

# **Two case studies on the non-local conditioning of variation**

Laurel MacKenzie and Meredith Tamminga  
University of Manchester  
University of Pennsylvania

LSA 2013 Annual Meeting  
January 3, 2013

# Overview

How are variable phenomena represented in the linguistic systems of individuals?

Similarity of variable processes to categorical rules  $\longrightarrow$  variation inside the grammar

Dissimilarity of variable processes to categorical rules  $\longrightarrow$  variation outside the grammar

# **Inherent variability & variable rules**

## **inherent variability**

“the hypothesis that the human language faculty necessarily accommodates and generates variation, and that the workings of grammar have a quantitative, noncategorical, and nondeterministic component”

Guy & Boberg (1997:149), paraphrasing WLH

# **Inherent variability & variable rules**

## **inherent variability**

“the hypothesis that the human language faculty necessarily accommodates and generates variation, and that the workings of grammar have a quantitative, noncategorical, and nondeterministic component”

Guy & Boberg (1997:149), paraphrasing WLH

## **variable rules**

“enlargement of the concept ‘rule of grammar’”

Labov (1969:737)

# Guy and Boberg and the OCP

Guy & Boberg's proposal:

“a unified probabilistic grammar that accounts for both” categorical and probabilistic alternations (p. 150)

# Guy and Boberg and the OCP

Guy & Boberg's proposal:

“a unified probabilistic grammar that accounts for both” categorical and probabilistic alternations (p. 150)

Their motivation:

conditions on variable *t/d*-deletion resemble the effects of the Obligatory Contour Principle

deletion rate: /**nt**/ > /**st**/ = /**pt**/ > /**ft**/ > /**lt**/

← phonological similarity to /t/

# Guy and Boberg and the OCP

Guy's interpretation of this finding:

Separating variation (performance) from grammar (competence) would necessitate two separate versions of the OCP.

It is likely that many constraints on categorical processes would have “separate but equal performance twin[s]” in this way.

(Guy 1997:134)

# Guy and Boberg and the OCP

Guy's interpretation of this finding:

Separating variation (performance) from grammar (competence) would necessitate two separate versions of the OCP.

It is likely that many constraints on categorical processes would have “separate but equal performance twin[s]” in this way.

(Guy 1997:134)

This would result in “considerable duplication of formal machinery.” (Coetzee & Pater 2011:406)



# Variation in phonological theory

“...the prospects of variation in mainstream generative phonology have changed dramatically. It now occupies a central place in the study of phonology, and to some extent dictates the architecture of phonological grammar”

(Coetzee & Kawahara 2012)

# Variation in phonological theory

## “grammatical overreach”:

“if these purely grammatical models are accounting nearly perfectly for the data, then grammar is doing more than its fair share”

(Coetzee & Kawahara 2010)

They consider the role of frequency; we pursue two other cases of extragrammatical variability:

- Subject length effects
- Persistence effects

# Auxiliary contraction

*is*

Yeah, **Salzburg's** nice. **Austria's** nice.

**Europe is** nice! (sw\_1151)

*has*

Oh, I'm sure **it's** been done. I'm sure **it has**  
been done. (sw\_1060)

*will*

If I walk, **it'll** be ten degrees warmer, but **it**  
**will** last twenty minutes. (sw\_1146)

# Auxiliary contraction

*is*

Yeah, **Salzburg's** nice. **Austria's** nice.

**Europe** **is** nice! (sw\_1151)

*has*

Oh, I'm sure **it's** been done. I'm sure **it has** been done. (sw\_1060)

*will*

If I walk, **it'll** be ten degrees warmer, but **it will** last twenty minutes. (sw\_1146)

# Data sources

- The Switchboard corpus (Godfrey et al., 1992)
- The Fisher corpus (Cieri et al., 2004)
  - 5-minute telephone conversations between strangers on a given topic
- The Philadelphia Neighborhood Corpus (Labov & Rosenfelder, 2011)
  - Sociolinguistic interviews carried out by Penn Linguistics students

# Coding

## dependent variable

	contracted	uncontracted
<i>is</i>	[z], [s]	[ɪz], [əz]
<i>has</i>	[z], [s]	[hæz], [həz], [əz]
<i>will</i>	[əl]	[wɪl], [wəl]

(MacKenzie 2012)

# Coding

## independent variables

length of subject in orthographic words

<u>Salzburg's</u> nice	1
<u>The real estate out here's</u> been pretty good	4
<u>About the only thing I can do mechanically with</u> <u>a, a car</u> is put gas in it	12

# Coding

## **independent variables**

length of subject in orthographic words  
*is* only: preceding vowel vs. consonant



# Coding

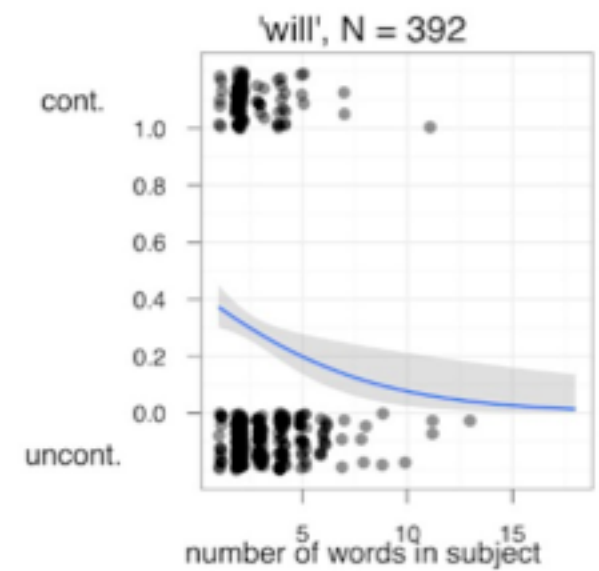
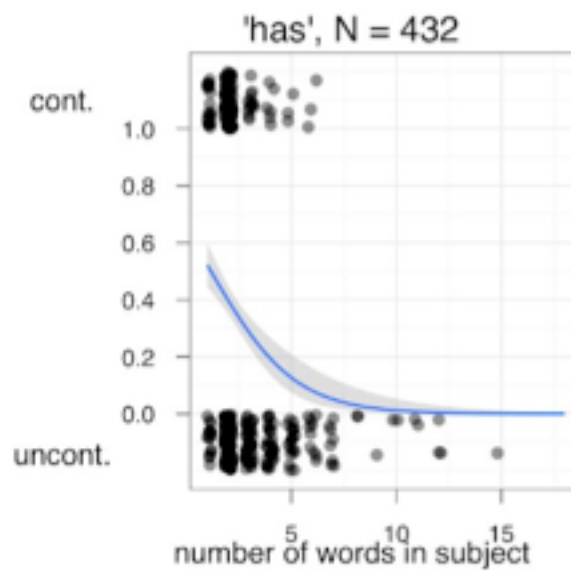
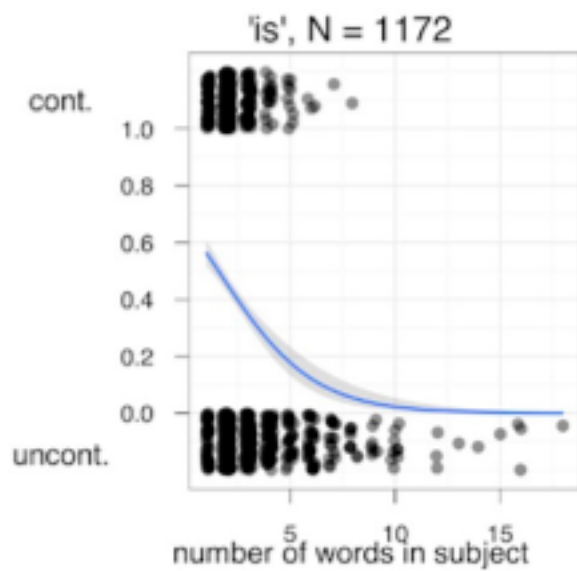
## **independent variables**

length of subject in orthographic words

*is only:* preceding vowel vs. consonant

*is only:* following grammatical class

# Subject length effect



# Subject length effect: implications

Some conditions on contraction do resemble conditions on categorical alternations

- e.g. **preceding segment**: compare Korean allomorphy

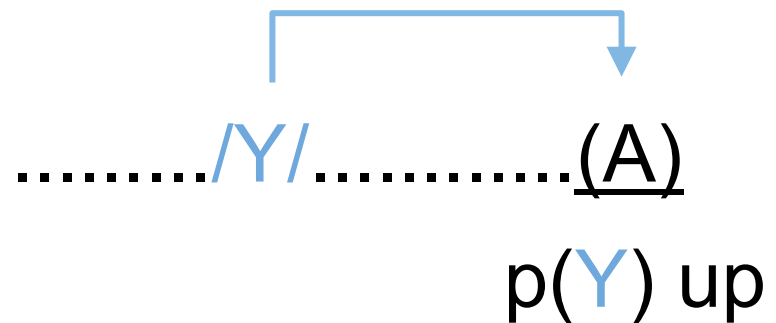
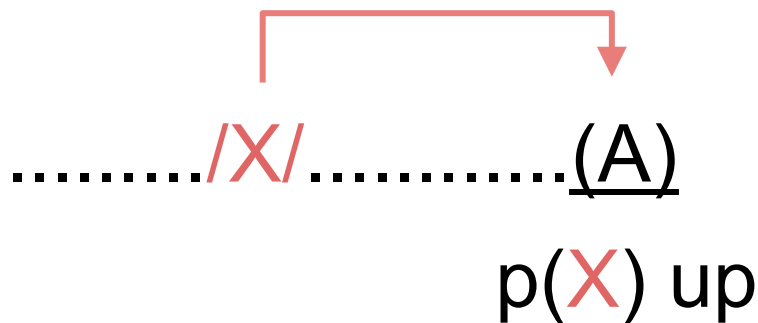
But, **subject length is different**:

- “Grammars can't count”: categorical alternations don't make reference to quantities larger than 2 (Selkirk 1986)
- Yet auxiliary realization appears to be sensitive to precise subject word count

# Persistence

Tendency for a recently-used linguistic form to be used again

Variable (A) with two variants /X/ and /Y/:



# Data source

Subset of the PNC: 42 white speakers

<b>Birth year</b>	<b>Female</b>	<b>Male</b>
Before 1930	5	5
1930–1959	11	10
After 1959	5	6

Both DH and ING known to be stable in Philadelphia

(Labov 2001)

# Variables

ING: alternation between unstressed /ɪŋ/ and /ɪn/ (*working/workin'*)

- proper nouns excluded

DH: alternation between fricative /ð/ and stop /d/ word-initially (*this/dis*)

- intermediate affricate variant included with fricative
- deletions excluded (*'em*)
- lexical item *the* excluded
- neutralized following apical stops

# Coding

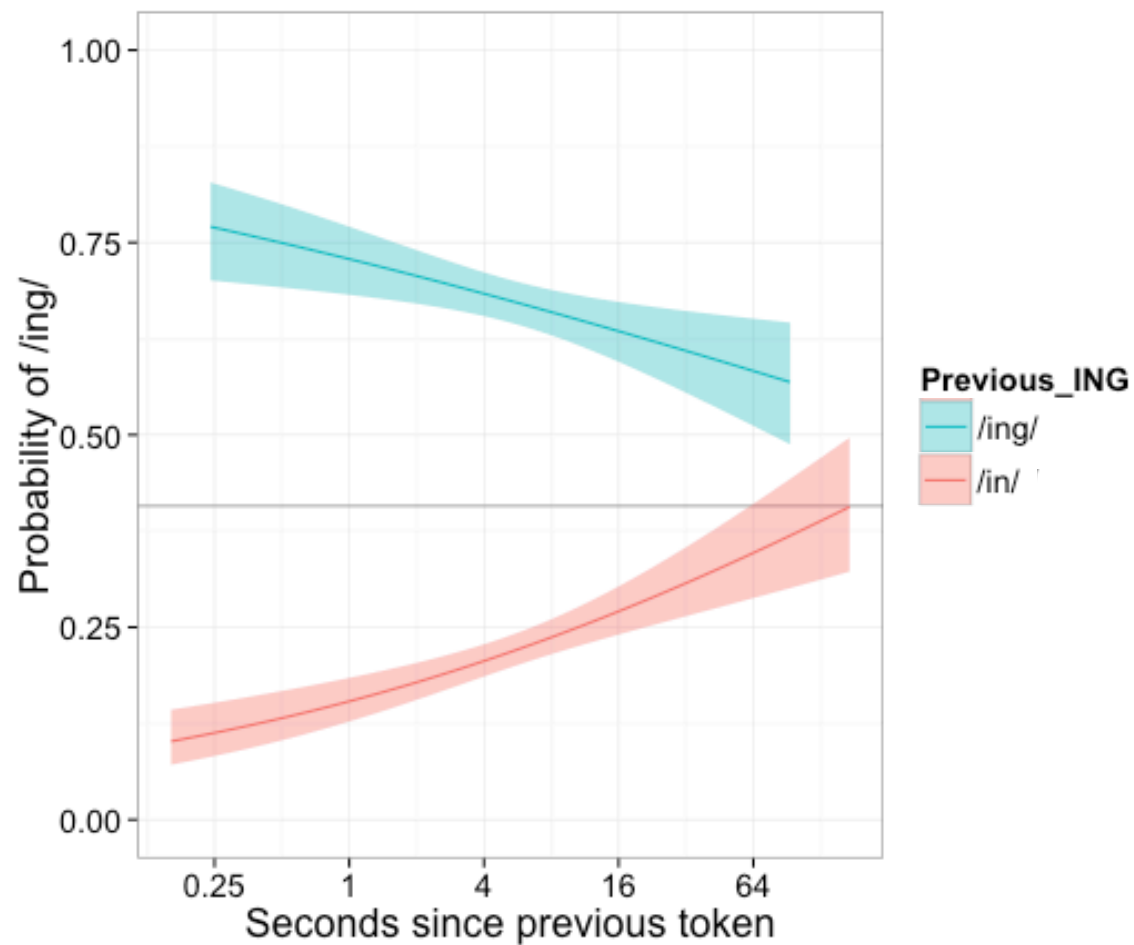
Each token coded for value of previous token

Distance from previous token measured in seconds and log-transformed

Previous tokens not coded across interruption by interlocuter

# Persistence effect on ING

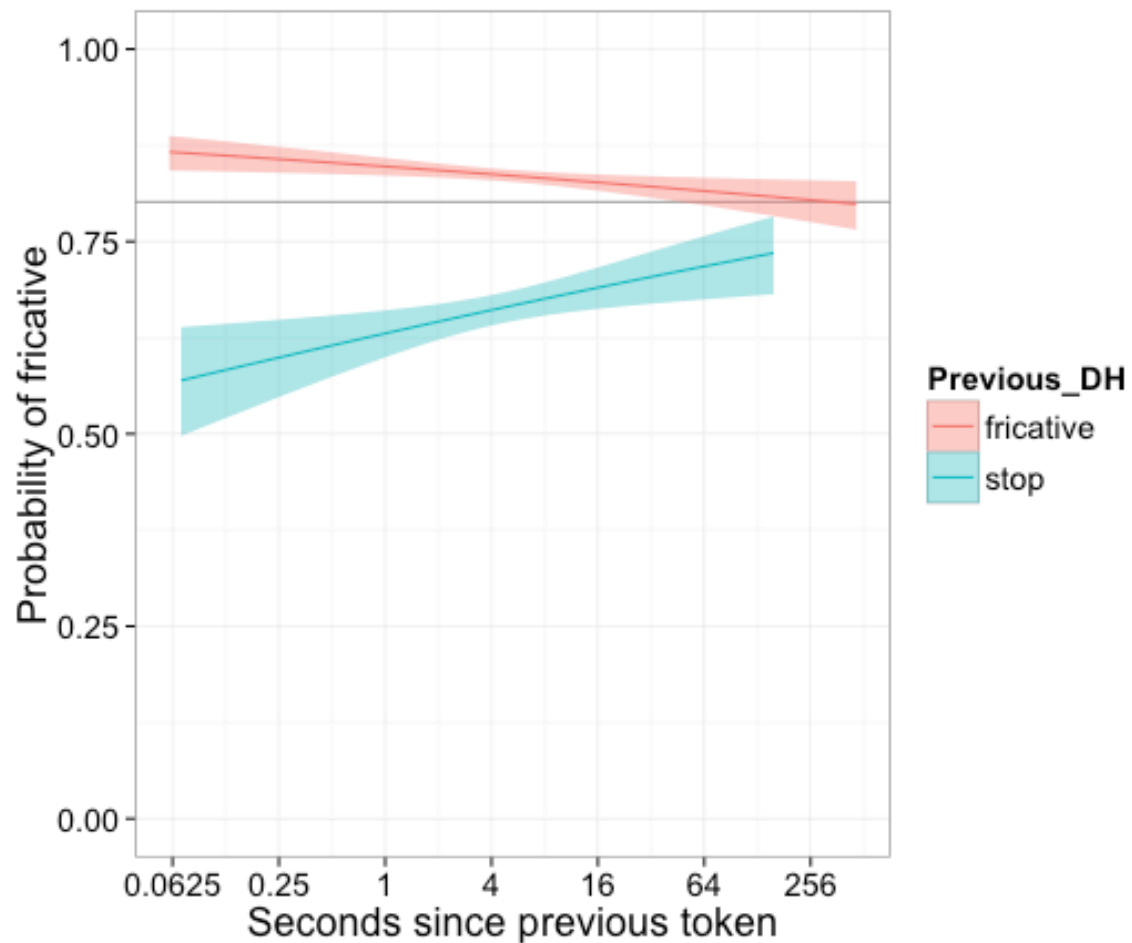
N = 2671





# Persistence effect on DH

N = 11,172



# Persistence effect: implications

Like contraction, ING and DH conditioned by linguistic factors in ways that look like categorical rules

e.g. **following segment**: compare Yiddish voicing

But again, **persistence is different**:

- Conditions on allomorphy and phonological rules are locally-constrained (Embick 2010)
- Highly non-local; in effect for over a minute

# Extragrammatical variation

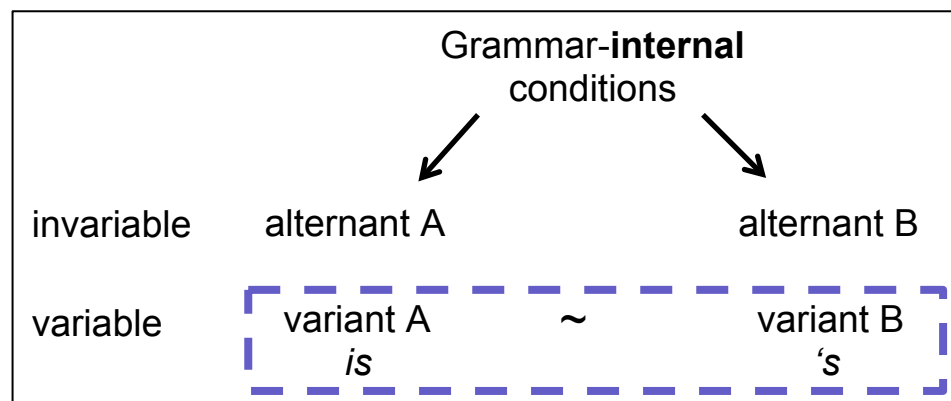
Subject length effect would require grammar to count

Persistence effect would require grammar to have a memory

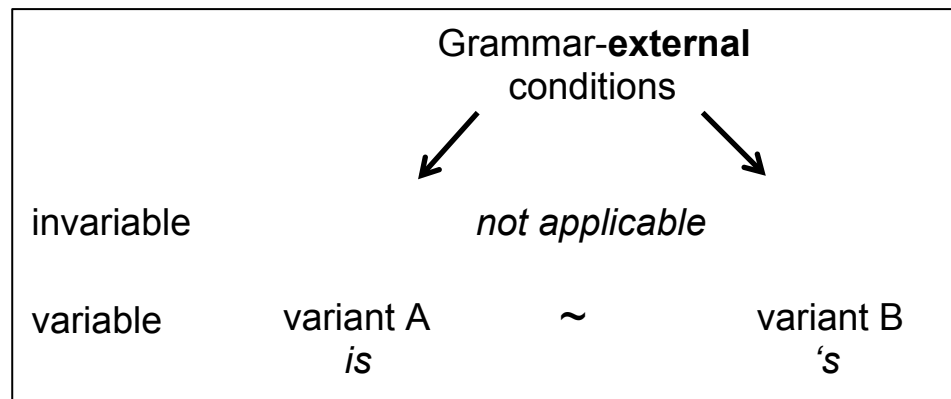
Would need to constrain grammar to **not** allow such effects to operate on categorical processes if they were represented grammar-internally

# Modeling variation

## 1 Grammar



## 2 Use



# Conclusion

Surface probabilities reflect variation originating within and outside of the grammar.

Grammatical architectures must still be structured to allow variation.

Sociolinguistic models should distinguish between different types of variability.

**Thank you!**