The phonological lexicon, usage factors, and rates of change: evidence from Manchester English

Ricardo Bermúdez-Otero[1]

with George Bailey,[1] Maciej Baranowski,[1] and Danielle Turton[2]

University of Manchester,[1] *Newcastle University*[2]

ConSOLE XXV, Leipzig, 4 Jan 2017

Economic & Social Research Council
Grant RES-000-22-4242

Preview

The problem

Phonetic implementation has been found to be sensitive to **gradient usage-related** properties of lexical items: e.g.

- token frequency
- neighbourhood density

This is a problem for classical **modular** architectures of grammar:

these assume that lexical phonological representations consist solely of **categorical** information.

Two types of response:

- *moderate* e.g. **gradient symbolic activation**
  lexical representations are categorical but gradiently activated
- *radical* e.g. **exemplar theory**
  lexical representations contain fine phonetic detail

The case of **lexical token frequency:**

Synchronic observation

High-frequency words exhibit **more lenition** than low-frequency words.

Diachronic predictions

- Radical approaches (e.g. exemplar theory)

  The erosive effects of use are directly registered in phonetically detailed lexical representations, creating a word-specific feedback loop.

  Hence, words that are used more often **change at a faster rate**.

- Moderate approach (e.g. gradient symbolic activation)

  The gradient effects of usage factors are not stored: there is no word-specific feedback loop.

  High- and low-frequency words **change at the same rate**, producing a **constant rate effect** ('CRE', Kroch 1989).

Our study

We trace the rise of /t/-glottalling in apparent time in Manchester English:

- data from 62 speakers born between 1926 and 1986.

We define the variable as the realization of non-foot-initial /t/ either with or without audible oral constriction:

- oral [t], [ˀt], [ʰt], [ʰtʰ], etc.
- glottal [ʔ]

We distinguish two sub-environments:

- word-final *ma***t** 7056 tokens
- word-medial intervocalic *ma***tt***er* 2131 tokens

Today we focus on the intervocalic cases.

Our findings

As in many previous studies of lenition and coarticulation, we find a strong effect of **lexical token frequency**:

At every point in apparent time, the probability of /t/-glottalling is higher in high-frequency words.

But the data are also consistent with a **constant rate effect**:

We find no evidence against the assertion that /t/-glottalling increases at the same rate in high- and low-frequency words.

We show this by means of two tests.

Test 1: comparison of two **curve-fitting models**

(a) independent logistic curves for 3 frequency bins (Kroch 1989)

(b) one single logistic curve for all words, plus a time-invariant bias for each frequency bin (Kauhanen & Walkden 2015)

Model (b), though more constrained and inherently CRE-compliant, fits the data just as well as model (a).

Test 2: generalized mixed-effects **logistic regression**

- The best model does not include a frequency:birthyear interaction (with frequency entered as a continuous variable).
- Such an interaction would be expected if the effect of frequency changed across apparent time.

The implications

Our findings challenge **exemplar theory**:

- No evidence that high-frequency words change faster.
- So: no need for the lexical representations of individual words to register subcategorical effects of more frequent use.

Our findings support **expanded modular architectures**:

- /t/-glottalling advances as the probability of application of the glottalling rule increases across generations.
- Frequency effects arise from orthogonal time-invariant mechanisms (cascading activation / listener modelling / etc).

Theoretical background:

The phonological lexicon
and the phonetic effects of usage factors

The classical modular feedforward architecture of grammar

Underlying representation (discrete)

↓ phonological rules

Surface representation (discrete)

↓ phonetic rules

Auditory and articulatory representations (continuous)

- Phonological lexical representations consist of discrete categories.
- No ‘fine’ (gradient, subcategorical) phonetic detail in the lexicon.

Argument 1: the double articulation of language

The modular architecture captures the insight that phonology is a **discrete combinatorial system**:

- an arbitrarily large number of signifiers is set up through the recombination of a small number of discrete meaningless units
- signifiers do not have holistic phonetic properties

In contrast,

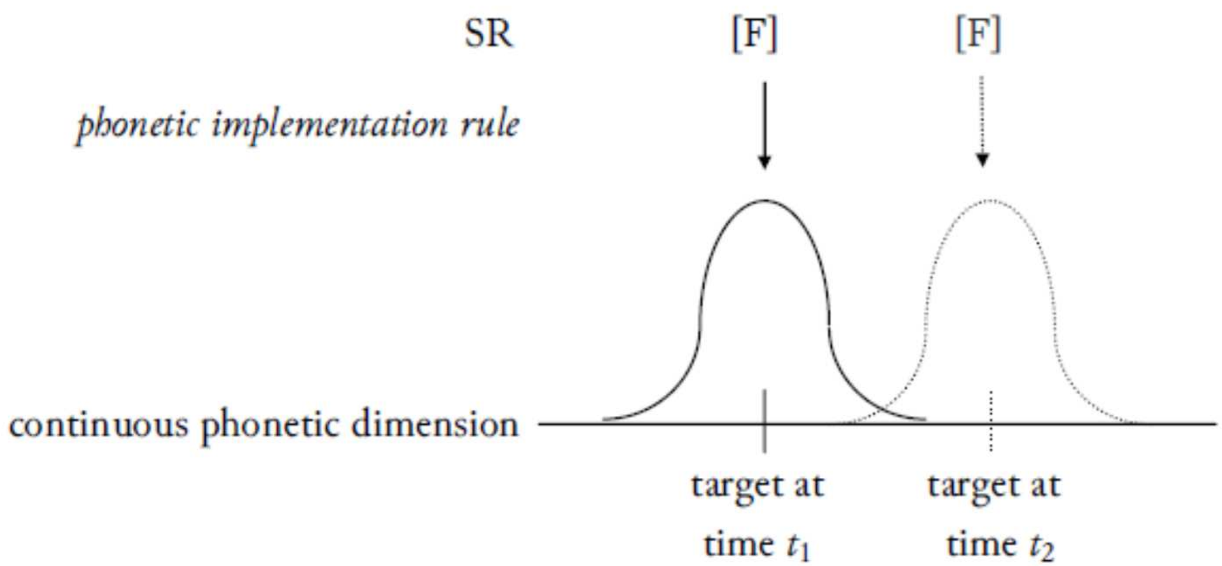
if a communication system relies on holistic signals,

then parsing error imposes a tight upper bound on the number of possible signals (Nowak *et al*. 1999).

Argument 2: neogrammarian change

Modularity explains the existence of neogrammarian change, i.e. **phonetically gradient but lexically regular change.**

- Neogrammarian change affects the implementation rules assigning phonetic targets to discrete categories in phonetic representations.



- Tokens of the same category receive the same phonetic target regardless of the word in which they occur.

> Das bewegungsgefühl bildet sich ja nicht für jedes einzelne wort besonders, sondern überall, wo in der rede die gleichen elemente widerkehren, wird ihre erzeugung auch durch das gleiche bewegungsgefühl geregelt.
>
> (Paul 1886[1880]: 62)

> "A motory sensation [*sc.* ‛articulatory programme'] does not form itself specially for every word, but in every case where the same elements [*sc.* ‛combinatorial units'] recur in speech their production is guided by the same motory sensation."

The challenge: phonetic effects of usage factors

Gradient usage-related lexical properties
- token frequency
- neighbourhood density
- contextual predictability

↓ affect

fine phonetic detail
- duration
- gesture amplitude
- coarticulation(gestural overlap)

The classical modular feedforward architecture of grammar cannot explain this observation:

- there is **no phonetic detail** in underlying or surface representations
- there is **no lexical information** in surface representations

The case of lexical token frequency

High-frequency words are **hypoarticulated**:

- shorter duration (Whalen 1991, Gahl 2008)
- more vowel centralization (Wright 2003, Dinkin 2008)
- more coarticulatory nasalization (Zellou & Tamminga 2014)

etc

The case of neighbourhood density

Neighbourhood density = number of phonologically similar words weighted by frequency.

Words in high-density neighbourhoods are **hyperarticulated**:

- less vowel centralization (Wrigth 2003)
- longer VOT in voiceless plosives (Baese-Berk & Goldrick 2009)

etc

A moderate response: gradient symbolic activation

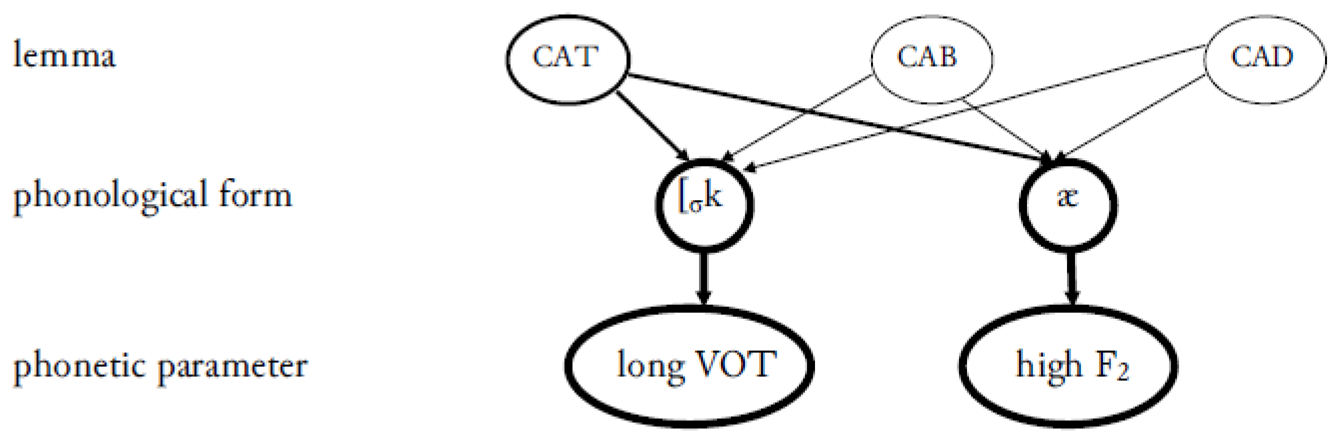
Key ideas

- The modular architecture is correct at Marr's (1982: 25) **computational** level of description.
- The phonetic effects of usage factors arise at Marr's **algorithmic** (processing) level.
- Lexical representations consist solely of **discrete categories**: there is no fine phonetic detail in the lexicon.
- But discrete symbolic representations can be **gradiently activated**.
- Gradient activation **cascades** from higher to lower levels of representation before processing is complete at the higher levels.

Synthesis of classical symbolic and connectionist approaches to cognition (e.g. Smolensky *et al* 2014, Smolensky & Goldrick 2016).

Application to neighbourhood density effects
(Baese-Berk & Goldrick 2009)

Activation cascades from lemmas through phonological forms to phonetic parameters even before lemma selection is complete.



[Line thickness represents activation strength.]

A radical response: exemplar theory

Key ideas

- The lexicon contains exemplar clouds, i.e. collections of **episodic memory traces** containing fine phonetic detail.
- Usage causes lexical representations to be **constantly updated** as old exemplars decay and new exemplars are added to the cloud.
- The phonetic effects of usage factors are thus directly registered in the lexicon.

A few references on relatively ‘pure’ or ‘hard’ exemplar theory:

Goldinger (1998)
Hawkins (2003)
Johnson (2006)
Wade & Möbius (2010)

Lexical frequency effects in exemplar theory

The existence of neogrammarian change is problematic for exemplar theory because exemplar clouds encode word-specific phonetic properties.

Indeed, Bybee (1998, 2002) and Pierrehumbert (2001, 2002) assert that **no change is truly neogrammarian**.

- *Key claim*

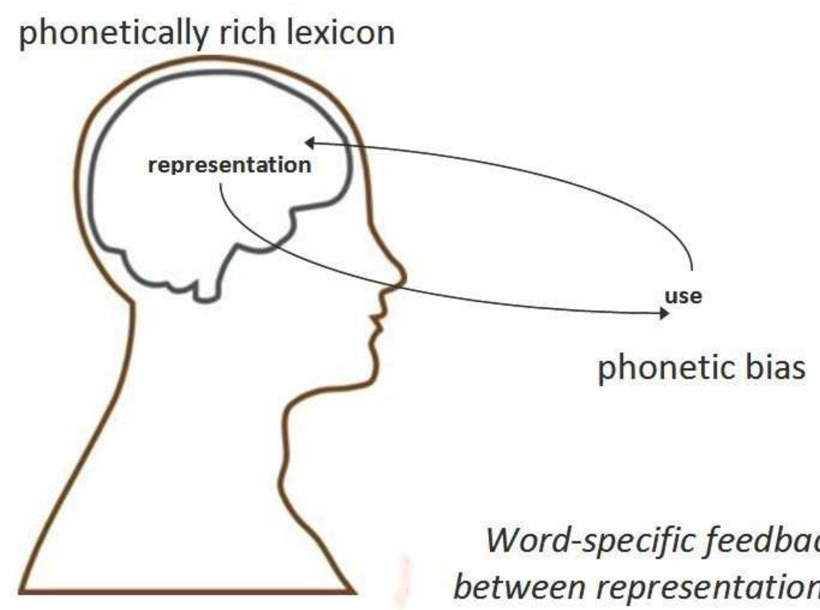
  In diachronic changes involving phonetic reduction (lenition, coarticulation), high-frequency words

  are ahead synchronically

  because they **change faster diachronically**.

Postulated mechanism

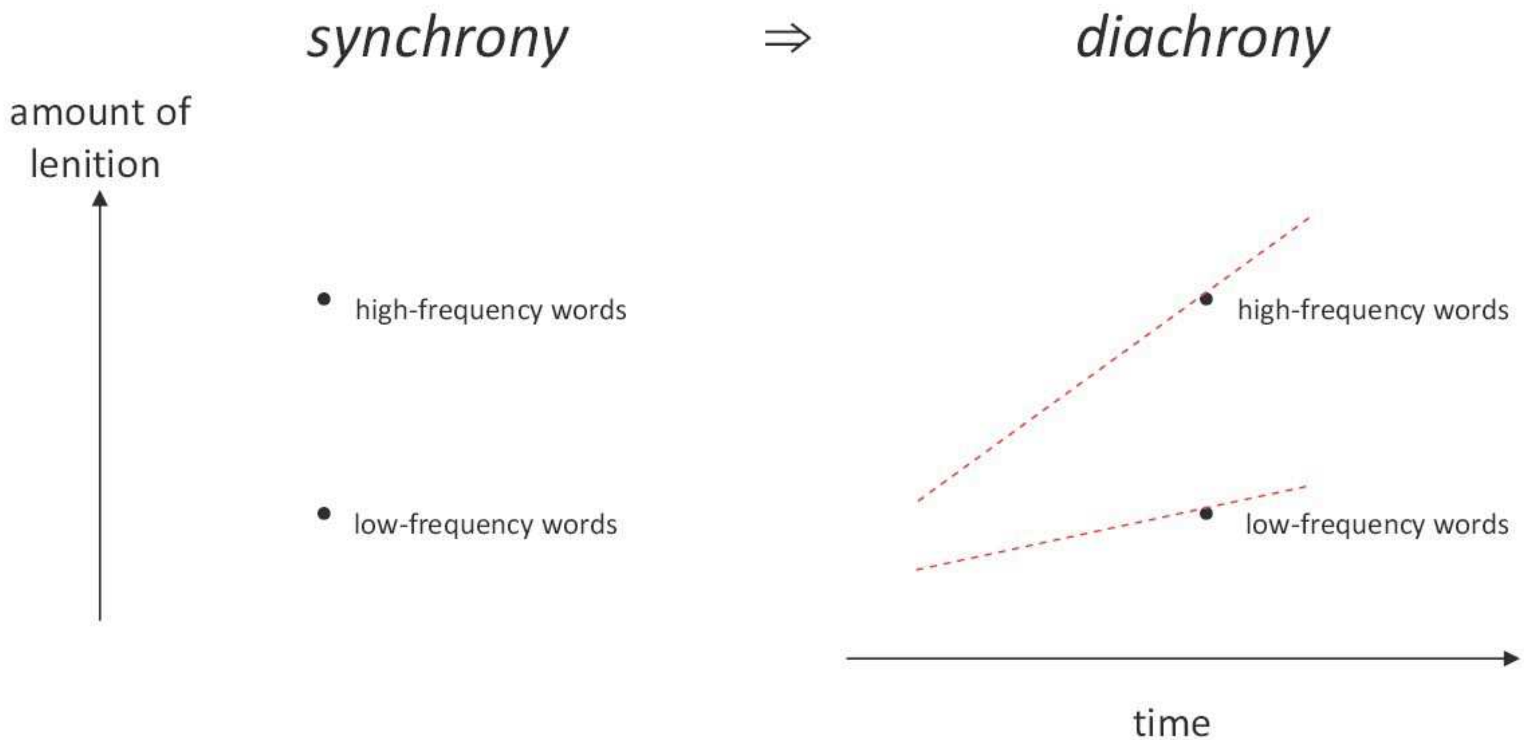
- High-frequency words undergo greater exposure to reductive phonetic biases during language use.
- The effect of these biases is registered in phonetically detailed lexical representations (e.g. exemplar clouds).
- Thus, the feedback loop between mental representation and use is **routed through words, not through phonological categories**.



Word-specific feedback loop between representation and use

Predicted diachronic trajectory

High-frequency words exhibit more lenition synchronically because, diachronically, they undergo leniting changes faster.



Problems for the exemplar-theoretic account

Logic

More synchronic lenition does not entail faster diachronic change.

Theory

The predictions of exemplar theory for rates of change are unclear: e.g. the inertia of a large exemplar cloud could cancel out the effects of greater exposure to phonetic bias (Sóskuthy 2014).

Data

Only one empirical study reports high-frequency words changing faster: Hay & Foulkes (2016) – **but their data are unreliable**.

Hay & Foulkes' (2016) data: [t] > [d] in New Zealand

collected in the 1940s
birthyears from 1862 to 1900

c. 50 years apart in collection
36 years' gap in apparent time

collected in the 1990s
birthyears from 1936 to 1982

Mobile Unit Corpus

833 tokens

Canterbury Corpus

1,846 tokens

Pre-Labovian: slow speech

No change in apparent time
despite 50%+ rate of [d] in males!

No effect of token frequency

HETEROGENEOUS CORPORA

Post-Labovian: normal speech

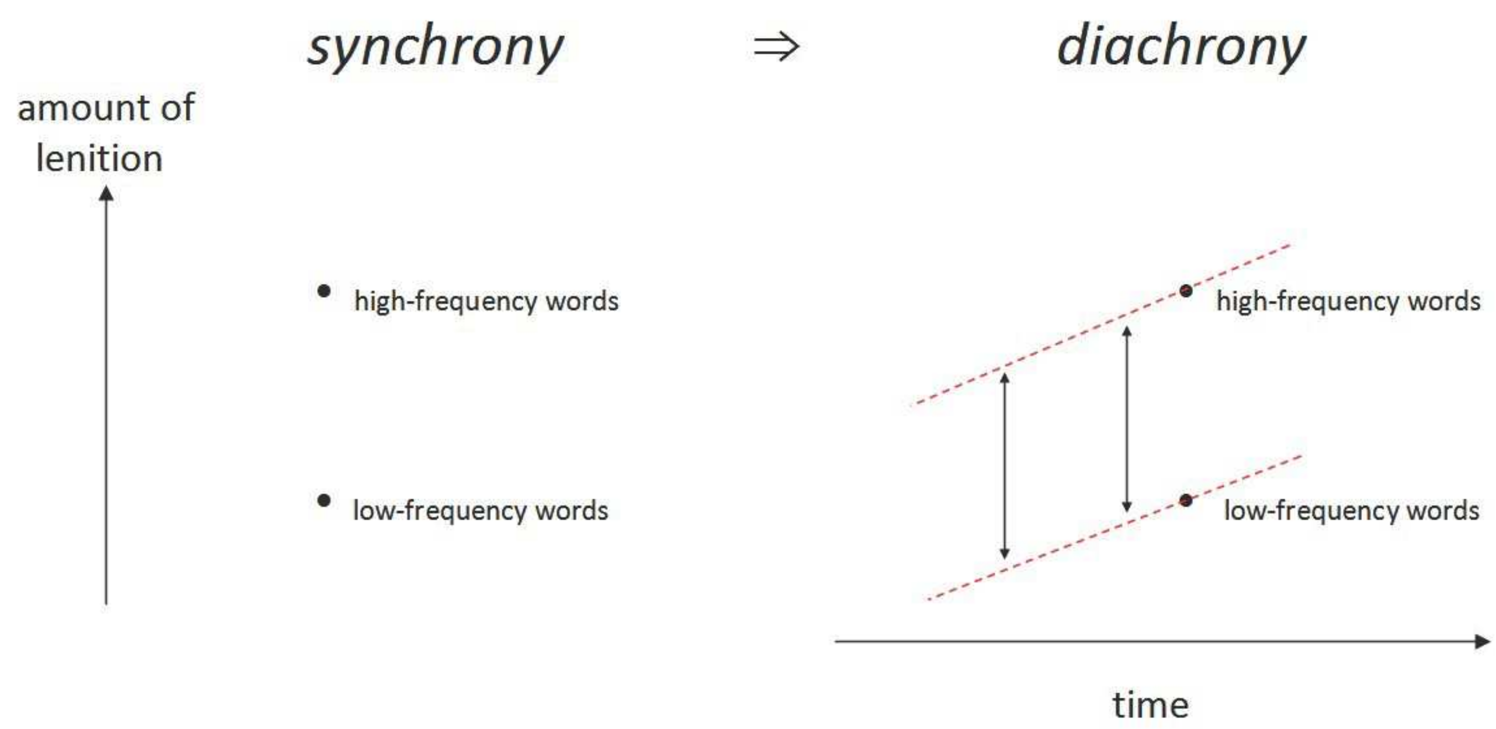
Change in apparent time

Effect of token frequency
but no interaction with time!

The putative interaction between token frequency and time arises only by extrapolating across a 36-year gap between two highly heterogeneous corpora.

Alternative hypothesis

High-frequency words are ahead of low-frequency ones synchronically, but change in parallel diachronically.



Hay et al. (2015) acknowledge that this is a logical possibility, which they label ‘a stationary frequency effect’.

In fact, it would be just a case of a widespread phenomenon: the **constant rate effect** (Kroch 1989)

- widely attested in syntactic change
  (Pintzuk 2003, Zimmermann 2015)
- shown to arise in phonological changes too
  (Fruehwald et al. 2013)
- detected in frequency-conditioned coarticulatory change in one previous study
  (Zellou & Tamminga 2014)
    although Zellou & Tamminga do not cite Kroch or use the term ‘constant rate effect’.

Methodological issues:

Testing for
constant rate effects

The constant rate effect: one concept, several operationalizations

The concept

If a single variable rule $\mathfrak{R}$ realizes /A/ as [B] in environments E1 and E2,

- and the frequency of application of $\mathfrak{R}$ rises through time,
- and E1 is subject to a time-invariant bias in favour of [B],
- and E2 is subject to a time-invariant bias against [B],

then, in a sense to be made precise, it should be the case that

- [B] will appear **‘earlier’** in the favouring environment E1,
- [B] will appear **‘later’** in the disfavouring environment E2,
- but the occurrence of [B] will grow **‘at the same rate’** in E1 and E2 since this growth is determined by the application frequency of $\mathfrak{R}$.

Particular theories of the CRE must:

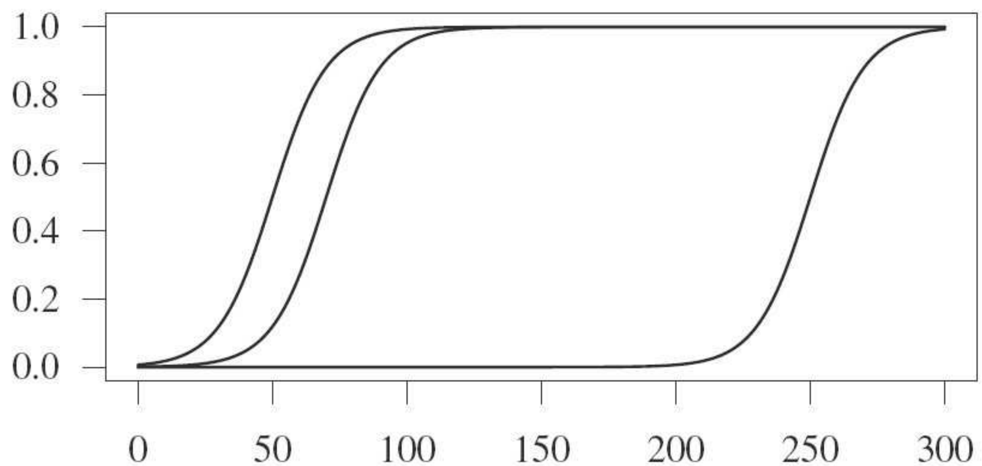
- provide operational definitions of ‘earlier’, ‘later’, and ‘same rate’,
- and specify the nature and source of the biases.

Kroch's operationalization
Independent logistic curves are fitted to each environment:

$$p_t = \frac{e^{s(t-k)}}{1+e^{s(t-k)}} = \left(1+e^{-s(t-k)}\right)^{-1}$$

A constant rate effect is said to occur iff the curves have:

- **different intercepts** (*k*)
- but **identical slopes** (*s*)



Drawbacks of Kroch's operationalization

- In sound change, growth curves are often not logistic (e.g. nonmonotonic change in Zellou & Tamminga 2014).
- The criterion of identical slopes and different intercepts is a guess; it is not derived mathematically from first principles.
- Available significance tests for slope differences are suspect (Kauhanen p.c.).
- Needs categorically defined contexts (e.g., for our purposes, discrete frequency bins, instead of frequency continuum).
- Does not limit the time lag between change in favouring and disfavouring contexts (Kauhanen & Walkden 2015).
- Demonstrably yields some false positives, i.e. mislabels some pseudo-CREs as true CREs (e.g. Corley 2014).

Kauhanen & Walkden's operationalization

- Explicit mathematical modelling of the interaction between a single logistic growth curve and time-invariant contextual biases: the CRE is 'built in'.
- Sets stringent limitations on the time lag between contexts.
- Known to avoid false positives incurred by Kroch's procedure.

But:

- still based on a logistic curve,
- still requires categorically defined contexts.

Another popular operationalization

Generalized mixed-effects **logistic or linear regression**.

A constant rate effect is deemed to have occurred if

a context:time interaction is not significant,
consistent with the effect of context being stable across time.
E.g. Fruehwald et al. (2013), Zellou & Tamminga (2014).

Problems:

- Absence of evidence for different rates ≠ evidence for identical rates.
- As corpus size grows, so does the risk of incorrectly rejecting a CRE.
- The procedure demonstrably rejects some true CREs:

  many time series generated by Kauhanen & Walkden's intrinsically CRE-compliant model exhibit significant context:time interactions.

The methodological challenge summarized

There is no problem-free operationalization of the concept 'rate of change'.

Hence, it is technically impossible to demonstrate that high- and low-frequency words change at the same rate.

Our response

We assume the hypothesis that frequency effects do cause CREs.

We subject this hypothesis to two empirical tests that

- are stringent
- have complementary strengths and weaknesses.

If the CRE hypothesis survives both tests, we conclude that

we have strong epistemic reasons to retain the CRE hypothesis because it has greater empirical content than the alternative.

Test 1

- Fit independent logistic curves to each frequency bin à la Kroch.
- Fit an intrinsically CRE-compliant model à la Kauhanen & Walkden.
- Compare the error of the two models.
- If the error is the same, the data do not require independent logistic curves for different frequency bins.

Test 2

- Generalized mixed-effects logistic regression with frequency as a continuous variable (no binning).
- If the optimal model does not include a frequency:birthyear interaction, the data are consistent with the effect of frequency remaining stable across apparent time.

The results of both of our tests turn out to be consistent with a frequency-driven CRE in Manchester /t/-glottalling.

Data

The phenomenon: /t/-glottalling in Manchester

- Phonological process whereby /t/ is replaced by a glottal stop in non-foot-initial position:

  i.e. **full glottal replacement.**

- Reported for accents all over the UK in recent years:

  e.g. Williams & Kerswill (1999), Foulkes & Docherty (1999).

- **Manchester** (Baranowski & Turton 2015)

  The change is more vigorous **word-medially**:

  e.g. *pre**tt**y, wa**t**er*, etc.

Speakers

- ESRC funded project on Manchester English:

  sociolinguistic interviews collected by local fieldworkers and students.

- Subset of **62 speakers**, born **between 1926 and 1986**.

  All grew up in Manchester (within M60); born to local parents.

  Ethnicity: white British.

  Socio-economic status: 5 levels, based on occupation (lower-working, upper-working, lower-middle, middle-middle, upper-middle)

Coding

- **Dependent variable**
  - oral: [t], [ˀt], [ʰt], [ʰtʰ], etc.
  - glottal: [ʔ]

- **Independent variables**
  - year of birth
  - gender
  - social class (5 occupational levels)
  - word frequency: Zipf-scaled frequency from **SUBTLEX-UK** (van Heuven *et al* 2014)

- **Random effects**
  - word
  - speaker

Results

Steps in the analysis

1) **Data visualization** (LOESS curves in R)

2) **Statistical tests**

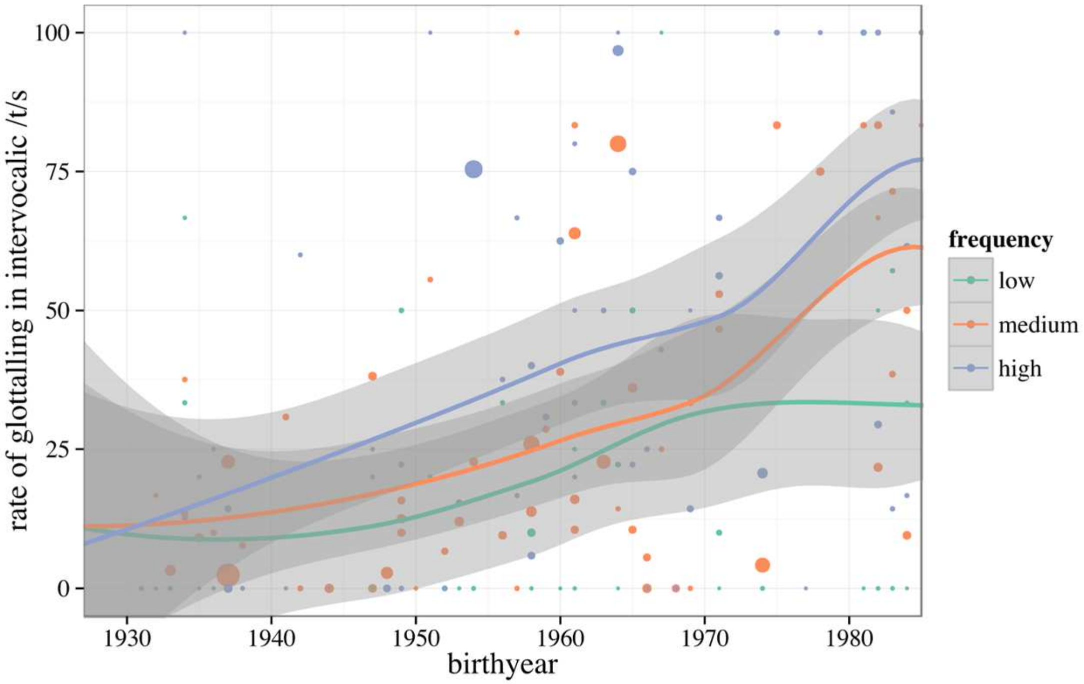
    2.1) Comparison of curve-fitting models

    - independent logistic curves (Kroch 1989)
    - single logistic + context biases (Kauhanen & Walkden 2015)

    2.2) Mixed-effects logistic regression

LOESS curves

- LOESS (locally-weighted polynomial regression) smoothing curves in R applied to each context for all intervocalic tokens (N = 2131)
- Visual inspection of curves suggests roughly parallel trajectories of change, although the low-frequency bin seems to stabilise around 1970
- However, largely overlapping 95% confidence intervals suggest that one should not read too much into these curves
- This method is also prone to smoothing artefacts



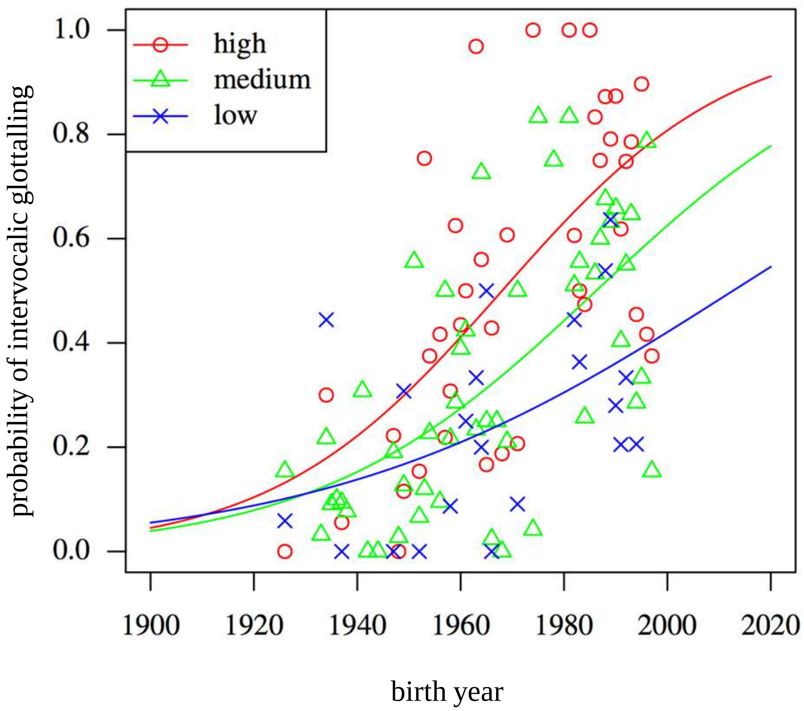
Independent logistic curves

- This method mirrors the Krochian procedure of fitting independent logistic curves to each context (Kroch 1989)

- Each year is represented by three points, reflecting the average rate of glottalling in the low-, medium-, and high-frequency contexts

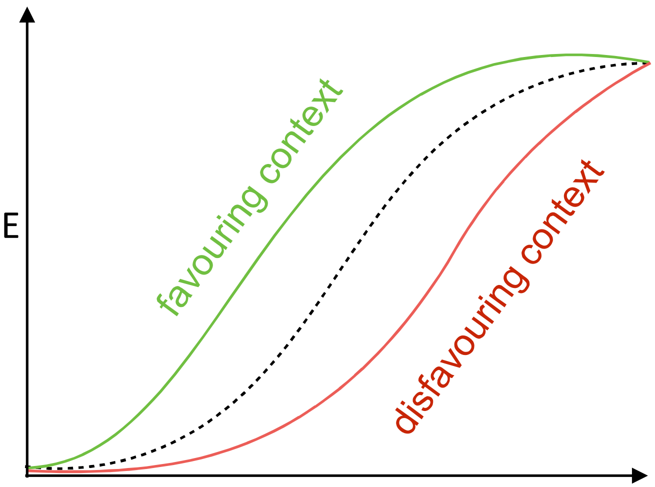
- Fairly parallel slopes here are suggestive of a CRE. However, significance tests for slope comparisons are poorly understood.

- For this reason, we are not using Kroch's model alone, but as a term of comparison with a model with a single logistic.

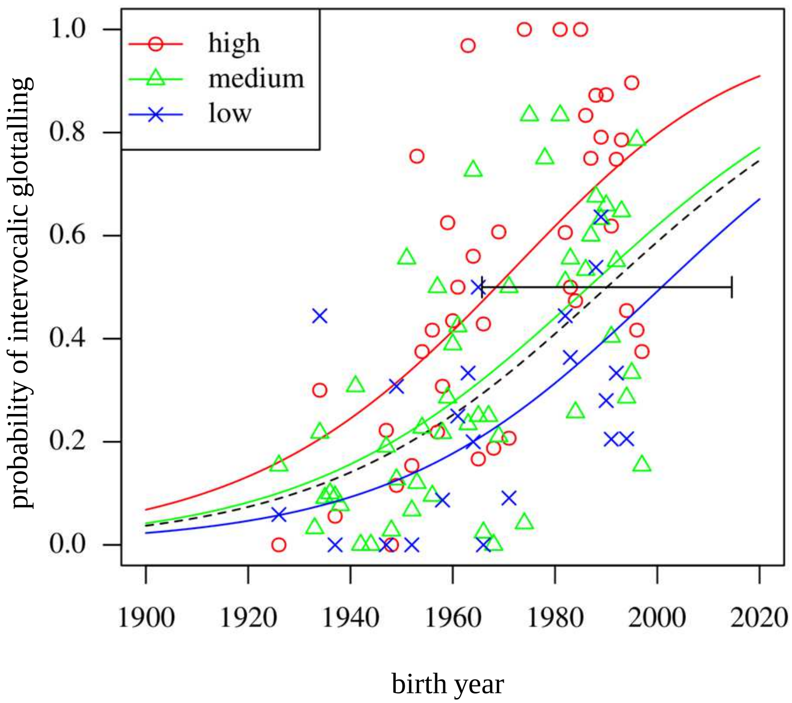


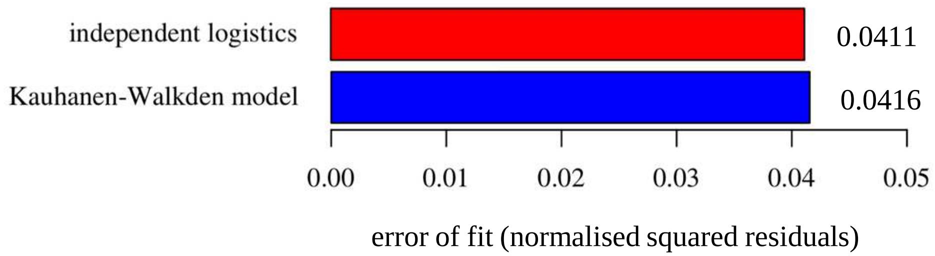
The Kauhanen & Walkden model

- This model has been developed by colleagues at the University of Manchester (Kauhanen & Walkden 2015)
- Calculates an overall rate of change across all contexts, modelled by a logistic function (denoted by the dashed line)
- Generates context-specific curves by applying a time-invariant bias either for or against the change depending on the favourability of the context
- The curves fitted by this model are constrained *a priori* to produce a CRE
- Can be used as a diagnostic by comparing the error rates of these CRE-constrained curves against the earlier model that plots independent logistic curves for each context
- This procedure also imposes a time constraint between the point at which the change reaches its fastest rate in the favouring and disfavouring contexts
- If the data falls within this time window, and the fit of this more constrained model leads to no increase in error compared to the Krochian model, this would be highly suggestive of a CRE



The Kauhanen & Walkden model





- Our data satisfies both criteria:
  - falling within the permissible time separation as indicated by the horizontal line
  - showing a comparable error rate between the two models (0.0416, cf. 0.0411)
- Fitting this more constrained model, with the CRE built in, leads to no increase in error over a model with completely independent logistic curves

Logistic regression

- These models do not rely on arbitrary binning; *frequency* is entered as a continuous predictor.

- If frequency effects were to change over time, this would be reflected in an interaction between *frequency* and *birth year* in the optimal model

Best model: t ~ birth year + social class + gender + word frequency + (1 | word) + (1 | speaker)

	Estimate	Standard Error	z-value	p-value
(intercept)	-9.76	2.13	-4.58	< 0.001
birth year	0.09	0.02	4.99	< 0.001
class: 1	-2.81	1.54	-1.82	0.07
class: 2	-3.29	1.53	-2.15	0.03
class: 3	-3.62	1.57	-2.31	0.02
class: 4	-3.97	1.56	-2.55	0.01
class: 5	-6.21	1.89	-3.29	< 0.001
gender: male	1.41	0.55	2.54	0.01
word frequency	1.09	0.22	4.96	< 0.001

- Adding an interaction into the model does not lead to a significant increase in fit (by ANOVA comparison, $p = 0.18$), and in fact makes a worse model based on AIC (1565.5, cf. 1565.3) and BIC (1633.2, cf. 1627.3)

Implications

Our findings suggest that the impact of lexical token frequency on sound change produces **constant rate effects**.

Challenge to **exemplar theory**:

- No evidence that frequent words change faster.
- Data consistent with the hypothesis that the feedback loop between mental representation and use is routed through phonological categories, and not through lexical items.

Paraphrasing Bloomfield,
phonemes (or allophones, or features) change, **not words**.

New addition to the list of problems for exemplar theory:

- difficulty in accounting for the existence of neogrammarian change

  (Pierrehumbert 2002)

- parsing challenge in matching auditory input to holistic fine-grained targets

  (German, Carlson & Pierrehumbert 2013: 230)

- failure to account for instantaneous generalization of phonetic learning across the lexicon

  (McQueen et al. 2006, Cutler et al. 2010)

- little evidence that episodic detail primes word recognition under naturalistic conditions

  (McLennan 2007: 68, Hanique et al. 2013)

- negative results in shadowing tasks

  (Mitterer & Ernestus 2008)

etc.

The discovery of frequency-driven constant rate effects calls for ***expanded modular architectures****:*

- Sound change advances as the rate of application of a rule increases across generations.
- The same rule applies to high- and low-frequency words, resulting in equal rates of change.
- Stationary frequency effects are triggered by an **orthogonal time-invariant mechanism.**

Does **symbolic gradient computation** provide the right account of frequency effects?

Possibly: it meets the criterion of time-invariance.

However, there are problems:

- There is no fully developed account of frequency effects from this perspective.
- Neighbourhood-density effects and frequency effects pattern differently, and so must involve different mechanisms (Munson 2007; Goldrick *et al* 2011: 69).

There are other approaches meeting the criterion of time-invariance: e.g. **listener modelling** (Lindblom 1990).

Thank you!

We are grateful to
Henri Kauhanen and George Walkden
for helpful comments and suggestions.

We gratefully acknowledge the support of the UK's
Economic & Social Research Council
(grant RES-000-22-4242).

References

Baese-Berk, Melissa & Matthew Goldrick. 2009. Mechanisms of interaction in speech production. *Language and Cognitive Processes* 24 (4), 527-54.

Baranowski, Maciej & Danielle Turton. 2015. Manchester English. In Raymond Hickey (ed.), *Researching Northern English*, 293-316. Amsterdam/Philadelphia: John Benjamins.

Bybee, Joan. 1998. The phonology of the lexicon: evidence from lexical diffusion. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language*, 65-85. Stanford: CSLI Publications.

Bybee, Joan. 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14 (3), 261-90.

Corley, Kerry. 2014. *The Constant Rate Effect in syntactic change: empirical fact or "lies, damned lies, and statistics"?* Undergraduate dissertation, University of Cambridge.

Cutler, Anne, Jason Eisner, James M. McQueen & Dennis Norris. 2010. How abstract phonemic categories are necessary for coping with speaker-related variation. In Cécile Fougeron, Barbara Kühnert, Mariapaola D'Imperio & Nathalie Vallée (eds.), *Laboratory Phonology 10* (Phonology and Phonetics 4-4), 91-111. Berlin: De Gruyter Mouton.

Dinkin, Aaron J. 2008. The real effect of word frequency on phonetic variation. In Joshua Tauberer, Aviad Eilam & Laurel MacKenzie (eds.), *Proceedings of the 31st Annual Penn Linguistics Colloquium* (University of Pennsylvania Working Papers in Linguistics 14.1), 97-106. Philadelphia: Penn Linguistics Club.

Fruehwald, Josef, Jonathan Gress-Wright & Joel Wallenberg. 2013. Phonological rule change: the Constant Rate Effect. In Seda Kan, Clair Moore-Cantwell & Robert Staubs (eds.), *NELS 40: Proceedings of the 40th Annual Meeting of the North East Linguistic Society*, vol. 1, 219-30. Amherst, Massachusetts: Graduate Linguistics Student Association, University of Massachusetts.

Gahl, Susanne. 2008. *Time* and *thyme* are not homophones: the effect of lemma frequency on word durations in spontaneous speech. *Language* 84 (3), 474-96.

References

German, James S., Katy Carlson & Janet B. Pierrehumbert. 2013. Reassignment of consonant allophones in rapid dialect acquisition. *Journal of Phonetics* 41 (3-4), 228-48.

Goldinger, Stephen D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105 (2), 251-79.

Hanique, Iris, Ellen Aalders & Mirjam Ernestus. 2013. How robust are exemplar effects in word comprehension? *The Mental Lexicon* 8 (3), 269 – 94.

Hawkins, Sarah. 2003. Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics* 31, 373-405.

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.

Hay, Jennifer & Paul Foulkes. 2016. The evolution of medial /t/ over real and remembered time. *Language* 92 (2), 298-330.

Johnson, Keith. 2006. Resonance in an exemplar-based lexicon: the emergence of social identity and phonology. *Journal of Phonetics* 34 (4), 485-99.

Kauhanen, Henri & George Walkden. 2015. Deriving the Constant Rate Effect. Ms, University of Manchester.

Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1 (3), 199-244.

Lindblom, Björn. 1990. Explaining phonetic variation: a sketch of the H&H theory. In William J. Hardcastle & Alain Marchal (eds.), *Speech production and speech modelling*, 403-39. Dordrecht: Kluwer.

McLennan, Conor T. 2007. Challenges facing a complementary-systems approach to abstract and episodic speech perception. *Proceedings of the16th International Congress of Phonetic Sciences (ICPhS XVI), Saarbrücken, 6-10 August 2007*, 67-70.

References

McQueen, James M., Anne Cutler & Dennis Norris. 2006. Phonological abstraction in the mental lexicon. *Cognitive Science* 30 (6), 1113 - 26.

Mitterer, Holger & Mirjam Ernestus. 2008. The link between speech perception and production is phonological and abstract: Evidence from the shadowing task. *Cognition* 109 (1), 168-73.

Myers, James & Yingshing Li. 2009. Lexical frequency effects in Taiwan Southern Min syllable contraction. *Journal of Phonetics* 37 (2), 212-30.

Nowak, Martin A., David C. Krakauer & Andreas Dress. 1999. An error limit for the evolution of language. *Proceedings of the Royal Society of London, Series B (Biological Sciences)* 266, 2131-6.

Paul, Hermann. 1886. *Principien der Sprachgeschichte,* 2nd edn. Halle: Max Niemeyer. 1st edn 1880.

Pierrehumbert, Janet. 2001. Exemplar dynamics: word frequency, lenition and contrast. In Joan L. Bybee & Paul Hopper (eds.), *Frequency and the emergence of language structure*, 137-57. Amsterdam: John Benjamins.

Pierrehumbert, Janet. 2002. Word-specific phonetics. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory Phonology 7*, 101-39. Berlin: Mouton de Gruyter.

Pintzuk, Susan. 2003. Variationist approaches to syntactic change. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics*, 509-28. Oxford: Blackwell.

Smolensky, Paul & Matthew Goldrick. 2016. Gradient symbolic representations in grammar: the case of French liaison. Ms, Johns Hopkins University and Northwestern University. Available as ROA 1286 at the Rutgers Optimality Archive, http://roa.rutgers.edu.

Smolensky, Paul, Matthew Goldrick & Donald Mathis. 2014. Optimization and quantization in gradient symbol systems: a framework for integrating the continuous and the discrete in cognition. *Cognitive Science* 38 (6), 1102-38.

Sóskuthy, Márton. 2014. Explaining lexical frequency effects: a critique and an alternative account. Paper presented at Sound Change in Interacting Human Systems, 3rd Biennial Workshop on Sound Change, University of California, Berkeley, 29 May 2014.

References

Wade, Travis & Bernd Möbius. 2010. Detailed phonetic memory for multi-word and part-word sequences. *Laboratory Phonology* 1 (2), 283.

Whalen, D. H. 1991. Infrequent words are longer in duration than frequent words. *The Journal of the Acoustical Society of America* 90 (4), 2311.

Wright, Richard. 2003. Factors of lexical competition in vowel articulation. In John Local, Richard Ogden & Rosalind Temple (eds.), *Phonetic interpretation: Papers in Laboratory Phonology VI*, 75-87. Oxford: Oxford University Press.

Zellou, Georgia & Meredith Tamminga. 2014. Nasal coarticulation changes over time in Philadelphia English. *Journal of Phonetics* 47, 18-35.

Zimmermann, Richard. 2015. A syntactic change with lots of data: the rise of *do*-support with possessive 'have' in American English. Paper presented at the Linguistics and English Language Research Seminar, University of Manchester, 3 November 2015.