Edinburgh Theoretical Historical Linguistics Summer School
July 2026

Frequency effects in phonological change favour formal phonology 2

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

Some issues to be touched on...

1. Are there diachronic frequency effects? (and if so: *of what kinds*?)
2. Are there non-diachronic frequency effects?
3. What does it mean if there *are* frequency effects?
   = what is the nature of phonology?

Frequency effects and phonology: what is predicted?

During the first session, we saw the following claims:

> «the frequency of use of lexical items determines the extent to which they are affected by change at the phonological level»
>
> - N-changes show high-frequency effects
> - A-changes show low-frequency effects
>
> «frequency effects are most compatible with usage-based phonology»

In this session, we will see that things are more complicated than that

- some of these claims are false, and others are too simplistic

We will consider four more precise (hence testable) predictions, and the extent to which UBP and FP are compatible with them:

1. Diachronic low-frequency effects should exist in A-changes
2. Diachronic high-frequency effects should *always*/*ever* exist in N-changes
3. Synchronic high-frequency effects should exist in speech

[4. Frequency effects (like all phonological generalisations) should always be gradient]

1. Do diachronic low-frequency effects exist in A-changes?

A-changes are not rare, and we have seen that there are claims that they are subject to low-frequency effects.

Which predictions do the two key models that we are considering make on this issue?

Diachronic low-frequency effects should exist in A-changes

- UBP – yes
- FP – ~~no~~ yes

Bybee (2006) argues that there is an explanatory link between low frequency effects, A-changes, and exemplar-based UBP:

- "frequency strengthens the memory representations of words or phrases making them easier to access whole and thus less likely to be subject to analogical reformation"
  - this is seen as a 'conserving effect of high-frequency' – bigger exemplar clouds are more robust ('entrenched'), and are thus resistant to change of this type
  - low-frequency words are more likely to undergo changes of this type on a UBP perspective as they are seen as lacking entrenched exemplar clouds

The kind of claim is often encountered as evidence that UBP is superior to FP

- however, the frequency patterning of A-changes also fits with the predictions of FP

How can low-frequency effects fit with FP?

The existence of any kind of frequency effect is sometimes painted as a problem for formal models of phonology

- formal models are, in fact, fully compatible with low-frequency effects, if the role of acquisition in change is recognised

English preterites

If learners are more likely to hear *ic drāf* 'I drove' than *ic bād* 'I bided', because *drāf* is more frequent than *bād*...

- then they are more likely to fix irregular *drāf/drove* in their grammar as the past tense of *drīfan/drive* than they are to fix irregular *bād/bode* in their grammar as the past tense of *bīdan/bide*
  - if speakers do not acquire an irregular past tense for *bīdan/bide* because they don't get enough input to do so (because the word is less frequent), then its past tense will be derived regularly by the rule for past tense formation (= 'add *-ed*')
  - acquiring an 'irregular' past tense is an extra learning task – FP would predict that considerable input (PLD) will be needed to succeed in it
- it is more likely that learners will fail in this extra learning task with low-frequency verbs than with high-frequency verbs = FP predicts low-frequency effects in A-changes

An FP account of verbs forming their preterite in an 'irregular' way:

		present	past
I drive	*I drove*	aɪ	oʊ
I write	*I wrote*	aɪ	oʊ
I shoot	*I shot*	uː	ɒ
I choose	*I chose*	uː	oʊ
I know	*I knew*	oʊ	ɪʊ
I grow	*I grew*	oʊ	ɪʊ

Such forms are not derived by rules, but have more than one UR for the morpheme ('suppletion')

'use letters to record language'
VERB
/raɪt/
/roʊt/past

Or /rait/ and a morpheme-specific Vocabulary Insertion rule

'Elsewhere' ordering and blocking accounts for all preterite formation:

	heaved	*heaped*	*heated*	*wrote*
	/hiːv+PAST/	/hiːp+PAST/	/hiːt+PAST/	/raɪt+PAST/
specific PAST	—	—	—	/roʊt/
regular PAST	/hiːv+d/	/hiːp+d/	/hiːt+d/	—
UR	/hiːv+d/	/hiːp+d/	/hiːt+d/	/roʊt/
Epenthesis	—	—	hiːtɪd	—
Assimilation	—	hiːpt	—	—
SR	[hiːvd]	[hiːpt]	[hiːtɪd]	[roʊt]

The change is a kind of simplification:

'bear, dwell' VERB /baɪd/ /boʊd/past	>	'bear, dwell' VERB /baɪd/

Simplification of structures is predicted in diachronic change – it represents the failure in the extra learning task that complexity requires

- FP is indeed compatible with the 'conserving effect of frequency' in A-changes

Diatonic Stress Shift fits with the predictions of an FP model, too

- DSS involves the creation of diatonic pairs from monotonic pairs, in words with lower token frequency before those with higher token frequency
  - this is expected on an FP model

convíct$_N$	~	*convíct*$_V$	>	*cónvict*$_N$	~	*convíct*$_V$
recórd$_N$	~	*recórd*$_V$	>	*récord*$_N$	~	*recórd*$_V$

The change involved in the creation of diatonic pairs through a change in the N

- $\sigma\acute{\sigma}_N > \acute{\sigma}\sigma_N$

Diatones exemplify some of the basic patterns of English stress:

- a final syllable is typically unstressed in nouns *hággis, áccent, júngle, númget*
  - = extrametricality, high-ranked NONFINALITY *etc*
  - non-final stress is assigned in noun by rule
  - final stress in nouns is exceptional = it needs to be stored in an UR (e.g., /akáʊnt/)
- a heavy final syllable is unproblematically stressed in verbs
  - final stress is assigned in verbs by rule *defý, obéy, chastíse, perváde*
  - it does not need to be stored in an UR

The pre-change situation was: (exemplified by *discount*)

V /dɪskaʊnt/ → σσ́ (a verb with heavy final syllable gets accent)

N /dɪskáʊnt/ → σσ́ (nouns normally do not have final stress, so it must be lexically stored)

The post-change situation was:

V /dɪskaʊnt/ → σσ́ (a verb with heavy final syllable gets accent)

N /dɪskaʊnt/ → σ(σ) → σ́(σ) (stress assigned to the only non-extrametrical syllable)
(noun final extrametricality)

Diatonic Stress Shift involves individual words submitting to the general pattern of English phonology, removing the exception of noun-forms allowing final stress

- this is, again, a simplification of the lexically stored form
  - learners need strong evidence for exception markers (lexical storage of stress), as they involve an extra learning task
  - low frequency words are less likely to provide that evidence, so are more likely to be (re)analysed in acquisition as non-exceptional
    – in which case the noun will receive non-initial stress

So:

- both UBP and FP predict low-frequency effects in A-changes

2. Do diachronic high-frequency effects *always* exist in N-changes?

The existence of diachronic high frequency effects in N-changes is more controversial than what we have seen for A-changes

- this links with the 'Neogrammarian' Exceptionlessness Hypothesis about N-change (the 'EH' is also call the 'regularity principle' in English)

We have seen that the UBP model is associated with the idea that N-changes should show high frequency effects – we will see *why* imminently

- this is a claim that the EH/RH is false – it is a claim that segments do not change in phonological environments all at the same time

○ the UBP assumption is really that *all* change is implemented by 'lexical diffusion'

We have also seen that the FP model is not compatible with this: if N-changes are changes in the 'rule component' ('phonological computation'), we expect exceptionlessness.

The two models thus make different predictions on this question:

Diachronic high-frequency effects should *always* exist in N-changes
• UBP – yes
• FP – no

Bybee (2001) argues for a predictive link between high frequency and high-frequency effects in UBP

> If sound changes are the result of phonetic processes that apply in real time as words are used, then those words that are used more often have more opportunity to be affected by phonetic processes. If representations are changed gradually, with each token of use having a potential effect on representation, then words of high frequency will change at a faster rate than will words of low frequency.[3] The streamlining of high-frequency words and phrases has the effect of automatizing production. Any motor activity that is repeated often becomes more efficient.

Bybee is defining 'sound changes' as "the result of phonetic pressures that apply in real time as words are used"

- this is controversial – in FP, 'rules' can result from other things, too, although the phonologisation of phonetic pressures is widely acknowledged

Relatedly, Pierrehumbert (2002) writes, about UBP and exemplars, that

> "Any systematic bias on the allophonic outcome would incrementally impact high frequency words. In short the model is applicable to any Neogrammarian sound change..."

Tamminga (2014) expands on how Pierrehumbert (2002) sets this out:

> Pierrehumbert explicitly extends the claim that frequent words lead sound change to any kind of gradient phonetic change, stating that "any systematic bias on the allophonic outcome would incrementally impact high frequency words at a greater rate than low frequency words" (2002:118). Just as frequent words that undergo reduction in speech should end up being more reduced in the phonetics inherent to their representation, frequent words that are undergoing non-reductive sound change (for example, the raising of /ey/ along the front diagonal in Philadelphia (Labov et al., 2013)) should accumulate advanced tokens more quickly than their less-frequent counterparts.

The origin of N-changes is often related to phonetic pressures, which can become phonologised into synchronic (and later diachronic) phonological patterns

- assuming all these connections, UBP predicts that high frequency effects should be omnipresent in N-changes
- there is no reason why any such change should *not* be affected by frequency

The existence of high frequency effects in N-changes has indeed been asserted in Bybee and Phillips' (and some others') work

- but do they *always* exist in N-changes?
  - [do they *ever*?]

Labov has long battled with the extent to which phonological change is exceptionless

- since 1981 he has argued that a distinct type of change (which I have been calling 'N-changes') are indeed exceptionless, patterning like synchronic rules
  - his final word on the subject is Labov (2020), where he sought to discover in full detail, with all statistical power, if a change can be shown to be exceptionless:

> How then would such an empirical study establish that any one sound change was progressing in a manner that affected each and every member of the word class according to its phonetic environment? First of all, we want to study a change in progress, rather than a stable distribution that we infer must be the result of change some time in the past (as in Bybee 2002). It should be a unidirectional change, in which the whole community participates.

Labov's (2020) conclusion is:

> We have examined the raising of /eyC/ as a prototypical sound change in a continuous phonetic space, below the level of conscious awareness, which has continued in the same direction for over a century. It was selected in order to test the Neogrammarian principle that change affects all words in which the target phoneme is found. The extended time period of 110 years, the volume of data, and the monotonic incrementation of /eyC/ have facilitated the comparison of the rates of change of individual words at different stages. A few candidates for lexical effects were detected in the course of this study, but by the end none were found to have escaped the tyranny of phonetic constraints.

So: No. N-changes definitely do not always show high-frequency effects.

Can we consider the question quantitatively?

- UBP predicts that changes can stop at any point in their spread through the lexicon, so we might expect 1% of changes to stop at 1% of the lexicon, 2% of changes to stop at 2% of the lexicon *etc*.

Ringe & Eska (2013) considered this point, and say:

> Regular sound change is the norm; in fact, the regularity of sound change is statistically overwhelming. The following crude experiment gives a good idea of the numbers involved. The first 200 words of the glossary in an Old English (OE) textbook, Moore and Knott 1955, which survive in Modern English (ModE) were compared to their contemporary reflexes.[3] The shapes of at least 88 percent of the modern words inspected can be derived from those of the OE words *entirely* by regular sound changes and known morphological changes; thus the incidence of apparent phonological irregularity is not more than 12 percent over the past thousand years as measured by listed lexemes.
>
> it is actually surprising that these numbers should be so low, since standard English is the result of prolonged dialect contact in London from the thirteenth to the seventeenth centuries, and borrowing between dialects is a principal cause of phonological irregularities.

There *are* cases of change which show signs of being N-changes, but which nonetheless also show signs of having truly lexical exceptions

- these changes are often claimed to be due to 'lexical diffusion'
  - it may be that we simply haven't found the phonological pattern yet, or that the evidence that could allow us to recognise the pattern has been lost
- or, on a UBP perspective, they could be evidence for true lexical diffusion – changes which have 'stopped' before they worked their way through the whole lexicon
  - if so, we can expect them to show a high-frequency pattern: the low-frequency forms should be those that keep the pre-change state

One change notorious for having 'exceptions' is English 'BATH-broadening'

- a > aː in Early/Late Modern English (there has been a later backing, too: aː > ɑː)
- before voiceless fricatives (*bath, grass* but *mass*) and nasals (*dance, plant* but *ant*)

Trapateau (2008) considers the evidence for this and concludes:

> A common pattern emerges from both graphs: high-frequency words do not necessarily favour lengthening; low-frequency words do not show any clear tendency towards lengthening. Word frequency per se does not seem to have a clear effect when combined with the two broad phonetic environments (pre-fricative and pre-nasal contexts).

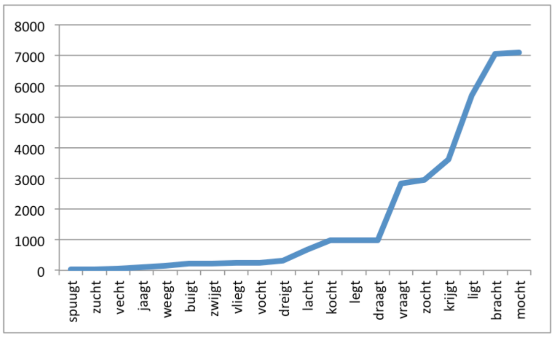
While the precise nature of 'lexical diffusion' is mysterious, the evidence for frequency effects in such changes does not look good.

What about Coronal Stop Deletion?

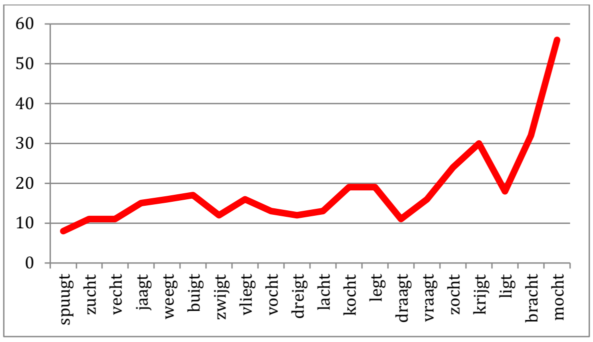
Let's return to the graphs summarising the Dutch case that Phillips (2006) sees as a change of the type t > Ø / x__#

- they seem to make a strong case for the existence of a high-frequency effect

Token frequency



Likelihood of absence of [t]



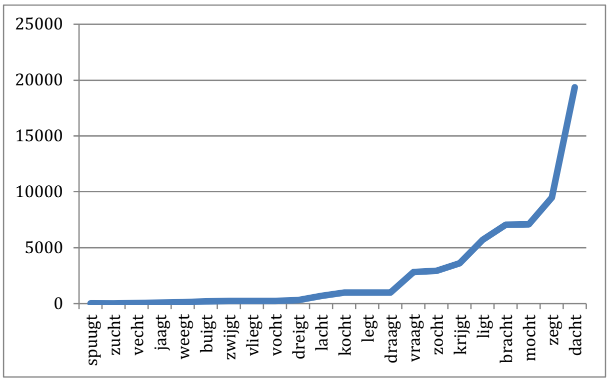
Token frequency and likelihood of deletion increase in these graphs very similarly

- given the interpretation of numbers in such situations, we cannot expect a perfect fit

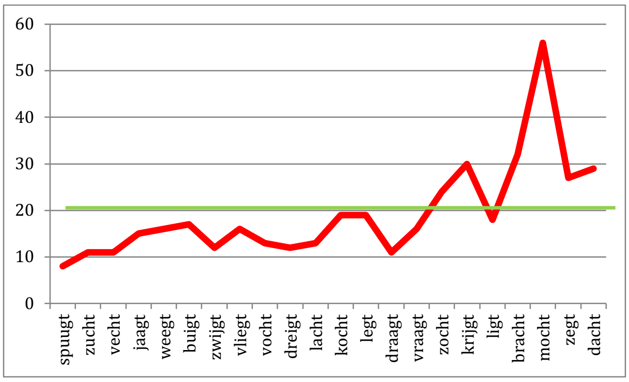
But, in fact, I overplayed the case in the graphs by including only the first 20 words

- if we add in all the words that Phillips gives numbers for, the situation is different

Token frequency



Likelihood of absence of [t]



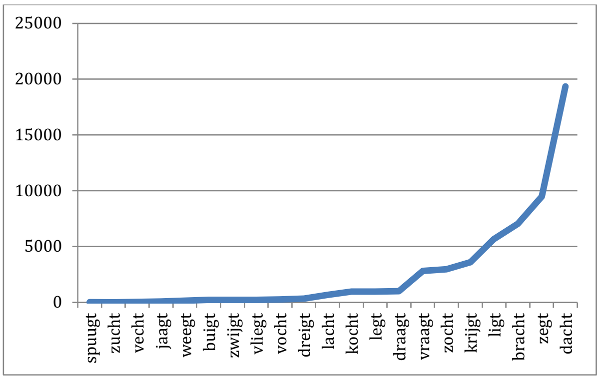
It's not completely sure what the 'likelihood of absence of [t]' graph is actually showing

- there is a peak for *mocht*, and lots of variation between 10% and 30% (= ——)
  - *mocht* is the only function word in the set – function words often have different prosody to lexical words, and this could explain its unusual behaviour
  - function words are typically unstressed – unlike lexical words – and unstressed forms are more likely to be subject to deletion
  - it is not clear that the variation is significantly related to frequency

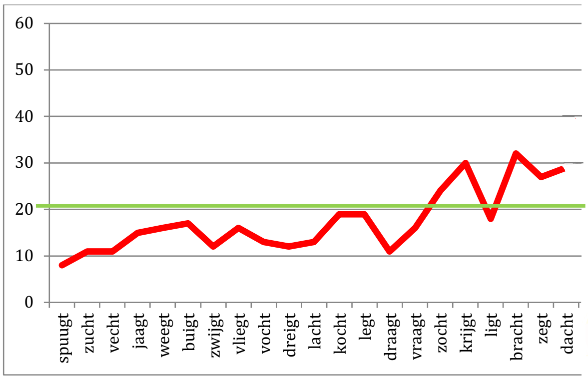
If we remove *mocht*, the evidence for a frequency effect is much less compelling

- the likelihood of coronal-stop deletion does not fit well with the gradient of the increase in frequency of use

Token frequency



Likelihood of absence of [t]



And also... CSD-type phenomena are generally regarded as cases of stable variation

- there may be no frequency effect here, but if there *is*, it is not a diachronic fact
  – the amount of 'deletion' is not described as increasing over time
  – so: is this a synchronic frequency effect?
- ○ *if* the frequency effect is a synchronic fact, it is a facet of variation
  – variation can be modelled in UBP and FP

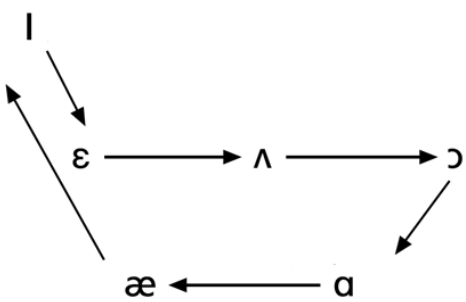
Dinkin (2008) investigated the existence of frequency effects in N-changes carefully

- Dinkin, Aaron. 2008. The real effect of word frequency on phonetic variation. *University of Pennsylvania Working Papers in Linguistics* 14, 97-106.

Dinkin (2008) conducted a detailed consideration of a change which is in progress in US English: the Northern Cities Vowel Shift

- this can affect /ɪ, ɛ, æ, ʌ/ (and other vowels at its most extreme)
- ○ Dinkin analysed measurements of a large number of tokens of these vowels

One representation of aspects of the NCVS is as follows:



A view of the Northern Cities Shift
(based on Labov 2010:15)

Assuming that this is indeed a case of change...

- UBP predicts frequency effects will be found in its patterning
- FP predicts the opposite

Dinkin finds that some small high frequency effect seems to exist for /ɪ/, as in the table below – and he find similar results for /ɛ/

- high-frequency words which contain these vowels move a little more in the direction of the shift than low-frequency words

Table 1 shows the results for /i/. The multiple regression found eleven phonetic variables plus the Top-5000 frequency variable as having statistically significant effects on backness of /i/: other things being equal, an /i/-word among the 5000 most frequent words of the Brown Corpus was on average about 60 Hz backer than a less frequent word. Since /i/ is being backed in the Northern Cities Shift, this is consistent with the Exemplar Theory claim that more frequent words will lead sound changes. Note, however, that word frequency has a smaller effect than any phonetic variable.

variable	coefficient	variable	coefficient
onset cluster	–489 Hz	labial onset	–119 Hz
liquid onset	–423 Hz	complex coda	–84 Hz
apical onset	–167 Hz	apical coda	–71 Hz
palatal onset	–151 Hz	/l/ coda	–69 Hz
nasal coda	+136 Hz	polysyllable	–66 Hz
labial coda	–122 Hz	**Top 5000**	**–57 Hz**

$p < .01\%$ $n = 2492$ constant = 2147 Hz $r^2 = 32\%$

Table 1: effects of frequency and phonetic variables on /i/ in the North.

However: Dinkin finds the opposite effect for /æ/, as in the table below, and for /ʌ/ (and also /ʊ/ which is not involved in the NCVS)

- high-frequency words which contain these vowels move a little less in the direction of the shift than low-frequency words

However, when we move on to /æ/, the Exemplar Theory prediction breaks down. On Table 3, we see that tokens of /æ/ in the top 5000 words are backer than less frequent words, which is contrary to the Northern Cities Shift.

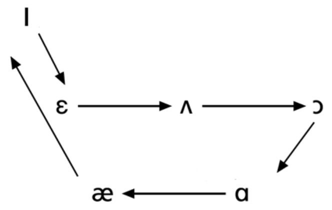
variable	coefficient	variable	coefficient
nasal coda	+275 Hz	stop coda	+94 Hz
velar coda	–207 Hz	labdent. coda	–79 Hz
apical coda	–152 Hz	voiced coda	+75 Hz
liquid onset	–134 Hz	apical onset	–63 Hz
onset cluser	–123 Hz	complex coda	+42 Hz
labial coda	–123 Hz	**Top 5000**	**–23 Hz**
polysyllable	–99 Hz		

$p \leq .01\%$ $n = 5091$ constant = 2058 Hz $r^2 = 30\%$

Table 3: effects of frequency and phonetic variables on /æ/ in the North.

This seems paradoxical, but Dinkin shows that it is understandable along the following lines:

- the NCVS changes that affect /ɪ, ɛ/ involve centralisation
- the NCVS changes that affect /æ, ʌ/ involve movements away from the centre



The frequency effects that Dinkin observes have nothing to do with the Northern Cities Vowel Shift changes

- the frequency effects are all due to a separate slight centralisation effect on all segments in frequent words
  - this is ‘overlaid’ on any effects of the NCVS
  - it makes sense as a synchronic ‘reduction’ – there does seem to be a synchronic frequency effects in this phenomenon (we return to this later...)
- NB: multiple factors can affect the pronunciation of forms coming out of our mouths
- the NCVS changes are not affected by token frequency, but synchronic ‘reduction’ is

So: **Are there diachronic phonological high frequency effects?**

- it looks like: no? but... there *are* synchronic high frequency effects?

3. Do synchronic high-frequency effects exist?

We have seen that:

- conceptually, both high-frequency effects and low-frequency effects are possible
  - high- *or* low-frequency words might engage most in a phonological phenomenon

Also, we have seen that there is a potential difference between:

- whole word effects
- segmental-category type effects – we have seen evidence against the existence of frequency effects of this kind: their existence is not proven

In connection with this,

- Kiparsky (2016) distinguishes between “an imperceptible phonetic effect of a few milliseconds, or neutralization to a categorically distinct pronunciation”
  - what *are* the former kind of phenomenon? – we shall consider them now

Which predictions do the two models make on this question?

Synchronic high-frequency effects should exist in speech

- UBP – yes
- FP – ~~no~~ yes

As we saw in session 1, Lavoie (2002) has shown that:

- the duration of *for* and *four* can be the same
  - but *for* – a more frequent word – often has a shorter realisation
  - *four* – a less frequent word – never has the shortest realisations
- the differences in duration here are measured in milliseconds
  - this is synchronic variation in pronunciation

We have just seen Dinkin (2008) show that vowels are centralised synchronically more in frequent words than in infrequent words

- the differences in here are measured in 10s of Hz
  - this is synchronic variation in pronunciation

There has been other serious work focusing on imperceptible phonetic effects of milliseconds' difference in the duration of words.

Gahl (2008) was one of the central papers in establishing this line of work

- this focuses on the pronunciation of whole words

The measurements involved consider the duration of chunks of speech

- such durations are massively variable (as we saw before)

This strand of work argues that more frequent words are reduced more than less frequent words.

As Gahl (2008) points out, this should mean that words which are typically transcribed as 'the same' will be pronounced differently if they differ in token frequency

time	[tʰaɪm]	– high frequency = more likely to reduce	1958.6*pmill*
thyme	[tʰaɪm]	– low frequency = less likely to reduce	0.20*pmill*
laps	[laps]	– high frequency = more likely to reduce	3.27*pmill*
lapse	[laps]	– low frequency = less likely to reduce	1.37*pmill*
steak	[steɪk]	– equally likely to reduce	16.24*pmill*
stake	[steɪk]	– equally likely to reduce	16.61*pmill*

This is reminiscent of Lavoie's (2002) work that we saw in session 1:

for	[fɔː]	– high frequency = more likely to reduce	6895.10*pmill* (log 3.92)
four	[fɔː]	– low frequency = less likely to reduce	255.78*pmill* (log 3.69)

But – with words like *time* and *thyme* – Gahl compared lexical words, avoiding the possibility that function-word prosody (e.g., stresslessness) drives the results.

Gahl (2008) controls for a range of factors in a corpus-based study and argues that frequency is, indeed, a key factor in word duration

- a summary of Gahl's (2008) results (cleaned up by Lohmann 2018) is given here

FIXED-EFFECT PREDICTORS	COEFF	*SE*	*t*
(intercept)	−0.5247	0.104	−5.07
Speaking rate	−0.0492	0.020	−2.42
Duration of low-frequency homophone	0.2141	0.040	5.42
Bigram probability	−0.0171	0.005	−3.21
M-score (grapheme-phoneme probability)	−0.2213	0.073	−3.02
Noun proportion	0.1034	0.024	4.29
Pause ratio	0.2813	0.137	2.06
Logged token frequency of word	−0.0297	0.001	−4.43

TABLE 1. Summary of the regression model fitted to log-transformed average durations of the high-frequency homophones in the data set (see Gahl 2008:486).

Gahl (2008) concludes:

> Crucially for the current study, the log frequency of a word was a significant predictor of word duration when all other factors were controlled for: as frequency increases, word duration decreases, when other factors are held constant.

Gahl's (2008) work uses statistical analysis

- some work on frequency effects (such as that by Bybee, for example) has been criticised for not considering the statistical analysis of its numerical results
- ○ however, Lohmann (2018) shows that the precise statistical methods that Gahl uses are faulty, although he does still find a duration difference between low-and-high-frequency pairs like *time-thyme*.

The details of this kind of study are complex, with tiny differences, needing complex and sometimes controversial statistical interpretation

- the effect is small, but it looks like there can be a correlation between the frequency of a word and its phonetic duration

One key issue with this kind of result, however is summarised by Lohmann (2017)

- studies of this kind of thing do not always find any effect

> In previous corpus-based research, effects of lemma frequency on duration were tested on content word homophones, e.g. *thyme* versus *time* (Gahl 2008, 2009), and function word homophones (Jurafsky et al. 2002, Jurafsky 2003). While in Gahl's studies a positive finding for lemma frequency is reported, Jurafsky et al. fail to find empirical support for such an effect. There is one corpus-based study testing frequency effects on N/V homophones in child-directed speech (Conwell 2016), which reports a marginally significant effect of lemma frequency on word duration.

So: Are there synchronic high-frequency effects?

- it looks like: yes? sometimes?
  - in whole word effects = 'tiny-word-based effects'
- the qualifications of the 'yes' here are all relevant to understanding what is going on

Why is it relevant that the effects are 'whole-word' effects?

- are words phonological objects?
  - FP-type phonological generalisations typically involve features or segments or syllabic constituents or feet *etc*
  - the frequency effects considered here are not effects in processes affecting such units
  - if frequency effects in processes affecting that kind of unit are *not* robustly in existence, that might give us pause: are they truly *phonological* effects?

Why is it relevant that the effects are 'tiny' (c. 20 ms)?

- UBP is perhaps more obviously compatible with this kind of effect than FP, but we should, again, pause: what does UBP really predict in this connection?
  - all phonological objects in UBP are gradient, so these frequency effects should be gradient, too

time	[tʰaɪm]	– high frequency = more likely to reduce	1958.6*pmill*
thyme	[tʰaɪm]	– low frequency = less likely to reduce	0.20*pmill*

On the basis of this magnitude of frequency difference, *time* should be many times shorter in duration than *thyme*! (1000s of times shorter? three-quarters as short?)

- this magnitude of effect is not reported in the literature
  - why not? this is what UBP seems to actually predict
  - the effects in this type of 'reduction-like' synchronic high-frequency effects are indeed quite tiny (10s of milliseconds or Hertz)

Can this kind of synchronic high-frequency effect be accommodated in a model where phonology works as imagined in FP?

- multiple factors can affect the pronunciation of forms coming out of our mouths

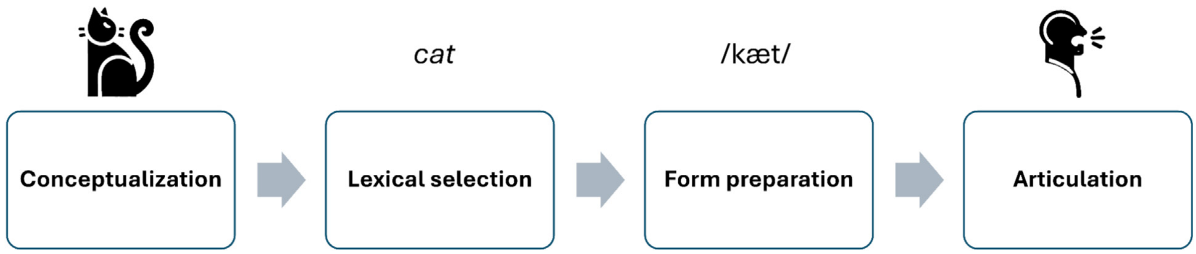
Something must account for synchronic high-frequency effects

- but is it phonology?

FP needs and expects there to be a module (or modules) that take care of speech planning and speech production

- these do something completely different to phonology, so can be expected to be of a different nature
  - phonology accounts for phonological processes, neutralisations, inventories, opacity, phonotactics, stress assignment, tone patterning, *etc*
  - speech planning and production prepare the body for articulation and execute that articulation

Zeng et al. (2025) set out "four identifiable stages" of the "speech production process":



Goldrick & Blumstein (2006) talk of: *'phonological' planning* *articulatory implementation*

Goldrick & Blumstein (2006) say, further:

> Most theories of single-word production assume three general stages of post-semantic processing. We refer to the first stage as *lexical selection*; it involves the selection of a word to express a nonverbal concept (e.g., selecting CALF to express "young bovine"). This lexical representation serves as input to *phonological planning* processes which specify the appropriate sound sequence for this word (e.g., monosyllable, onset: /k/, nucleus: /æ/, coda: /f/). Finally, *articulatory implementation* processes execute the specified sound sequence (e.g., for onset /k/, elevating the tongue body to create a closure on the soft palate while abducting the vocal folds).

'Phonological planning' is not part of phonology, but is the part of speech planning which takes a phonological representation and prepares for speech.

What kind of factors affect speech planning and production?

- many factors affect this process
  - these depend on context, and have nothing to do with phonology

Conscious knowledge of the import of utterances can have an impact

- Gahl (2008) discusses material which demonstrates this:

> Guion (1995) found that pairs of homophones differed in duration when the words were embedded in constructed sentence pairs (such as *We'll* ***need*** *the watch for a few hours*, *We'll* ***knead*** *the dough for five minutes*). When the same words were read in generic carrier phrases (*Say . . . to me again*), however, there was no significant difference in duration. Another study that did not report any significant durational differences between homophone pairs is Cohn et al. 2005a,b, which tested words in lists, as well as in constructed sentences, some of which were the same as in Guion's study.

It seems that what is 'reduced' (spoken more quickly) is a whole utterance

- why? because the utterance *we'll need...* is more predictable and accessible
  - the utterance *we'll knead...* is more surprising in most contexts and thus needs more time (for the listener?) to process (so speakers give them that time)
- a prediction on this basis is: *we'll* will be shortened more in *we'll need the watch for a few hours* than *we'll knead the dough for five minutes*, too

Gahl (2008) continues:

> In this connection, it is interesting to note that Guion (1995) found durational differences when homophones were embedded in full sentences, but not in generic carrier phrases. This is consistent with a pattern in which experimental stimuli in meaningful contexts display variation resembling that found in the corpus, whereas stimuli in generic carrier phrases pattern more like word lists, which do not reliably give rise to differences in homophone durations.

A prediction on this basis is: bakers in a bakery will say *we'll knead the dough for five minutes* faster (so words will have shorter durations) than non-bakers on the street

- the context of an utterance is crucial to how fast we can say it (and real-world context is surely not a phonological factor)
  - this feeds into the planning of the articulation of an utterance
- words said outside of a context (e.g., in a generic carrier phrase) will not have this effect of understandability-in-context overlaid on them during speech planning
  - there will be more influence from phonological structures in that case
  - context-type-factors can 'hide' the effect of phonology

These kinds of factors are very 'unreliable' – we might expect them to sometimes have an effect and sometimes not

- we find this in synchronic high-frequency effects: they are only present sometimes

Another possible cause of effects in speech planning and production is that people get 'better' at doing things the more they do them

- speakers of English will have said †wɒʔt^{h}aɪmɪzɪʔ† many times, and will have said †adəsprɪgəft^{h}aɪm† rarely

○ speakers will be more fluent in articulating the gestural scores that our articulators use to implement †wɒʔt^{h}aɪmɪzɪʔ† in speech – more fluent = quicker

○ that includes the sequence †t^{h}aɪm† – a linguist can segment that from utterances, and, if taken from †wɒʔt^{h}aɪmɪzɪʔ†, it will be quicker than †t^{h}aɪm† from †adəsprɪgəft^{h}aɪm†

Statistically sophisticated work (e.g., 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

○ this is simply like getting used to doing the actions that we use to drive a car

We get better at the 'speech production' of some gestural scores – better = quicker.

If this approach is right...

- pairs of words like *time* and *thyme* are homophones at the end of phonology: [t^{h}aɪm]
- however, sometimes they will not be pronounced with the same duration, due to factors related to speech planning and production

When UBP analysts acknowledge that we still need to be able to account for all the things that phonology has typically accounted for, it is quite common to find claims like this:

- from Pierrehumbert (2006)

> the traditional arguments for a phonological level of representation are valid. We need phonology to explain basic findings in the structure of the lexicon, psycholinguistics, and historical change. The first and simplest exemplar models provided a valuable challenge to the field by revealing the weak points of neo-generative approaches to sociophonetic variation. However, the future lies with hybrid models, which have multiple levels of representation (like neo-generative models) while also having explicit mechanisms for statistical learning and situational indexing (like exemplar models).

Do we need 'hybrid' models of phonology combining FP-like analysis with exemplar-type UBP analysis?

- or do we just need phonology + speech-planning-and-production?

○ we need phonology + speech-planning-and-production anyway to account for patterns in data – why add in exemplars?

What about the claims that...

There is no easy way to express the fact that these phenomena seem to affect some words more than others in a formal/generative grammar

- Pierrehumbert (2001) claims that this could have substantial implications:

> These results challenge standard models of phonology and phonetics at two levels. First, in all standard models, the lexicon is distinguished from the phonological grammar. The exact phonetic details of a word's pronunciation arise because the word is retrieved from the lexicon, and processed by the rules or constraints of the grammar whose result (the surface phonological form of the word) is fed to a phonetic implementation component. The phonetic implementation component computes the articulatory and/or acoustic goals which actualize the word as speech. The phonetic implementation component applies in exactly the same way to all surface phonological representations, and the outcome depends solely on the categories and prosodic structures displayed in those representations. As a result, there is no way in which the phonetic implementation can apply differently to some words than to others. If a phonetic implementation rule is variable and gradient, then the same probability distribution of outcomes would arise for all words which meet the structural description of the rule.

This is a strawman

- (informed) FP does not claim that only phonology determines what comes out of speakers' mouths
  - speech-planning-and-production is complex, taking phonological representations as a starting point, but also being influenced by context in multiple ways

Du & Durvasula (2024) are clear about this:

> We suspect that the *common strawman view of discrete representations* arose from a misunderstanding of the classic claim that phonology is a *feedforward system*, which means that there is no feedback from performance systems back to phonology. However, this term has been misunderstood to mean that *only* the output of phonology affects phonetics, that is, performance is sensitive only to the output of phonology. In this latter view, generative phonology is viewed as a sort of production system, which is contrary to the original claims of generative phonologists.

What should we conclude?

In terms of diachrony, frequency effects are a red-herring

- they either don't exist or are expected by both approaches
  - there is not firm evidence for segmental-type high-frequency effects

There *are* diachronic low-frequency effects

- but this doesn't prove anything in terms of the nature of phonology

The evidence for diachronic high-frequency effects is *not compelling*, and there and there is evidence against them

- so this doesn't prove anything in terms of the nature of phonology

In terms of synchrony, frequency effects are complicated

- synchronic frequency effects seem robustly to exist in speech, but are not supportive of UBP, and are a facet of speech, not phonology

What should we conclude?

The two basic approaches to phonology make different predictions in terms of what should occur in the relationship between frequency and phonology

1. Diachronic low-frequency effects should exist in A-changes
   - UBP – yes ✓
   - FP – yes ✓
2. Diachronic high-frequency effects should *always*/*ever* exist in N-changes
   - UBP – yes ✗
   - FP – no ✓
3. Synchronic high-frequency effects should exist in speech
   - UBP – yes ✓ (?)
   - FP – yes ✓

= Frequency effects in phonological change favour formal phonology

Number of valid predictions:

UBP: 2 (1½?)
FP: 3

Number of falsified predictions:

UBP: 1
FP: 0