



Carnegie Mellon University
Language
Technologies
Institute

11-324/11-624/11-724 Human Language for AI

Linguistic and Social Variables in Language Variation

David R. Mortensen

November 17, 2021

Language Technologies Institute
Carnegie Mellon University

Learning Objectives I

At the end of this lecture, students will have a basic understanding of, and know multiple examples of, several types of linguistic variation:

- Phonetic variation
- Phonological variation
- Morphological variation
- Lexical variation
- Syntactic variation
- Discourse variation
- Pragmatic variation

Learning Objectives II

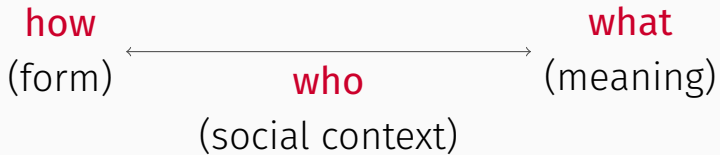
At the end of this lecture, students will understand how multiple social variables interact with linguistic variables in linguistic variation:

- Place
- Power
- Class
- Formality
- Race

- Ethnicity
- Gender
- Age

Students should be able to formulate testable hypotheses about linguistic behavior given social identity.

Students will also understand the major methodologies used in variationist sociolinguistics.



When you talk, read, and write, you are not just communicating *propositional* meaning; you are communicating *social* meaning. One way you do this is through **linguistic variation**.

Phonetic Variation

Difference is how phonemes are realized. Part of “accent.”

In Martha’s Vineyard (an island off the coast of the US state of Massachusetts) many people pronounce the diphthongs /aj/ and /aw/ differently than in Standard American English and the English of mainland Massachusetts.

	Std Am	Martha’s Vineyard
bite	bajt	bəjt
bide	bajd	bəjd
bout	bawt	bəwt
pouch	p ^h awtʃ	p ^h əwtʃ

No splits or mergers.

Phonological Variation

Phonological variation is different from phonetic variation in that the system of contrasts is restructured. Consider the following song:

*Oh, they had to carry Harry to the ferry,
And the ferry carried Harry to the shore;
And the reason that they had to carry Harry to the ferry
Was that Harry couldn't carry any more.*

This only makes sense as a song if *carry*, *Harry*, and *ferry* all rhyme, which they do in certain innovative dialects of North American English and do not in more conservative dialects. Since this involves MERGERS, it is **phonological variation**.

A similar type of merger has happened in the English of Pittsburgh, Pennsylvania, and also in the Great Basin of the western US. In these dialects, /i/ and /ɪ/ have merged to /ɪ/ before /l/ so that *still* and *steel* are pronounced the same.

Morphological Variation

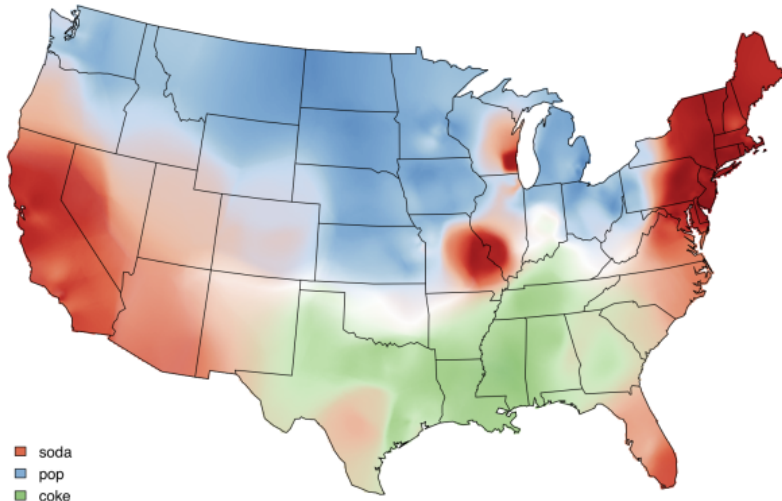
Sociolects may differ in the morphological constructions they include. For example, consider the *eye-fuck* construction. There are many compounds with this structure, but they are largely restricted to a particular social context.

There is also paradigmatic variation, as in English *to see*:

	DIALECT A	DIALECT B
INFINITIVE	see	see
NON-PAST	see	see
PAST	saw	seen
PRESENT PARTICIPLE	seeing	seeing
PAST PARTICIPLE	seen	seen

Lexical Variation

What is your generic term for a sweetened, carbonated beverage?



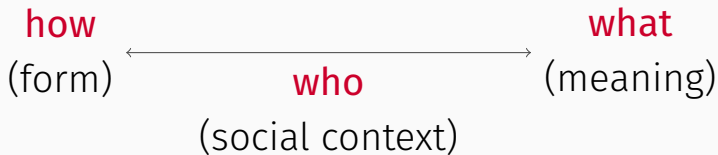
Map by Joshua Katz, Department of Statistics, NC State University
Based on survey data from Bert Vaux, Department of Linguistics, University of Cambridge

- (1) a. Most babies like cuddled.
b. Most babies like to be cuddled.
- (2) a. The grass needs cut.
b. The grass needs cutting.
c. The grass needs to be cut.

There is socially conditioned variation in how discourse is structured. For example, in some social contexts, speakers may structure their discourse as personal narratives or factual assertions. What other examples of discourse variation can you think of?

Consider our discussion of cohesion in different social settings.

- There is great socially-conditioned variation in politeness strategies (e.g. positive politeness versus negative politeness)
 - **positive politeness** make the hearer feel good about themselves
 - **negative politeness** avoid imposition on the hearer
- What about turn-taking in conversation?
- What about other speech acts?

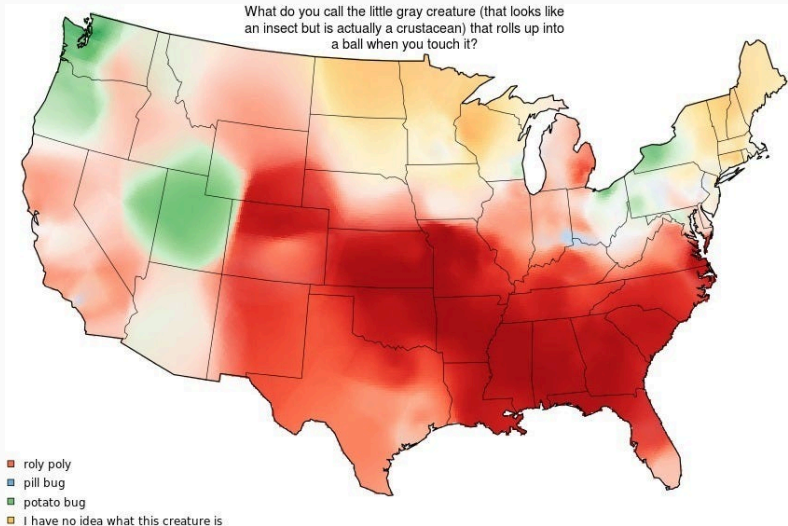


Dahntahn [dantan] says, “I’m from Pittsburgh.”

You talk in ways that express your local or regional identity.

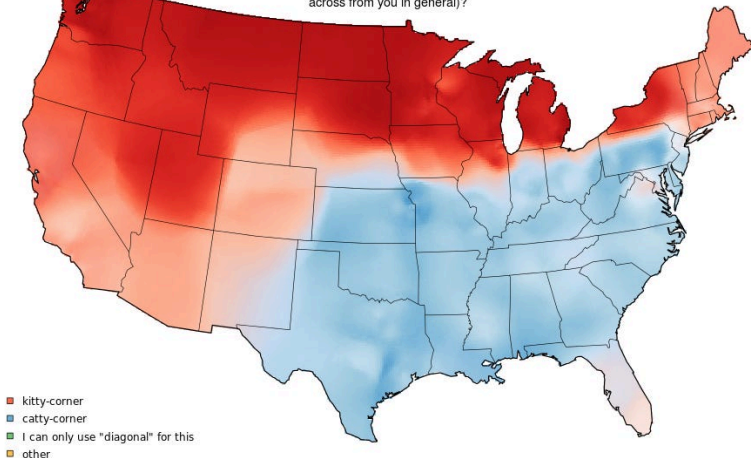
Language and Place

What do you call the little gray creature (that looks like an insect but is actually a crustacean) that rolls up into a ball when you touch it?



Language and Place

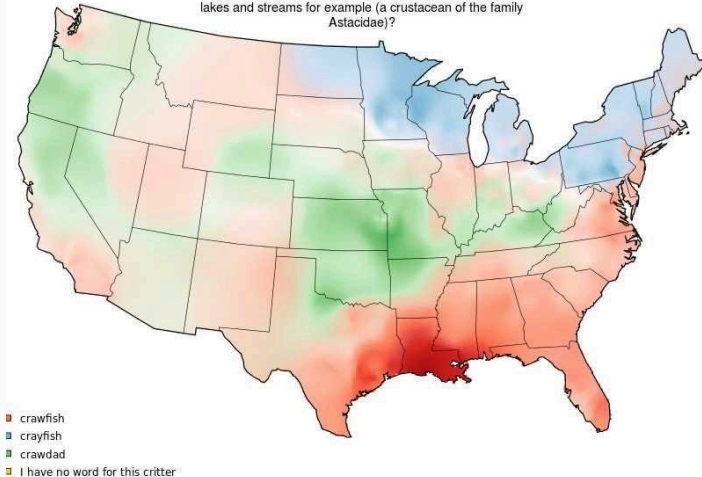
What term do you use to refer to something that is across both streets from you at an intersection (or diagonally across from you in general)?



Joshua Katz, Dept of Statistics, NC State University

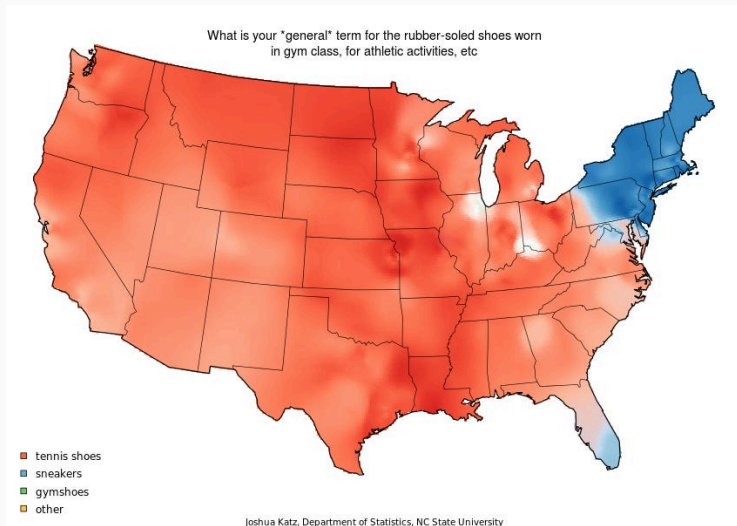
Language and Place

What do you call the miniature lobster that one finds in lakes and streams for example (a crustacean of the family Astacidae)?

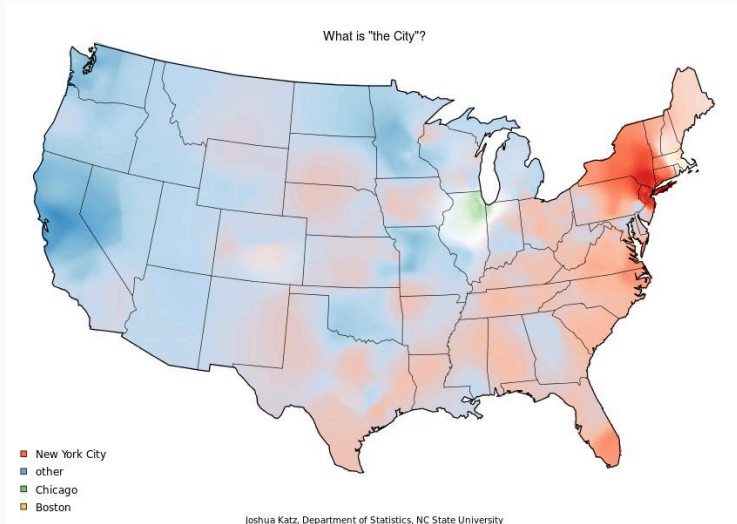


Joshua Katz, Department of Statistics, NC State University

Language and Place



Language and Place



Martha's Vineyard

In William Labov's famous study of /aj/ and /aw/, he found that the variants with a schwa nucleus (like [t^həjm] for *time* or [t^həwn] for *town*) was associated with identification with Martha's Vineyard.

- Identifying with Martha's Vineyard more than Massachusetts
- Living away from the tourist areas
- Being a fisherman, a traditional role of people on Martha's Vineyard
- Being male (see language and gender, below)

Relative power is also reflected in linguistic form. Examples?

- Politeness strategies
- Honorific systems

Closely related to power is class. Language varies according to class identity. Example: British English (as exemplified in *Downton Abbey*, reality, etc.)

- Working class dialects (*harm* and *arm* merge)
- Middle class dialects
- Upper class dialects

- Speech and writing differ according to level of formality
- An email message to a professor might differ from a text message to a friend
 - I do not know
 - I don't know
 - idk
- Familiar pronouns

Language and Race

- While Britain is characterized by class divisions in language (and society), the United States is characterized by racial divisions
- Race and language have been studied extensively in the US
- AAVE and SWE have become increasingly racialized language varieties, even though the dialect features in the southern US were originally more regional than racial
- AAVE/BE differs from Standard American English in its phonology, morphology, syntax, pragmatics, and discourse patterns

- Just as language varies with race, it also varies with ethnic identity.
- Chicano English or Hispanic Vernacular English is an example of this.
- It is not the same as L2 English—many of its speakers are native speakers of English

- There are many stereotypes about how genders differ linguistically
- Many of these are **not** true (for example, that women talk more)
- However, there are real ways in which people communicate their gender identity through linguistic form.
- One example: even when you control for vocal-tract length, **female vowels spaces tend to be more dispersed than male vowel spaces**
- Another example: in North America, **women use more first-person pronouns than men.**

Language and Age

- If there were not linguistic differences between generations, language would not change.
- However, language and age is very difficult to study.
- In most cultures, there is a small, finite number of genders; societies usually define a finite number of racial and ethnic categories.
- **Age is continuous.**
- Furthermore, **chronological age doesn't always line up with generational identity.**

How do we study this?

- Surveys
- Corpus studies
- Experimental studies

Questions?