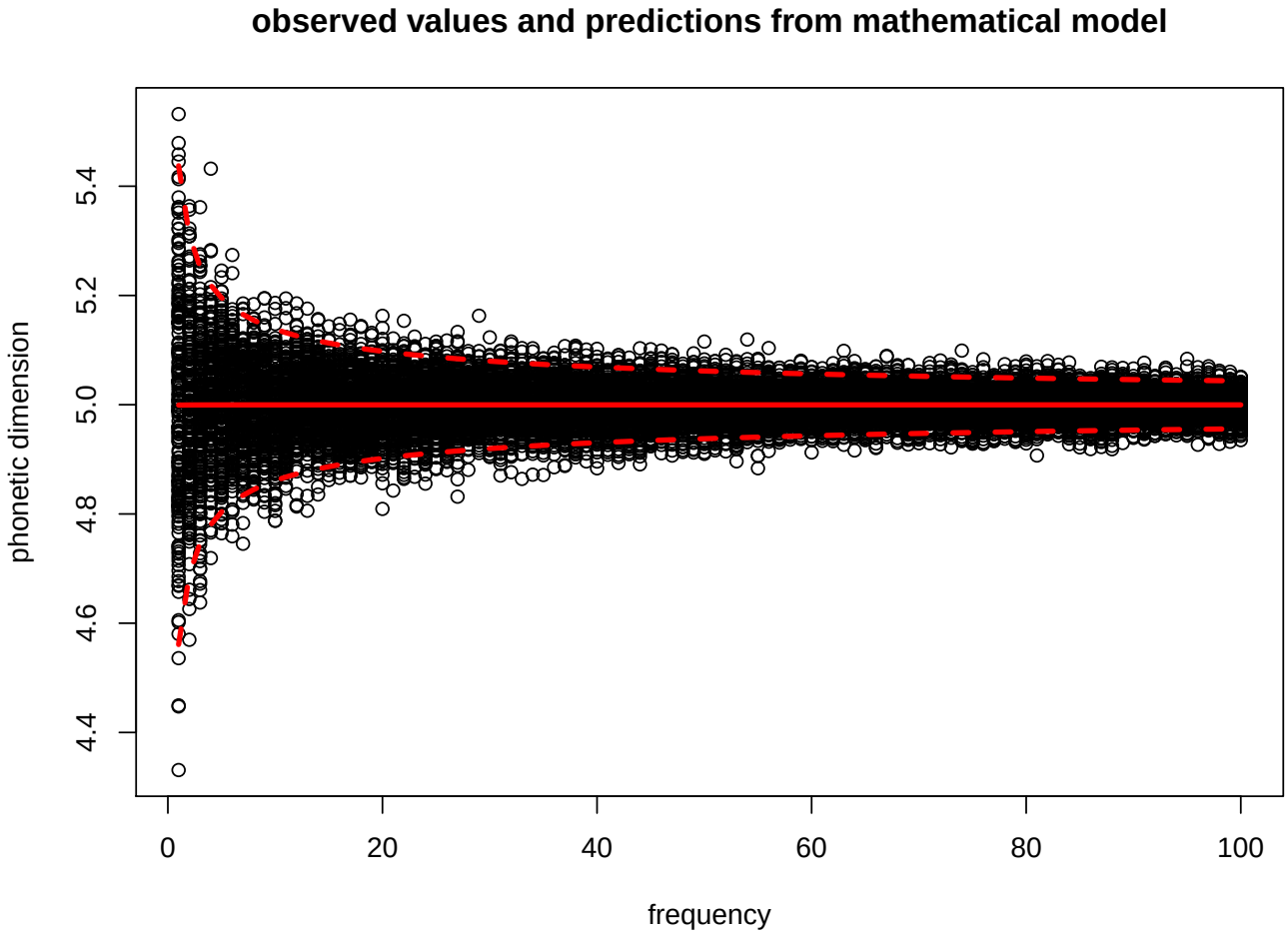
Simulation architecture

- a single phonetic dimension
- each word initialised at 0
- 20,000 simulated word categories
- word frequency varies between 1 and 100 per time unit
- constant positive bias
- 500 time units for each word
- looking at means for each word

Exposure

Results

- word frequency vs word means after 500 time units



Exposure

Results

1. frequency has no effect on the expected values of the means
   - the effect of MASS cancels out the effect of exposure
2. the mean values vary more for low frequency words
   - consequence of the algebra of random variables...
   - this is an empirically testable prediction!

Outline

Independence

Prediction

- instances of the same sound category in different words are not completely independent
- high-frequency forms more independent than low-frequency forms (cf. Bybee (2001))
  - $\rightarrow$ they resist analogical change more
  - $\rightarrow$ they can also stray further from the rest of the category?

Independence

Simulation architecture

- modelling framework: simplified version of Pierrehumbert (2002)
- each word represented separately
- category representation, bias and update same as previously
- sampling is different:
  - samples from a weighted mixture distribution
  - frequent words $\rightarrow$ heavier weight
  - target word $\rightarrow$ heavier weight

Independence

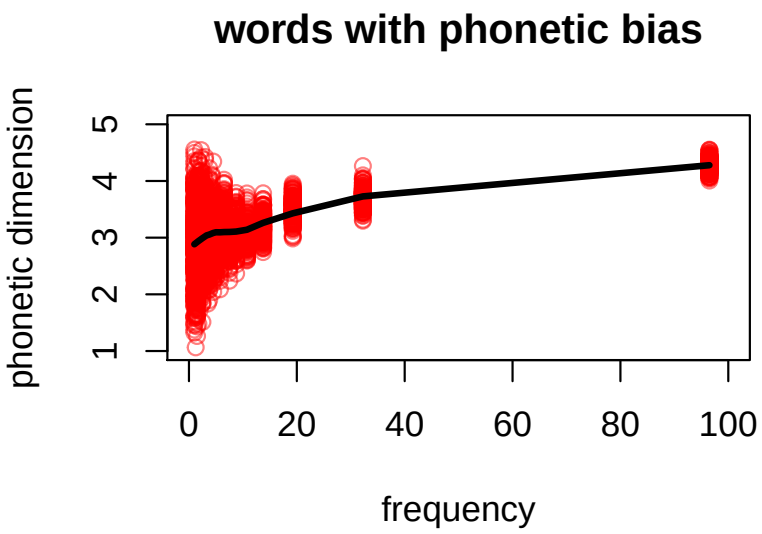
Simulation architecture

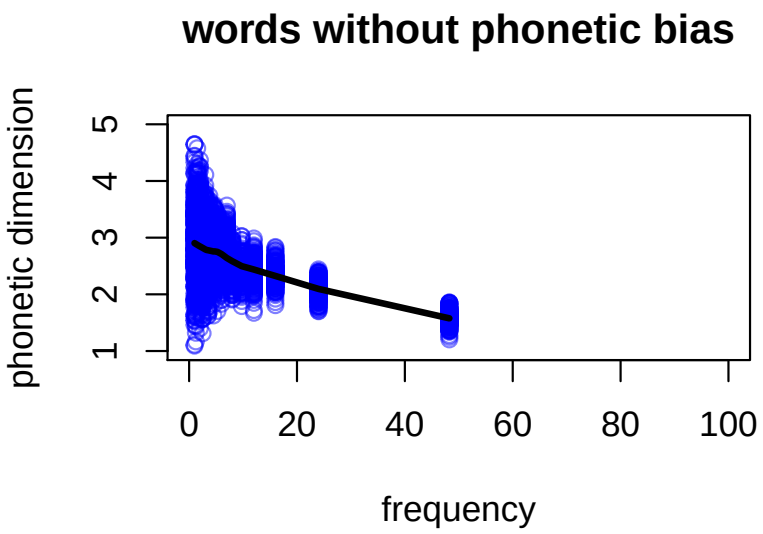
- same as before, except:
- only half of the words are affected by the bias
  - otherwise, results would be identical to prev. sim.
  - similar to e.g. /u/-fronting: *tube* → biased; *cool* → not biased
- 100 word categories (frequencies: Zipf distribution)
- simulation repeated 100 times with same parameters
- looking at one word from each frequency bin

Independence

Results

- word frequency vs word means
  (pooled results from 200 simulations)





Independence

Results

1. frequency has a positive effect for biased words (*tube*-type)
2. frequency has a negative effect for non-biased words (*cool*-type)
3. the mean values vary more for low frequency words in both groups

Outline

Conclusion

- careful modelling is crucial to unpack the predictions of theories of sound change
- exposure:
  - no effect on expected value of mean
  - the evolution of infrequent words is less predictable
- independence:
  - positive frequency effect in trigger environment
  - negative frequency effect in elsewhere environment
  - the evolution of infrequent words is less predictable

Bybee, J. L. (2001). Phonology and language use. Cambridge University Press, Cambridge.

Jurafsky, D., Bell, A., Gregory, M., and Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In Bybee, J. L. and Hopper, P., editors, Frequency and the emergence of linguistic structure, pages 229–254. John Benjamins, Amsterdam.

Phillips, B. S. (2006). Word Frequency and Lexical Diffusion. Palgrave Macmillan, Basingstoke, UK & New York, NY.

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In Bybee, J. L. and Hopper, P., editors, Frequency effects and the emergence of lexical structure, pages 137–157. John Benjamins, Amsterdam.

Pierrehumbert, J. B. (2002). Word-specific phonetics. In Gussenhoven, C. and Warner, N., editors, Laboratory phonology, Vol. VII. Mouton de Gruyter, Berlin.

Schuchardt, H. (1885/1972). On sound laws: Against the Neogrammarians. In Vennemann, T. and Wilbur, T. H., editors, Schuchardt, the Neogrammarians, and the transformational

theory of phonological change (Linguistische Forschungen 26), pages 29–72. Athenaum, Frankfurt am Main.

Conclusion

Simulation architecture for independence simulations