Deconstructing TD deletion

Meredith Tamminga
University of Pennsylvania

Josef Fruehwald
University of Edinburgh

NWAV 42 – Oct. 19, 2013

Introduction

Argument: "TD deletion" is not a unified phenomenon but rather the output of three distinct processes

Two recurring questions:

- What is a variable?
- How can any particular variable be defined?

What is a variable?

A core concept in sociolinguistics

- Probabilistic grammatical operation?
- Decision from among a set of options?
- Merely an analytic heuristic?

What is a variable?

"Sociolinguistic variation analysis is concerned with choices speakers make among the alternatives available to them regardless of the structural provenience of those choices. In many cases… there is no motivation for tying variation to rules of grammar."

Fasold 1991 p. 12-13

How can variables be defined?

Where does one variable end and another begin?

How can variables be defined?

"…considerations of variable constraints may provide a principled basis for combining or separating linguistic processes."

Wolfram 1975 p. 71

Influential principle of isomorphic constraints

(see Tagliamonte 2002)

Our view

- Crucial to tie observed surface variation to underlying grammatical structures
- Two new types of quantitative evidence to assess potential structural relationships.
- Delineating the grammatical underpinnings of variation also relevant to social meaning

TD basics

Variable deletion of word-final coronal stops in consonant clusters

Grammatical status of the coronal stop:

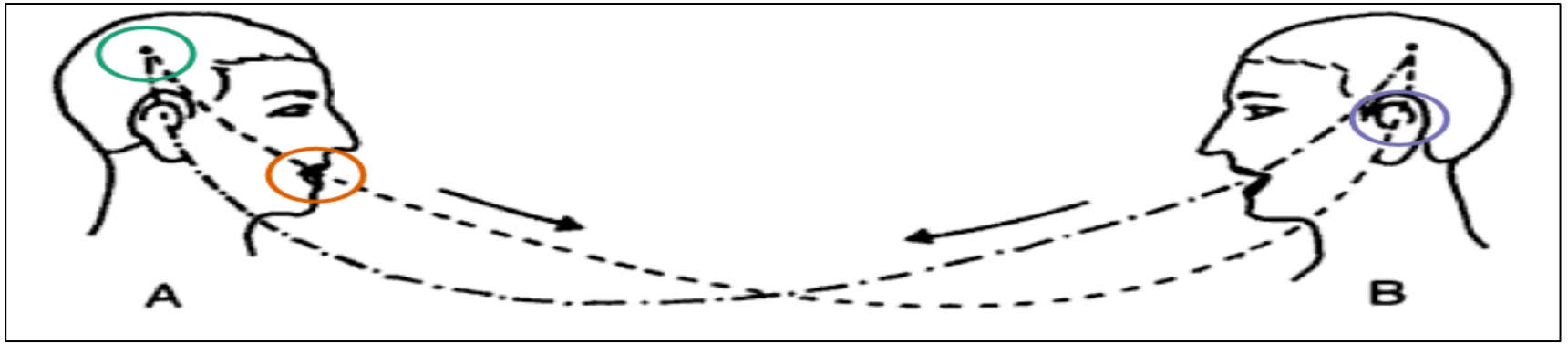
monomorpheme > semiweak > regular past

'fast' > 'kept' > 'worked'

TD history (abridged)

$$t,d \rightarrow \varnothing \; / \; \begin{bmatrix} \alpha \text{ cons} \\ \xi \text{ obs} \end{bmatrix} \gamma + \delta + \begin{bmatrix} \_\_\_\_\_ \\ \varepsilon \text{ voice} \end{bmatrix} \beta \sim V$$

work	**kep+t**	**fast**
work + ed	kept	fast
worked	kept	fast



Data

Buckeye Corpus (Pitt et al. 2007)

- 8,006 tokens
- 40 white speakers from Columbus

Philadelphia Neighborhood Corpus (subset from Labov & Rosenfelder 2011)

- 10,189 tokens
- 42 white speakers from South Philadelphia

Methods

- Buckeye tokens extracted using Python from hand-corrected phone tier
- PNC tokens coded auditorily using Praat script

Deletion rates	Monomorph	Semiweak	Regular past
Buckeye	53%	41%	23%
PNC	55%	45%	31%

Clustering

Clustering of individual speakers: how similar or different are deletion rates for individuals?

- If all TD deletion is the same, expect speaker differences to show consistent range across categories

Clustering

Clustering of individual tokens: how evenly dispersed are the variants in sequences?

- Variants generally facilitate their own reuse (Poplack 1980, Scherre & Naro 1991, Cameron 1992 *inter alia*)
- If all TD deletion is the same, expect tokens to show persistence that disregards categories

Speaker clustering

First source of evidence on unity of process: patterns in individuals' rates

TD deletion context: how do rates from different speakers cluster across grammatical categories?

Fit model to get individual speaker random intercepts within each category

Speaker clustering

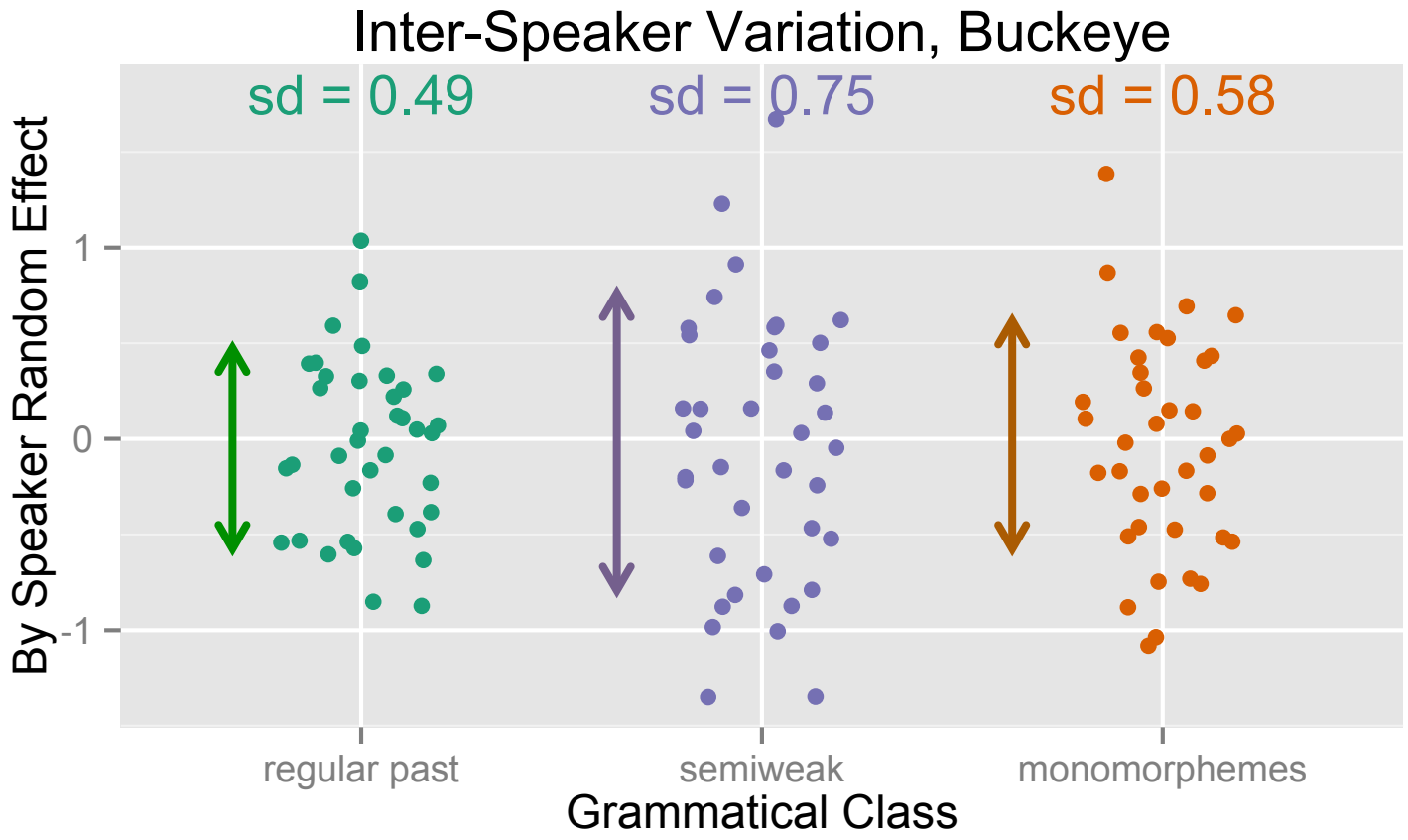
Fit mixed-effects logistic models using:

TD ~ Gram + PreSeg + FolSeg + Freq +
(-1+Gram|Speaker) + (1|Word)

Captures speaker-level random error for each of the grammatical categories separately

Are inter-speaker differences consistent across categories?

Speaker clustering



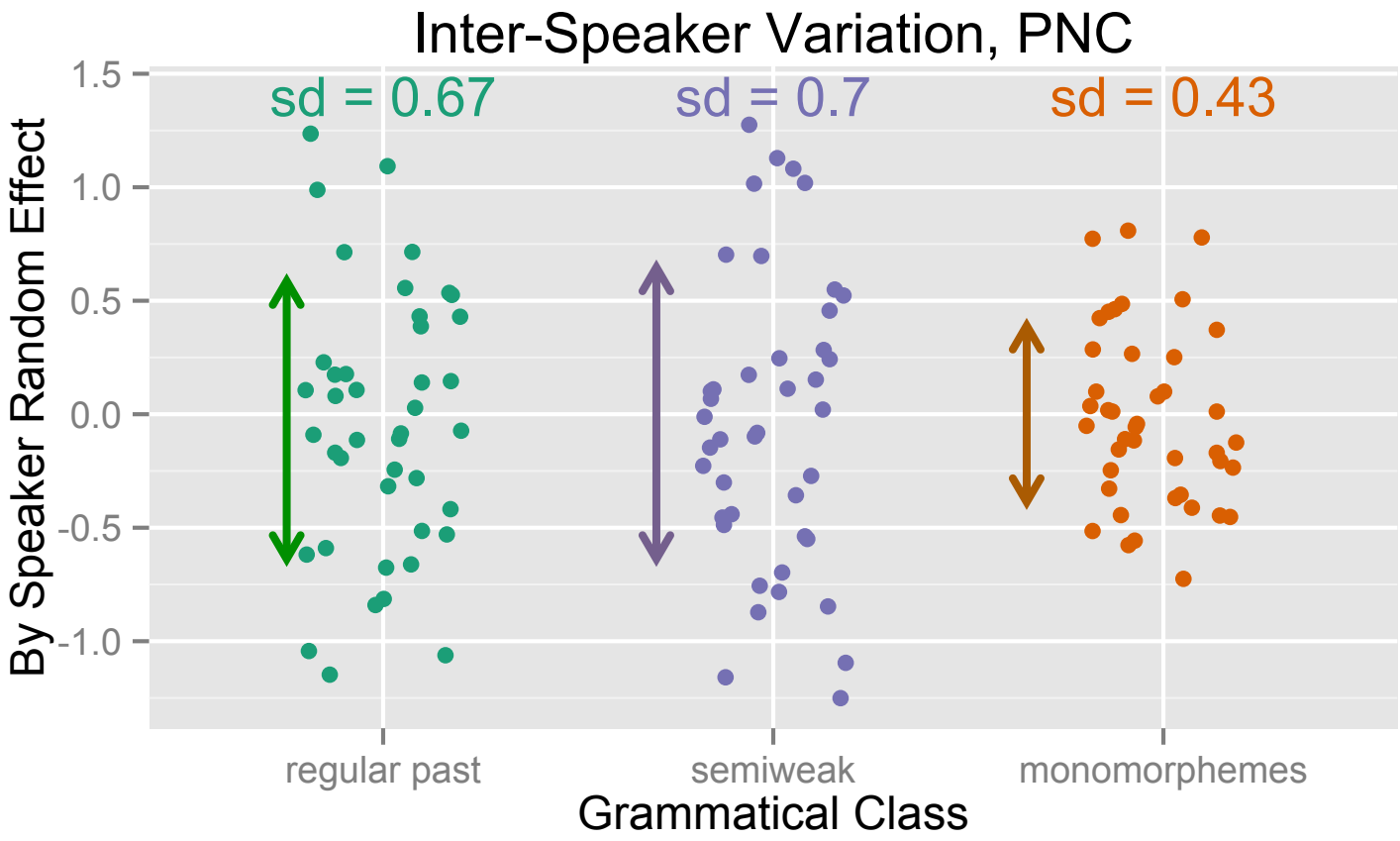
Speaker clustering

- Fruehwald & MacKenzie (2012): speakers more tightly clustered for regular past tense and monomorphemes than semiweak

- Fruehwald (2012): morphological variation in the semiweak verbs (like Guy and Boyd 1990) combined with the phonological variation across the board

$$keep + T_{past} \rightarrow kep \sim kept$$

Speaker clustering



Speaker clustering

Attribute wider variance to overlapping processes

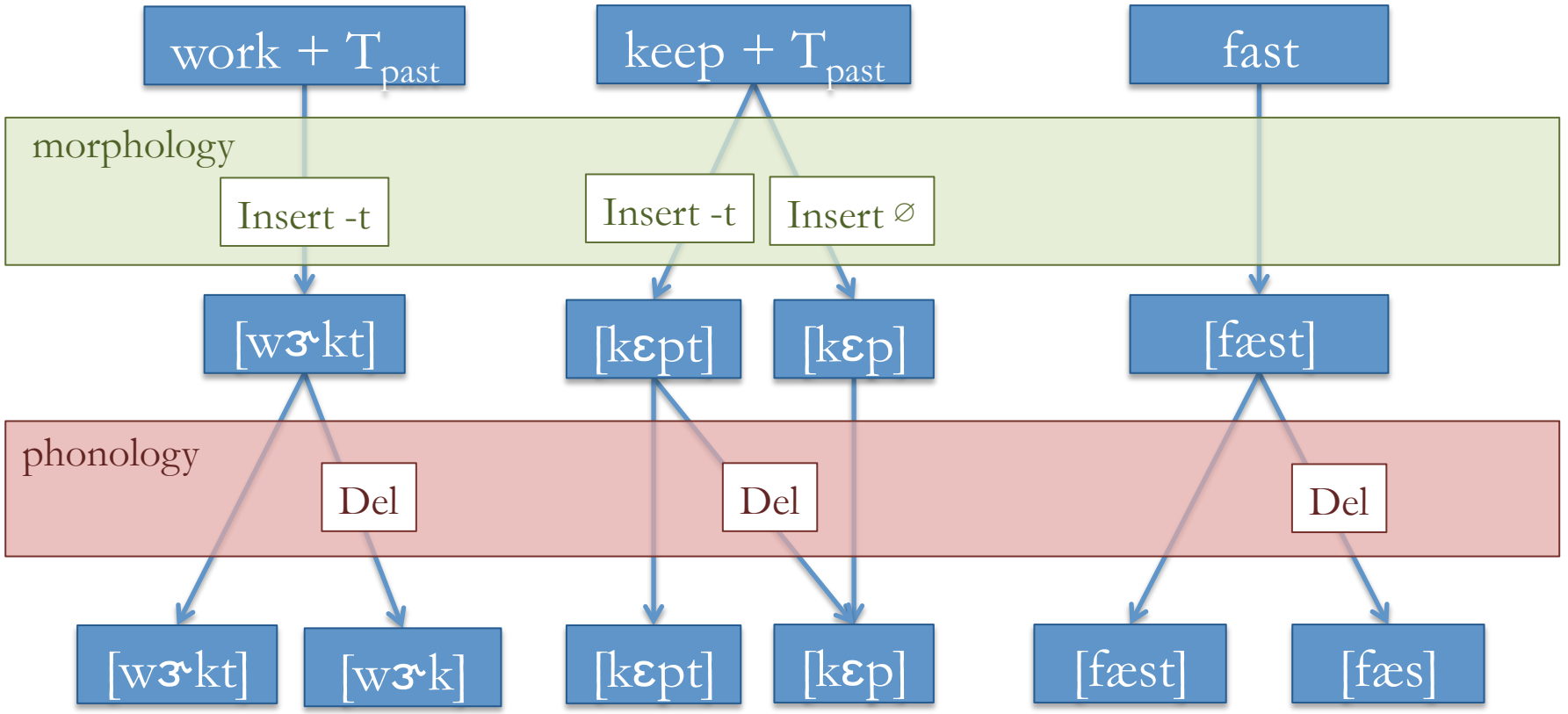
Semiweaks subject to two variable processes:

- Variable morphology (allomorphy):

  $T_{+past} \rightarrow -\varnothing$ or $T_{+past} \rightarrow -t$

- Variable phonological deletion

Deconstruction so far



Token clustering

Second source of evidence on unity of process: patterns in sequences of observations

In the TD context: how do tokens of the variants cluster across grammatical categories?

Code each token for values of previous token, by category, test for persistence

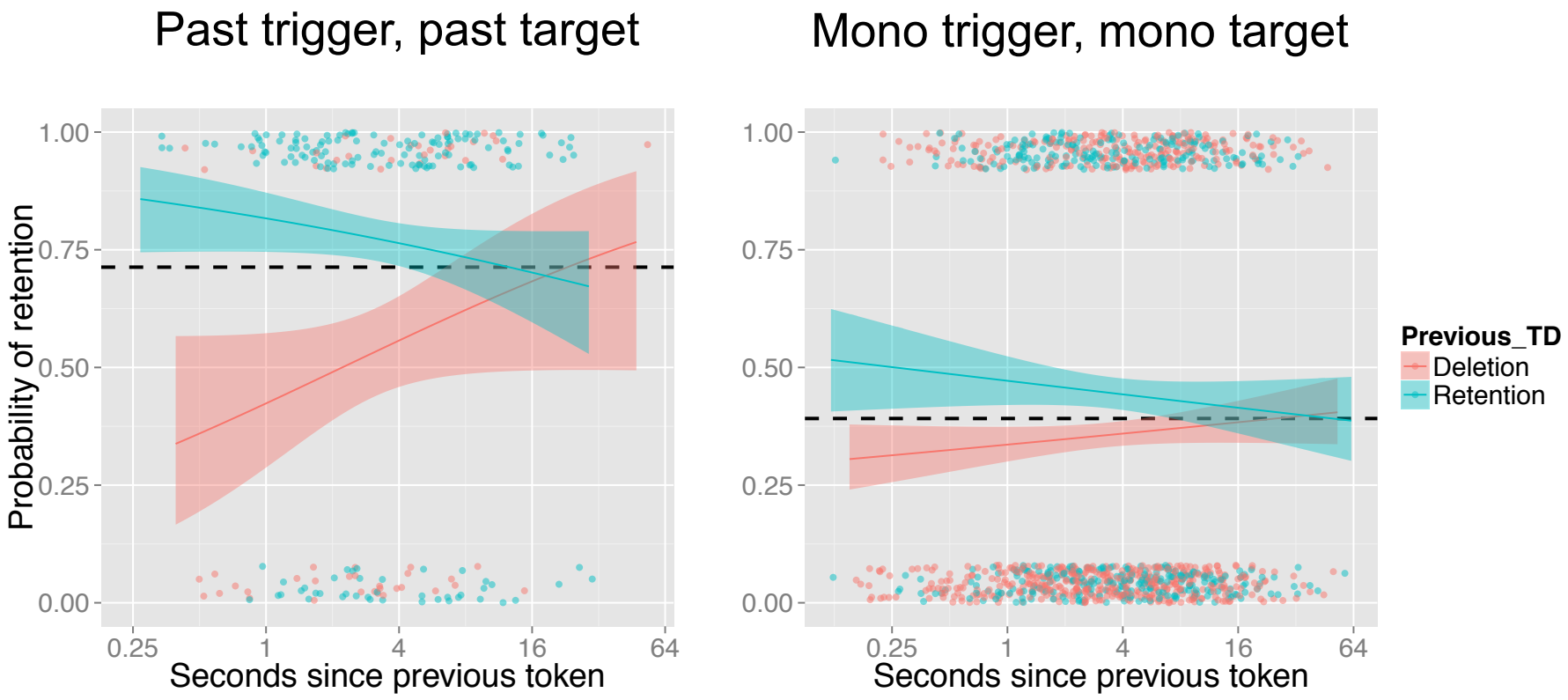
Token clustering

For each token, code the previous token's:

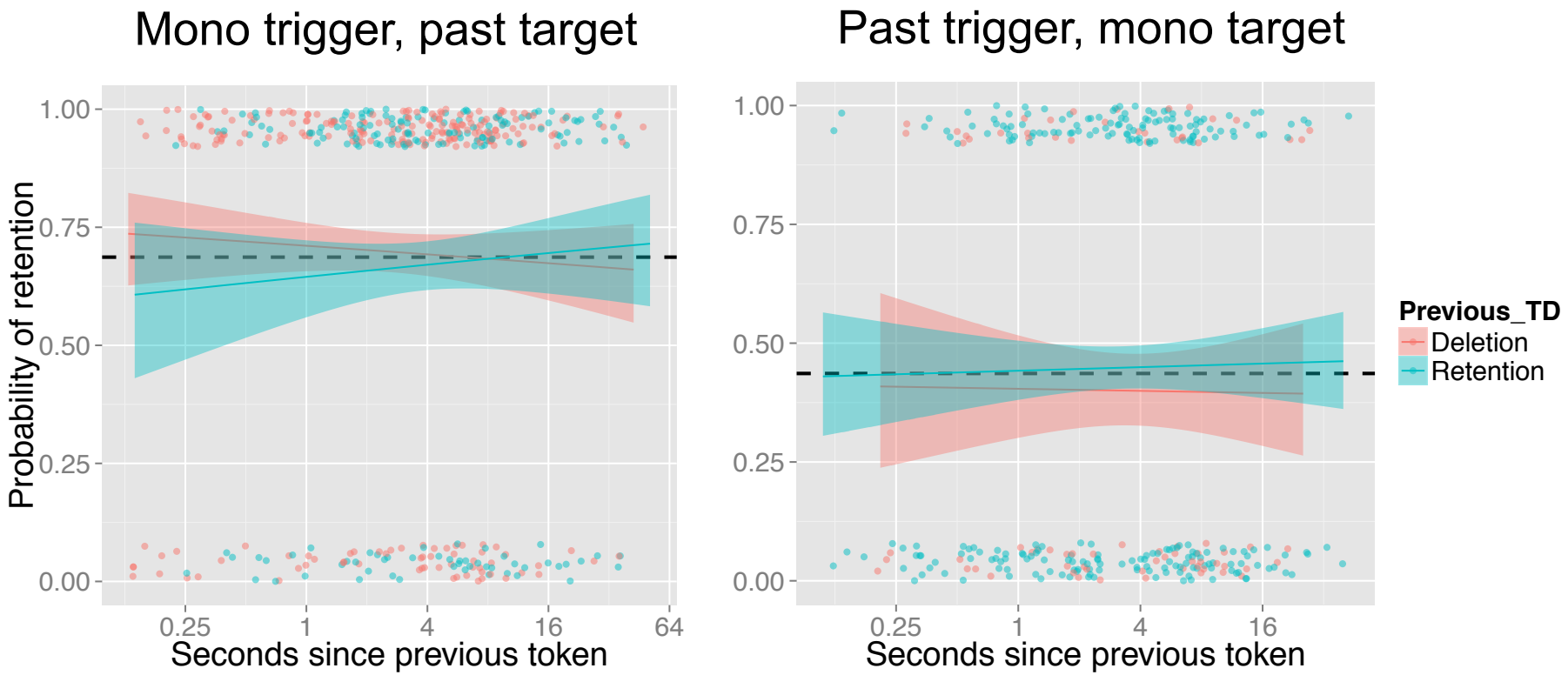
- variant (deletion or retention)
- grammatical status (monomorpheme or past tense)
- distance (in seconds)

Are tokens preceded by deletion more likely to be deleted, and vice versa?

Token clustering



Token clustering

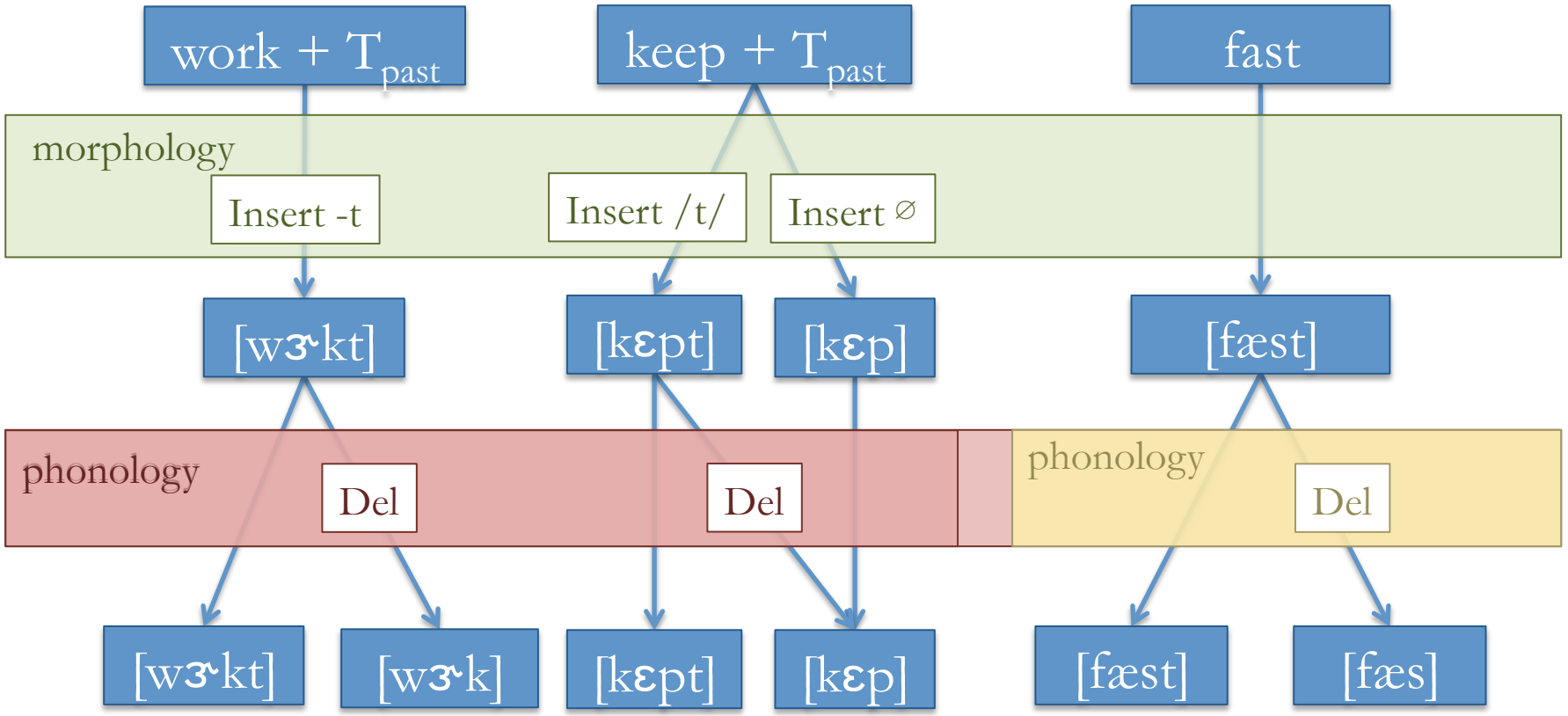


Token clustering

Grammatically-matched trigger/target pairs show persistence, but mismatched pairs don't.

Suggests distinct processes of TD deletion in monomorphemes and regular past tense verbs at the phonological level

The full deconstruction



Summary

Two results:

- Greater deletion rate variance across speakers in semiweaks than in monomorphemes or regular past tense verbs
- Grammatically matched trigger-target pairs show persistence, but mismatched ones don't

Summary

Three distinct sources of TD variability:

- Allomorphy in semiweaks
- Phonological deletion in all past tense forms
- Phonological deletion in monomorphemes

Discussion

Methodologically, we:

- Dispute the irrelevance of grammatical representations for defining the variable
- Offer two new types of quantitative evidence on unity of process in linguistic variation

These types of evidence can be considered in conjunction with constraint hierarchies.

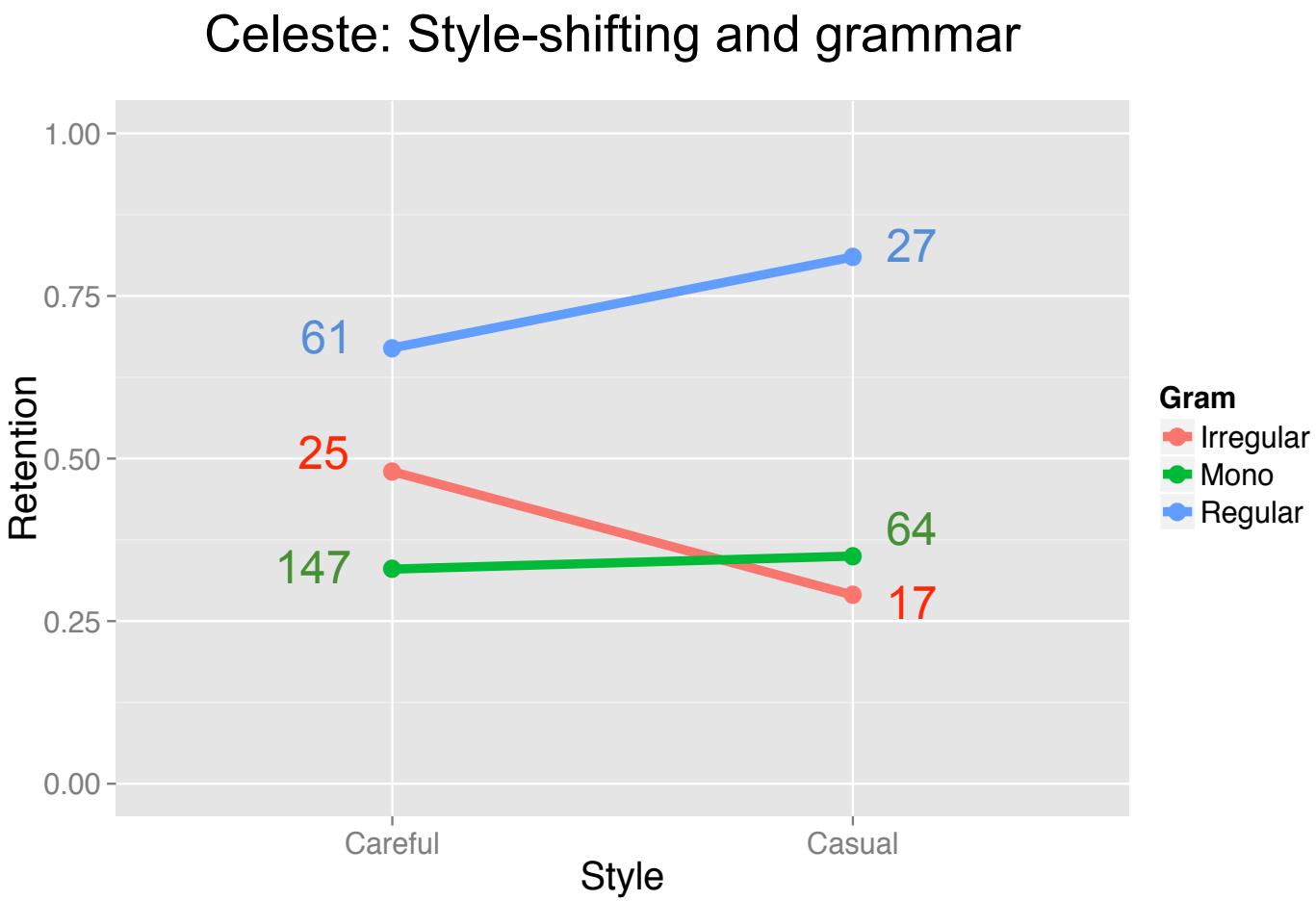
Discussion

Socio-stylistic payoff for attention to grammatical underpinnings of variation?

A sneak peek at some work in progress…

- Interview with Celeste, two distinct styles

Discussion


Celeste: Style-shifting and grammar
1.00
0.75
0.50
0.25
0.00
Retention
61
27
25
64
147
17
Careful
Casual
Style
Gram
Irregular
Mono
Regular

Conclusion

Variable phenomena that look similar on the surface may have different grammatical origins

Suggested model involves three processes

New types of quantitative evidence may lead to similar revisions of other variables

Thank you!

Thanks to Bill Labov, Dave Embick, Laurel MacKenzie, audiences at PLC 35 and NWAV 41

Contact us:

Meredith: tamminga@ling.upenn.edu

Joe: josef.frueh@ed.ac.uk

References

Bybee, J. 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. Language Variation and Change 14, 261-290.

Fasold, R. 1991. The quiet demise of variable rules. American Speech 66, 3-21.

Fruehwald, J. 2012. Redevelopment of a morphological class. Penn Working Papers in Linguistics 18.1, 77-86.

Fruehwald, J. & L. MacKenzie. 2012.

Guy, G. 1991. Explanation in variable phonology: An exponential model of morphological constraints. Language Variation and Change 3, 1-22.

Guy, G. 2007. Lexical exceptions in variable phonology. Penn Working Papers in Linguistics 13.2, 109-119.

Guy, G. & S. Boyd. 1990. The development of a morphological class. Language Variation and Change 2, 1-18.

Labov, W., P. Cohen, C. Robins & J. Lewis..1968. A study of the non-standard English of Negro and Puerto Rican speakers in New York City. Final report, Coorperative research Project 3288, Vol. I.

Tagliamonte, S. 2002. Comparative Sociolinguistics. In Chambers & Schilling (eds.), The Handbook of Language Variation and Change, Ch. 6. Wiley-Blackwell.

Weinreich, U. & W. Labov & M. Herzog. 1968. Empirical foundations for a theory of language change. In Lehmann & Yakov (eds.), Directions for Historical Linguistics: A Symposium, 97-195.

Wolfram, W. 1975. Variable constraints and rule relations. In Fasold & Shuy (eds.), Analyzing Variation in Language: Papers from the Second Colloquium of NWAV, 70-88. Georgetown University Press.