Variability in Intervocalic Stops: Production and Processing

Natasha Warner

and Benjamin V. Tucker

Department of Linguistics,

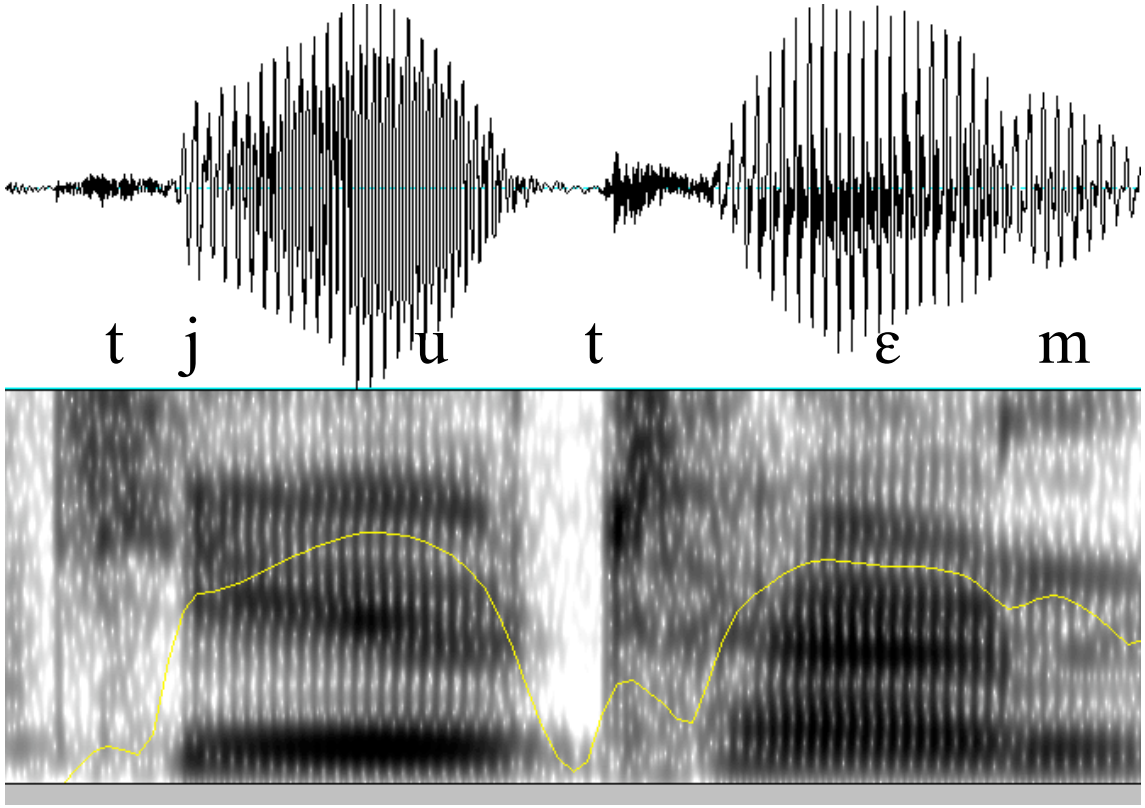
University of Arizona

Speech Variability

- Speech is rampantly variable: segments, syllables, entire words get reduced or deleted (but not always) (cf. many papers by Ernestus et al., Pluymaekers et al. 2005, Johnson 2004, Greenberg 1997)
- Stops can become approximants (vowel-like), vowels can become devoiced (fricative-like)
- Despite all this, we usually understand it all fine!
- How much variability comes from phonology, from systematic phonetic sources, from random variation?

Examples

- What does this say?
- "Do you have time to talk to me for a little while?"

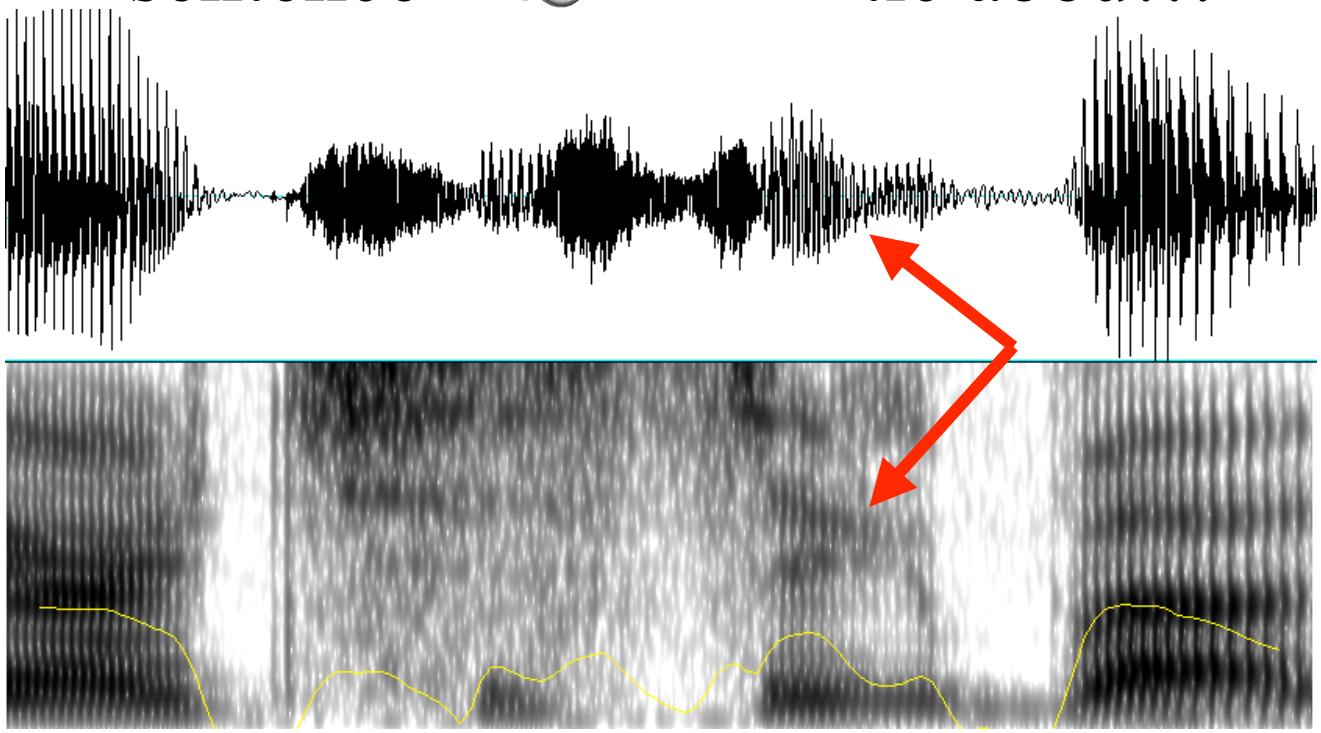


Do you have time…

Complete word "have" deleted

A voiceless stop doesn't have to be voiceless

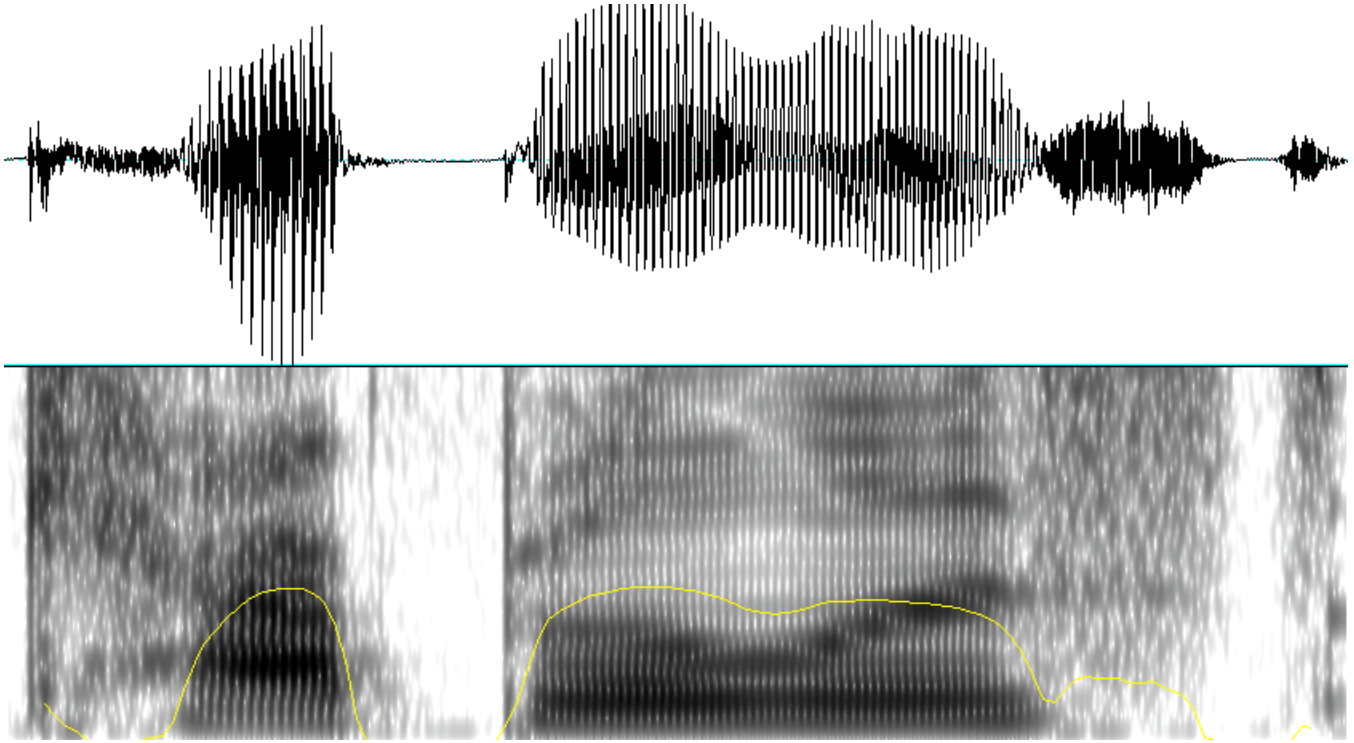
- "She's very artisti**c** about things" (list reading)
- sentence "tic abou…"



a ɹ t^h ɨ s (t) ɨ g ɨ b ʌ

A flap example

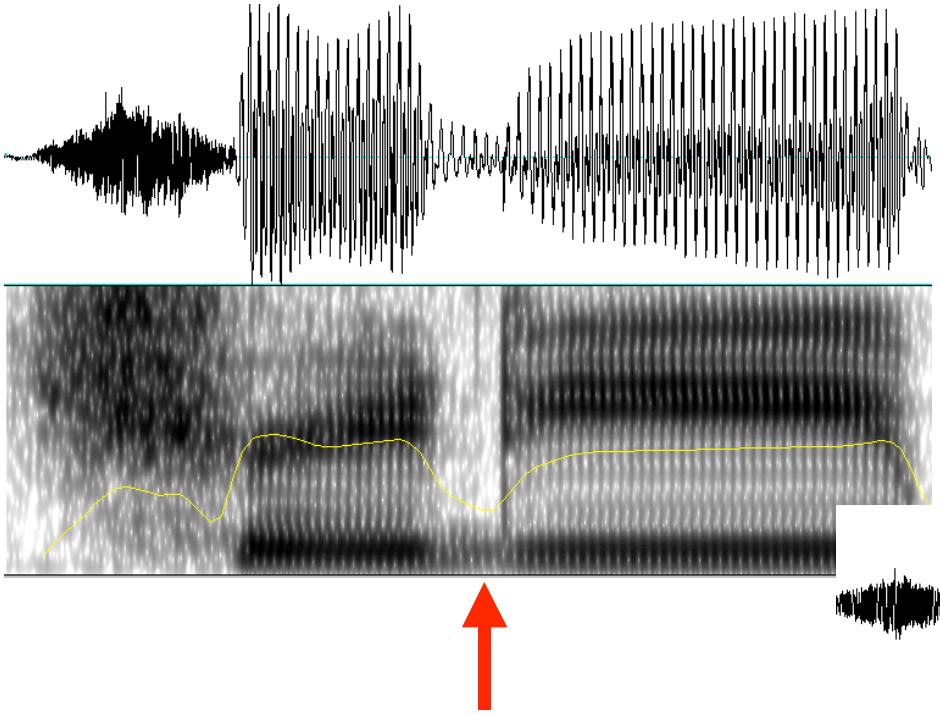
- Lest you think reduction only happens in casual, connected speech: "capi**t̲**alist"



kʰ æ p ɨ l ɪ s t

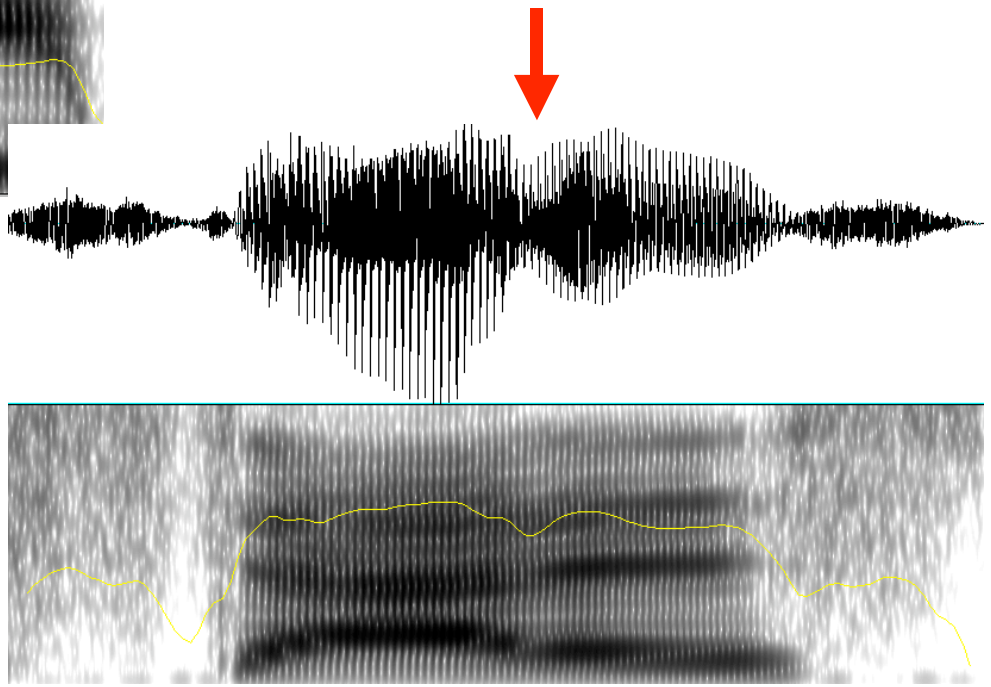
Isolated word list reading, in sound booth.

Main interest: flaps in comparison to other intervocalic stops



A clear flap:

"treaty"



A reduced one:

"status"

Flapping in Amer. English

- /t, d/ are traditionally said to become [ɾ] if intervocalic before unstressed syllables: butter, bottle, treaty, ladder, capitalist, …
- Even across word boundaries: but I, bad as it is…
- This seems to be pretty categorical, although not 100% (Patterson & Connine 2001)
- But there are claims that flapping is not a categorical phonological rule, but phonetic, gradient variability (Fukaya & Byrd 2005)

Phonetics and phonology in flapping

- "The underlying motivation for the phenomenon is a prosodic one that does not pick out a single place of articulation for a symbolic alternation" (Fukaya & Byrd 2005)
- They argue that general prosodic patterns lead to short articulations, which are perceived as a categorically different sound.

Our questions

- Does a categorical phonological rule apply to /t/ and /d/ (and not to /p, k, b, g/)?
- Is some phonetic variability systematic, and conditioned by word frequency, stress and segmental environment, speech style, etc.?
- How common is reduction?
- How do listeners understand reduced forms?
- Do listeners adjust their expectations for sounds based on speech style of the context?

Some things we're **not** asking

- Most past literature on flaps (Kahn 1976, Patterson & Connine 2005) focuses on **whether** /t, d/ flap in some environment. We're looking only at flapping environments, to see what happens **among** flaps.
- Past literature also compares /t, d/ to look for (in)complete neutralization. We compare /t, d/, but not with the purpose of finding differences that tiny.

Methods

- Intervocalic, pre-unstressed /p, t, k, b, d, g/
- 6 segmental environments and 2 stress environments:

Sample stimulus words by stop and stress

	Post-stress	Inter-unstress.		Post-stress	Inter-unstress.
/p/	appetite	precipice	/b/	inhibit	halibut
/t/	status	limited	/d/	credit	prejudice
/k/	recognize	applicable	/g/	magazine	esophagus

Sample stimulus words by segmental environment

Before schwa	status
Before syllabic /l/	cattle
Before /ɚ/	butter
Before full vowel /i/	pretty
After /r/	forty
Phrasal (Across word boundary)	write a letter

Materials 2

- 10 items in each of the 6 segmental environments x 6 phonemes x 2 stress environments, where possible within the lexicon
- Several combinations of factors don't (or rarely) occur in the inter-unstressed environment:

quadrupedal	[kwaˈdɹupəɾl̩]
synodal	[ˈsɪnəɾl̩]

And our students won't know these words anyway!

Subjects & Procedure

22 speakers recorded (7 analyzed so far)

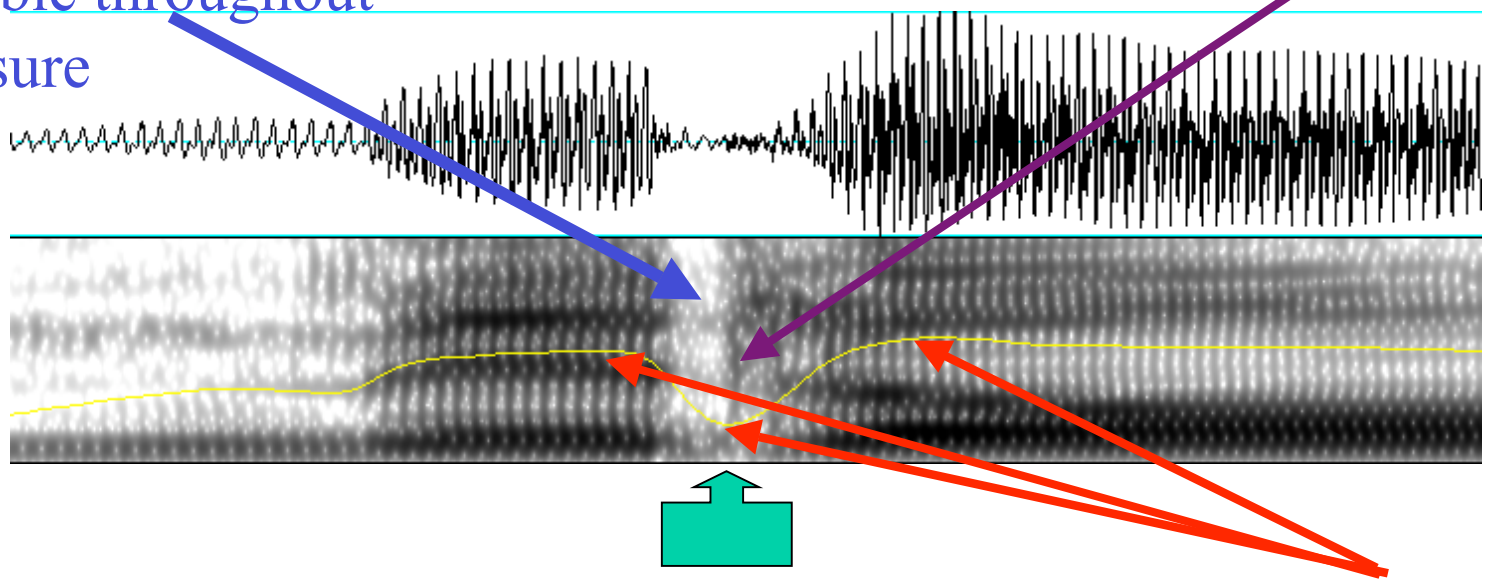
3 speech styles recorded

- open conversation, with friend or family, by phone (in sound booth)
- story reading (targets embedded in stories)
- isolated word list reading

Measurements

whether F2, F3 are visible throughout closure

whether a burst is present

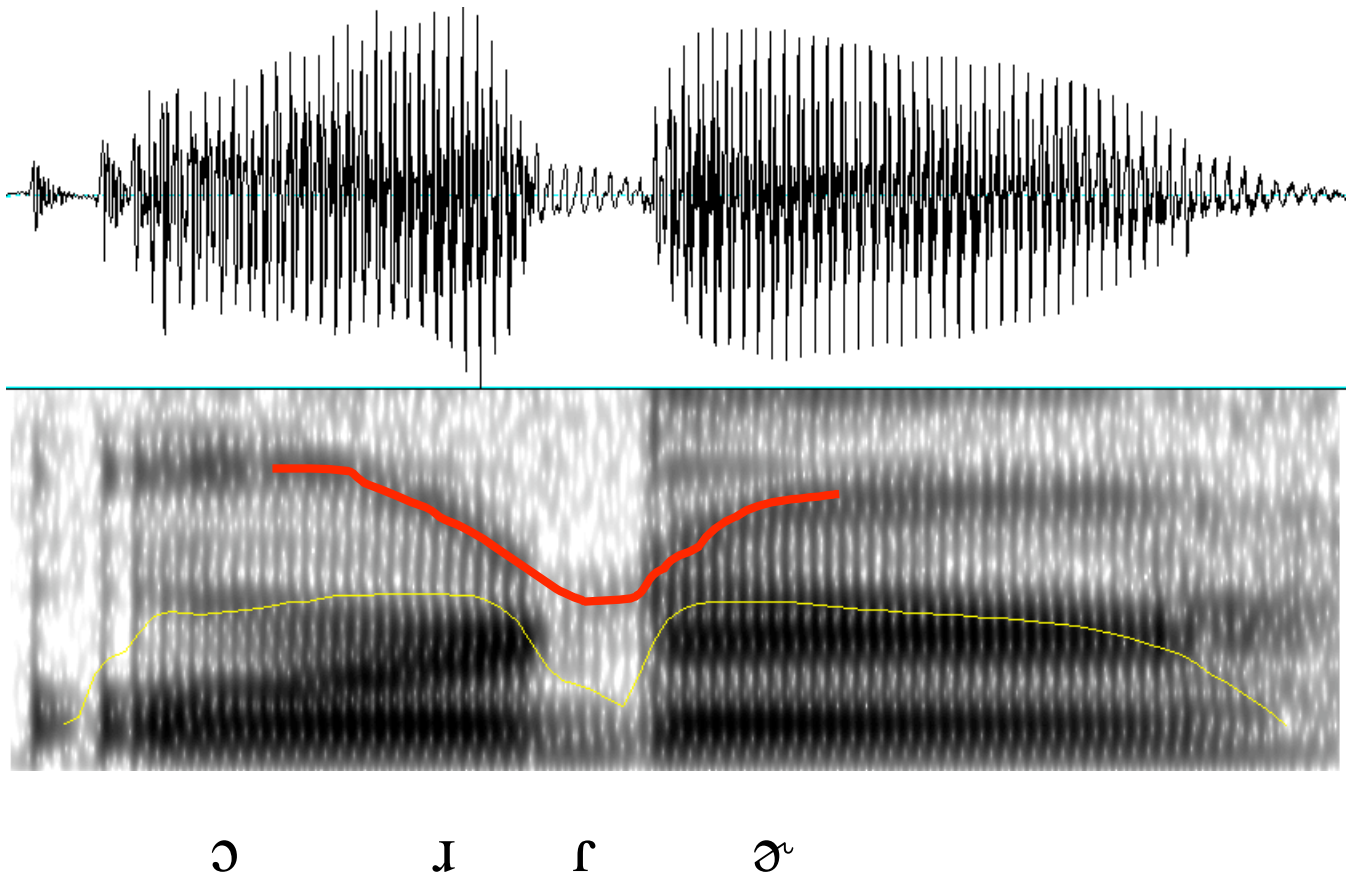


additional durations not reported here

- cons. duration
- cessation of voicing?

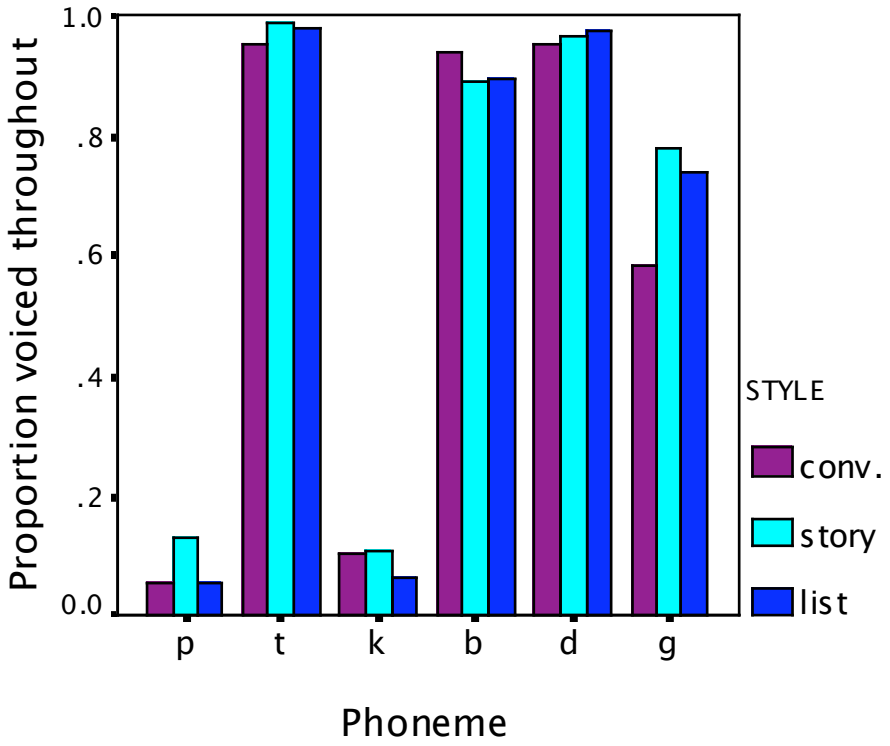
ratio of minimum intensity to average peak intensity of surrounding vowels

A surprising acoustic cue: F4

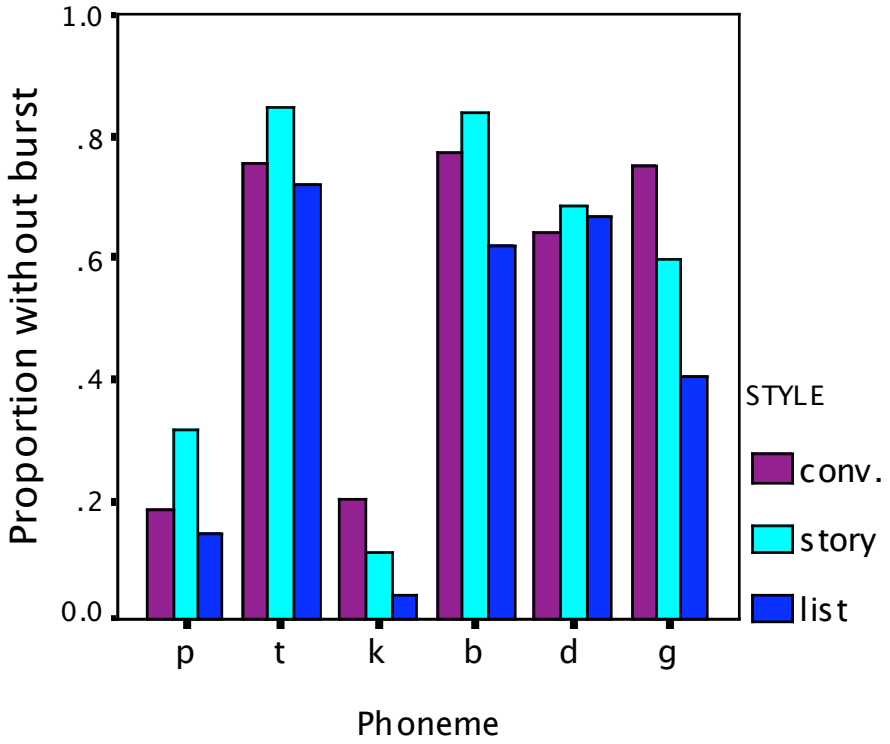


- Primarily around /r/'s
- F4 is hardly used for anything, except retro-flexes
- But this is timed to the flap, and occurs even for highly reduced tokens

Results: frequency of reduction

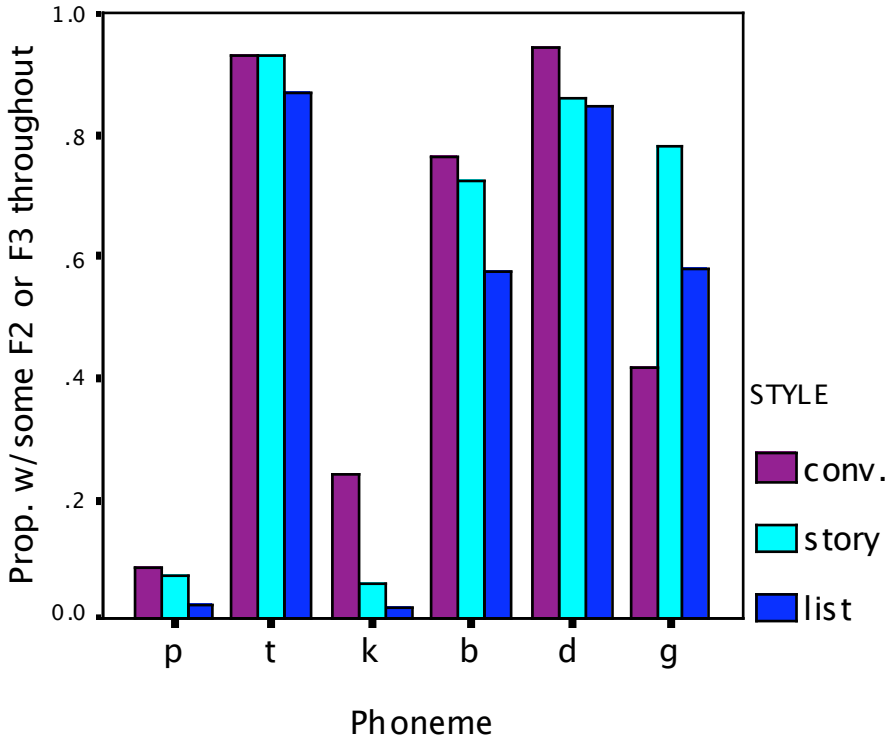


Clearly articulated stops would have bursts, and /p, k/ would be voiceless.

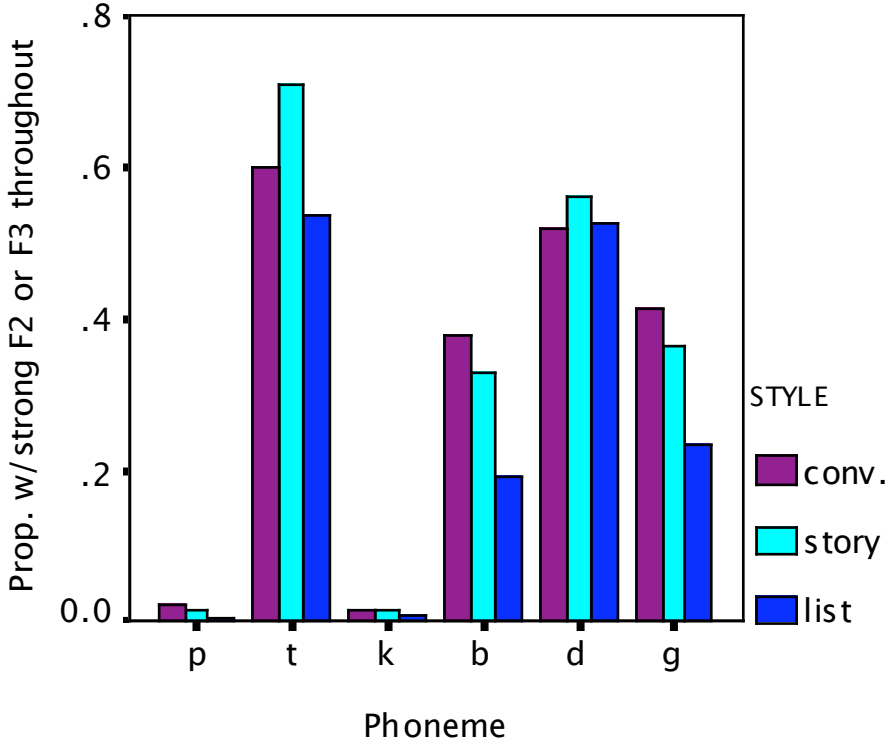


(For all measures except cons. dur., up is more approximant-like, down more stop-like.)

Frequency of reduction: formants

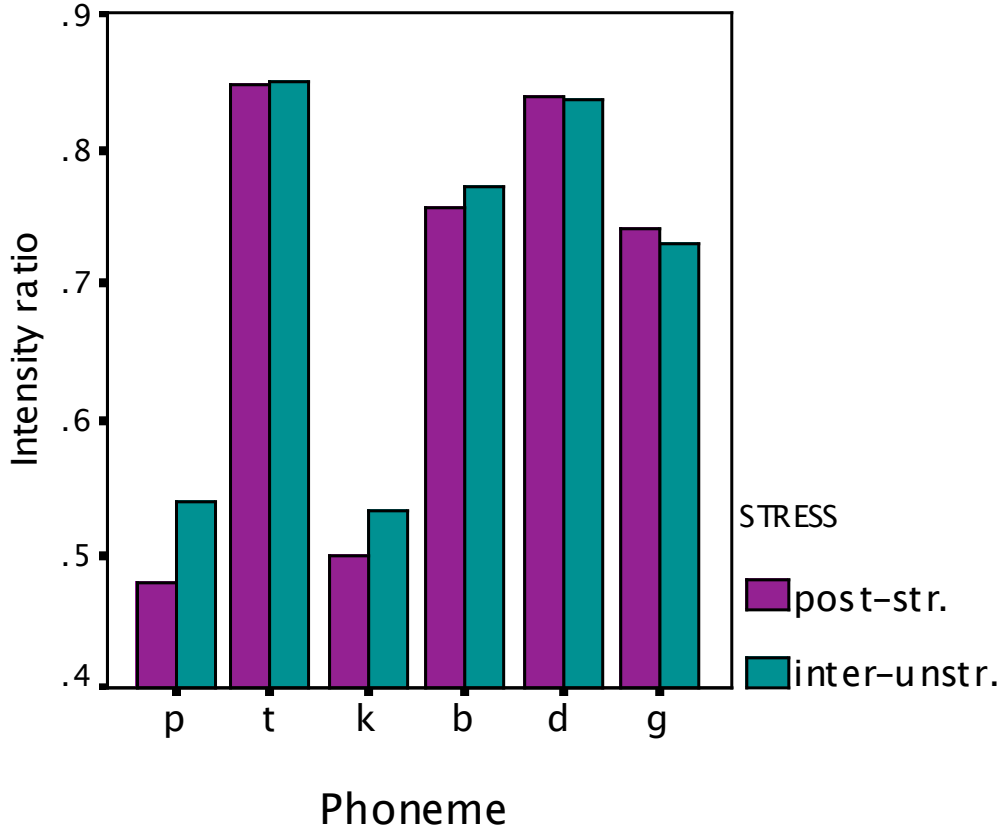


Clearly articulated stops wouldn't have formants.



•Conclusion: There is a lot of reduction in the data, in all speech styles.

Effects of stress environment



- All items are before unstressed syllables, but they can be either post-stress (e.g. 'city') or between unstressed (e.g. 'humanity')

•Result: inter-unstressed environment may be more reduced, but not significantly or consistently

Effects of speech style: deletions

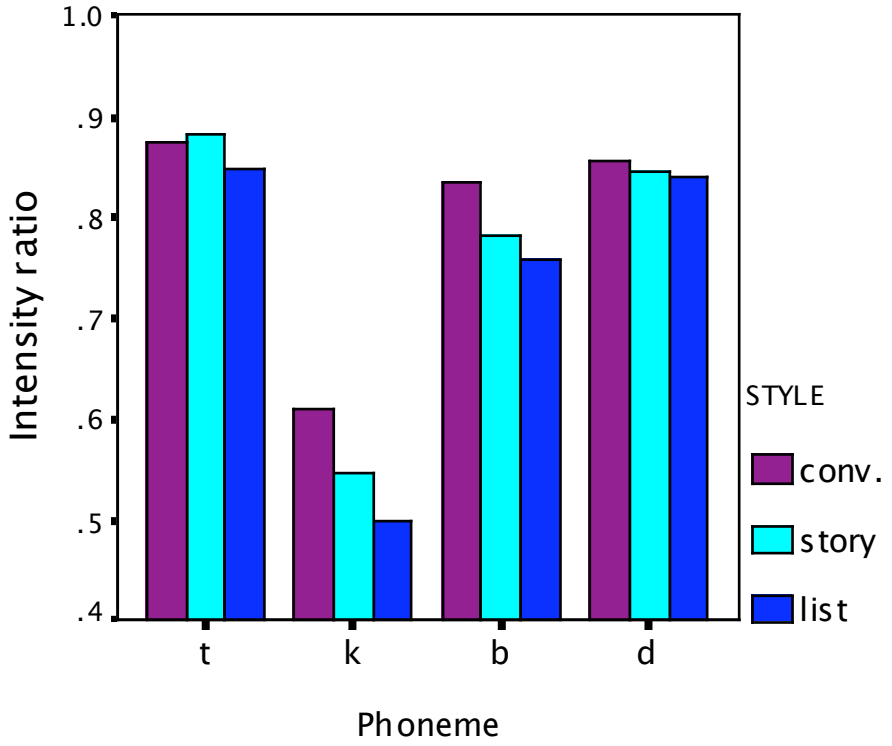
- In 86 out of 4726 stop tokens, the stop is so deleted we can't find any trace of it to measure.

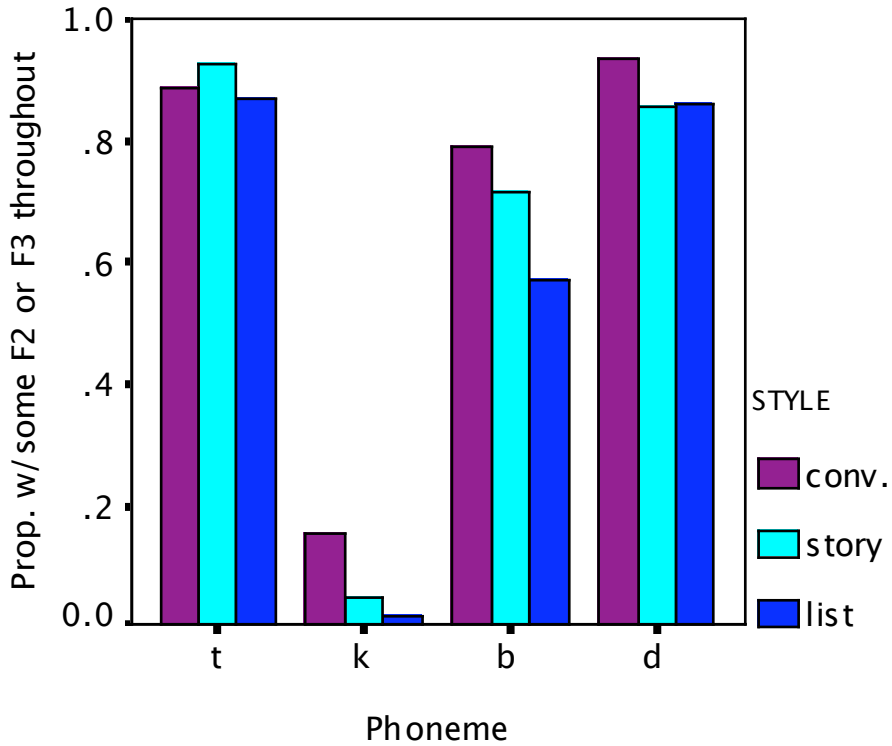
Number of tokens	Conver -sation	Story reading	List reading
deleted	48	25	13
not deleted	508	833	3299

•Complete deletions are rare (because we can label even highly reduced flaps), but significantly more likely in more casual speech.

Effects of speech style: reduction

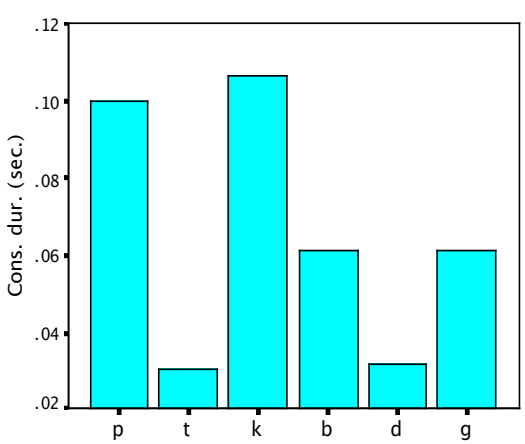
- More casual speech is significantly more reduced than careful speech on 3 measures.
- For some measures, there is less style effect for /t, d/, because of ceiling effects.

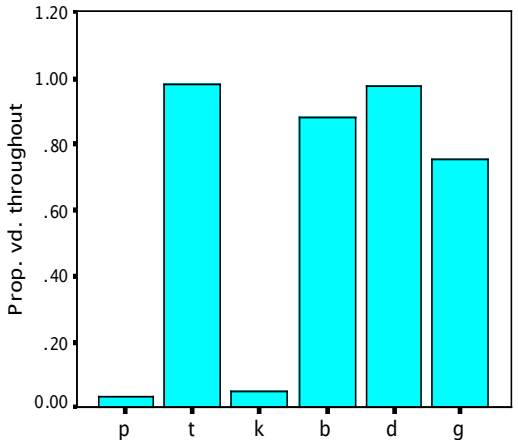


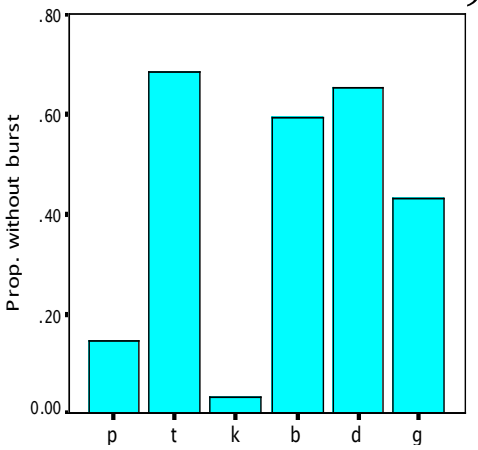


Effects of phoneme

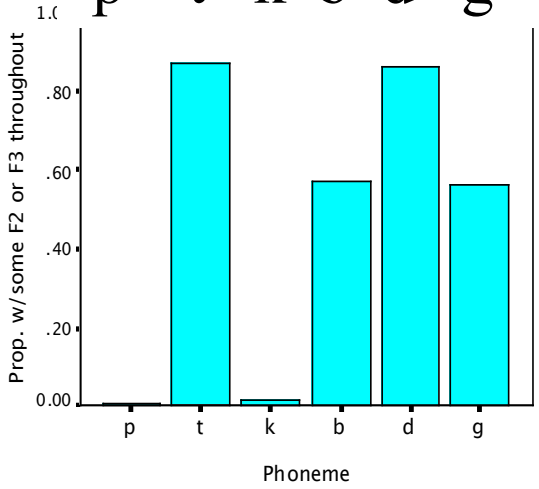
- /t/ behaves like a voiced stop (similar to /d/)
- /t, d/ are more approximant-like than /b, g/



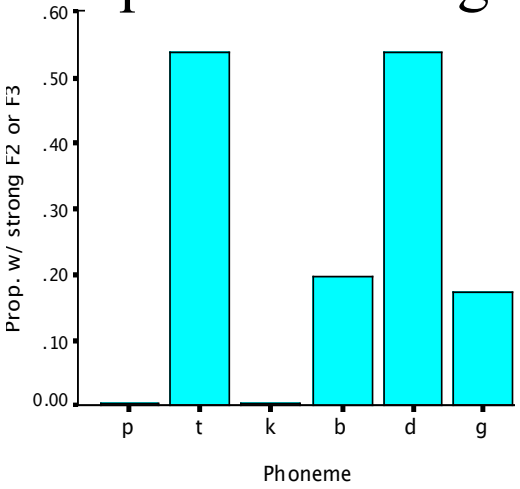




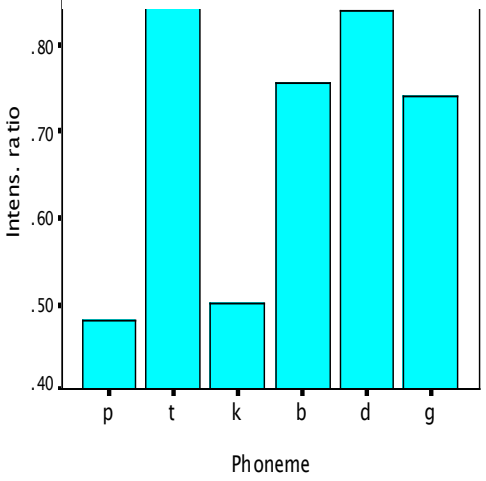
p t k b d g



p t k b d g



p t k b d g

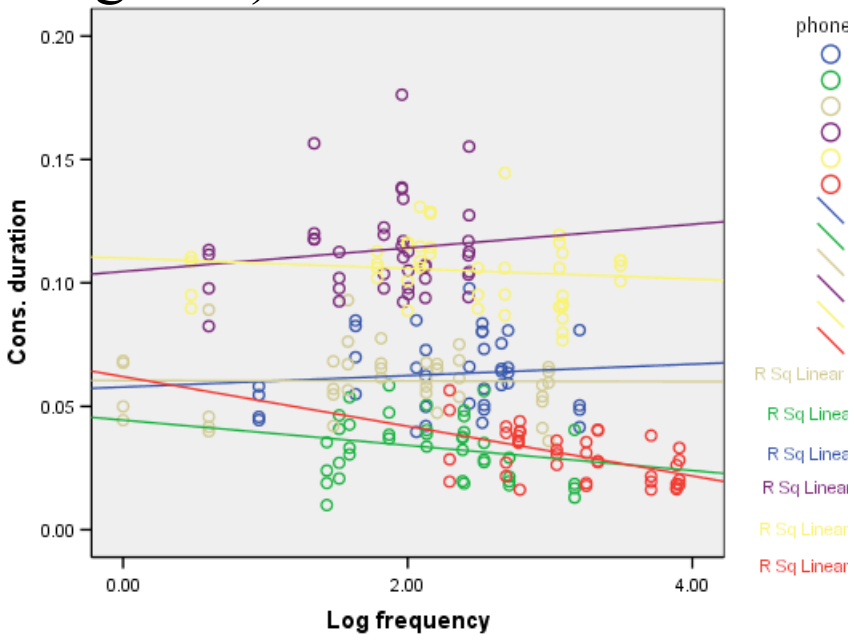


•/t, d/ are similar to each other

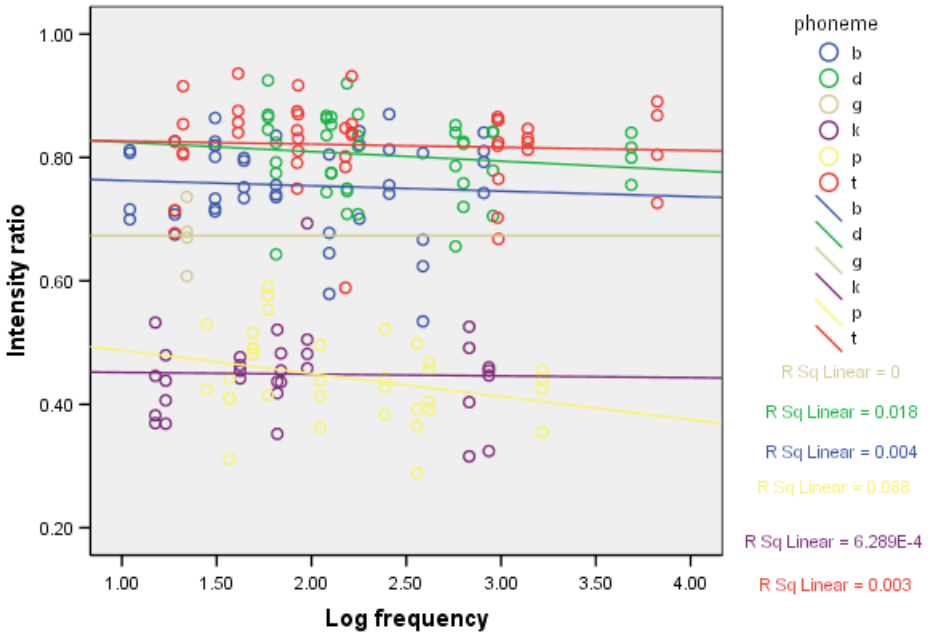
Results: Word frequency (4 S's)

- Frequencies from Celex and British Nat'l Corpus
- High frequency words not more reduced
- Patterson & Connine (2005): freq. effect on **whether** /t/ flaps

(should be negative)



(should be positive)



Intermed. summary: production Is there phonology?

- Since /t/ behaves like a voiced stop, there must at least be a phonological process applying to /t/ (cf. Zue & Laferriere 1979).
- Patterson & Connine (2005) show it affects /t/ in almost all cases: close to categorical.
- Our results show phonology puts /t/ in a different range from /p, k/: also categorical.
- Effects of phoneme are far larger than any other systematic effect in the experiment: categorical, phonological effects may be larger than gradient phonetic ones.

Does phonology affect /d/ too?

- Results show /d/ does not differ from /t/: they are similarly approximant-like on a wide range of measures. /d/ and /t/ both differ from /b/ and /g/.

➢ Therefore, the same (or a similar) phonological process must apply to /d/, too.

➢ It does not apply to any of /p, k, b, g/.

➢ (We didn't measure prec. vowel duration. We show that a phonological process affects /t/ and /d/, not that the result is identical.)

Is this articulatorily based?

- It could just be that the tongue tip can move faster than other articulators, leading to faster gestures and/or gestural overlap, and this is a purely phonetic effect.
- But other languages, and even British English, don't have flapping!
- The phonological aspect could certainly be derived from the articulatory facts, but has to be phonologized: an abstract process.

So is phonology everything?

- No! There is considerable gradient phonetic variability as well.
- Systematic variability: more reduction in casual speech, possibly more reduction for inter-unstressed stops.
- Substantial random variability as well.

Processing/perception

- How does all of this affect the listener?
- Auditory lexical decision and cross-modal identity priming using words produced with reduced vs. unreduced flap and /g/
- Do listeners adjust their expectations of what a stop or flap should sound like based on how casual/reduced the preceding context is?

Processing experiment 1: Auditory lexical decision

- Auditory lexical decision

  Subjects decide if an auditory stimulus is a word or not. RT's and error rates measured.

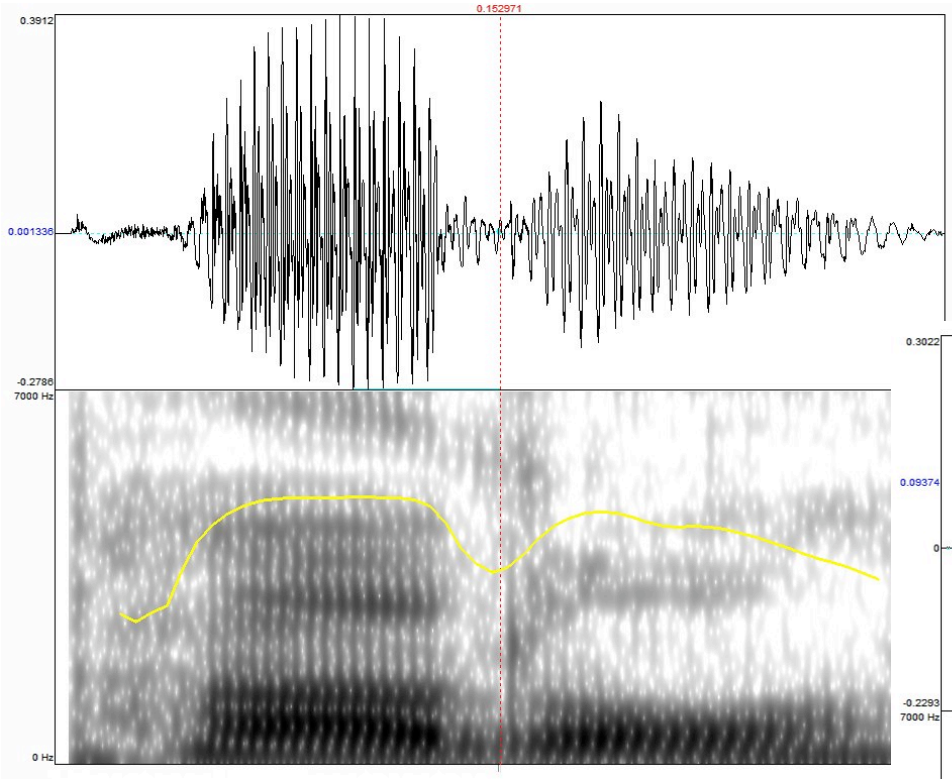
- Targets contain a flapped /d/ or a /g/ in similar environment:

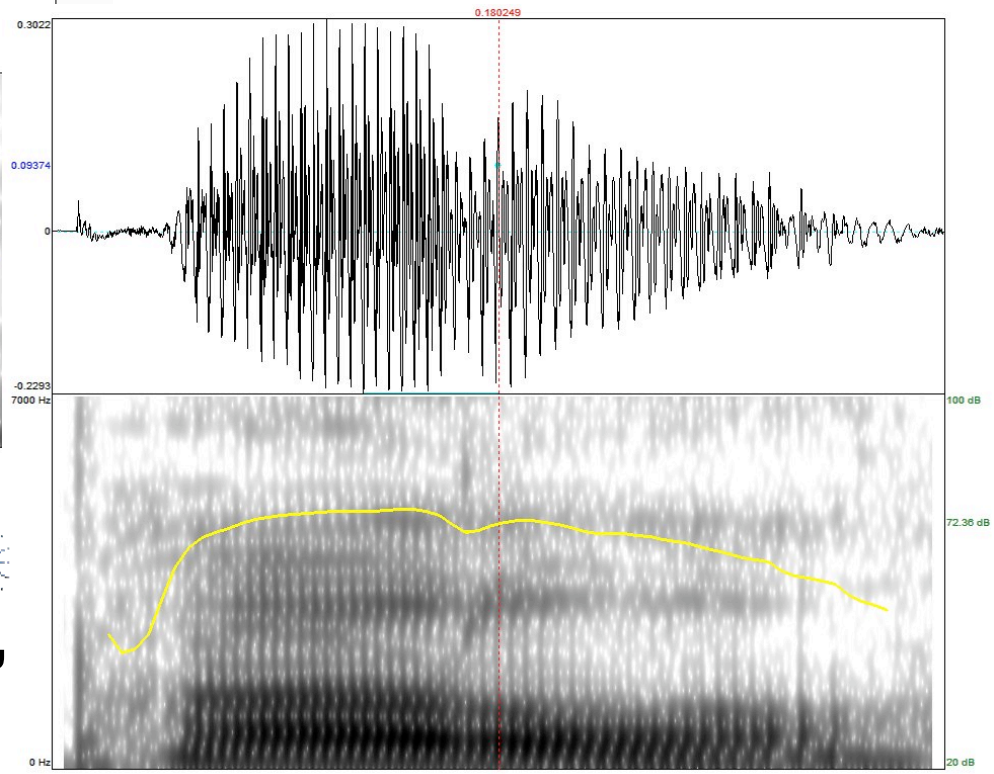
  "puddle" "baggy"

- Items are produced either with a reduced, approximant-like flap/stop or a clear flap/stop
- All target /d/'s are flapped: none are unnaturally clear

Sample items for /d/

Unreduced ‘puddle’

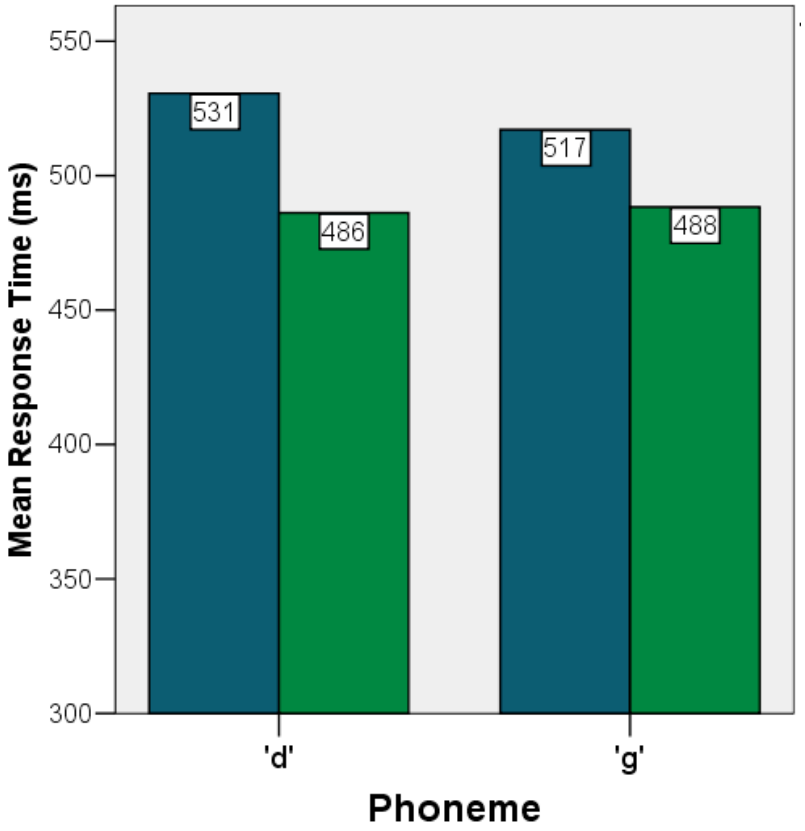


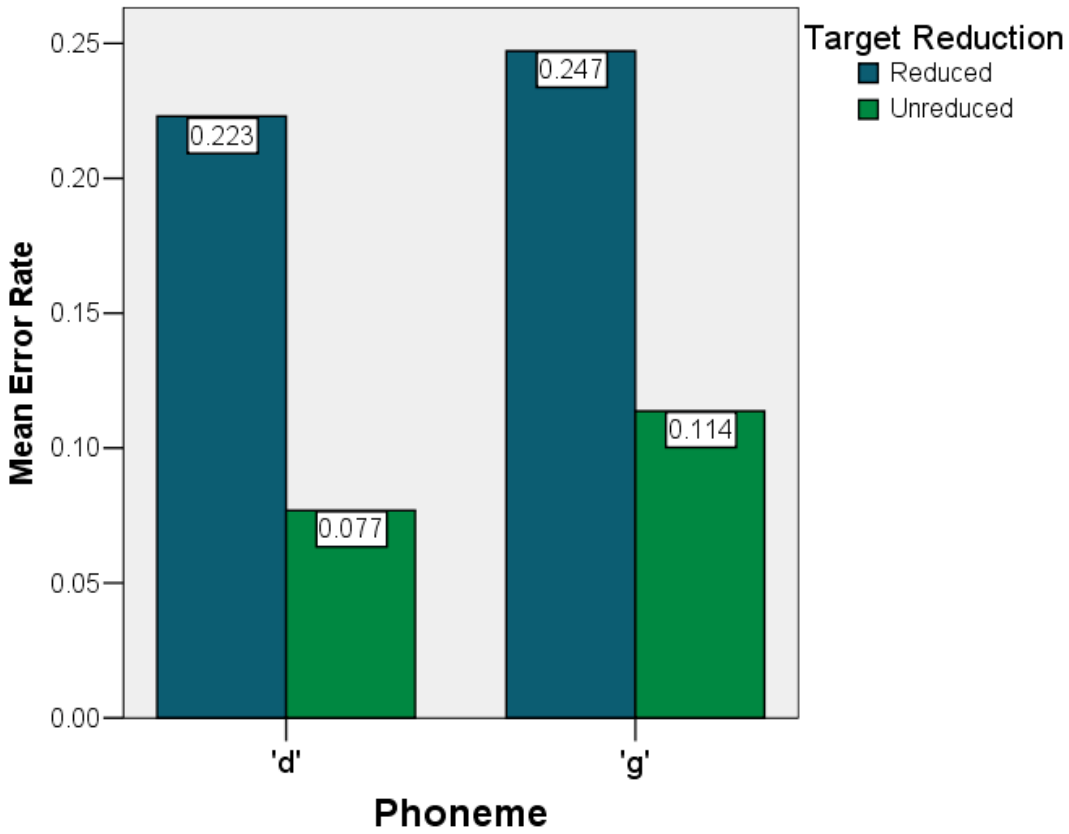


Reduced ‘puddle’

Results

Responses are significantly slower and less accurate for reduced targets.

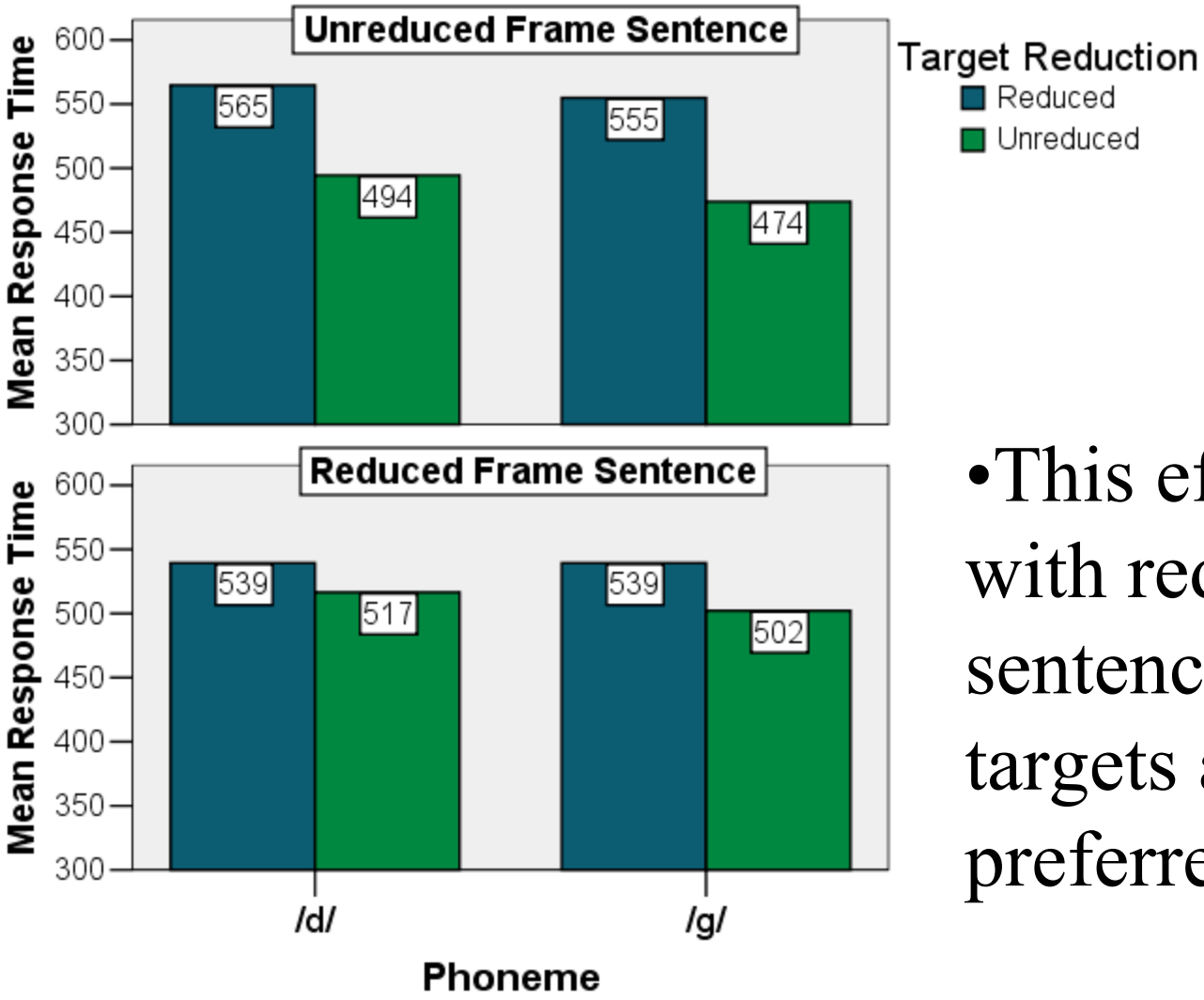




Processing Experiment 2: Lexical Decision in Context

- Recall how much easier it was to understand 'do you have time' with some context…
- Same procedure and items as Experiment 1 but with a preceding frame sentence as context.
- Context sentence: "A lot of the time he says ___" in either reduced or careful speech.
- Will listeners adjust their expectations about how sounds should be realized based on the degree of reduction in the preceding speech?

Results: Response Time



- Unreduced targets are significantly easier to recognize.
- This effect is smaller with reduced frame sentences: unreduced targets are not preferred by as much.

Results: Error Rate

- Unreduced targets are significantly easier to recognize.

- /d/ items follow the same pattern as the response times.
- /g/ items are doing the opposite.

Conclusions: processing

- It is easier to recognize clear than reduced words.
- $\Rightarrow$ Clarity of acoustic cues outweighs having heard reduced forms more often in natural speech.
- Listeners adjust how they expect a phoneme to be realized based on reduction of preceding context: the preference for unreduced speech isn't as strong after hearing reduced speech in the context

Conclusions: summary

- Intervocalic stops demonstrate a categorical, phonological, abstract effect on /t, d/ (flapping), as well as both systematic and random phonetic variability.
- Casual speech is more reduced than formal. Stress environment has limited effects.
- We understand each other despite a great deal of several types of variability, but we do understand clear speech more easily.
- Listeners adjust expectations based on the speech style of preceding context.

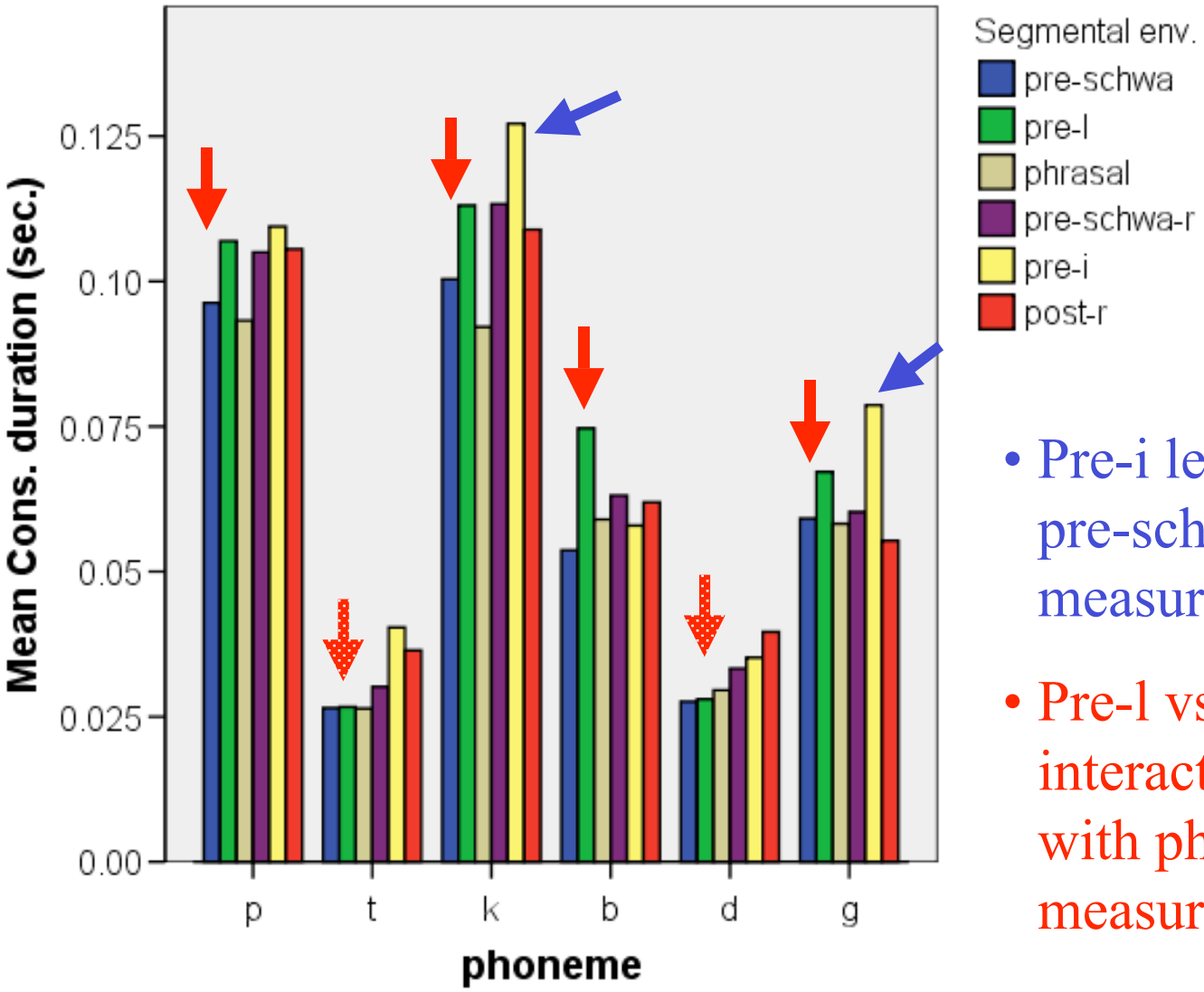
Effects of segmental environment

- Examined in word list reading, post-stress conditions only (full factorial design)
- Phoneme and segmental environment interact for most measures, but inconsistently
- Two interesting patterns:

➢ Stops appear to reduce less or differently before /i/ than elsewhere (because /i/ is peripheral?)

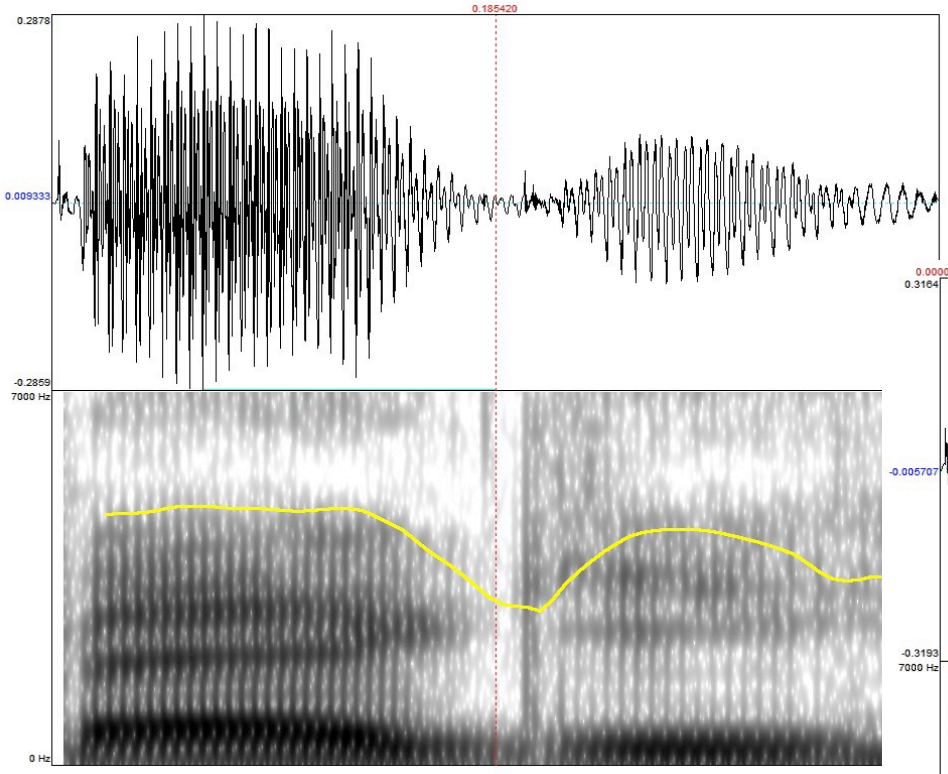
➢ /b, g/ appear to reduce less before [l̩] than elsewhere, while /t, d/ do not (shared pl. artic.)

Effects of segmental environment

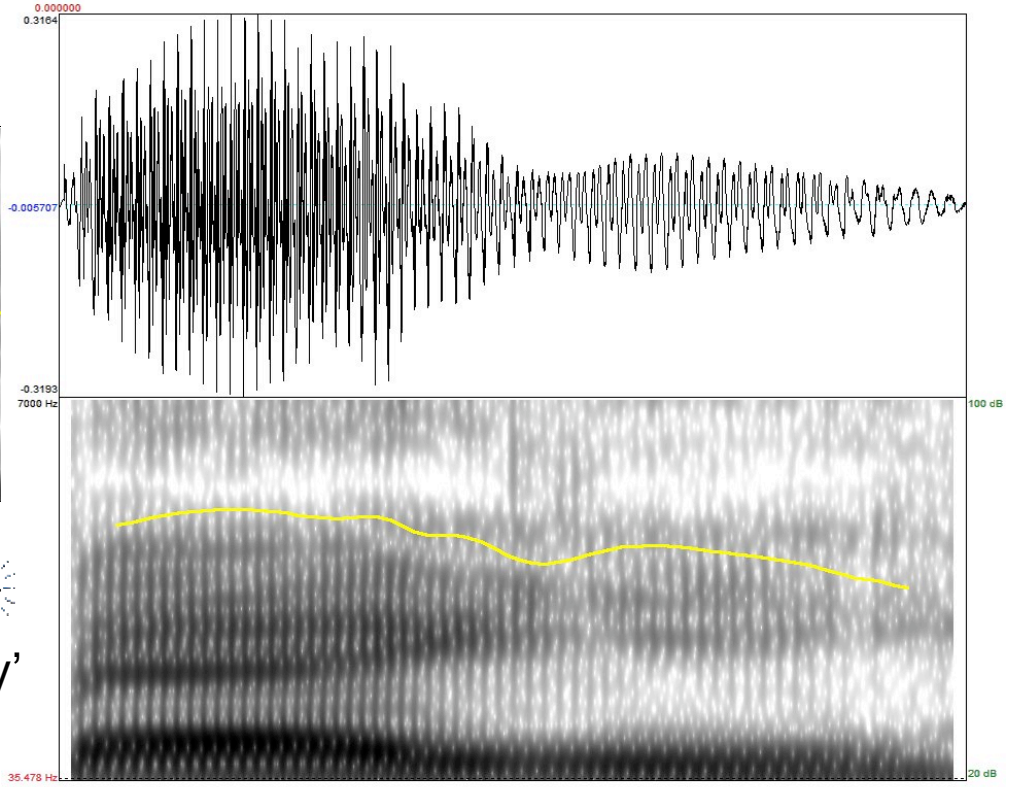


- Pre-i less reduced than pre-schwa for 4/6 measures
- Pre-l vs. pre-schwa only interacts significantly with phoneme for this measure

Method: Sample items /g/



Unreduced ‘baggy’



Reduced ‘baggy’