**EGG School, Banja Luka**
July/August 2018

# Does word frequency affect phonology? Reasons to be cautious... 4

Patrick Honeybone
University of Edinburgh
patrick.honeybone@ed.ac.uk

**The contents of this session**

1. Reasons to be very cautious about UBP
2. What are categorical frequency effects?
3. What is the final balance of predictions?

**How do things stand in terms of the full balance of predictions...?**

1. Low frequency ('frequency conserving') effects should exist
- UBP – yes ✓
- FP – yes ✓

2. High frequency effects should exist
- UBP – yes ✓ (?)
- FP – no × (?)

3. High frequency effects should *always* exist in 'natural' changes/rules
- UBP – yes ×
- FP – no ✓

4. Frequency effects, like all phonological generalisations, should always be gradient
- UBP – yes
- FP – no
- will it make all the difference...?

## *th* in Glasgow Scots

Like most varieties of British English, /θ/ in Glasgow is subject to '*th*-fronting', which involves the realisation as [f]:

- θ > f can be seen as being a change in progress
  - implemented as a variable ('optional') θ → f

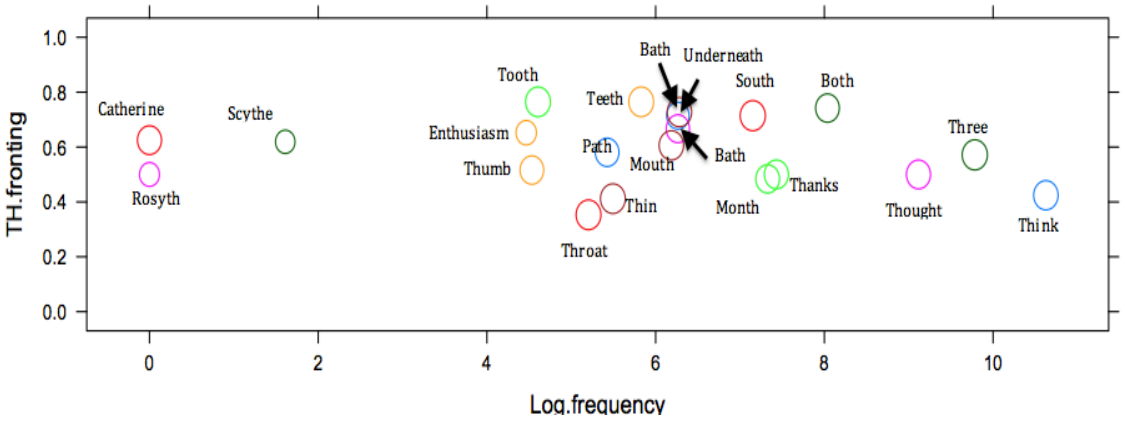
*think* [θɪŋk] ~ [fɪŋk]

*bath* [baθ] ~ [baf]

A number of studies have shown that there is no frequency effect in this phenomenon

- Clark & Trousdale (2009), Nielsen (2010), Schleef & Ramsammy (2013)
  - it is exceptionlessly conditioned by phonological environment
  - syllabic/word position: it is most likely in the word-final position
  - segmental neighbourhood: following [round] segments make it more likely

Nielsen (2010), interpreting data from Stuart-Smith for Glasgow, shows that all words are affected, and the differences are not statistically significant



However, Glasgow Scots (like other varieties of Scots) has a more complicated situation

- there is also what Clark & Trousdale (2009) call:
  - "the change from [θ] to [h] [which] involves a reduction (in the form of lenition)"

This means that forms of words with [h] can also occur:

*think* [θɪŋk] ~ [fɪŋk] ~ [hɪŋk]

Stuart-Smith & Timmins (2006) have considered the realisation of these forms in Glasgow in detail, counting the tokens of each type for all possible words in a large corpus their results show considerable variation...

In word-medial position:

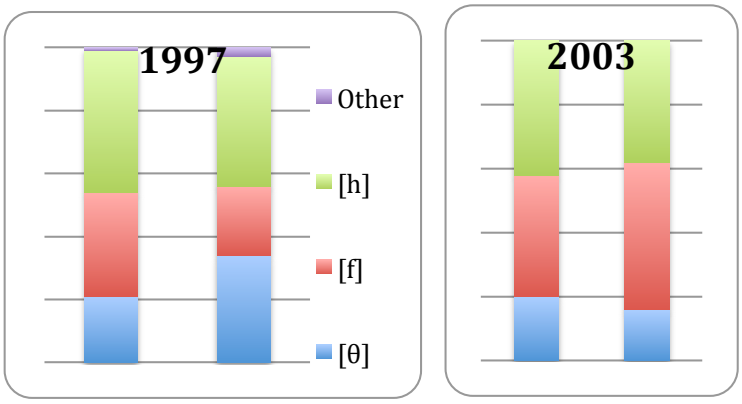
| *1997* | | | | | 2 | *2003* | | | |
|---|---|---|---|---|---|---|---|---|---|
| **word** | **[th]** | **[f]** | **[h]** | **Total** | **word** | **[th]** | **[f]** | **[h]** | **Total** |
| something | 2 | 2 | 31 | 36 | something | 5 | 3 | 153 | 161 |
| nothing | 3 | 7 | 21 | 31 | anything | 3 | 1 | 46 | 50 |
| anything | 4 | 3 | 15 | 23 | nothing | 3 | 1 | 40 | 44 |
| Catholics | 6 | 1 | | 7 | healthy | 3 | 39 | | 42 |
| everything | | 1 | 4 | 6 | everything | | | 21 | 21 |
| without | 3 | | | 3 | birthday | 3 | 10 | 1 | 14 |
| Catherine | | 2 | | 2 | without | | 9 | | 9 |
| menthol | | 1 | | 1 | maths | | 5 | | 5 |
| Strathmore | 1 | | | 1 | Samantha | 1 | 4 | | 5 |
| | | | | | Jonathon | 1 | 3 | | 4 |
| | | | | | Dorothy | | 2 | | 2 |
| | | | | | athletics | | 1 | | 1 |
| | | | | | birthdays | | 1 | | 1 |
| | | | | | Blackthorn | 1 | | | 1 |
| | | | | | Catherine | | 1 | | 1 |
| | | | | | enthusiasm | | 1 | | 1 |
| | | | | | Grath's | 1 | | | 1 |
| | | | | | months | 1 | | | 1 |
| | | | | | Southpark | 1 | | | 1 |

In word-initial position:

**Table 4.** Lexical distribution of main variants for /th/ in word-initial position in working-class adolescents in 1997 and 2003 in descending order of frequency.

| *1997* | | | | | *2003* | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **word** | **[th]** | **[f]** | **[h]** | **Total** | **word** | **[th]** | **[f]** | **[h]** | **Total** |
| think | 13 | 6 | 45 | 64 | think | 23 | 22 | 155 | 200 |
| thing | 4 | 5 | 19 | 30 | thing | 11 | 8 | 59 | 78 |
| thought | 4 | 22 | | 26 | three | 12 | 32 | 1 | 45 |
| three | 6 | 11 | | 17 | thought | 8 | 33 | | 41 |
| through | 9 | 5 | | 14 | things | 5 | 1 | 30 | 36 |
| thirty | 6 | 2 | | 8 | through | 10 | 21 | | 31 |
| thinks | 1 | | 6 | 7 | thingy | 6 | | 9 | 15 |
| thingmy | 3 | | 2 | 5 | thousand | 6 | 8 | | 14 |
| thank | 2 | | | 2 | thingwy | 2 | | 11 | 13 |
| third | | 2 | | 2 | thinking | 4 | 3 | 3 | 10 |
| thirteen | 2 | | | 2 | thirty | 7 | 2 | | 9 |
| thematic | 1 | | | 1 | third | 2 | 6 | | 8 |
| thieving | 1 | | | 1 | Thursday | 3 | 5 | | 8 |
| thinking | 1 | | | 1 | thingmy | 5 | | 2 | 7 |
| throat | | 1 | | 1 | thinks | | 1 | 5 | 6 |
| thrown | | 1 | | 1 | thank | 5 | | | 5 |
| Thursday | 1 | | | 1 | thingummy | 1 | 1 | 2 | 4 |
| | | | | | thistle | | 4 | | 4 |
| | | | | | threw | | 4 | | 4 |
| | | | | | thin | 1 | 2 | | 3 |
| | | | | | thirteen | 3 | | | 3 |
| | | | | | thingd | | | 2 | 2 |
| | | | | | thingwys | | | 2 | 2 |
| | | | | | throwing | | 2 | | 2 |
| | | | | | thanks | | 1 | | 1 |
| | | | | | theft | | 1 | | 1 |
| | | | | | therapist | | 1 | | 1 |
| | | | | | thingamajiggy | 1 | | | 1 |
| | | | | | thingy'd | | | 1 | 1 |
| | | | | | thinked | 1 | | | 1 |
| | | | | | thinness | | 1 | | 1 |
| | | | | | thong | | 1 | | 1 |
| | | | | | throat | | 1 | | 1 |

In comparing Stuart-Smith & Timmins' two corpora, a couple of things are clear:




- There's a much higher ratio of [f] recorded in 2003 than in 1997
- The proportion of [h] has not changed between 1997 and 2003
- also, there's something odd in terms of where [h] can occur...

In word-initial position:

**Table 4.** Lexical distribution of main variants for /th/ in word-initial position in working-class adolescents in 1997 and 2003 in descending order of frequency.

| *1997* | | | | | *2003* | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **word** | **[th]** | **[f]** | **[h]** | **Total** | **word** | **[th]** | **[f]** | **[h]** | **Total** |
| think | 13 | 6 | 45 | 64 | think | 23 | 22 | 155 | 200 |
| thing | 4 | 5 | 19 | 30 | thing | 11 | 8 | 59 | 78 |
| thought | 4 | 22 | | 26 | three | 12 | 32 | 1 | 45 |
| three | 6 | 11 | | 17 | thought | 8 | 33 | | 41 |
| through | 9 | 5 | | 14 | things | 5 | 1 | 30 | 36 |
| thirty | 6 | 2 | | 8 | through | 10 | 21 | | 31 |
| thinks | 1 | | 6 | 7 | thingy | 6 | | 9 | 15 |
| thingmy | 3 | | 2 | 5 | thousand | 6 | 8 | | 14 |
| thank | 2 | | | 2 | thingwy | 2 | | 11 | 13 |
| third | | 2 | | 2 | thinking | 4 | 3 | 3 | 10 |
| thirteen | 2 | | | 2 | thirty | 7 | 2 | | 9 |
| thematic | 1 | | | 1 | third | 2 | 6 | | 8 |
| thieving | 1 | | | 1 | Thursday | 3 | 5 | | 8 |
| thinking | 1 | | | 1 | thingmy | 5 | | 2 | 7 |
| throat | | 1 | | 1 | thinks | | 1 | 5 | 6 |
| thrown | | 1 | | 1 | thank | 5 | | | 5 |
| Thursday | 1 | | | 1 | thingummy | 1 | 1 | 2 | 4 |
| | | | | | thistle | | 4 | | 4 |
| | | | | | threw | | 4 | | 4 |
| | | | | | thin | 1 | 2 | | 3 |
| | | | | | thirteen | 3 | | | 3 |
| | | | | | thingd | | | 2 | 2 |
| | | | | | thingwys | | | 2 | 2 |
| | | | | | throwing | | 2 | | 2 |
| | | | | | thanks | | 1 | | 1 |
| | | | | | theft | | 1 | | 1 |
| | | | | | therapist | | 1 | | 1 |
| | | | | | thingamajiggy | 1 | | | 1 |
| | | | | | thingy'd | | | 1 | 1 |
| | | | | | thinked | 1 | | | 1 |
| | | | | | thinness | | 1 | | 1 |
| | | | | | thong | | 1 | | 1 |
| | | | | | throat | | 1 | | 1 |

One thing is noticeable about the occurrence of [h]: it only occurs in a few morphemes:

- practically only *think* and *thing* (+*three*)

If the data is reanalysed on the basis of the morphemes involved, the picture becomes clearer...

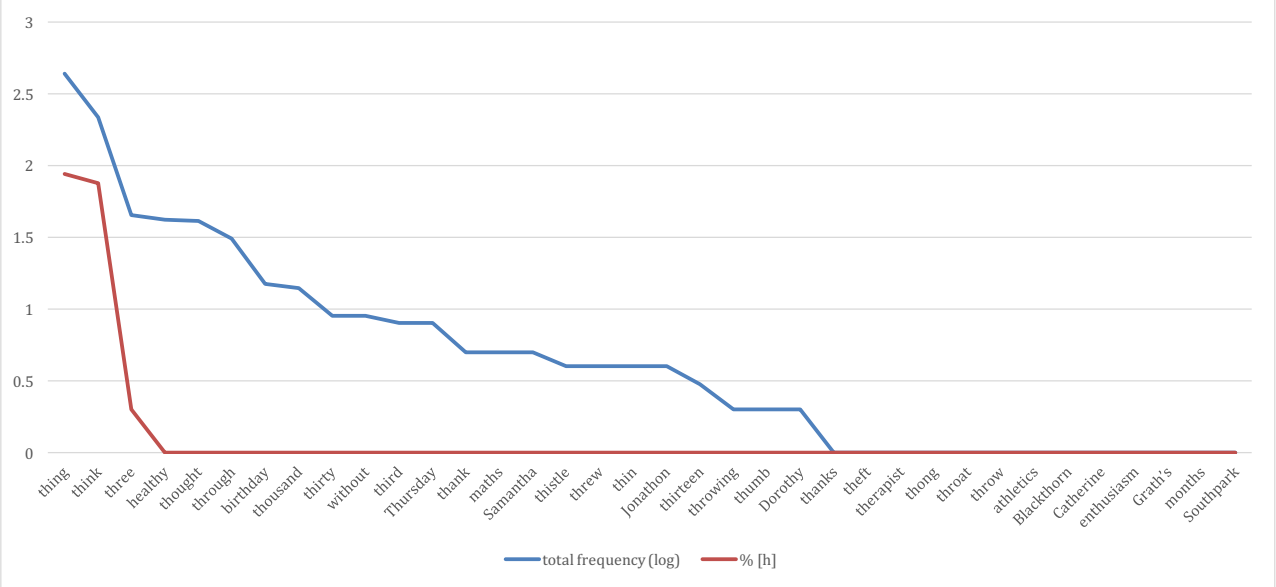
| | h | total | % |
|---|---|---|---|
| *thing* | 378 | 435 | 87 |
| *think* | 163 | 217 | 75 |
| *three* | 1 | 45 | 2 |
| *healthy* | 0 | 42 | 0 |
| *thought* | 0 | 41 | 0 |
| *through* | 0 | 31 | 0 |
| *birthday* | ~~1~~ | 15 | 0 |
| *thousand* | 0 | 14 | 0 |
| *thirty* | 0 | 9 | 0 |
| *without* | 0 | 9 | 0 |
| *third* | 0 | 8 | 0 |
| *Thursday* | 0 | 8 | 0 |
| *thank* | 0 | 5 | 0 |
| *maths* | 0 | 5 | 0 |
| *Samantha* | 0 | 5 | 0 |
| *thistle* | 0 | 4 | 0 |
| *threw* | 0 | 4 | 0 |
| *thin* | 0 | 4 | 0 |
| *Jonathon* | 0 | 4 | 0 |
| *thirteen* | 0 | 3 | 0 |
| *throwing* | 0 | 2 | 0 |
| *thumb* | 0 | 2 | 0 |
| *Dorothy* | 0 | 2 | 0 |
| *thanks* | 0 | 1 | 0 |
| *theft* | 0 | 1 | 0 |
| *therapist* | 0 | 1 | 0 |
| *thong* | 0 | 1 | 0 |
| *throat* | 0 | 1 | 0 |
| *throw* | 0 | 1 | 0 |
| *athletics* | 0 | 1 | 0 |
| *Blackthorn* | 0 | 1 | 0 |
| *Catherine* | 0 | 1 | 0 |
| *enthusiasm* | 0 | 1 | 0 |
| *Grath's* | 0 | 1 | 0 |
| *months* | 0 | 1 | 0 |
| *Southpark* | 0 | 1 | 0 |

## What's going on in Glasgow *th*?

There is a categorical frequency effect in terms of the occurrence of [h]

- only three morphemes allow [h] and they are the most frequent
  - but there's no 'gradient tail' to this, which would be expected in a UBP model

This chart shows (log) frequency and % [h] in words – they don't fit...




This shows that there is a small group of words which can feature [h], and they all are indeed frequent

- however, no other word can feature [h] at all
  - UBP expects patterns to be gradient because exemplars are not categorial things (unlike URs in autonomous phonology), so words like *thought* and *through* should show *some* occurrences of [h] – word-frequency is gradient
  - UBP expects the close correlation of frequency and gradience of occurrence that we saw in our first consideration of Coronal Stop Deletion
- this pattern is a categorical frequency effect which indicates that the words which can feature [h] must be marked in the lexicon in some way as being able to partake in the pattern, while others are not
  - this is more compatible with FP, where URs have categorical properties which can be marked exceptionally in URs, than it is with UBP, where URs do not exist

The diachrony of the situation explains all this, as shown below...

The words below refer to whole classes of lexical items:

- *hat* = *hat, hot, howl, ahead, Bahamas, alcohol, adulthood, behave, unhinged* etc.
- *think* = *thing, think* **only** (along with other frequent words, for some speakers: *three*?)
- *thin* = *thistle, thin, through, birthday, healthy, Samantha* etc.

At stage I, there is nothing to see...

| | URs with /h/ | URs with /θ/ |
|---|---|---|
| STAGE I | *hat* /h/ : [h] | *think, thin* /θ/ : [θ] |

At stage II, a (variable) rule of θ → h is innovated

- this is a normal kind of debuccalisation, although is it inhibited in coda position
- as it is variable, if it doesn't apply, /θ/ surfaces as [θ]
  - there is dialectal evidence that the process was once quite widespread

| | *think* | | *bath* | |
|---|---|---|---|---|
| | /θɪnk/ | /θɪnk/ | /baθ/ | /baθ/ |
| θ → h | hɪnk | — | bah | — |
| *CODA-h | — | — | * | — |
| | [hɪŋk] | [θɪŋk] | * | [baθ] |

| | URs with /h/ | URs with /θ/ |
|---|---|---|
| STAGE I | *hat* /h/ : [h] | *think, thin* /θ/ : [θ] |
| STAGE II NEW PROCESS | *hat* /h/ : [h] | *think, thin* /θ/ → [h] (variable)<br>– if this doesn't apply, /θ/ : [θ] |

There is then a substantial reanalysis – crucially before the variation is lost

- this involves the loss of θ → h
- and a reanalysis of the URs involved
  - crucially, the variation is retained by assuming that variable words have two URs:

    *thing* = /θɪŋ, hɪŋ/ – either UR can be used

Importantly for our purposes, having two URs is unusual.

NB: it is not unprecedented (remember this...?)

- ‹indefinite article› = /ə, ən/
- ‹THIRD PERSON oblique PRONOUN› = /ðəm, əm/

MacKenzie (2013) argues for the following, among others:

*would* = /wʊd$_{PW}$, d/
*should* = /ʃʊd$_{PW}$, ʃd/
*have* = /hav$_{PW}$, v/

Each word can have its own tendency in terms of when each of the URs are employed, just as in the following:

*either* /aɪðə, iðə/
*neither* /naɪðə, niðə/

The words below refer to whole classes of lexical items:

- *hat* = *hat, hot, howl, ahead, Bahamas, alcohol, adulthood, behave, unhinged* etc.
- *think* = *thing, think* **only** (along with other frequent words, for some speakers: *three*?)
- *thin* = *thistle, thin, through, birthday, healthy, Samantha* etc.
- green = a lexical entry with two URs
- blue = synchronic process

| | URs with /h/ | URs with /θ/ |
|---|---|---|
| STAGE I | *hat* /h/ : [h] | *think, thin* /θ/ : [θ] |
| STAGE II NEW PROCESS | *hat* /h/ : [h] | *think, thin* /θ/ → [h] (variable) – if this doesn't apply, /θ/ : [θ] |
| STAGE III REANALYSIS + LOSS OF θ → h | *hat, think, thin* /h/ : [h] | *think, thin* /θ/ : [θ] |

It is uneconomical for a lexical item to have two URs, and we might imagine that a pressure to assume simple lexical entries, with one UR will militate against them

- lexical entries with two URs will therefore be under diachronic pressure to simplify and lose one UR
- which lexical entries are most likely to retain the two URs...?
  - those which learners have lots of experience of = high frequency words
- which UR is likely to be lost?
  - there would be pressure from English, which is spoken all through the environment where Scots is spoken, to retain the *θ*-ful UR

The words below refer to whole classes of lexical items:

- *hat* = *hat, hot, howl, ahead, Bahamas, alcohol, adulthood, behave, unhinged* etc.
- *think* = *thing, think* **only** (along with other frequent words, for some speakers: *three*?)
- *thin* = *thistle, thin, through, birthday, healthy, Samantha* etc.
- green = a lexical entry with two URs
- blue = synchronic process

| | URs with /h/ | URs with /θ/ |
|---|---|---|
| STAGE I | *hat* /h/ : [h] | *think, thin* /θ/ : [θ] |
| STAGE II NEW PROCESS | *hat* /h/ : [h] | *think, thin* /θ/ → [h] (variable)<br>– if this doesn't apply, /θ/ : [θ] |
| STAGE III REANALYSIS + LOSS OF θ → h | *hat, think, thin* /h/ : [h] | *think, thin* /θ/ : [θ] |
| STAGE IV MOST LEXICAL ENTRIES SIMPLIFY | *hat, think* /h/ : [h] | *think, thin* /θ/ : [θ] |

All this means that there is a low frequency effect in which (high) frequency 'conserves' the lexical entries with two URs

- only frequent words have retained two URs since the reanalysis of the θ → h process
- words which took part in the θ → h process but cannot now have [h] have lost the *h*-ful UR through diachronic change
  - subsequent generations of children 'misfixed' the URs, performing an act of 'simplification'/'regularisation' – the low frequency effect is predictable
  - this leaves only the *θ*-ful UR

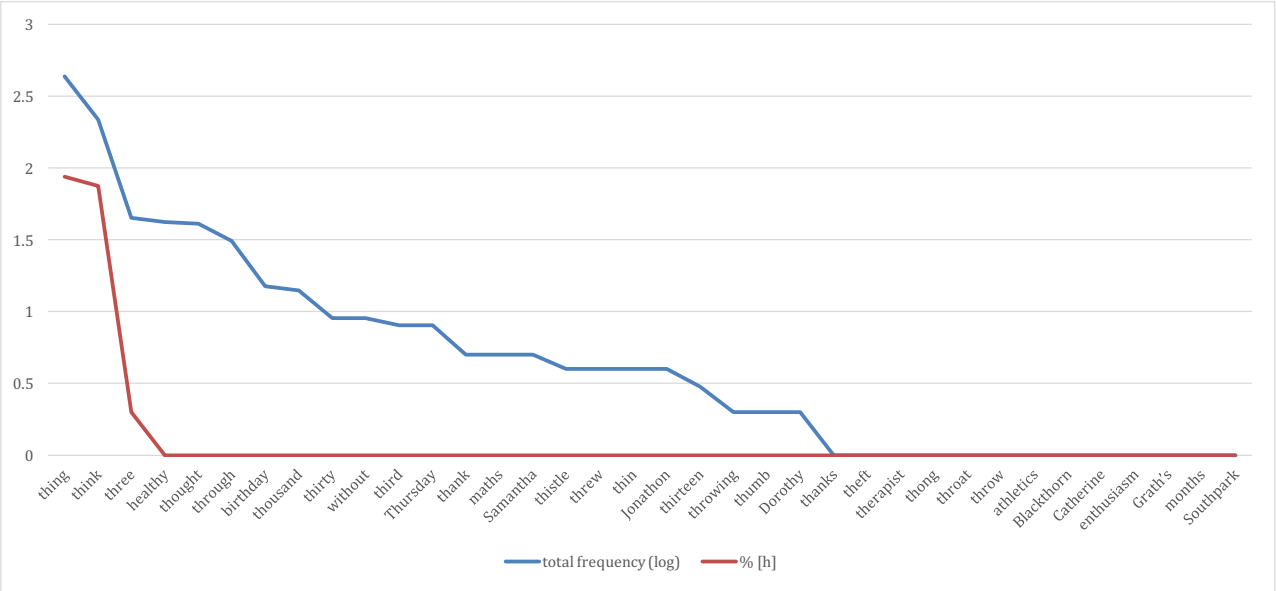
The frequency effect identified here has these properties:

- is categorical because it relies on two distinct phonological objects: URs
- it is a diachronic effect
- it is a low frequency effect: low frequency words do not support the positing of two URs by children acquiring the language

NB: as we have seen, low frequency effects are entirely compatible with FP

NB: as we have seen, this is a categorical frequency effect, which are predicted by FP

- UBP predicts that the gradual decrease in lexical token frequency should be accompanied by a gradual decrease proportion of [h]-ful pronunciations
  - this prediction is not met



## But are categorical frequency effects really a thing?

How about T-to-R? What is T-to-R...? Here's an example, from Stainforth, South Yorkshire:

*We crack on and we get it done – the quicker we get it done, we can get off.*

[gɛɹɪt] [gɛɹɪt] [gɛɹɒf]

So… something *t*-related is realised as something *r*-like

- the environment involves at least V_V

One key characteristic of T-to-R is:

T-to-R is neutralising and structure-preserving in terms of the segments involved

- it neutralises a dialect's /t : r/ contrast ([gɛɹɪt] ~ [fɛɹɪt] = *get it* ~ *ferret*)
- mostly this entails that both are realised as [ɹ]
- the realisation of /t/ as a flap [ɾ] only counts as T-to-R if the underlying rhotic /r/ is commonly realised as [ɾ] in the variety
  - (as in Liverpool and the West Midlands of England)

Wells (1982) gave the phenomenon the name 'T-to-R'

- he gives the following informal rule:

t → r / [short V] __ #[V]

*Where* is T-to-R?

● ● = attested

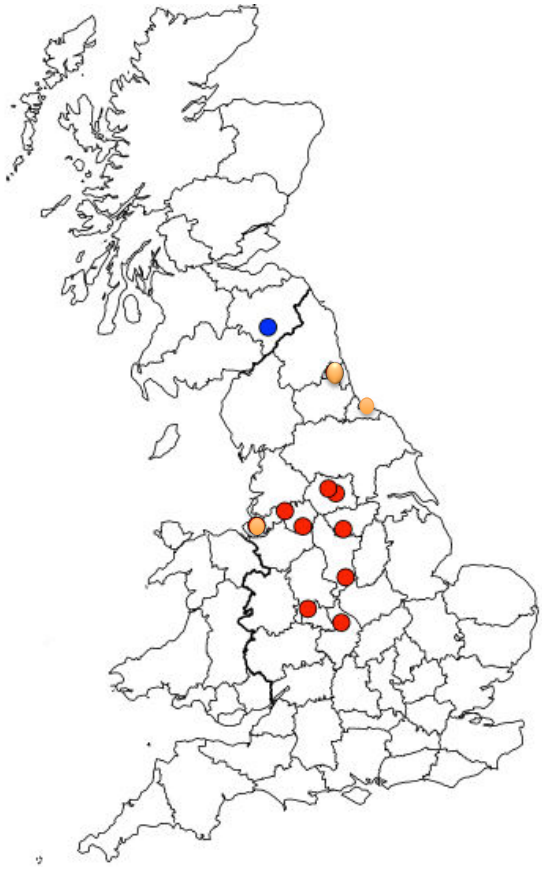
● = shown to be absent

T-to-R has been investigated in a number of places in some detail

- I have results from:

●

  - Newcastle upon Tyne
  - Liverpool
  - Redcar



There have been serious investigations of T-to-R using corpora of natural speech

Clark, L. & Watson, K. (2011) 'Testing claims of a usage-based phonology with Liverpool English *t*-to-*r*.' *English Language and Linguistics* 15, 523-47.

- WC speakers born in the early 1900s; 4 female + 4 male informants
  - 669 tokens of potential T-to-R sites
  - 330 tokens of rhotics in these environments

T-to-R it is fundamentally a cross-lexical (= post-lexical) phenomenon

Clark & Watson (2011)

- "In Liverpool English intervocalic (t) contexts, *t*-to-*r* occurred … word-medially in only 1.5 per cent of all instances (always in the word *whatever*)."

Thus, in its normal patterning, the second vowel is supplied by a following vowel-initial word, so there can be such alternations as:

| | | | |
|---|---|---|---|
| shut | [ʃʊt] | shut up | [ʃʊɹʊp] |
| *get* | [gɛt] | *get off* | [gɛɹɒp] |

T-to-R is definitely a variable phenomenon: forms with obstruents are also possible

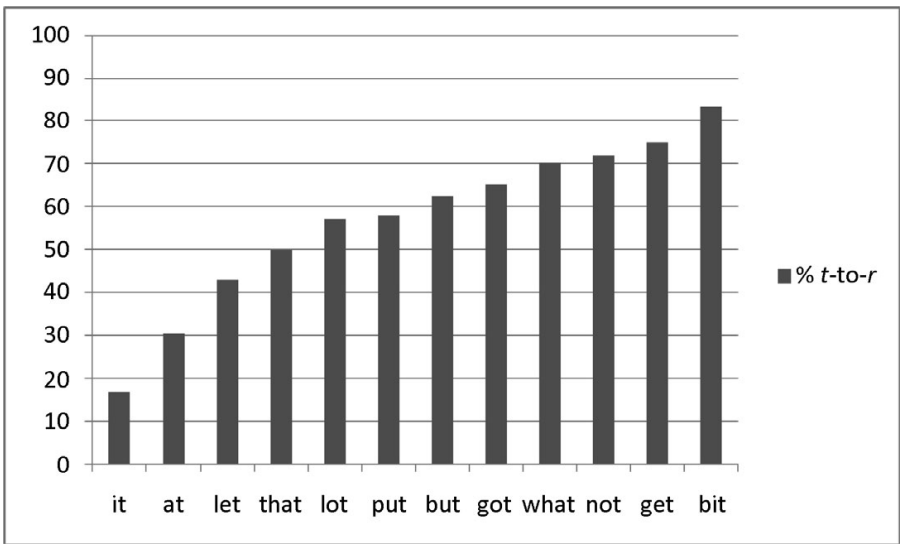
Docherty, Foulkes, Milroy, Milroy & Walshaw (1997) also show this in a data from three speakers (A, B, C) from a Newcastle upon Tyne corpus:

| Speaker | Glottalised | [ɹ] |
|---|---|---|
| A | got a nice jacket<br>got a dark red car | got a little bow |
| B | got a real monkey | got it<br>got a big black dog |
| C | get out | got an accent<br>put in |

**T-to-R: corpora are not enough**

Clark & Watson's (2011) 330 tokens of T-to-R all come from 12 words

- this confirms that T-to-R is variable: no word has 100% rhotic realisations




It has been claimed that T-to-R only occurs in such words – it never occurs in *other* words

- however, "corpora cannot attest the absence of something" (Scheer 2013)

Could it simply be that Clark & Watson's corpus is too small?

- can T-to-R *really* occur in *all* words? perhaps other words occur rarely and, as T-to-R is optional, no corpus has yet caught it in them?

No! In work based at Edinburgh, we have shown that:

T-to-R is lexically-conditioned – it can occur in certain words, but is not possible in other words which have essentially the same phonological form

This can be shown through the elicitation of intuition-based grammaticality judgements

- using a methodology developed for an investigation in Newcastle upon Tyne:
  - Buchstaller, I., Corrigan, K., Holmberg, A., Honeybone, P. & Maguire, W. (2013) 'T-to-R and the Northern Subject Rule: questionnaire-based spatial, social and structural linguistics.' *English Language and Linguistics* 17, 85-128.
- and also applied in Liverpool:
  - Caffrey, Catherine (2011) 'T-to-R in Liverpool English.' MA (Hons) dissertation, University of Edinburgh.
- and in Redcar:

Scanlon, Bobby (2015) 'What do speakers know? The case of T-to-R in Redcar.'

The intuition investigations rely on the fact that speakers of T-to-R dialects are aware of T-to-R, and can thus recognise it when spelled

- after a contextualisation, the informants were asked questions like the following:

Can you pronounce *not* with an *r*?

For example, can you say: *Oh no - norragain!*
[normal spelling: Oh no - not again!]

**1**----------**2**----------**3**

1. I would never pronounce this word with an *r*.
2. I can sometimes pronounce this word with an *r*, but I wouldn't do it very often.
3. It would be normal for me to pronounce this word with an *r*.

Can you pronounce *knot* with an *r*?

For example, can you say: *Oh no - he's tied it in a knorragain!*
[normal spelling: Oh no - he's tied it in a knot again!]

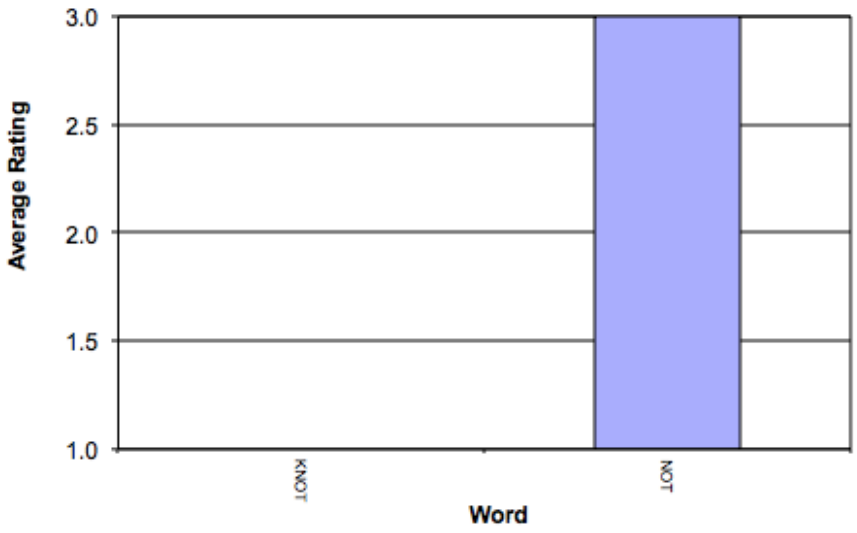
**1**----------**2**----------**3**

This allowed us to test if T-to-R is possible with a range of phonological similar words

- eg, *not, knot, got, dot, that, cat, get, wet*

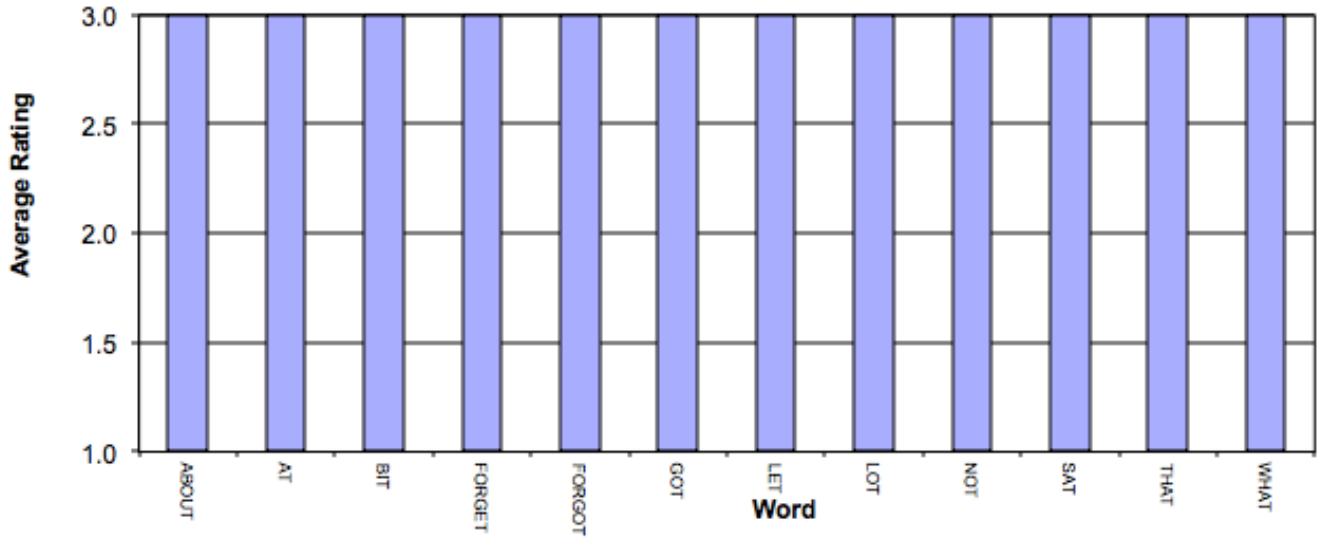
The results are clear…

Liverpool: *not* vs *knot*




Liverpool informants were happy to use score 3.

Liverpool: all words with an average of 3



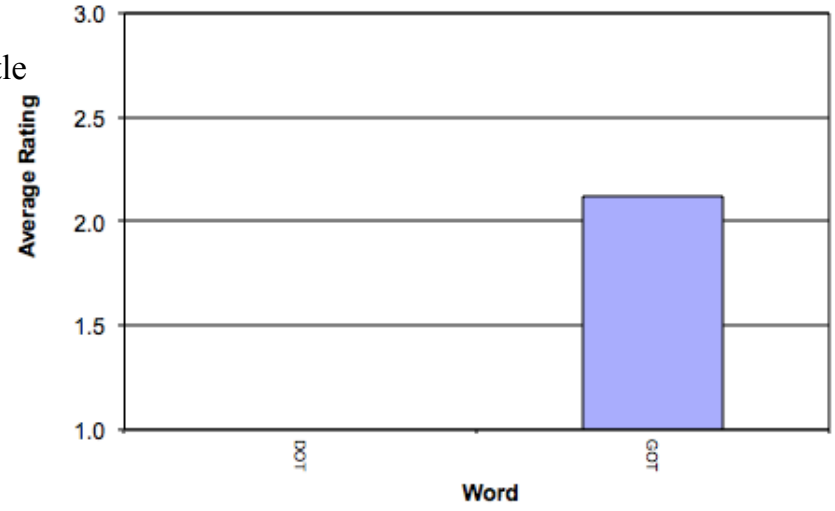

This fits well with Clark & Watson's list: *at, bit, got, let, lot, not, that, what*

- *get* (not asked); *but* = 2.9; *put* = 2.7; *it* = 2.6
- a few other words are also at 3: *sat, forgot, forget, about*

The results are clear in Newcastle, too…

*dot* vs *got*

- *not* and *knot* were used in priming sentences in Newcastle
  - (in Liverpool, *what* was used as an example sentence)

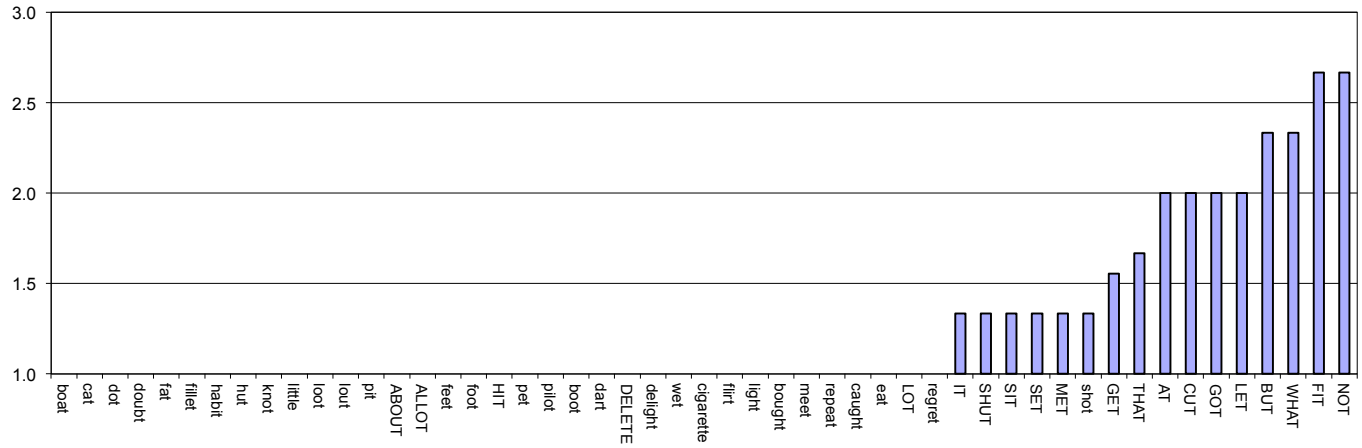



Newcastle informants were not so keen to use score 3

- accepted words typically scored an average of around 2

The full Newcastle results show that most words with final /t/ do not allow T-to-R

- this also fits well with Clark & Watson's list
  - T-to-R is also possible in Newcastle with *at*, *got*, *let*, *not*, *that*, *what*, *get*, *but*, *it*
  - and also with a few other words, but not *many* others




T-to-R is thus indeed clearly lexically conditioned

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *get it* | [gɛɹɪt] | *that is* | [ðaɹɪz] | *but it* | [bʊɹɪt] | *not a* | [nɒɹə] |
| *wet it* | *[wɛɹɪt] | *fat is* | *[faɹɪz] | *foot it* | *[fʊɹɪt] | *knot a* | *[nɒɹə] |

T-to-R is, however, unquestionably productive.

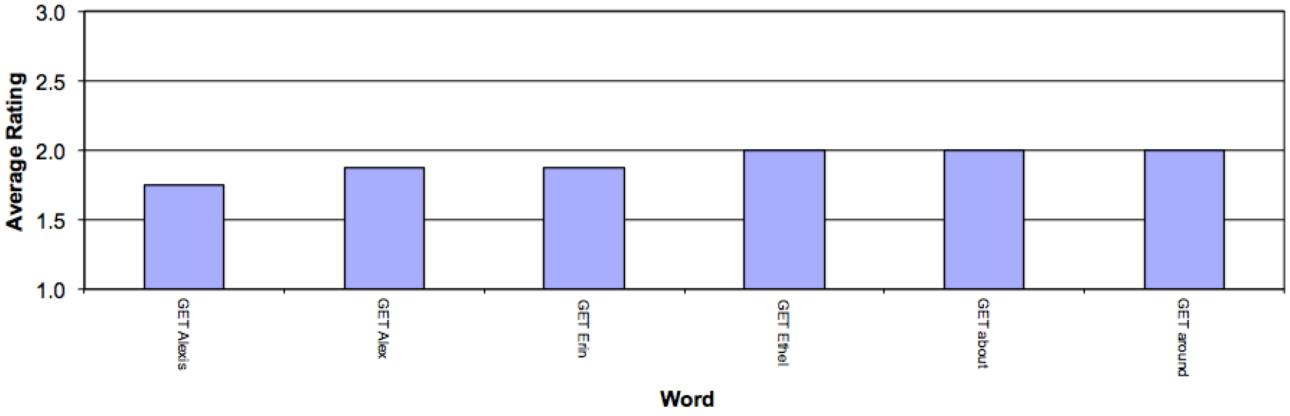
Could it be simply that there are lexically stored chunks which have an underived rhotic?

- *get it* [gɛɹɪt] ← /gɛrɪt/
- *not a* [nɒɹə] ← /nɒrə/

○ could all the forms where a rhotic occurs have become univerbated?

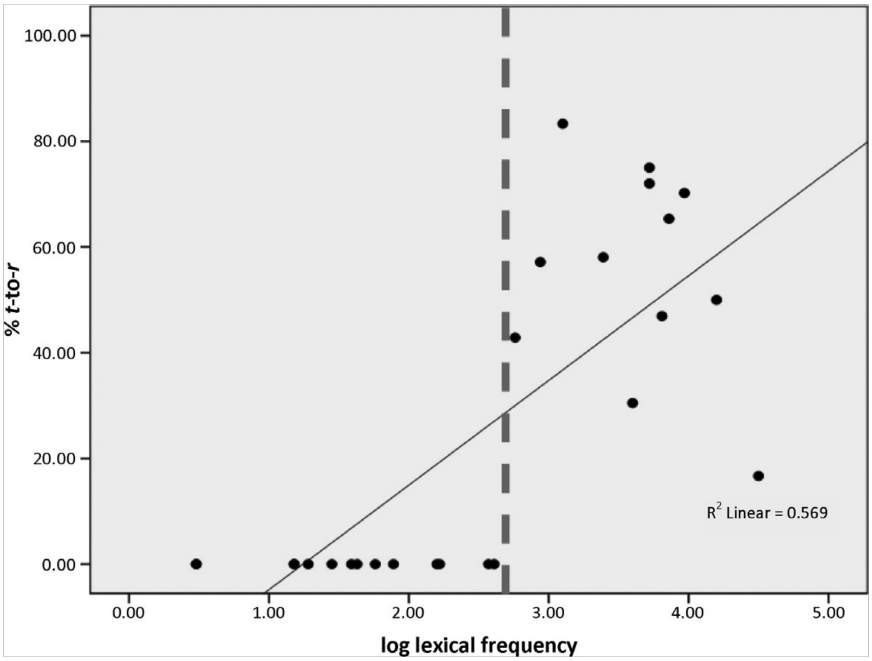
- no

We tested for this in Newcastle: *get Ethel* is just as possible as *get about*



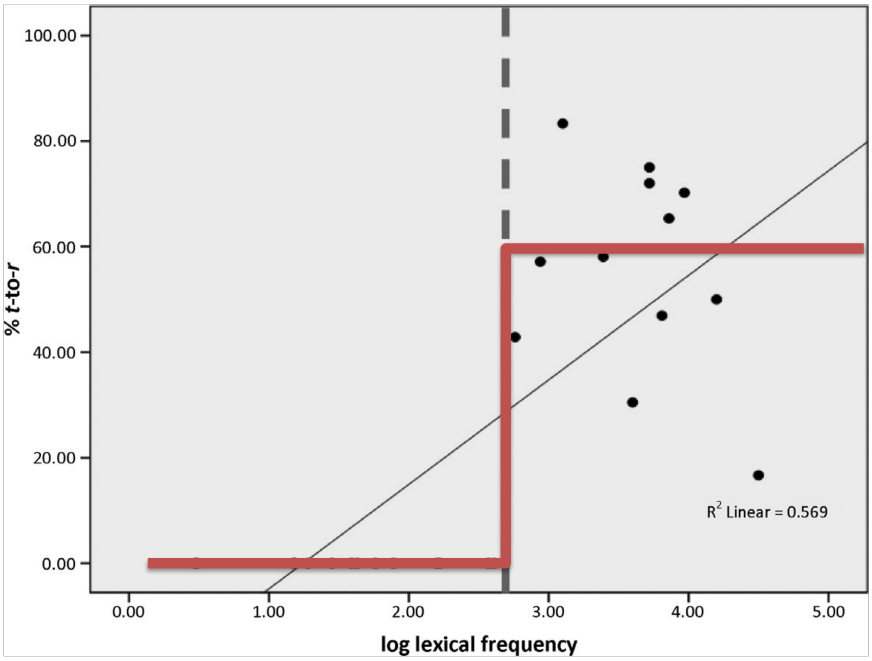

And... important for our purposes: T-to-R shows a frequency effect

Clark & Watson (2011) considered the interaction of T-to-R and word frequency




The x-axis shows the (log) frequency of all words in their corpus with: [...V̆t#V...]

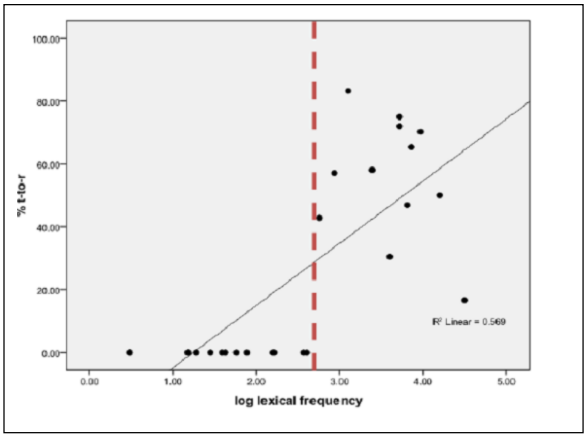
And this too is a *categorical* frequency effect




The x-axis shows the (log) frequency of all words in their corpus with: [...V̆t#V...]

There is a small group of words which can undergo T-to-R, and they all are indeed frequent

- however, no other words undergo T-to-R at all
  - word frequency increases gradiently
  - but T-to-R is categorical
  - a word either allows it or not
  - (and words allow it to different extents)
- these are not the predictions of a Usage-Based frequency-type approach




This pattern in T-to-R is a categorical frequency effect

- the words which can undergo T-to-R all are indeed frequent
  - but it is not clear that frequency is driving the realisation of the process
  - why doesn't the amount of T-to-R increase gradiently, with frequency?
  - specific words must be marked categorically as able to partake in the pattern, and these words are frequent
  - while other words must be marked categorically as not able to partake in the pattern, and these words are less frequent
  - how does a categorical frequency effect like that come about...? with all the other characteristics of T-to-R...?

## T-to-R: a diachronic explanation for the categorical frequency effect

T-to-R *does* make sense, but only if we consider its diachrony

- there is reliable data for earlier stages of relevant varieties which gives us the vital clue for the origins of T-to-R:
  - Ellis, A.J. (1889) *On early English pronunciation*. Part V: *The existing phonology of English dialects compared with that of West Saxon speech*. London: Trübner & Co.
  - Wright, J. (1892) *A grammar of the dialect of Windhill in the West Riding of Yorkshire.* London: English Dialect Society.
  - Broadbent (2008) considers this material and provides important leads

- this material describes of a phenomenon which provided the kick-off for T-to-R
  - the discussion focuses on West Yorkshire English (around Leeds-Bradford)
  - this account can be transferred to other varieties which also feature T-to-R

- T-to-R is due to a number of reanalyses of this phenomenon, involving IV stages

Stage I: speech current in the relevant area in (pre-)mid-19th century

- as described in Ellis (1889)

Ellis (1889) writes that, in West Yorkshire:

- "t, d preceding a vowel and after a short vowel becomes very vulgarly (r)"
  - what was (r)? there is good evidence, that at earlier stages in the history of English, including in Northern varieties around this time, the rhotic was typically [ɾ]
- Broadbent (2008) interprets this to be the description of a flapping process like this:
  - t,d → ɾ / V̆_V

Ellis (1889) mentions that this can occur in all sorts of words, such as these:

*at, sat, chat, what, spat, cat, hat* [past of hit]
*ad, glad, swaddy* [soldier], *bad, shadow*,
*dad, mad, et* [past of eat], *let, get, met, wet, set*
*wed, led, bled, sled* [slipper],
*it, hit, sit, flit, split, little, bit, hiddy* [to hide]
*smiddy* [smith], *did, bid,*
*ot* [hot], *got, cot, spot, od* [hold], *sod, not, modern*
*but, foot, shut, glutton, mutty* [calf], *mud* [might]
*good, stood, huddle, budding, sud* [should]

Stage I: this means that at this point, the following forms occurred:

| *not* | *knot* | *nod* | *rot* | |
|---|---|---|---|---|
| /nɒt/ | /nɒt/ | /nɒd/ | /rɒt/ | |
| [nɒt] | [nɒt] | [nɒd] | [ɾɒt] | |
| *not the...* | *knot the...* | *nod the...* | *rot the...* | |
| [nɒtðə] | [nɒtðə] | [nɒdðə] | [ɾɒtðə] | |
| *not a...* | *knot a...* | *nod a...* | *rot a...* | |
| [nɒɾə] | [nɒɾə] | [nɒɾə] | [ɾɒɾə] | t,d → ɾ / V̆_V |

Stage II: during the mid-late 19th century a major reanalysis occurred

- this reanalysis involved both Underlying Representations and the process
  - it produced a quite different system, but the surface forms did not change much

- the fundamental change was to reanalyse the words that underwent the flapping process as possessing two URs – eg, /nɒt$_{PW}$, nɒt/ 'not'
  - one UR was prosodically weak and therefore cliticised onto the following word
  - the other UR continued to project its own prosodic word
  - either could be used, giving variability

- the flapping rule was reanalysed so that it applies when the prosodically weak UR was used, requiring a clitic-boundary between the flanking vowels, within the Clitic Group
  - t,d → ɾ / V_ )$_{CL}$V  t,d → ɾ / {...V_ )$_{CL}$V...}$_{CG}$

English allows for lexical items with two URs:
*would* = /wʊd$_{PW}$, d/
*should* = /ʃʊd$_{PW}$, ʃd/
*them* = /ðɛm$_{PW}$, m/

Northern English allows right-leaning cliticisation:
Definite Article Reduction: *t'other, t'internet*

Stage II: this gives the following forms at this point

- the surface change is that the underlined forms become possible

| 'not' | 'knot' | 'nod' | 'rot' | |
|---|---|---|---|---|
| /nɒt$_{PW}$, nɒt/ | /nɒt$_{PW}$, nɒt/ | /nɒd$_{PW}$, nɒd/ | /rɒt$_{PW}$, rɒt/ | |
| *not the...* | *knot the...* | *nod the...* | *rot the...* | |
| [nɒtðə, nɒtðə] | [nɒtðə, nɒtðə] | [nɒdðə, nɒdðə] | [ɾɒtðə, ɾɒtðə] | |
| *not a...* | *knot a...* | *nod a...* | *rot a...* | |
| [nɒtə, nɒɾə] | [nɒtə, nɒɾə] | [nɒdə, nɒɾə] | [ɾɒtə, ɾɒɾə] | t,d → ɾ / {...V_ )$_{CL}$V...}$_{CG}$ |

Stage III: after this, there has been a gradual loss of 'exceptional' prosodically-weak URs

- the situation in stage II seems likely to be unstable
  - it is unusual/uneconomical for a lexical item to have two URs
  - such lexical entries are likely to simplify in acquisition over time
  - learners have rationalised the language and only acquired the 'normal' prosodically-strong URs for these lexical entries
- this is expected unless a lexical entry with two URs is reinforced in the data that the acquirer hears by exceptionally high frequency
  - words which originally took part in the flapping process but are now 'non-T-to-R' have lost the prosodically weak UR through diachronic change in acquisition
- NB: the rule was reinterpreted to only apply to /t/ – more on this below

Stage III: this gives the following forms at this point

| 'not' | 'knot' | 'nod' | 'rot' | |
|---|---|---|---|---|
| /nɒtPW, nɒt/ | /nɒtPW/ | /nɒdPW/ | /rɒtPW/ | |
| *not the...* | *knot the...* | *nod the...* | *rot the...* | |
| [nɒtðə, nɒtðə] | [nɒtðə] | [nɒdðə] | [ɾɒtðə] | |
| *not a...* | *knot a...* | *nod a...* | *rot a...* | |
| [nɒtə, nɒɾə] | [nɒtə] | [nɒdə] | [ɾɒtə] | $t \rightarrow ɾ \; / \; \{...V\_\_ )_{CL} V...\}_{CG}$ |

Why have all the words with /d/ stopped taking part?

- this is exactly what is predicted by this approach…
- words with final-*d* are less frequent than words with final-*t*

| Word | FrSp |
|---|---|
| it | 24508 |
| that | 14252 |
| what | 7313 |
| that | 7246 |
| but | 6366 |
| got | 5025 |
| not | 4693 |
| at | 4115 |
| get | 3464 |
| did | 3368 |
| would | 3278 |
| had | 2835 |
| said | 2685 |
| could | 1949 |
| put | 1640 |
| good | 1566 |
| should | 1173 |
| hundred | 1023 |
| God | 305 |
| bad | 296 |

BNC spoken word frequencies:

- the 10 most frequent words with – V̆t#
- and 10 most frequent words with – V̆d#

FrSp =
frequency (per million words) in speech

yellow = words ending in –*t*
red = words ending in –*d*

Stage IV: at some point there also was an approximantisation of /r/ in most varieties

- = ɾ > ɹ
  - this takes the output of flapping process with it
  - it is possible because the process was neutralising, and its output identical with the surface rhotic → identified with it
  - the precise and relative chronology of this is not important, and perhaps unknowable, but it does not interfere with the two previous developments

If we assume that it occurred last, it had the following effect

- forms which change are underlined

| ‘not’ | ‘knot’ | ‘nod’ | ‘rot’ |
|---|---|---|---|
| /nɒt$_{PW}$, nɒt/ | /nɒt$_{PW}$/ | /nɒd$_{PW}$/ | /rɒt$_{PW}$/ |
| *not the…* | *knot the…* | *nod the…* | *rot the…* |
| [nɒtðə, nɒtðə] | [nɒtðə] | [nɒdðə] | [ɹ̲ɒtðə] |
| *not a…* | *knot a…* | *nod a…* | *rot a…* |
| [nɒtə, nɒɹ̲ə] | [nɒtə] | [nɒdə] | [ɹ̲ɒtə] |

t → ɹ̲ / {…V_ )$_{CL}$V…}$_{CG}$

T-to-R words now have 2 URs in their lexical entry, one is marked as prosodically weak

- non-T-to-R words only have one 'normal' UR

*knot* = /nɒt$_{PW}$/ not a T-to-R word
*not* = /nɒt$_{PW}$, nɒt/ a T-to-R word – the weak form needs a host to the right

This accounts for the lexical-conditioning: only some words have 2 URs.

Either UR can be used by a speaker – this accounts for the variability

- different words have different (non-phonological) patterns of UR choice
  - when the prosodically weak UR is chosen, the T-to-R process applies

    t → ɹ / {...V__ )$_{CL}$V...}$_{CG}$

- the process always applies when a weak form is used – it is not variable
- the process is postlexical/'late phonology': T-to-R is not part of lexical phonology

When the prosodically-weak UR is used, structures like the following occur, involving the cliticisation to the following P-word, and fitting the structural description of the process:

*get Alex* {(gɛt)$_{CL}$alɛks}$_{CG}$ → {(gɛɹ)$_{CL}$alɛks}$_{CG}$ ≠ {gɛt}$_{PW}${alɛks}$_{PW}$
*not about* {(nɒt)$_{CL}$əbaʊt}$_{CG}$ → {(nɒɹ)$_{CL}$əbaʊt}$_{CG}$ ≠ {nɒt}$_{PW}${əbaʊt}$_{PW}$

T-to-R involves a well-behaved exceptionless process.

T-to-R confirms that there can be 'categorical frequency effects'

- they are categorical because they rely on two distinct phonological objects: URs
  - they have 'sharp edges', and thus fit well with a formal approach to phonology
  - no single word, other than those considered, shows even one occurrence of T-to-R
  - this doesn't fit well with the gradient expectations of Usage-Based Phonology
  - they are diachronic, low frequency effects, where frequency conserves: low frequency words lose their weak form as there is not enough evidence in the PLD to support the positing of two URs by children acquiring the dialects in question

## What's the final score?

Where are we left in terms of the predictions made about frequency effects?

1. Low frequency effects should exist
- UBP – yes ✓
- FP – yes ✓

2. High frequency effects should exist
- UBP – yes ✓(?)
- FP – no ✗ (?)

3. High frequency effects should *always* exist in 'natural' changes/rules
- UBP – yes ✗
- FP – no ✓

4. Frequency effects, like all phonological generalisations, should always be gradient
- UBP – yes ✗
- FP – no ✓

Formal, generative models come out better than UBP when confronted with the full range of predictions regarding frequency effects!

- FP 3, UBP 2
  - but what about 2? FP isn't completely off the hook...

Something must account for those high frequency effects that truly exist

- but is it phonology?
  - or an aspect of speech production?
  - FP needs and expects there to be a module that takes care of speech production
- is it that we get 'better' or quicker at saying ˈwɒʔtʰaɪmɪzɪʔˈ than ˈadəsprɪgəftʰaɪmˈ?
  - or rather: at articulating the gestural scores that our articulators use to implement these whole utterances in speech?

Recent, statistically sophisticated work (Maslowski 2015) has shown that:

- 'reduction' effects show up in nonce words after 3 pronunciations
- they are the same after 3 pronunciations as after 50 pronunciations
- no difference in reduction occurs in the speech of people who have heard a nonce word 3 times or 50 times – only if they articulate it
  - this is not a frequency effect, but an effect of getting used to pronouncing things
  - it is simply like getting used to doing the actions that we use to drive a car?

If this is right, it could lead to a situation where FP gets 4 out of 4...