Universität Stuttgart

IRIS

Language in Social Context: Bridging NLP and Sociolinguistics

Agnieszka Faleńska
Filip Miletić

Day 4: Language Variation in Interaction

Image credit: DALLE (2024)

Recap: Demographic Factors of Language Variation

- **Geographic origin** affects language use due to distance & barriers
- **Socioeconomic status** exhibits systematic patterns (e.g. hypercorrection), but its operationalization is challenging and theory-dependent
- **Gender** is also linked to replicated findings (e.g. women conforming to overt norms), with a very recent shift to interaction-based and identity-inclusive approaches
- **Age** may reflect community-specific patterns and diachronic change

- **Social media data** captures patterns similar to real-life communication, but with variability across domains, topics, and languages
- **Methodological trade-offs** ⇒ need for complementary perspectives

Day 4: Outline

Adapting to interlocutors

Accommodation theory (Giles, 1973; Bourhis et al., 2007)

Audience design theory (Bell, 1984; Hay et al., 1999; Renn & Terry, 2008)

Social meaning of variation

Indexicality (Eckert, 2008; Beaton & Washington, 2015)

Personae (D'Onofrio, 2020; Podesva, 2011)

NLP for linguistic accommodation

Adapting to interlocutors

Background

- So far, **intra-speaker (stylistic) variation** addressed via formality in interview tasks ⇒ attention to speech (Labov, 1972)
- Communication accommodation theory: adjusting to the addressee via convergence vs. divergence (Giles, 1973)
- Audience design theory (Bell, 1984)

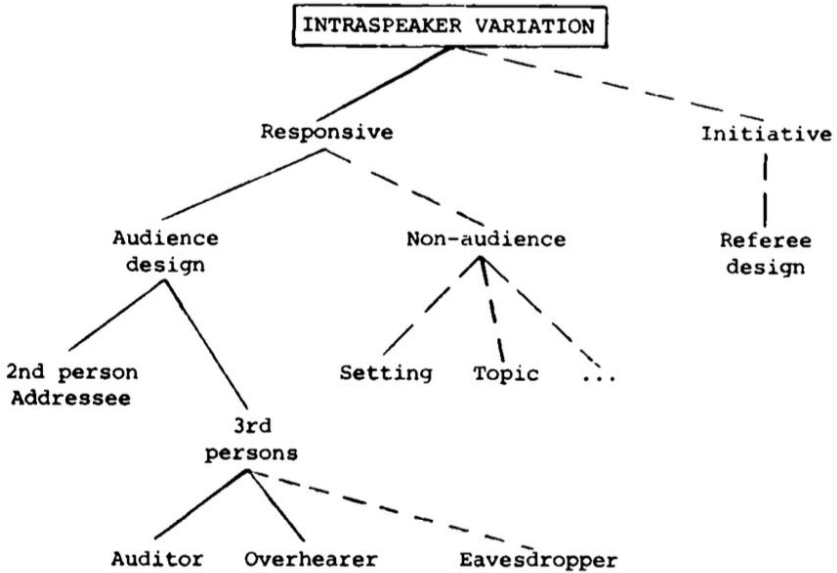


FIGURE 6: Style as audience design.

Figure: Bell (1984: 162)

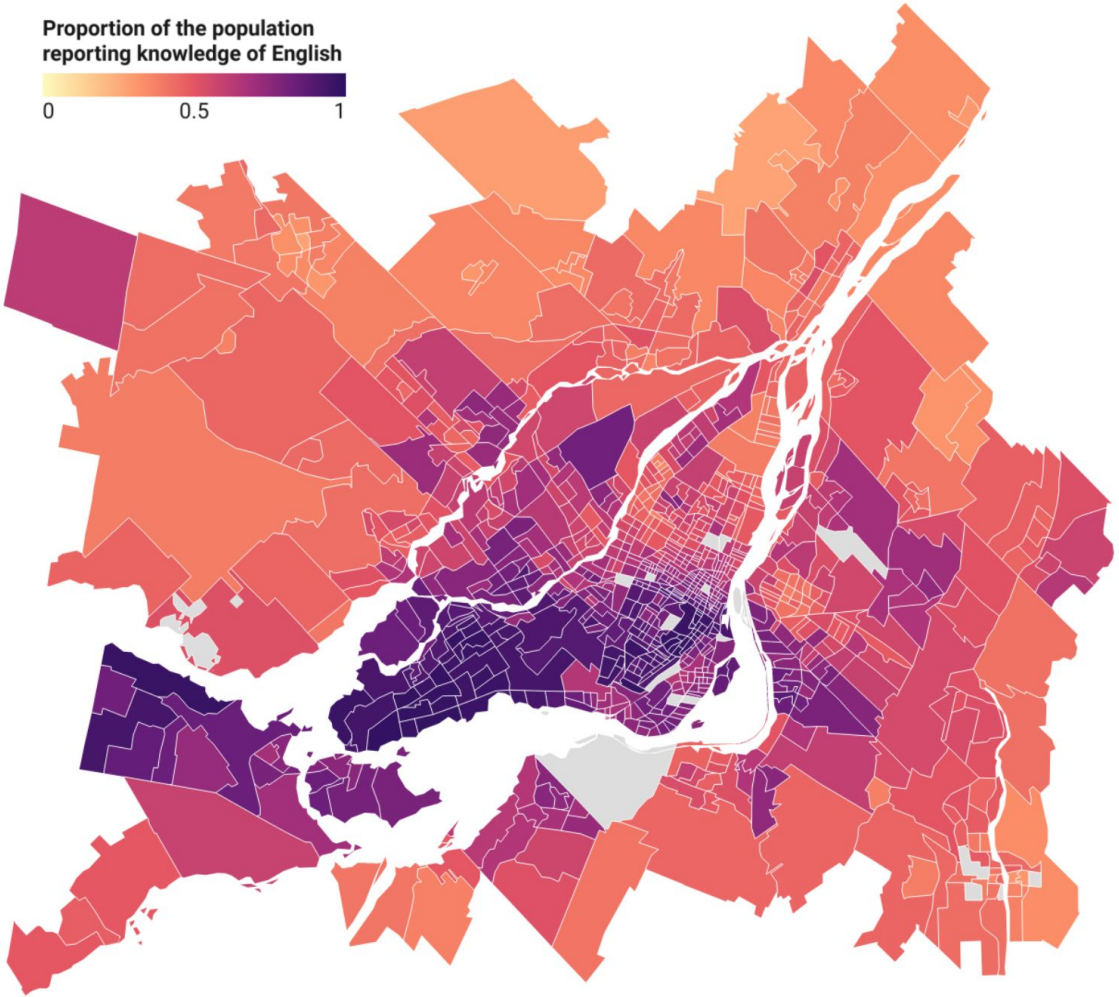
Bourhis et al. (2007): Accommodation in Montreal

- Historically diglossic use of English and French in Quebec
  - French speakers constitute a demographic majority
  - English speakers in political and economic power until early 20th c.
- Laws regulating language use, most notably Bill 101 (1977)
- Four **field studies** assessing **accommodation in language choice**

Bourhis et al. (2007)

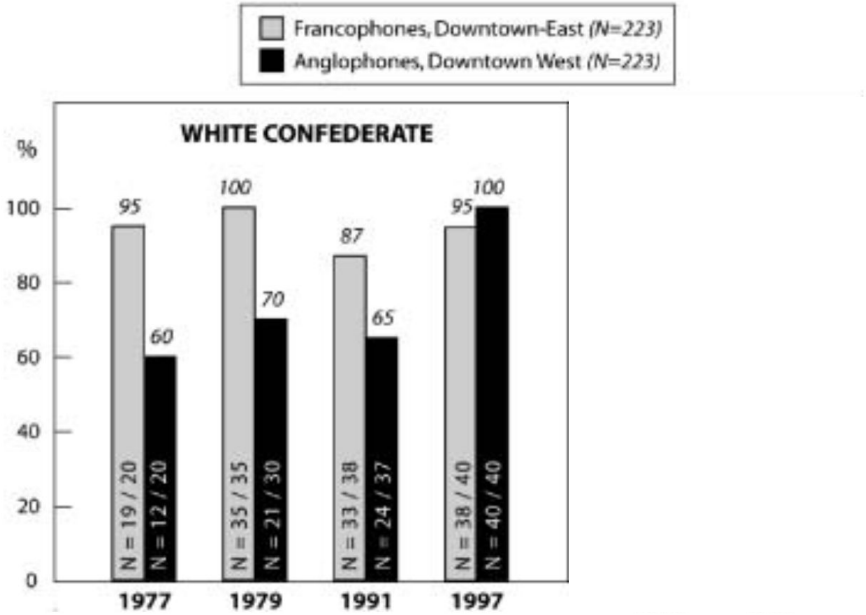
1977 east/west downtown
1979 + McGill
+ Université de Montréal
1991 + Black and White fieldworkers
1997

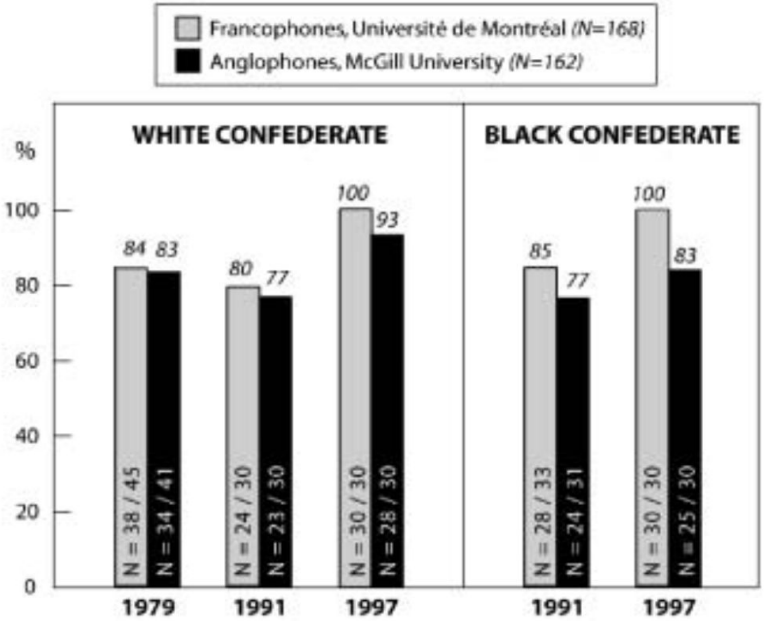
- Pedestrians asked for directions in French or in English
- Dependent variable: language used in response
- Convergence: use of pedestrian's second language



Map source: Miletić (2022: 48) – Map data: Statistics Canada (2017) – Map created with Datawrapper

Bourhis et al. (2007): Accommodation in Montreal





Hay et al. (1999): Oprah and /ay/

- An auditory–acoustic analysis of **a single speaker** to investigate the effect of **referee design** on monophthongization of /ay/
- **Variable:** realization of /ay/, often [a:] in Southern US and AAE
- **Target speaker:** Oprah Winfrey
- **Data:** Segments where Oprah is facing the camera and referring to a third person, usually an upcoming guest
- **Explanatory factors:**
  - Ethnicity of the referee
  - Lexical frequency

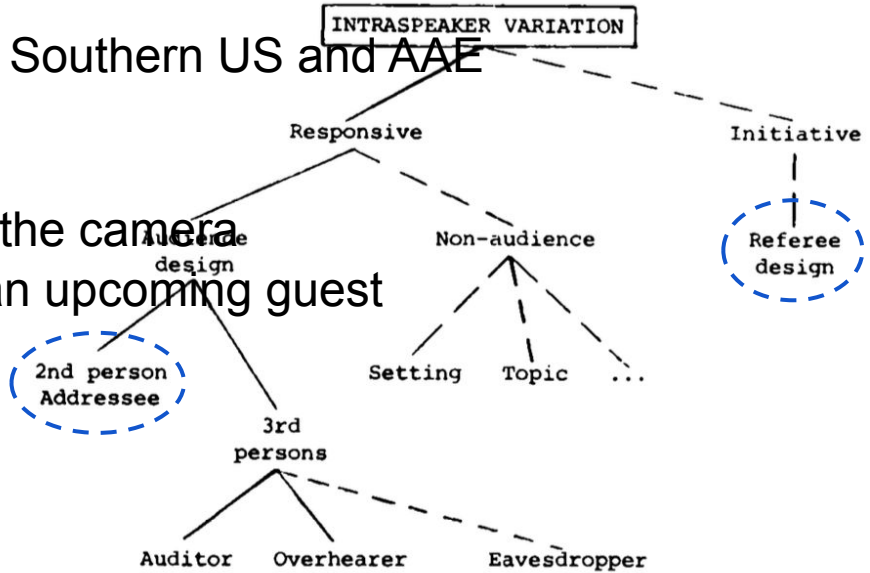


FIGURE 6: Style as audience design.

Figure: Bell (1984: 162)

Hay et al. (1999): Oprah and /ay/

rate of monophthongization

	diphthong	monophthong		total
AA referee	55	33	38%	88
non-AA referee	96	11	10%	107
total	151	44		195

Table 1. Ethnicity of referee and monophthongization

.	diphthong	monophthong		total
Frequent	105	47	31%	152
Not frequent	66	11	14%	77
Total	171	58		229

Table 2: Lexical frequency and monophthongization.

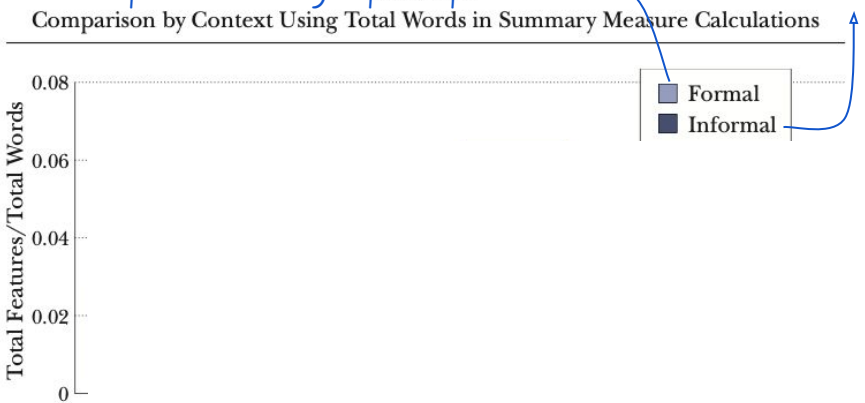
Renn & Terry (2009): Style shift in AAVE adolescents

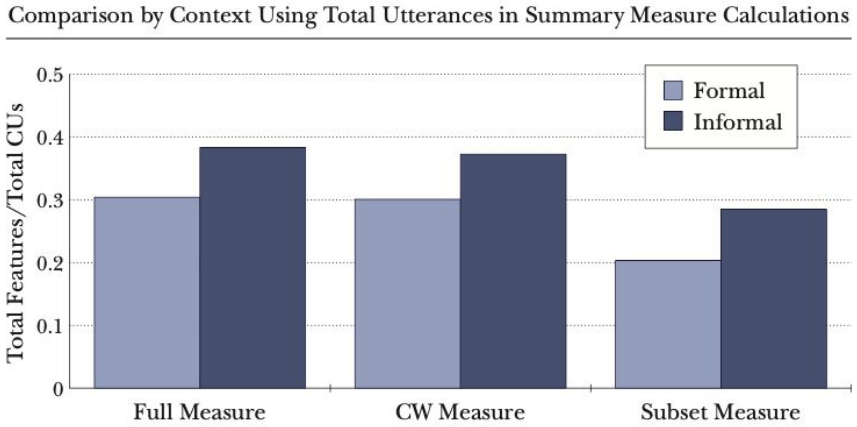
- Aggregate quantitative variables for style shift among AAVE speakers
- Participants: 108 African American boys and girls, 11–13 years old

Third-person singular *-s* absence	*She like him* for *She likes him*
Multiple negation	*They didn't do nothing*
Ain't for *Is not*	*The cars ain't gonna move*

simulated speech to a group of parents

free discussion





Takeaways

- Intraspeaker variation was initially analyzed in terms of attention to speech depending on degrees of formality
- But there are clear patterns of variation which reflect adaptation to a much broader range of factors
  - Interlocutors (⇒ Accommodation Theory)
  - Other aspects of the communicative situation, e.g. referee (⇒ Audience Design)
- This requires a different study design which may not scale easily

Social meaning of variation

Background

Variation and the indexical field (Eckert, 2008)

- Most work interprets differences in the use of sociolinguistic variables as a consequence of a standard demographic category (age, gender etc.)
- But variables index demographic categories indirectly and associatively via **orders of indexicality** (Silverstein, 2003)
- E.g. Southern US accent ⇒ from the south ⇒ "redneck" stereotype (Eckert, 2019)

social meaning

- **Personae:** "social types that are recognizably linked with ways of being and speaking" (D'Onofrio, 2020)

Beaton & Washington (2015): Slurs and the indexical field

- Qualitative **meta-discursive analysis** of a written corpus
  to define the **indexical field** of a reclaimed slur in Brazilian Portuguese
- **Target lexical item:** *favelado* ‘slum dweller’
  - associated with socially imposed connotation
  - ≠ *morador de favela* lit. ‘inhabitant of the slum’ ⇒ only referential
- **Data:**
  - academic and institutional documents
  - online discussions

Beaton & Washington (2015): Slurs and the indexical field

I am *favelado*,
but I am not a drug dealer!

So we have to show
that we are *favelado*
but have good manners.

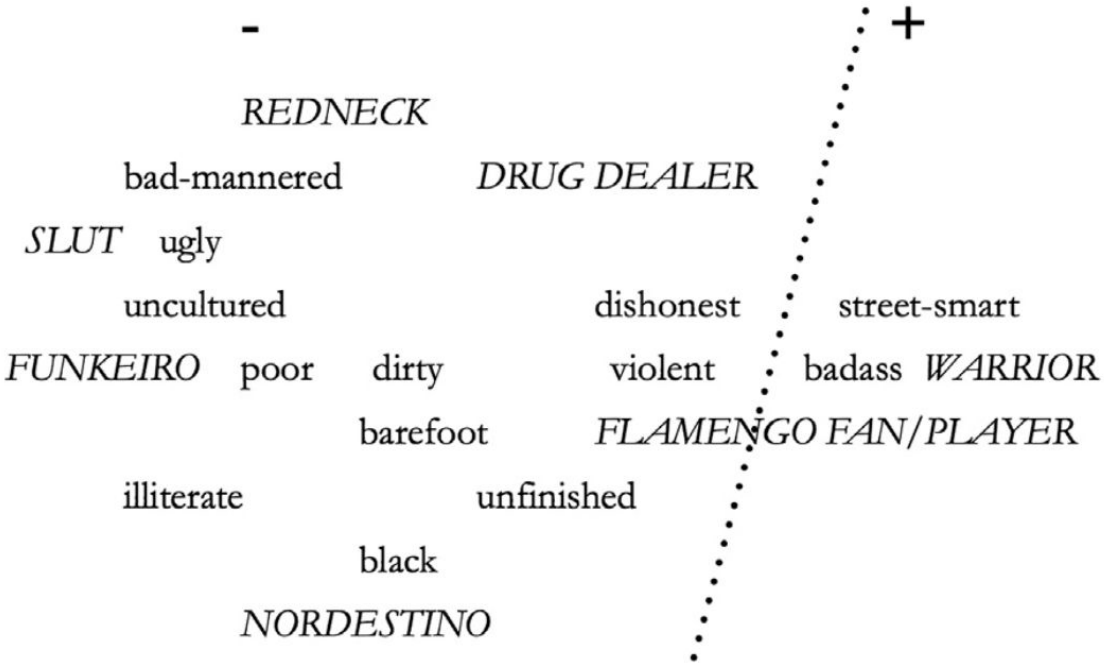
restriction of indexical domain

amelioration

I don't live in the favela… but I cheer for a team CALLED FLAMENGO…
And this team isn't *favelado*… and if it was, then why the prejudice?
And if being FLAMENGO is being FAVELADO, then that's what I am.

To be *favelado* is to be different, it is being a warrior in order to be a student, it is being a warrior to be a salaried worker, it is being a warrior in order to have rights, it is being a warrior in order to be "equal".

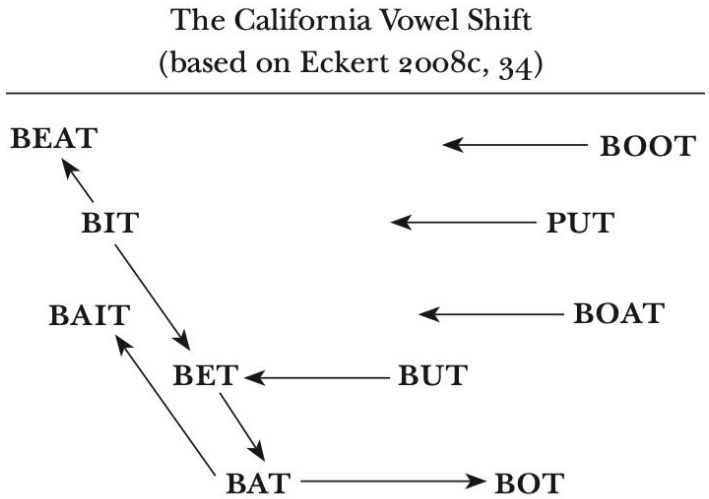
Beaton & Washington (2015): Slurs and the indexical field



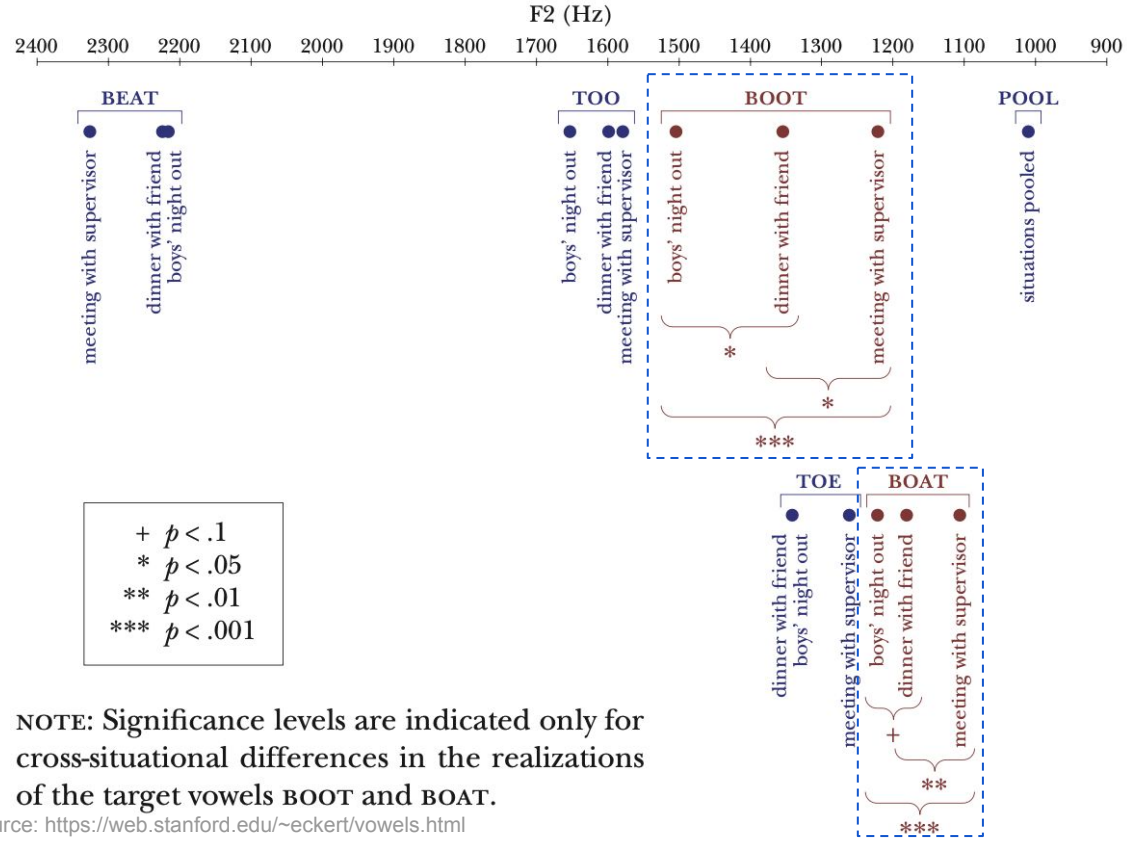
Indexical field for *favelado*. Lower case = permanent qualities; upper case = social types; dotted line separates positive and negative qualities and social types. (Figure 2, p. 16.)

Podesva (2011): California vowel shift & gay identity

- **Acoustic analysis of a single speaker** in different contexts to examine his use of pronunciation in **persona construction**
- **Speaker:** Regan, 34, gay Asian American living in San Francisco
- **Method:** recordings by Regan in 3 contexts
  - Meeting with supervisor
  - Dinner with friend
  - Boys' night out
- **Variable:** California Vowel Shift

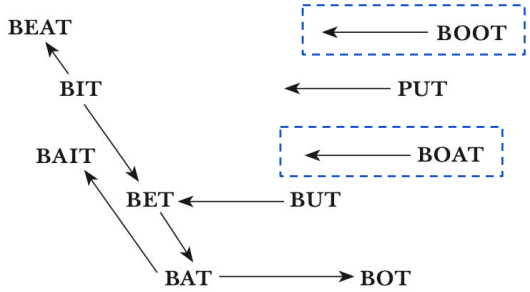


Podesva (2011): California vowel shift & gay identity



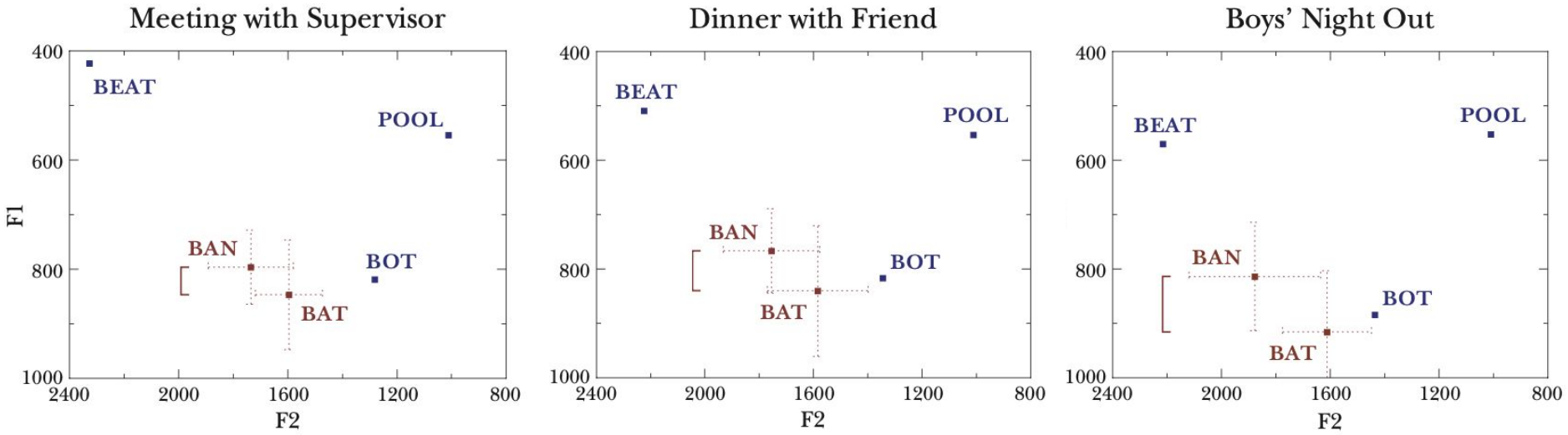
NOTE: Significance levels are indicated only for cross-situational differences in the realizations of the target vowels BOOT and BOAT.

fronting of
BOOT and BOAT

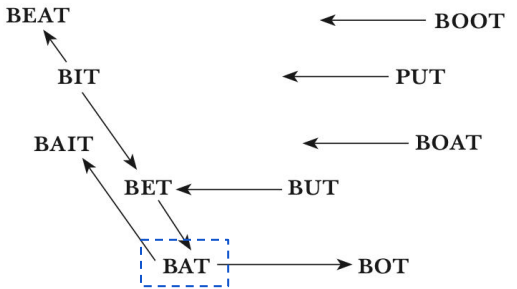


Recording source: https://web.stanford.edu/~eckert/vowels.html

Podesva (2011): California vowel shift & gay identity

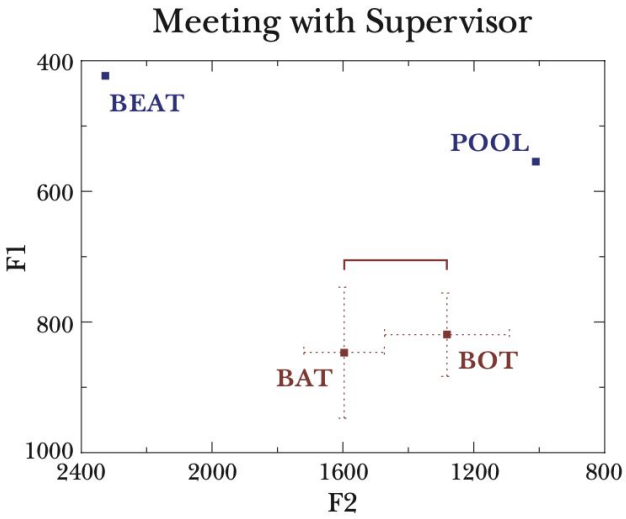


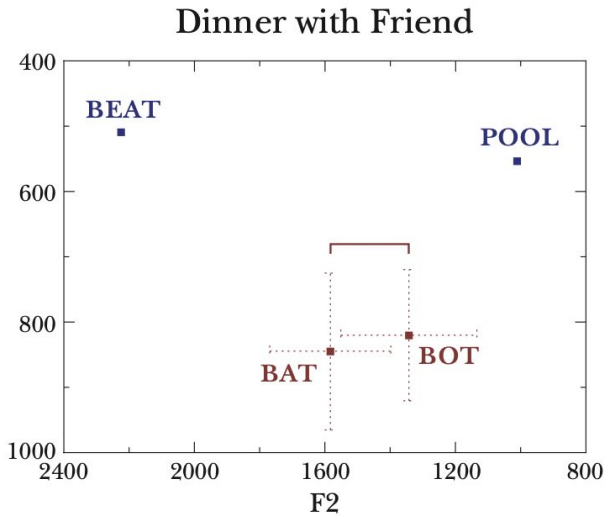
raising of BAN

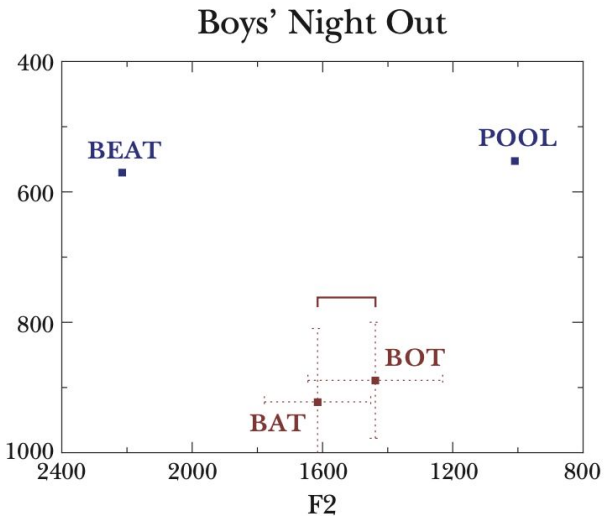


Recording source: https://web.stanford.edu/~eckert/vowels.html

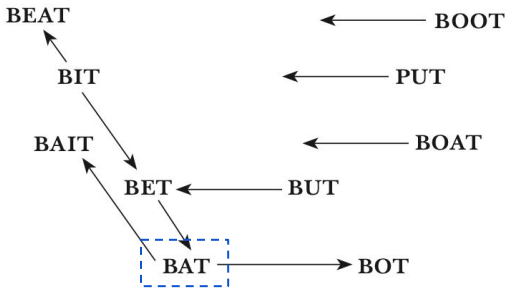
Podesva (2011): California vowel shift & gay identity







backing of BAT



Recording source: https://web.stanford.edu/~eckert/vowels.html

Podesva (2011): California vowel shift & gay identity

Ethnographic and interactional context

- The most fun and funny member of the group
- Focus on gay parties (topically and linguistically) ⇒ “gay partier” persona

California character types

- Character types linked to CA speech:
  valley girls, surfers, stoners, slackers
- Evocation of traits like “fun”, “laid-back”, “carefree”

Takeaways

- A sociolinguistic variable may stratify with multiple sociodemographic categories, so social meaning relates to what is common to those categories ⇒ style
- Analytical strategies: indexical fields, persona construction

"Because these macrosocial categories [such as class, age, and gender] are fundamental to the social order, they correlate regularly with linguistic variation. This is not because the categories themselves engage directly with linguistic practice, but because their intersections structure the conditions and everyday experiences of life on the ground, and variation takes on meaning in the local social practice that unfolds in response to these conditions." (Eckert, 2019: 751)

Photo credit:
https://web.stanford.edu/~eckert/

NLP for linguistic accommodation

So far…

- Small discussion groups and conversation pairs
- In-depth qualitative interpretation ⇒ informative, but how do we generalize?
- Let’s analyze language socialization through big data…
  - How to **measure** linguistic adaptation?
  - How to **operationalize** “an audience”?
  - What **triggers** language adaptation in social media?

How to measure linguistic adaptation?

Style accommodation

quantifier

A: At **least** you were outside.
B: It doesn't make **much** difference where you are [...]

⇢It doesn’t really matter where you are.

preposition coordination

A: **At** what time does your shop close?
B: **At** five o’clock

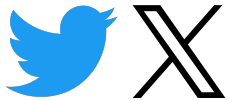
⇢ five o’clock

Source: Danescu-Niculescu-Mizil et al. (2011)

Linguistic adaptation in social media

Danescu-Niculescu-Mizil et al. (2011): “Mark My Words! Linguistic Style Accommodation in Social Media”

- 25% of users hold conversations on Twitter (Java et al. 2007)
- 37% of tweets is conversational (Ritter et al. 2010)
- Naturally occurring conversations
  - not face-to-face
  - not in real life
  - restriction to 140 characters
  - wide spectrum of relations between people



Source: Danescu-Niculescu-Mizil et al. (2011)

Linguistic adaptation on Twitter

- Data: 215,000 conversations (replies) between 2,200 pairs of users
- Statistics for a pair:

	Mean	Median	Min	Max
Number of conversations	98	60	1	1744
Average number of turns	2.7	2.6	2	16.8
Days of contact	270	257	1	886

- **Instant** accommodation:
  - occurs in one conversational turn to another
- **Style** accomodation

Source: Danescu-Niculescu-Mizil et al. (2011)

How to **measure style** on Twitter?

LIWC Linguistic Inquiry Word Count (Pennebaker et al. 2007)

- 60 psychologically meaningful categories and related dictionaries of words
- used in various psychologically-motivated applications, such as identifying social relations or analysing mental health

Focus on 14 strictly non-topical dimensions

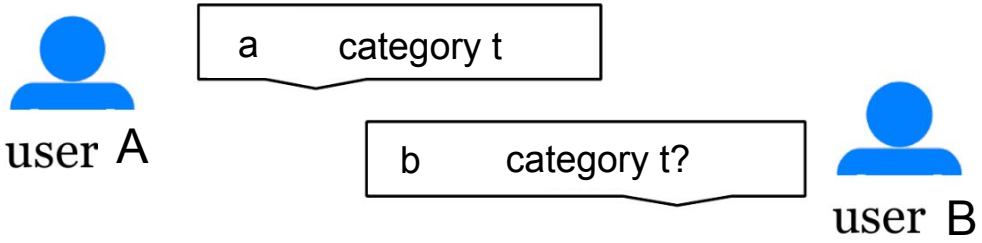
- A tweet exhibits a stylistic dimension if it contains at least one word from the LIWC vocabulary
- A tweet can exhibit multiple dimensions

Dimension	Examples	Size
Article	an, the	3
Certainty	always, never	83
Conjunction	but, whereas	28
Discrepancy	should, would	76
Exclusive	without, exclude	17
Inclusive	with, include	18
Indefinite pronoun	it, those	46
Negation	not, never	57
Preposition	to, with	60
Quantifier	few, much	89
Tentative	maybe, perhaps	155
1st person singular pronoun	I, me	12
1st person plural pronoun	we, us	12
2nd person pronoun	you, your	20

Source: Danescu-Niculescu-Mizil et al. (2011)

How to **measure style accommodation** on Twitter?

Probabilistic framework to measure convergence

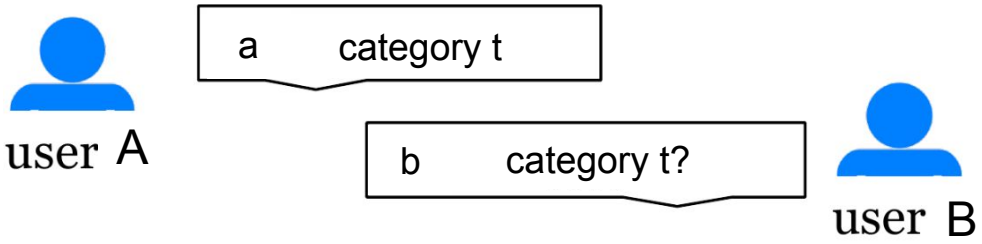


- B's reply to a: $b_{\hookrightarrow a}$
- Tweets with category t: a^t $b^t_{\hookrightarrow a}$

Source: Danescu-Niculescu-Mizil et al. (2011), images from slides by Dong Nguyen, 2017

How to **measure style accommodation** on Twitter?

Probabilistic framework to measure convergence



- Convergence of B to a:

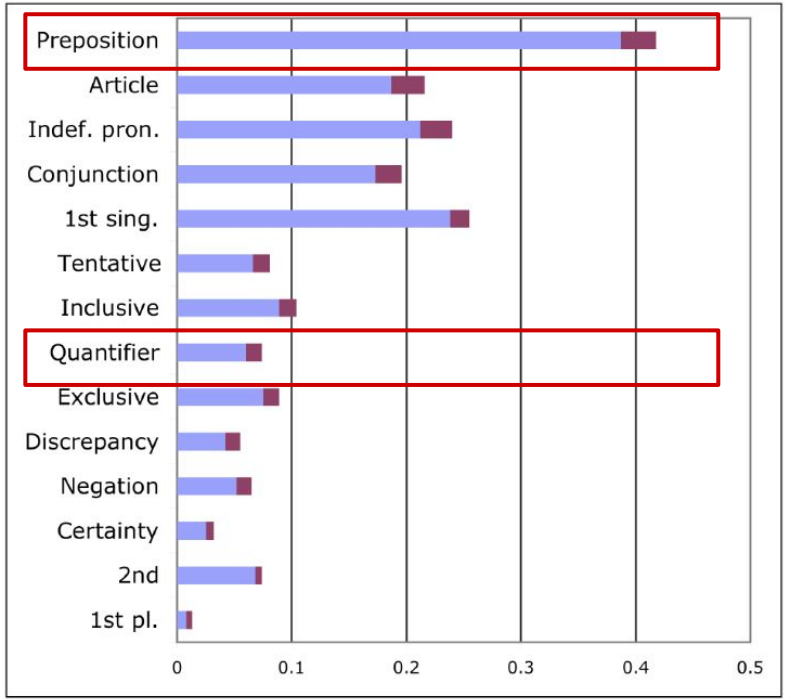
$$Conv_{A,B}(t) \overset{def}{=} P(b^t_{\hookrightarrow a} = 1 | a^t = 1) - P(b^t_{\hookrightarrow a} = 1)$$

- Overall convergence of t:

$$Conv(t) \overset{def}{=} E_{\text{pairs}(A,B)}(Conv_{A,B}(t))$$

Source: Source: Danescu-Niculescu-Mizil et al. (2011), images from slides by Dong Nguyen, 2017

Is there style accommodation on Twitter?



- blue $P(b^t_{\hookrightarrow a} = 1)$
- red $P(b^t_{\hookrightarrow a} = 1 | a^t = 1)$

Significant differences for almost all categories

Source: Danescu-Niculescu-Mizil et al. (2011)

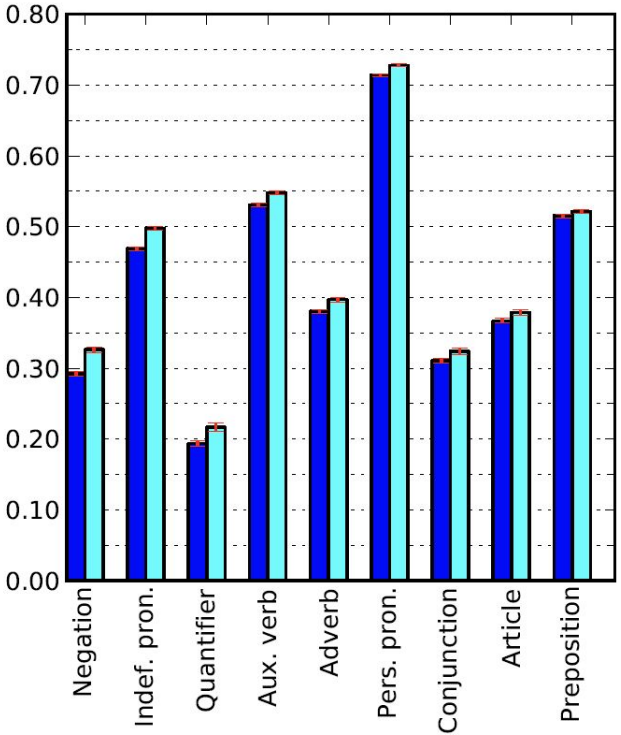
Is style accommodation a reflex?

Danescu-Niculescu-Mizil et al. (2011): “Chameleons in imagined conversations: A new approach to understanding coordination of linguistic style in dialogs”

- Is the convergence a **social strategy** or an ingrained **reflex** in humans?
- Study fictional movie dialogs → social benefits irrelevant to the writer
- Data:
  - movie scripts crawled from various web pages
  - character information from IMBD

Source: Danescu-Niculescu-Mizil et al. (2011), movie script icon from flaticon

Style accommodation in movie dialogs

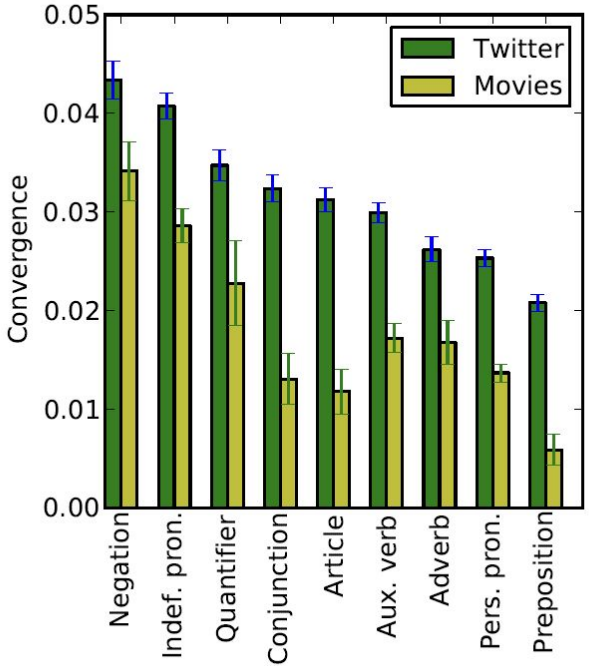


- dark $P(b^t_{\hookrightarrow a} = 1)$
- light $P(b^t_{\hookrightarrow a} = 1 | a^t = 1)$
- red error bars

Significant differences for almost all categories

Source: Danescu-Niculescu-Mizil et al. (2011)

Style accommodation in movie dialogs vs. Twitter



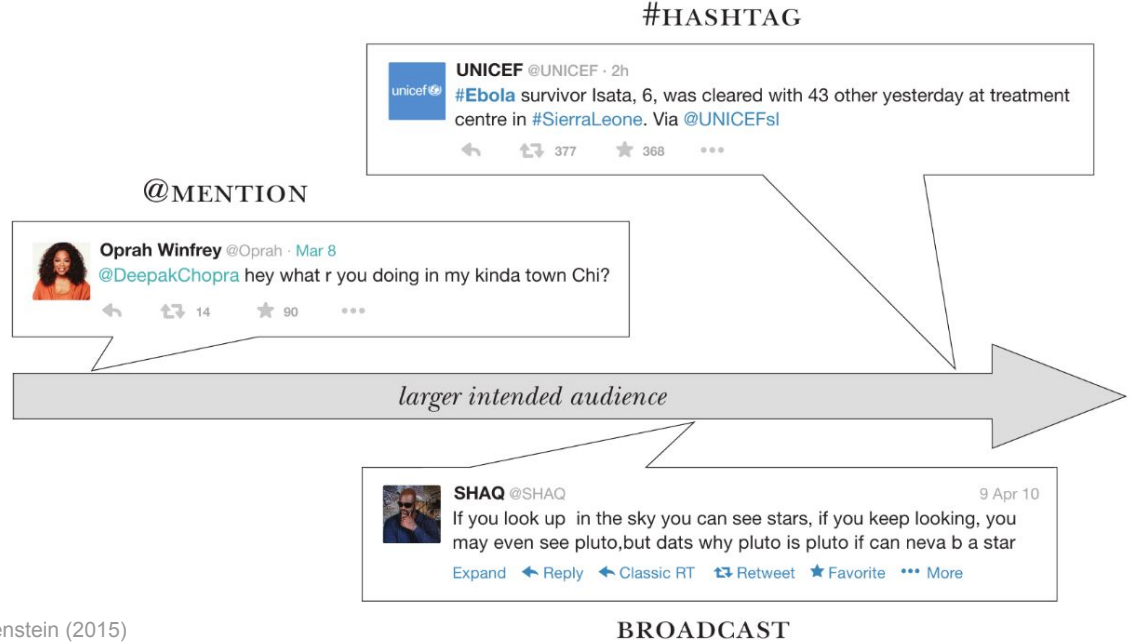
- Stronger convergence on Twitter
- General correspondence between feature categories
- Movie characters exhibit convergence in their linguistic styles

Source: Danescu-Niculescu-Mizil et al. (2011)

How to operationalize “an audience” in social media?

Imagined audience in social media

Marwick and boyd (2011) “(...) users, context collapse, and the imagined audience”



Source: Pavalanathan and Eisenstein (2015)

Adaptation to the imagined audience (Pavalanathan and Eisenstein, 2015)

Data: geotagged tweets from United States grouped according to MSAs

- Geographical variables
  - lexical items strongly associated with specific geographical regions
  - SAGE to find top 30 specific words for each MSA
  - 120 words, examples:

New York	lml, deadass, od, odee, werd, cud, nuttin, nicee, sed, lata, buggin, wrd, noe, wl, layin, okk, lols, lolrt, crazyy, sour, wid
Los Angeles	fasho, ahah, cuh, koo, cuhz, fkn, ahahah, ;o
Chicago	mfs, goofy, nbs, lbvs, bogus, 2ma, lbs, mf, ikr, lmmfao, hoop, crackin
Dallas	ion, nun, oomf, tf, (;, finna, dang, fa, (:, <<, >>, <–, .!, trippin, y'all

Source: Pavalanathan and Eisenstein (2015)

Adaptation to the imagined audience (Pavalanathan and Eisenstein, 2015)

Data: geotagged tweets from United States grouped according to MSAs

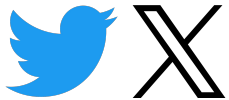
- “Tweetspeak” variables
  - Non-standard words frequently used on Twitter
  - Find 1,000 and filter out words from dictionary, entities, punctuation, special symbols, numbers, hashtags, and non-English words
  - 94 nonstandard terms, examples:

lol	:)	dat	lmaoo	ohh	dm	luv
im	jus	dnt	lmfao	wats	<3	pics
lmao	ppl	idk	hmm	ahh	didnt	ii
ya	lil	aww	w/	nah	naw	comin

Source: Pavalanathan and Eisenstein (2015)

Audience-based factors statistically associated with the use of the non-standard variables

- **Method**: logistic regression
- **Dependent variable**: was a specific linguistic variable used in a tweet?
- **Independent variables**: predictors characterizing the intended audience
  - #-init: Does the message begin with a hashtag?
    #-internal: Does the message contain a hashtag?
  - @-init: Does the message begin with a username mention?
    @-internal: Does the message contain a username mention?



Source: Pavalanathan and Eisenstein (2015)

Audience vs. **geographical** lexical variables

- Model tests nonstandard variables against broadcast messages
  - positive coefficient indicates greater tendency toward nonstandard variables
  - negative coefficient indicates an inhibition of nonstandard variables

Predictor	*Weight*	*Coefficient*	*95% Confidence Interval*	*Empirical Percentage*	*N*
Limited audience					
@-INIT	0.5701	0.2821**	[0.264, 0.300]	51.85	96,954
@-INTERNAL	0.5827	0.3340**	[0.299, 0.369]	56.41	15,494
Wide audience					
#-INIT	0.4004	–0.4037**	[–0.453, –0.355]	35.86	7,980
#-INTERNAL	0.4891	–0.0437**	[–0.076, –0.011]	50.40	16,937
Range	*18*				
TOTAL				50.00	224,000

*$p < .05$; **$p < .01$.

Source: Pavalanathan and Eisenstein (2015)

Audience vs. **tweetspeak** lexical variables

- Model tests nonstandard variables against broadcast messages
  - positive coefficient indicates greater tendency toward nonstandard variables
  - negative coefficient indicates an inhibition of nonstandard variables

Predictor	*Weight*	*Coefficient*	*95% Confidence Interval*	*Empirical Percentage*	*N*
Limited audience					
@-INIT	0.5997	0.4042**	[0.384, 0.425]	53.12	78,047
@-INTERNAL	0.5826	0.3333**	[0.294, 0.373]	54.12	12,076
Wide audience					
#-INIT	0.4079	–0.3728**	[–0.423, –0.323]	34.72	8,062
#-INTERNAL	0.4814	–0.0743**	[–0.108, –0.041]	49.10	16,472
Range	*19*				
TOTAL				50.00	188,000

*$p < .05$; **$p < .01$.

Source: Pavalanathan and Eisenstein (2015)

Adaptation to the imagined audience

Broader audience means…

- less non-standard variants (Pavalanathan and Eisenstein, 2015)
- less minority languages (Nguyen et al. 2015)

Source: slides by Dong Nguyen, 2017

Adaptation to the imagined audience

Broader audience means…

- less non-standard variants (Pavalanathan and Eisenstein, 2015)

- less minority languages (Nguyen et al. 2015)
  - users from Limburg use Limburgish (minority language) in
    33.8% of their tweets with a user mention
    28.6% of their tweets with a hashtag

- less identity markers (Shoemark et al. 2017)
  - when discussing the 2014 Scottish Independence Referendum users used
    more Scottish variants → expressing stronger Scottish linguistic identity
    less Scottish variants when addressing broader audience

What triggers language adaptation in social media?

Nguyen and Rosé (2011): “Language Use As a Reflection of **Socialization in Online Communities**”

- Data: breast cancer forum http://community.breastcancer.org/
- Users ask medical questions and/or look for social involvement

I am also new to the forum, but not new to bc, diagnosed last yr, [..] My follow-up with surgeon for reports is not until 8/9 over a week later. My husband too is so wonderful, only married a yr in May, 1 month before bc diagnosed, I could not get through this if it weren’t for him, never misses an appointment, [...] I wish everyone well. We will all survive.

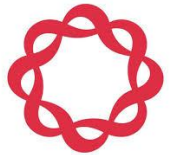
same user
4 years later

Oh Kim- sorry you have so much going on - and an idiot DH on top of it all. [..] Steph- vent away - that sucks - [..] XOXOXOXOXOXOXOX [..], quiet week- end kids went to DD’s & SIL on Friday evening, they take them to school [..], made an AM pop in as I am supposed to, SIL is an idiot but then you all know that.

Source: Nguyen and Rosé (2011)

Specific community norms and jargon

- Forum-specific jargon
  - terms cancer related: *x* (diagnosis), *rads* (radiation, radiotherapy).
  - forum-specific abbreviations: *dh* (dear husband), *dd* (dear daughter)
- Highly informal style
- High emotional involvement in the discussions
- Users frequently refer to other members by name or even as *sisters*



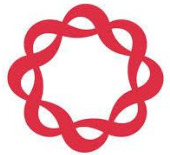
Source: Nguyen and Rosé (2011)

How to find out **when** users adapt to the norms?

Method:

- aggregate posts of all users by week (first week after registering, second…)
- look only at long-term participants
- filter high frequency words and compare distribution with KL divergence

$$KL(P,Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)}$$

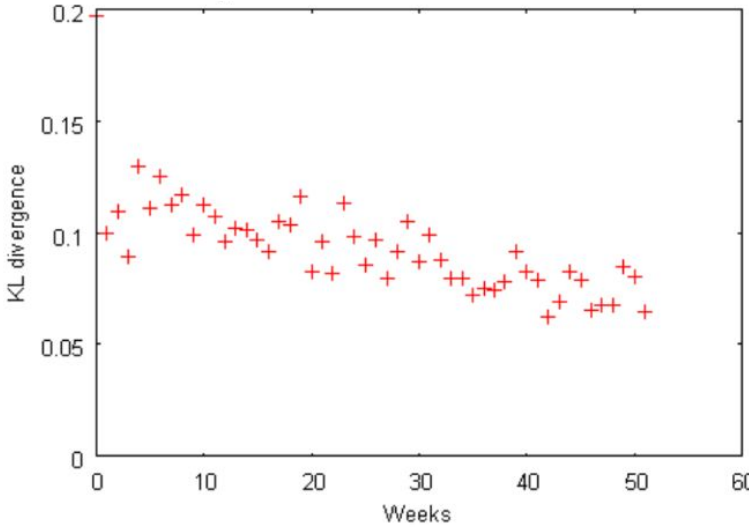


Source: Nguyen and Rosé (2011)

When do users adapt?

Compare the language of each week during the first year after registration with language **in the whole second year**

→ as time progresses, the word distributions look **more like** the distribution of the second year

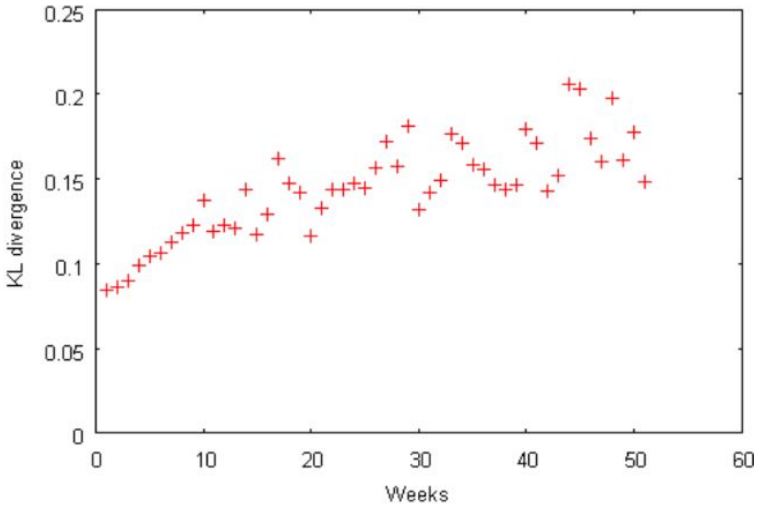


Source: Nguyen and Rosé (2011)

When do users adapt?

Compare the language of each week during the first year after registration with language from **the very first week**

→ as time progresses, the word distributions become **less similar** with the first week



Source: Nguyen and Rosé (2011)

Which linguistic features are being adapted?

- Text representations:
  - unigrams + bigrams for words
  - POS tags
  - LIWC categories
  - (...)

Would you know how to approach this?

- Modeling: linear regression + L1 regularization

Type	Short time members	Long time members
Abbreviations	Husband	My DD (Dear Daughter), Your PS (Plastic Surgeon)
Social networks		Facebook, fb
Greetings	Hi all	Hi girls, Hi gals
I versus other	LIWC-I, My, Me	LIWC-other, We, Sisters
Social support		Hugs, Condolences, So sorry
Thanking	Thanks, Thanx, Thx	
Forum		Bc org, On bco

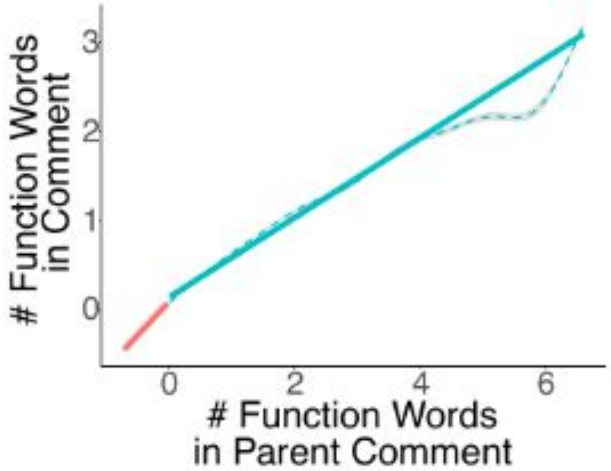
Source: Nguyen and Rosé (2011)

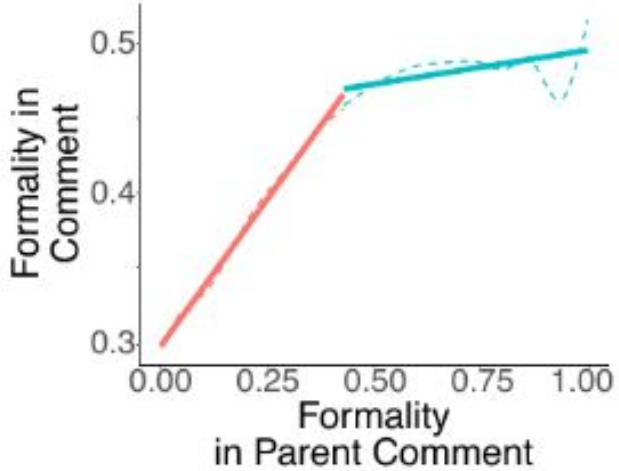
What triggers language adaptation in social media?

- Adaptation of community norms (Nguyen and Rosé, 2011)
  - less social distance
  - greater likability
- Fear from exclusion (Ananthasubramaniam et al., 2023)

Ananthasubramaniam et al., (2023): “Exploring Linguistic Style Matching in Online Communities”

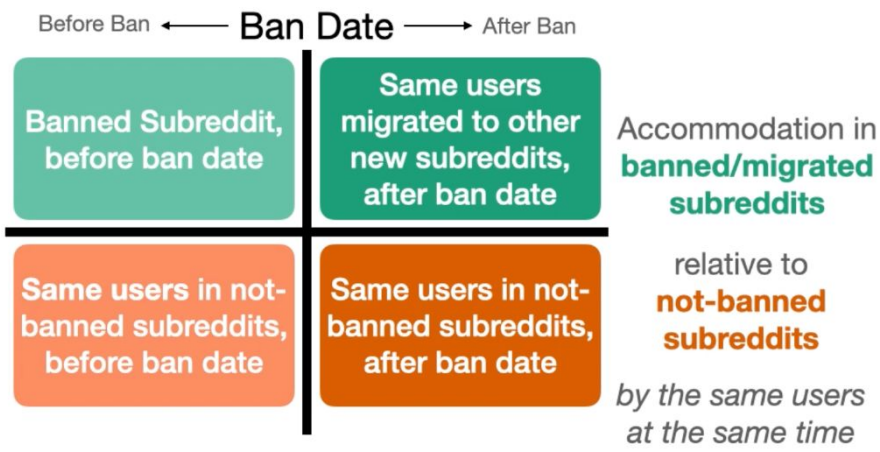
- Data: 17M Reddit conversations

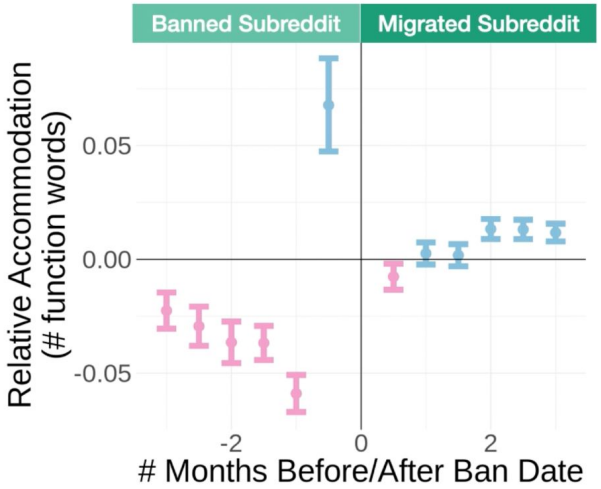




Source: Ananthasubramaniam et al., (2023), slides

Ananthasubramaniam et al., (2023): “Exploring Linguistic Style Matching in Online Communities”





Source: Ananthasubramaniam et al., (2023), slides

What triggers language adaptation in social media?

- Adaptation of community norms (Nguyen and Rosé, 2011)
- Fear from exclusion (Ananthasubramaniam et al., 2023)
- Power status (Danescu-Niculescu-Mizil et al., 2012)
  - discussions among Wikipedia editors
  - people coordinate more with interlocutors who have higher power
- Prestige (Hemphill and Otterbacher, 2012)
  - men tend to receive higher prestige (votes)
  - over time, women decreased their use of gendered stylistic characteristics (e.g., hedging)
- Level of involvement (Danescu-Niculescu-Mizil et al. 2013)
  - learning vs. conservative phase and leaving the community

Language, interaction, and social communities

- How to **measure** linguistic adaptation?
  - Example: instant, style accommodation
- How to **operationalize** “an audience”?
  - Imagined audience – hashtags and mentions
- What **triggers** language adaptation in social media?
  - Socialization in online communities

Practical exercise

Day 4: Style accommodation

- Today's goal: analyze style adaptation in an online forum
  - Open **Style Accommodation** notebook
  - Follow all the steps and fill in the missing pieces of code (marked with TODO)

  - Which features accommodate the strongest?

Takehomes

Recap: Language Variation in Interaction

- Language variation is influenced by & takes on meaning in a broader social context of interaction
- In sociolinguistics, stylistic variation has been explained as…
  - an effect of formality via attention to speech
  - an effect of the broader communicative situation
  - indexical positioning conveying social meaning
- Tricky to operationalize in social media, but possible through various proxy methods

Tomorrow…

Ethical issues

Standard practices in sociolinguistics (Milroy & Gordon, 2003; Childs et al., 2011)

Social media data (Williams et al., 2017; Fiesler & Proferes, 2018)

Development of NLP systems (Bender et al., 2021)

Harms & biases

Biases in representations (Knupleš et al., 2024)

Downstream biases (Lalor et al., 2022)

NLP applications

Personalization (Veronica Lynn et al., 2017; Rabinovich et al. 2017)

One size fits all? (Lucy et al. 2024)

Universität Stuttgart

IRIS

Thank you for your attention!

Image credit: DALLE (2024)