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?

Overview

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

Similarity of variable processes to categorical rules ⟶ variation inside the grammar

Overview

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

Similarity of variable processes to categorical rules ⟶ variation inside the grammar

Dissimilarity of variable processes to categorical rules ⟶ variation outside the grammar

Inherent variability & variable rules

inherent variability

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

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 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: /n**t**/ > /s**t**/ = /p**t**/ > /f**t**/ > /l**t**/

phonological similarity to /t/

Guy and Boberg and the OCP

Guy's interpretation of this finding:

Guy and Boberg and the OCP

Guy's interpretation of this finding:

> Separating variation from grammar would necessitate two separate versions of the OCP.

Guy and Boberg and the OCP

Guy's interpretation of this finding:

> Separating variation from grammar would necessitate two separate versions of the OCP.
>
> It is likely that many conditions 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 from grammar would necessitate two separate versions of the OCP.
>
> It is likely that many conditions 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

Variation in phonological theory

"variation… 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)

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:

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

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

Auxiliary contraction

is

Yeah, **Salzburg's** nice. **Austria's** nice.
Europe is nice! (sw_1151)

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)

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

Data sources

- The Switchboard corpus (Godfrey et al., 1992)

Data sources

- The Switchboard corpus (Godfrey et al., 1992)
- The Fisher corpus (Cieri et al., 2004)

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

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)

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

Subject length

coded in orthographic words

Subject length

coded in orthographic words

Salzburg's nice 1

Subject length

coded in orthographic words

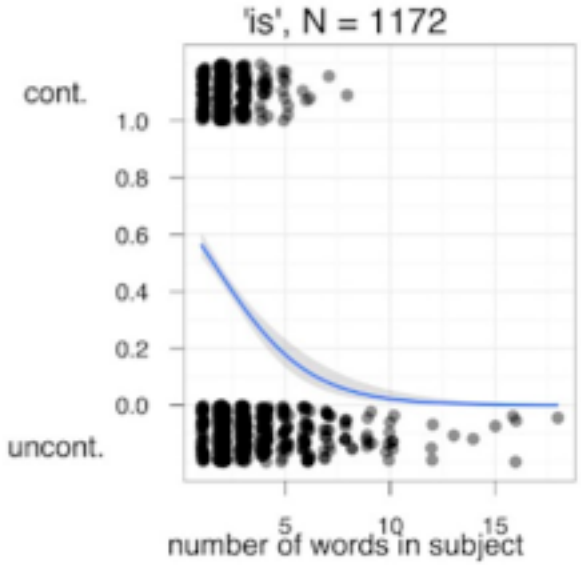
Salzburg's nice	1
The real estate out here's been pretty good	4

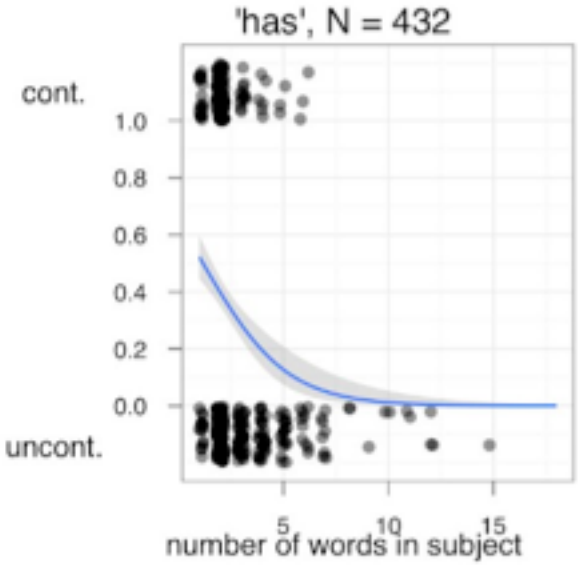
Subject length

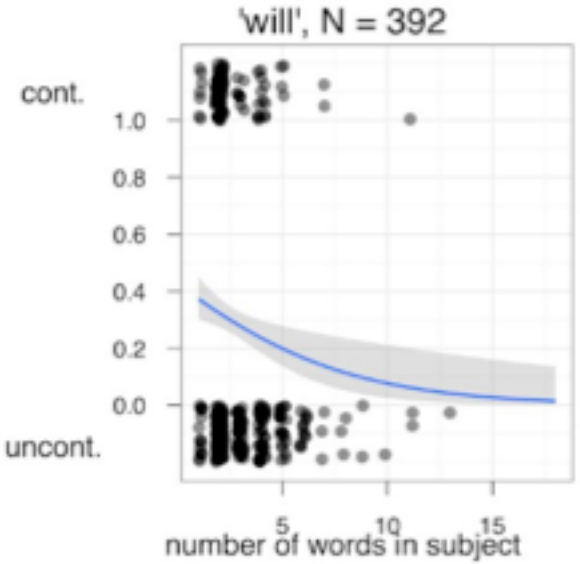
coded in orthographic words

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

Subject length effect







Subject length effect: implications

Subject length effect: implications

Some conditions on contraction do resemble conditions on categorical alternations

Subject length effect: implications

Some conditions on contraction do resemble conditions on categorical alternations

- e.g. **preceding segment**

Subject length effect: implications

Some conditions on contraction do resemble conditions on categorical alternations

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

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:**

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)

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

Persistence

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

Persistence

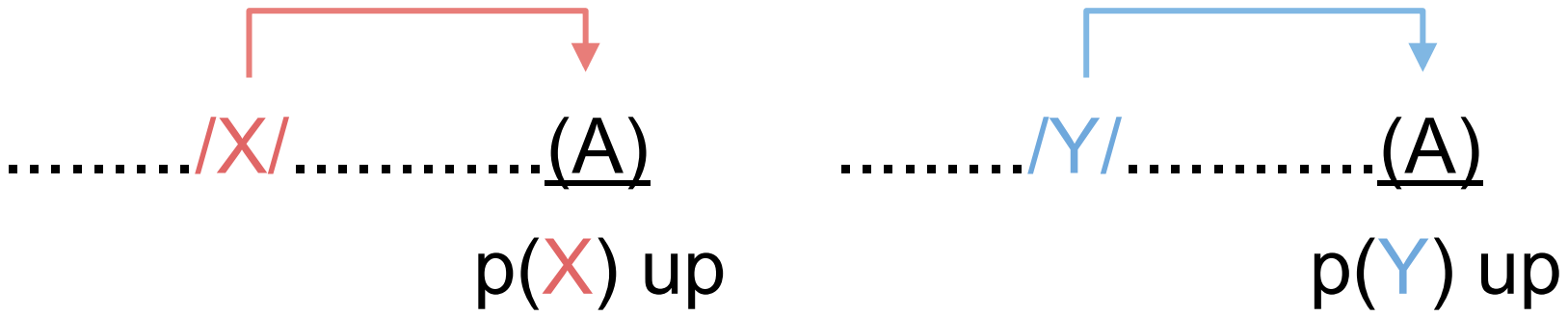
Tendency for a recently-used linguistic form to be used again

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

Persistence

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

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



Variables

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

Variables

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

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

Variables

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

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

Data source: 42 speakers from the PNC

Coding

Each token coded for value of previous token

Coding

Each token coded for value of previous token

Distance from previous token measured in seconds and log-transformed

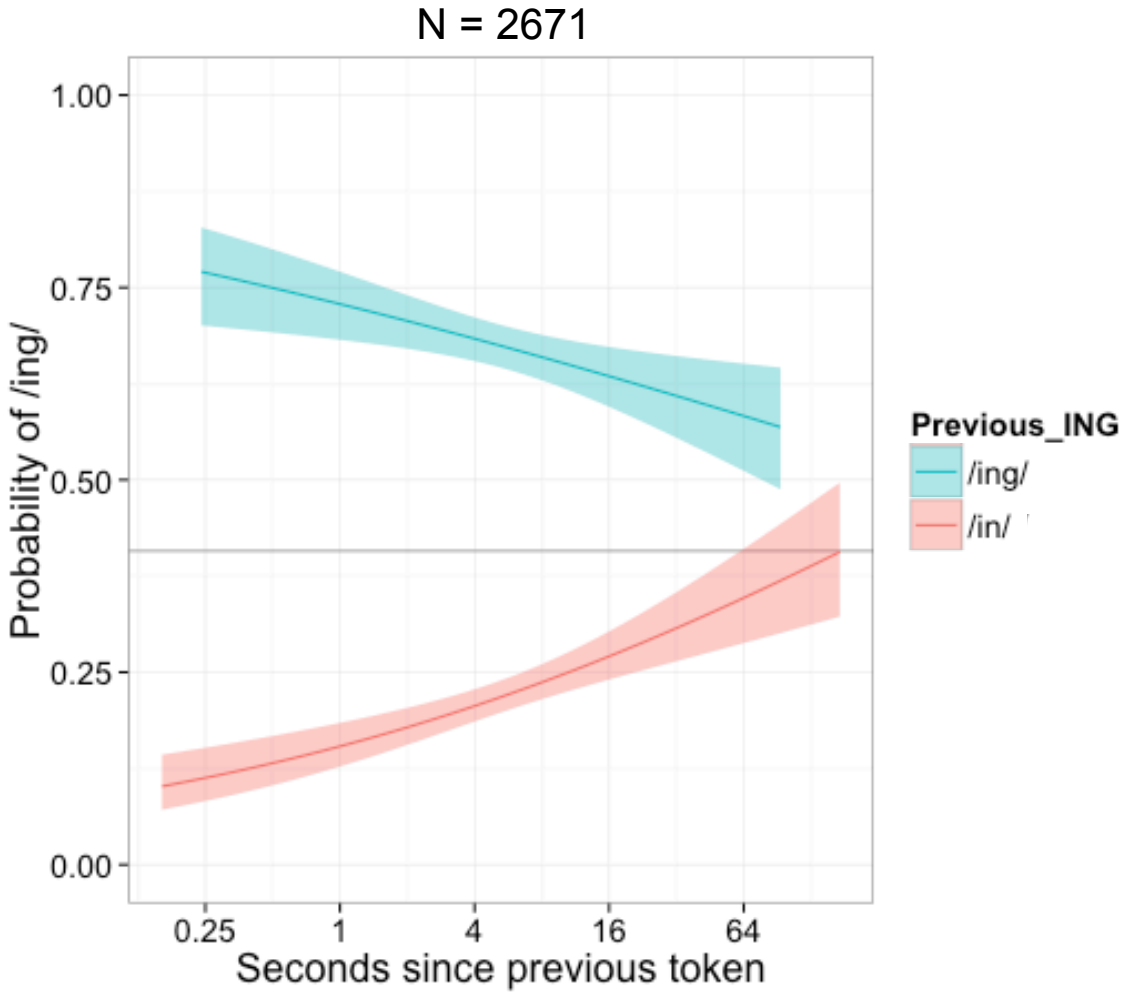
Coding

Each token coded for value of previous token

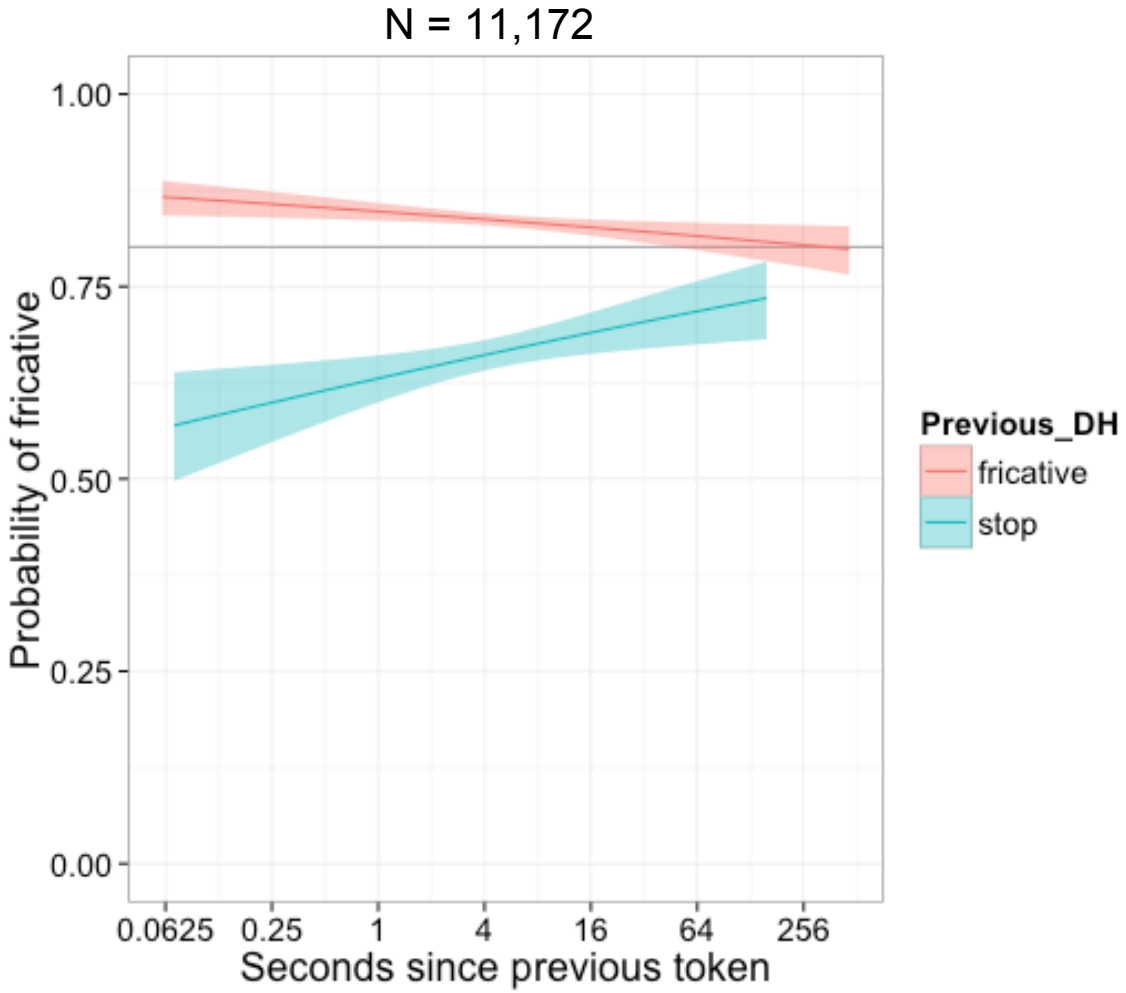
Distance from previous token measured in seconds and log-transformed

Previous tokens not coded across interruption by interlocuter

Persistence effect on ING



Persistence effect on DH



Persistence effect: implications

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

e.g. **following segment**

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 assimilation

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 assimilation

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

Extragrammatical variation

Subject length effect would require grammar to count

Extragrammatical variation

Subject length effect would require grammar to count

Persistence effect would require grammar to have a memory

Extragrammatical variation

Subject length effect would require grammar to count

Persistence effect would require grammar to have a memory

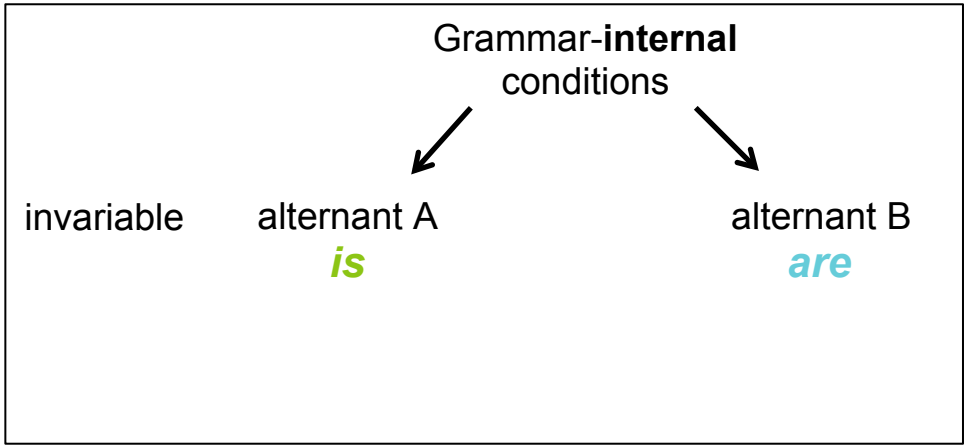
Would need to constrain grammar to **not** allow such effects to operate on categorical processes if they were represented grammar-internally

Modeling variation

1 Grammar

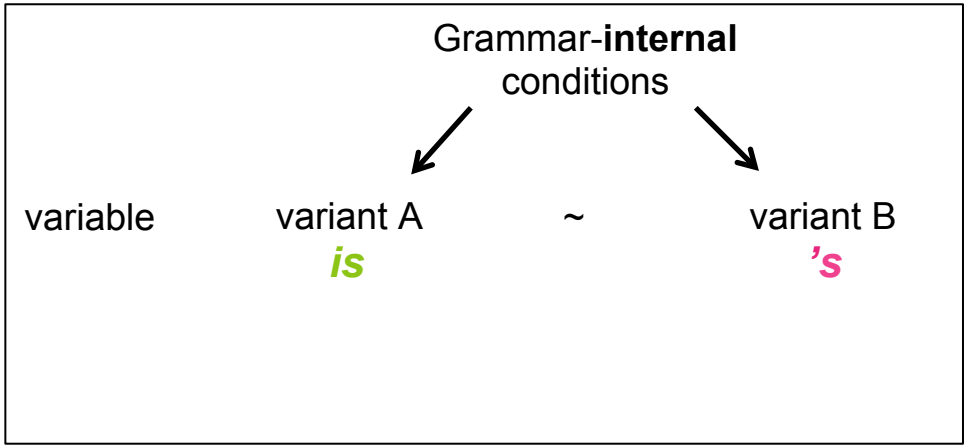
Modeling variation

1 Grammar



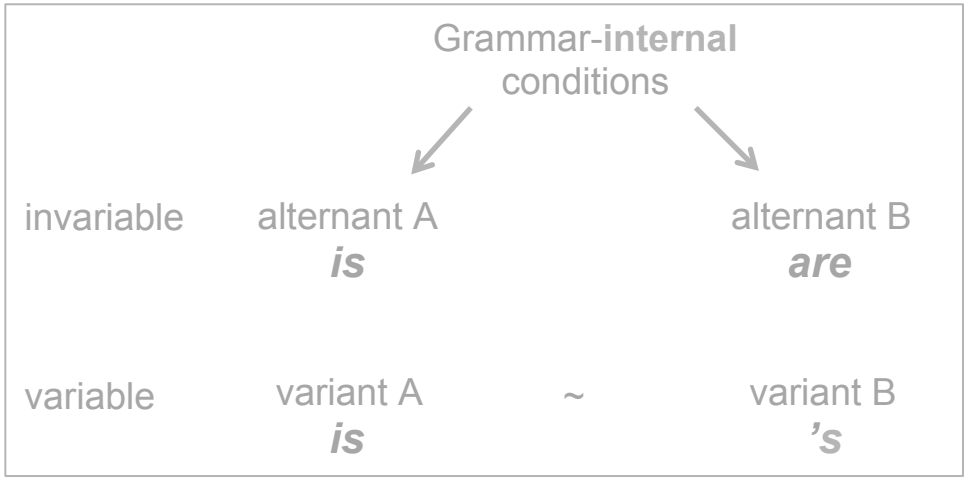
Modeling variation

1 Grammar



Modeling variation

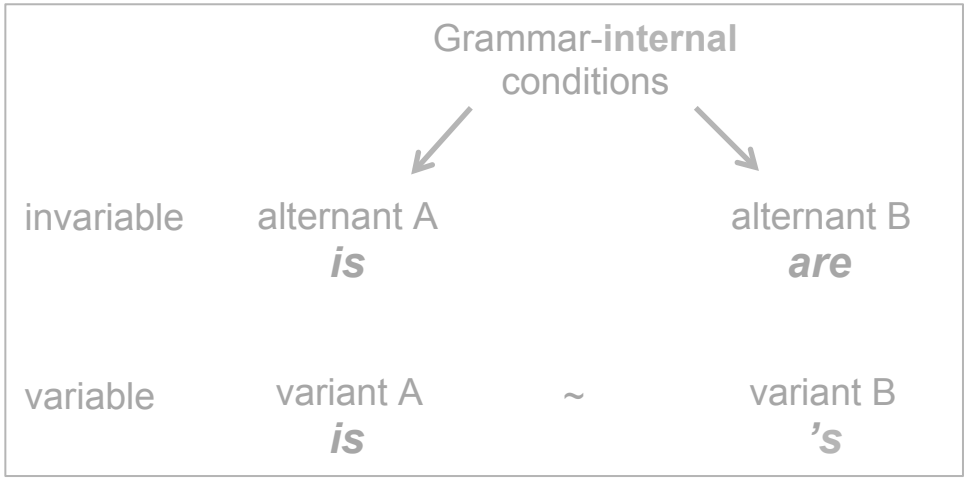
1 **Grammar**



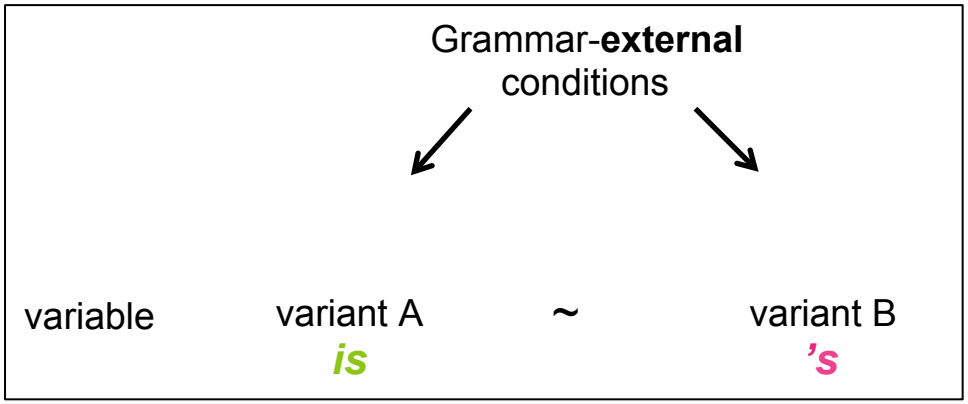
2 **Use**

Modeling variation

1 **Grammar**



2 **Use**



Conclusion

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

Conclusion

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

Grammatical architectures must still be structured to allow variation.

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!

Previous work

Demonstrated for many variables, but...

/s/-deletion, pro-drop:

tokens linked in grammar and discourse

Syntactic persistence:

debate over construct of “variable” in syntax

Variables here (ING, DH) never change referential meaning, each instance independent

Coding

Coding

dependent variable

Coding

dependent variable

contracted

uncontracted

Coding

dependent variable

	contracted	uncontracted
is	[z], [s]	[ɪz], [əz]

(MacKenzie 2012)

Coding

dependent variable

	contracted	uncontracted
is	[z], [s]	[ɪz], [əz]
has	[z], [s]	[hæz], [həz], [əz]

(MacKenzie 2012)

Coding

dependent variable

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

(MacKenzie 2012)