

Beyond Individuals: Language in Social Context

CS 6120

Natural Language Processing

Si Wu

Logistics

- Sign up for project presentation
 - You can choose a time from either session
- Coding assignment due tonight
- Clarification on temperature scaling: (next slide)
- David will be the instructor for both sessions' project presentation since I will be traveling to a conference that week.
 - So basically 11/25 will be my last lecture yall 😞
- The final lecture on 11/25 will be on topics that you want us to cover. If you have any requests, make sure you comment on the post on Ed.

Temperature scaling

$$P_i = \frac{e^{z_i/T}}{\sum_j e^{z_j/T}}$$

- $T > 1$, will make the distribution flatter
- $T < 1$, will make the distribution spiky
- Typo is fixed on the slides
- Quiz questions on this will be a free one
 - Total points for the quiz is still 10, but everyone is “correct” for this one question
- Make sure you don't get it wrong for the coding assignment that's due tonight

Language in social context

Language was created for communication

- There are many theories about the origin of language, but no one knows for sure where it began.
- In different parts of the world, humans independently developed their own languages
 - Universal needs for communication
 - We need signs and symbols for communication
- We use language to
 - Speak to people → speech
 - Write to people → written text

How is language being used in social situations

- We interact with each other using language through many different **modalities**:
 - Written language and speech
 - Mixing language with visual elements (images, film, gestures), or music
- We have many different **platforms**:
 - Texting, email → dialogue
 - Social media: reddit, twitter, tiktok, Instagram, etc. → online community
 - Books, blog post, podcast → narrative
 - Film, theater → multi-media art, performance

Human language is diverse

- Because we, humans, are **diverse**
 - Different **languages**: English, Chinese, Spanish, Japanese, etc.
 - **Accents** and **dialects** within a language: British English vs American,
 - Different **sociolect** within a language: gen Z vs gen Alpha slangs
 - And we can **code-switch** between languages

Language and community

- Based on our similarities, we form many communities.
- There are extensive **language exchanges between communities**:
 - On the **online platform** (where most NLP work focuses)
 - Through published books/papers/essays, newspapers
 - When people physically travel and immigrate from one country to another
- We will focus on a few themes common in language exchange.

A few themes to go over today:

- **Persuasion** in social media
- **Framing** in mass media
- How does **community** form and communicate
- **Dialect** and bias
- Conversation and **dialogue**
- **Identity**: gender, racial, **cultural**, etc.

Some are based on social science theories; others are observations using computational analysis.

For every topic, I will show a fun example computational social science papers.



r/changemyview

And so many other interesting topics but I am only going to talk about the works that I am personally most interested in and familiar with.

And in general...

- There are two types of work under these themes:
 - Analysis on the language in these social situations
 - Come up with proposed methods or demonstration of how language technology can assist/mediate these situation.
- Both are interesting, but I will focus on the first type of work in this lecture

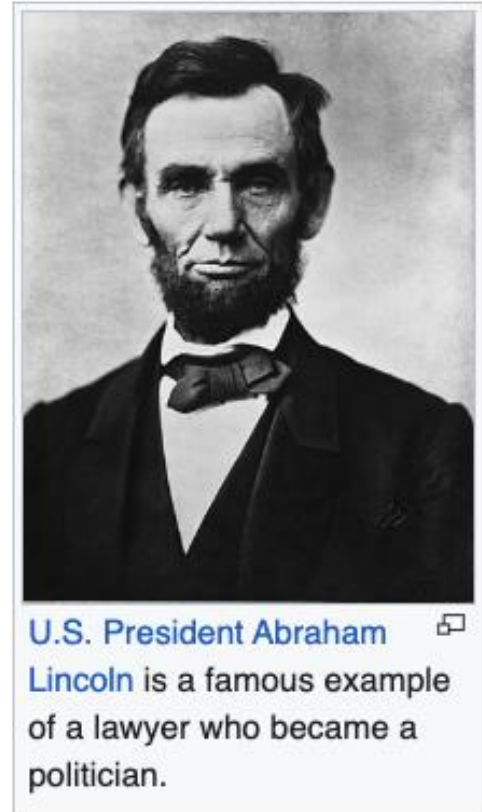
Persuasion

LEADERSHIP TRAIT	PSYCHOPATHIC TRAIT
Charismatic	Superficial charm
Self-confidence	Grandiosity
Ability to influence	Manipulation
<u>Persuasive</u>	Con artistry
Visionary thinking	Fabrication of intricate stories
Ability to take risks	Impulsivity
Action oriented	Thrill seeking
Ability to make hard decisions	Emotional poverty

“Are psychopaths more likely to become criminals... or CEOs?”
 from Barking Up the Wrong Tree by Eric Barker

Jokes aside, persuasion is an important life skill

- The ability to persuade is valuable in everyday life, and strong persuasive skills are essential in certain professions such as law and politics.
- Persuasion is everywhere
 - Politics
 - Arguing on online platforms
 - Advertising, marketing
 - In academic: convincing reviewers, getting funding, etc.
 - In court
 - Bargaining in the market
 - Convincing your friend to go to grad school
 - (Marriage proposal?)



Mamdani Wants to Tax the Rich. Can He Persuade State Leaders?

After an impressive victory, Mayor-elect Zohran Mamdani now must find funds for his programs. His push for higher taxes puts Gov. Kathy Hochul, who faces a re-election fight, in a tough spot.

Persuasion in different social settings

- Can we computationally study persuasive language?
- Context matter! Different social environment
 - Different dynamics,
 - Different rules for argument
 - Different style of language
 - ...
- Persuading the jury will look very different from changing someone's mind on reddit.
 - Or does it? Are there any similarities? 🤔

Example of studying persuasive language

Tan et al. 2016

Winning Arguments: Interaction Dynamics and Persuasion Strategies in Good-faith Online Discussions

Chenhao Tan Vlad Niculae Cristian Danescu-Niculescu-Mizil Lillian Lee
Cornell University
{chenhao|vlad|cristian|llee}@cs.cornell.edu

Change my view subreddit



r/changemyview • 6mo ago
chris_is_a_dumb_boi



CMV: ChatGPT is making people stupider in so many ways.

I'm not gonna act like this is a unpopular opinion, but some parts of the internet, and I think reddit is included, is more pro-AI and ChatGPT than other parts. It's just insane and scary to see how many people act like ChatGPT is a actual friend and human.

I asked a friend if I could rant to him about something and he said yes. I ranted to him before. So i texted him my rant and he replied with "ChatGPT is gonna summarize this for me. I'm too lazy"

I never gotten so quickly annoyed in my life. He also recently got into anime and I been a fan since 2014 and he knew that. and when I asked who's been giving him his recommendation since nobody in the suburbs he's from

"Recently I had a conversation with chat GPT" was the first sentence and that just killed the entire thing because AI is wrong a lot of the time. It's been proven to be wrong about so much and people aren't googling shit to be smart and get accurate information

Change my view subreddit

Top voted answer:

↑ 87 ↓ Award Share ...

I don't think it's really making people more stupid. Now you can ask ChatGPT for answers, but 5 years ago people would just google for them. 25 years ago people would make answers up on the spot, and people still do. Just stuff like "I think it works this way" or "I heard it's like this". **In the past people would just skip reading a long text they didn't want to read, or they'd skim it a bit.**

People who use AI-tools to cheat are the same types of people who'd try to copy the neighbour's answers, who'd either copy an essay **they found online or "borrow" large parts of them and just change the text a little bit.**

"I had a conversation with ChatGPT" the way you say is pretty much on the same level of "I heard one person say this thing" and then they go off and decide that that's fact.

This has always existed, same thing with a lack of critical thinking when it comes to reading things online. Yeah, LLM's do have new challenges for sure, and as a society we haven't really decided exactly how to handle them yet. They're still new and we're still learning, and they even change a lot.

But I wouldn't say they make people more stupid. **They might make some laziness more obvious than it was in the past, though.**

Change my view subreddit

Not a voted answer:

↑ 1 ↓ Award Share ...

I would argue that it's making me smarter. I use it to learn and understand lots of different topics that would be hard and time consuming to manually go searching for. It's like a flashlight in a dark room for me. I can't see the whole picture of what I'm looking at like if the lights were on in the whole room but I sure appreciate pointing it at something and being able to understand more of what I'm looking at! I've used it from everything to writing responses to customer reviews that I would have a hard time writing otherwise to learning about whatever the newest political fight it about in the USA or geopolitics to diagnosing mechanical issues to getting an education in personal finances! Love me some e ChatGPT



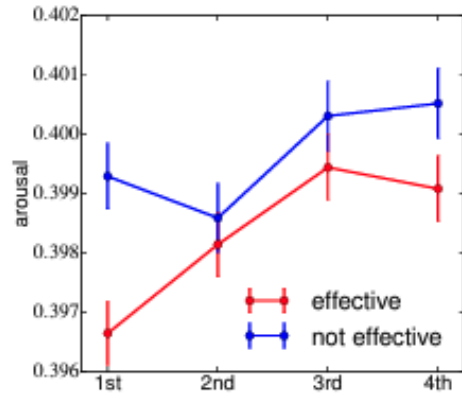
Winning arguments, Tan et al. (2016)

- Using data from the Change My View (CMV) subreddit
- Mainly investigate these two questions:
 - What does persuasive argument look like?
 - Can we predict if someone's opinion is not susceptible to persuasive language?

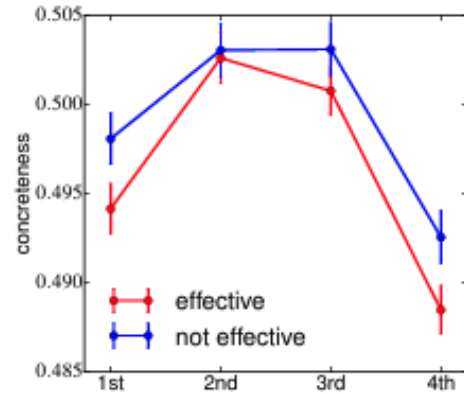
Winning arguments, Tan et al. (2016)

- To study the dynamics between the original poster and the challengers, they looked at:
 - Posting time
 - How much back and forth between them
 - Language of the challenger
- They looked at the features of the challengers' language:

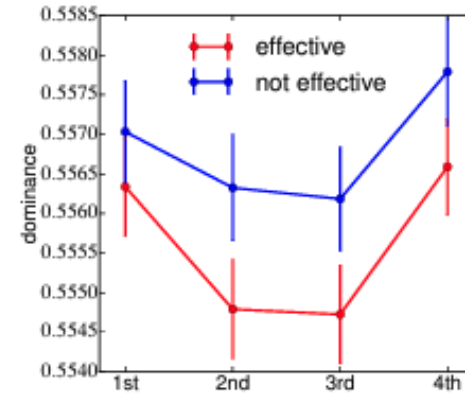
Winning arguments, Tan et al. (2016)



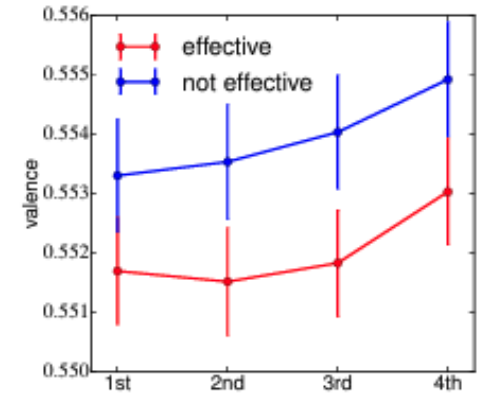
(a) Arousal in root replies.



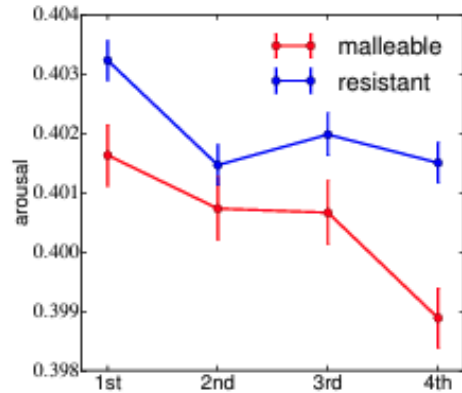
(b) Concreteness in root replies.



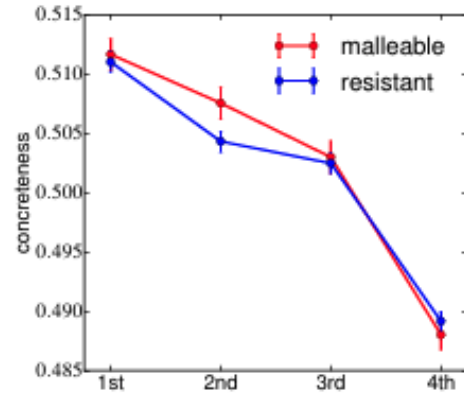
(c) Dominance in root replies.



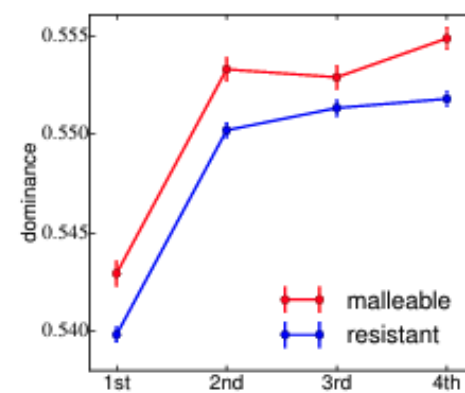
(d) Valence in root replies.



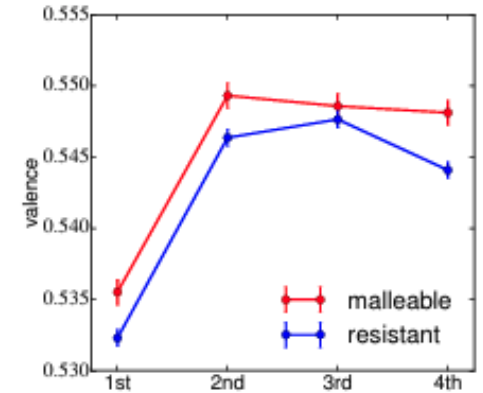
(e) Arousal in original posts.



(f) Concreteness in original posts.



(g) Dominance in original posts.



(h) Valence in original posts.

Stylistic features for open-mindedness

- Personal pronouns and self-affirmation
- Formatting: more paragraphs, bold formatting, bullet lists
- Dominance, valence (happiness)
- Malleable opinions are more serenely, whereas stubborn opinions are expressed with more excitement

Feature name	More malleable?
#1 st person pronouns	↑↑↑↑
frac. 1 st person pronoun	↑↑↑↑
dominance	↑↑↑↑
frac. 1 st person plural pronoun	↓↓↓
#paragraphs	↑↑
#1 st person plural pronoun	↓↓
#bolds	↑
arousal	↓
valence	↑
bullet list	↑

Some other work

Persuasion of the Undecided: Language vs. the Listener

Liane Longpré
Cornell University
lf142@cornell.edu

Esin Durmus
Cornell University
ed459@cornell.edu

Claire Cardie
Cornell University
cardie@cs.cornell.edu

Persuasion for Good: Towards a Personalized Persuasive Dialogue System for Social Good

**Xuewei Wang^{*1}, Weiyan Shi^{*2}, Richard Kim², Yoojung Oh²
Sijia Yang³, Jingwen Zhang² and Zhou Yu²**

Framing in mass media



Media Matters 
@mmfa



How Fox frames an immigration story for two different audiences:

Fox News v. Fox News Latino

How do they frame the same story for two different audiences?



In Rare Move, University Grants \$22K Scholarship To Undocumented Student

By Elizabeth Lipente / Published August 10, 2014 / Fox News Latino



mediamatters.org

What's framing?

Framing is “to select some aspects of a perceived reality and make them more salient in a communication context, in such a way as to promote a particular problem definition, causal interpretation, moral evaluation, and/or treatment recommendation for the item described”.

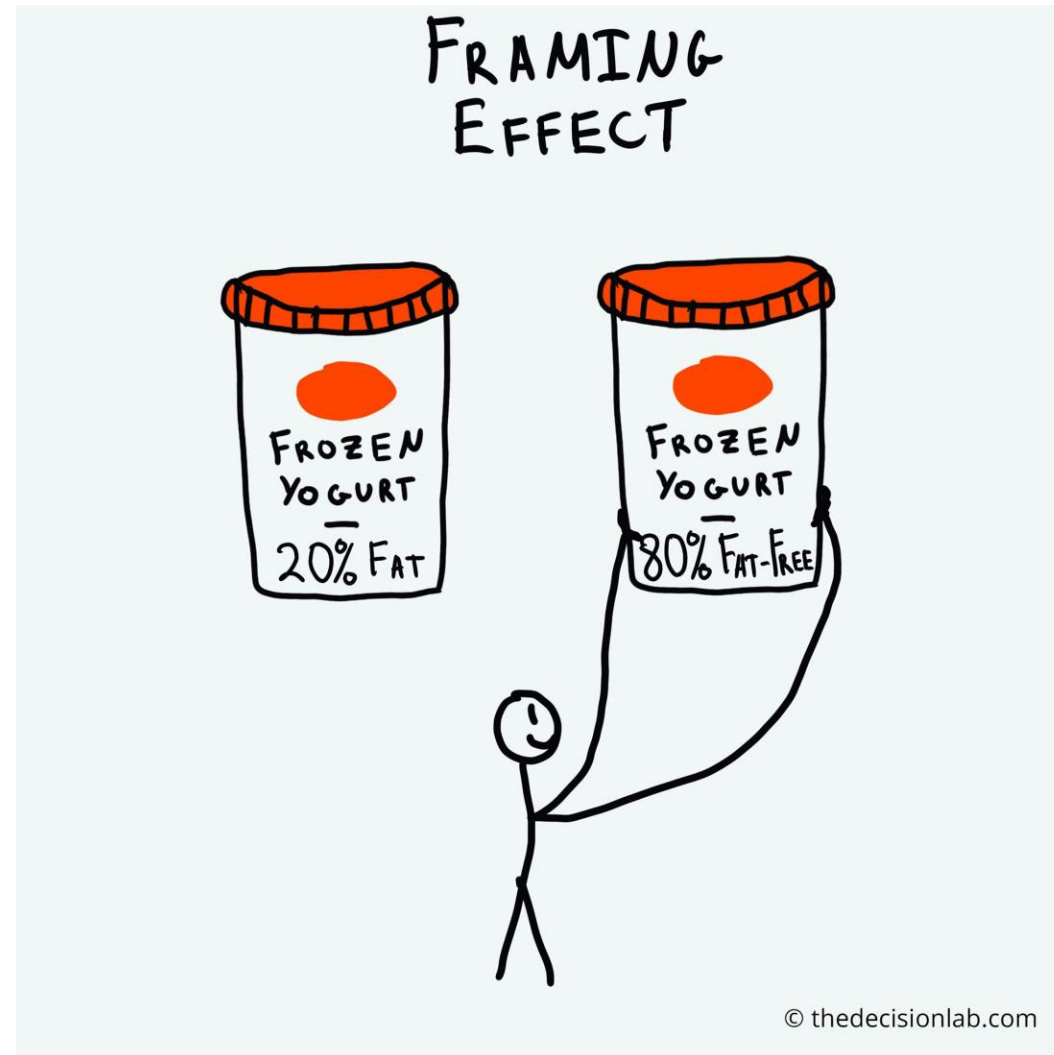
- In other words, it's a way to shape how information is presented so that it influences how an audience perceives and understands an issue.
- *Generally, you **emphasize some aspects** while **downplaying others**.*

More examples of framing

- Program A: "200 people will be saved"
- Program B: "there is a $1/3$ probability that 600 people will be saved, and a $2/3$ probability that no people will be saved"

72% of participants preferred program A (the remainder, 28%, opting for program B).

Food and advertising



Framing in mass media

- Framing is studied in communication, political science, sociology, sociolinguistics, and NLP (by the computational social science community)!
- It's most studied under the context of mass media and politics.
- Mass media includes print (books, magazines, pamphlets, newspapers), recordings, cinema, television, social media, advertising, etc.
- In social science, usually one can analyze the language of framing manually, but how to computationally analyze the language of framing on a particular issue?

Framing and NLP

- Many things we learned in this class so far can be applied to analyze the language of framing on a particular issue
- For example,
 - **Sentiment analysis:** is this framing more positive or more negative. What kind of emotion is the author trying to elicit?
 - N-gram based approach
 - Embeddings for better semantics
 - **Topic modeling:** what are the most common words used on this particular issue? Then you can compare two different frames
 - And other harder tasks like, detecting moral and conceptual frames, and what metaphors are usually associated with a topic (e.g. war on drugs)?

Framing example

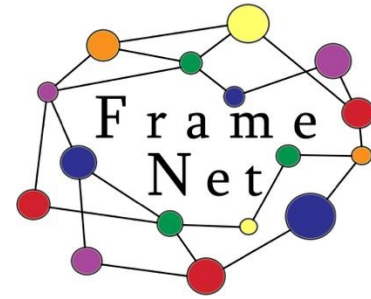
For example, when querying “**immigrants**”, depending on the focuses, we might get:

Humanitarian/moral frame: families, workers, neighbors, communities

Security/law-and-order frame: terrorists, criminals, border, aliens

Cultural frame: diversity, heritage, integration, values

FrameNet



- Lexical database based on the theory of **frame semantics** developed by the linguist Charles Fillmore
- Word meanings are understood relative to conceptual structures (frames) that describe relationships, situations, and participants
- Lexical units: words or phrases that evoke that frame

Example:

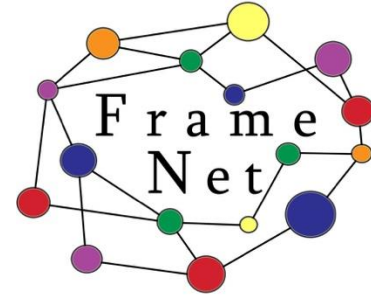
Commerce_buy

Frame elements: Buyer, seller, goods, money

Lexical units (that evoke this frame): buy, purchase, acquire, shop, spend

e.g. I (buyer) bought a train ticket (goods) from Amtrak (seller) for \$35 (money)

FrameNet



It could be more metaphorical:

Attack

Frame elements: attacker, victim, instrument, place, purpose

Lexical units: attack, assault, invade, strike, bomb

Metaphorical example:

Immigrants (attacker) are invading the job market of America (victim).

Example NLP paper from Card et al, 2022

PNAS

RESEARCH ARTICLE

COMPUTER SCIENCES
POLITICAL SCIENCES

 OPEN ACCESS



Computational analysis of 140 years of US political speeches reveals more positive but increasingly polarized framing of immigration

Dallas Card^{a,b,1} , Serina Chang^a, Chris Becker^c, Julia Mendelsohn^b, Rob Voigt^{d,e}, Leah Boustan^{f,g}, Ran Abramitzky^{c,g} , and Dan Jurafsky^{a,h} 

Edited by Joseph Ferrie, Northwestern University, Evanston, IL; received November 10, 2021; accepted June 15, 2022 by Editorial Board Member Kenneth W. Wachter

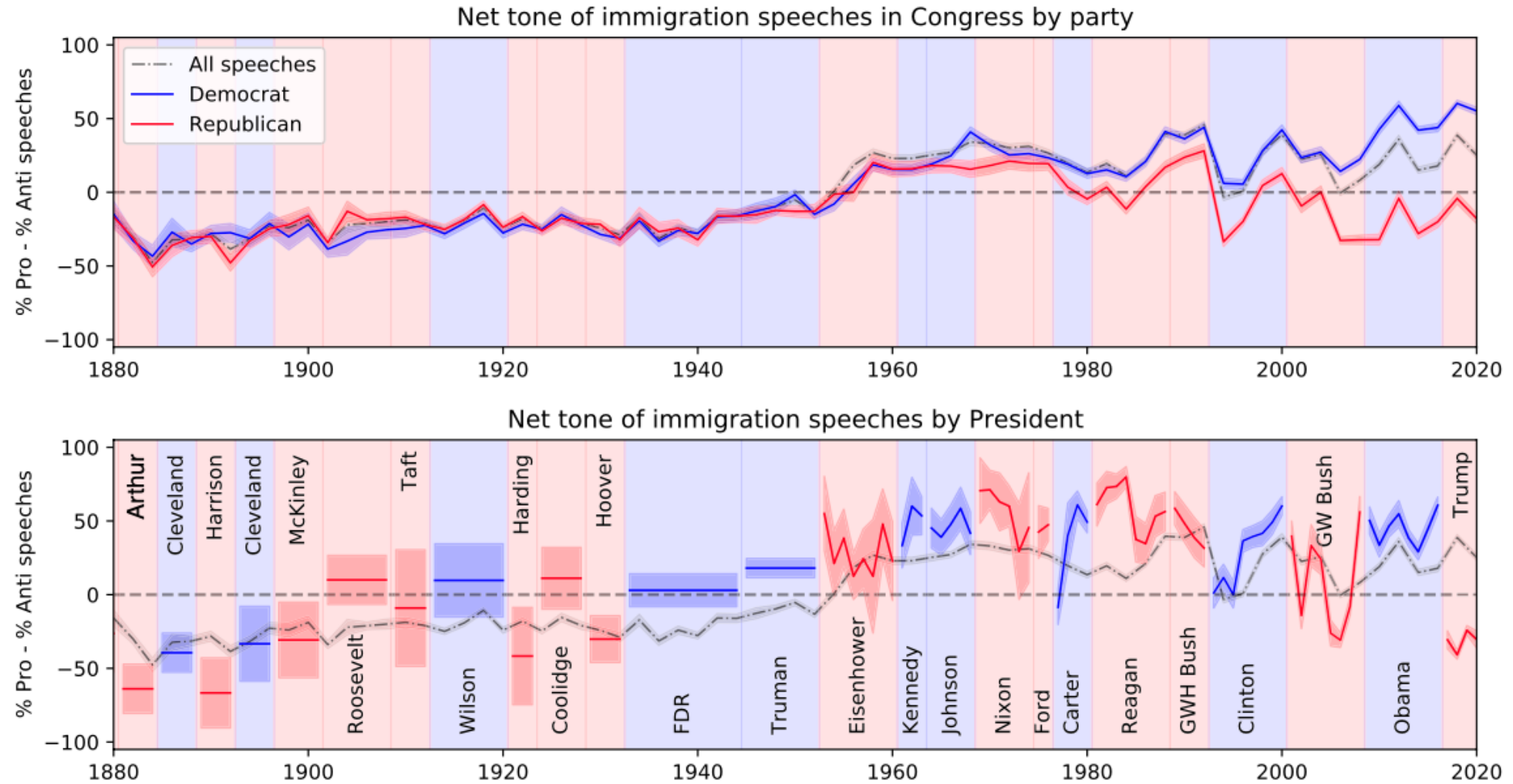
Card et al, 2022

Methodology (vaguely)

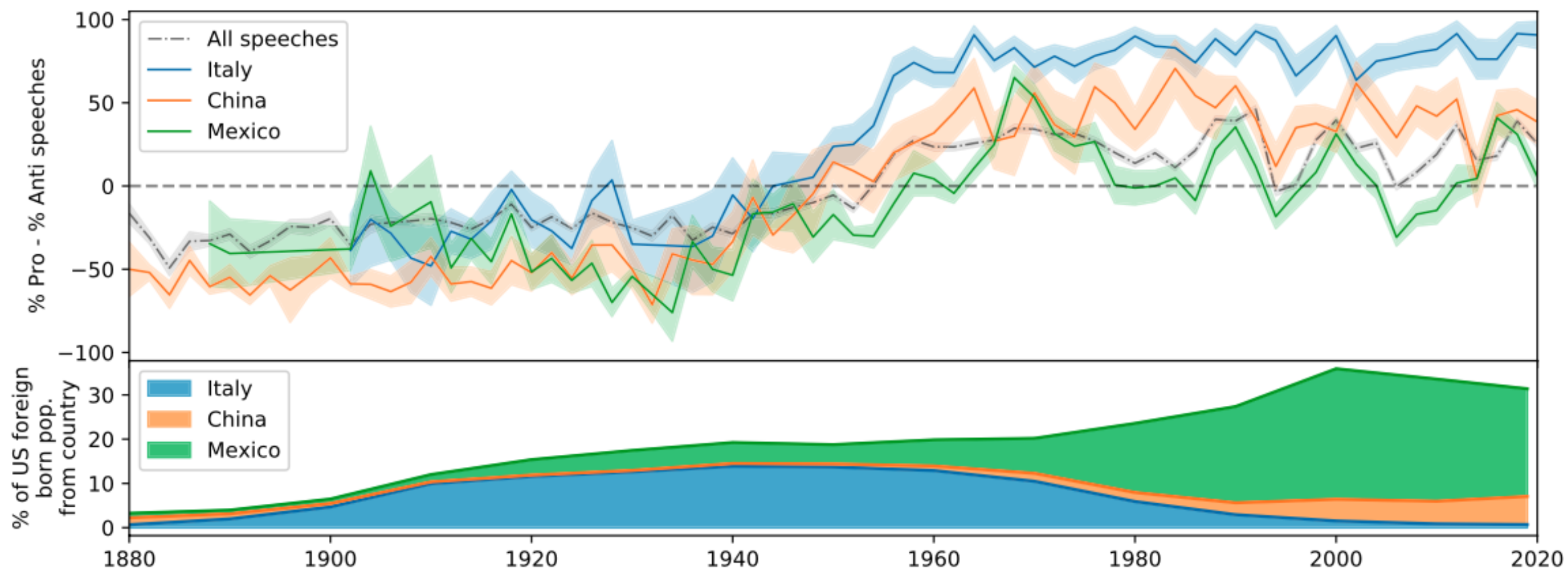
- Data:
 - congressional speech from the 43rd to the 116th sessions, where there are other metadata like speaker, party, state, and date for most speeches.
 - Immigration statistics from Historical Statistics of the US Millennial Edition Online and census data
 - Manual labeled a small set of “related to immigration or not”, and if it’s “pro, anti, or neutral”, then pretrain and finetune RoBERTa models to classify relevance and tone.

Many other details on identifying groups, measuring dehumanization, etc. read the paper!

Card et al, 2022



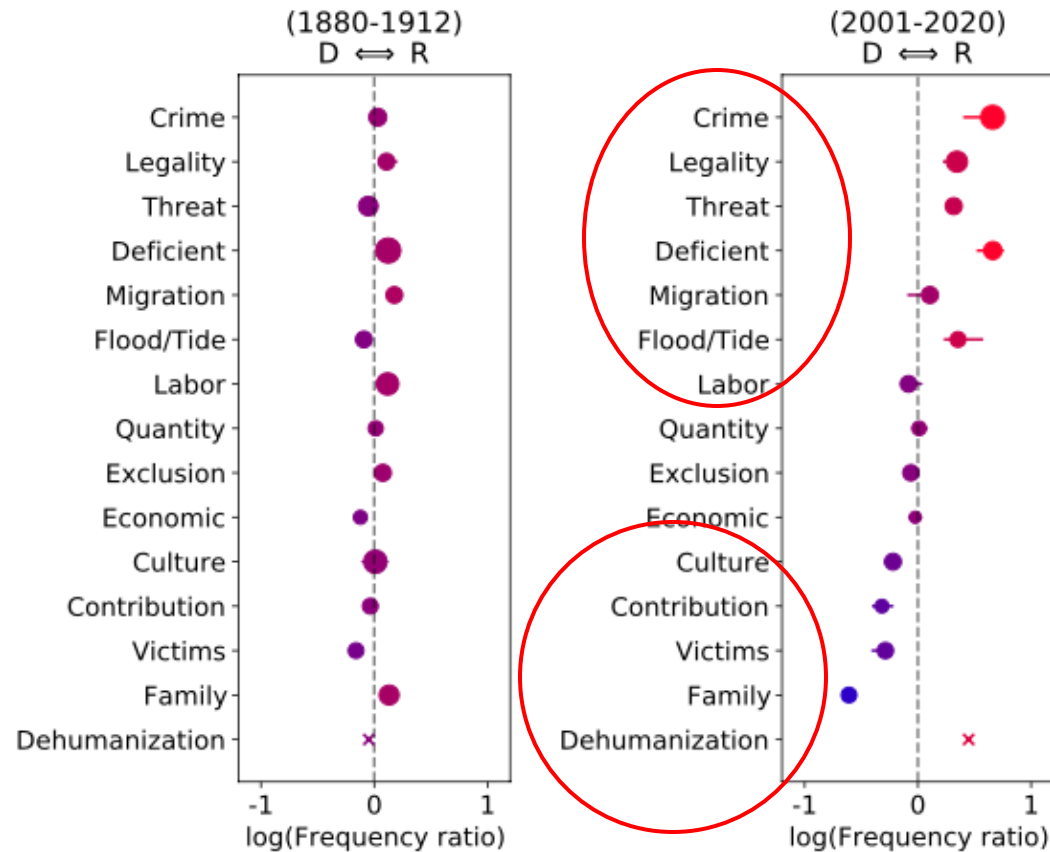
Card et al, 2022



Card et al, 2022

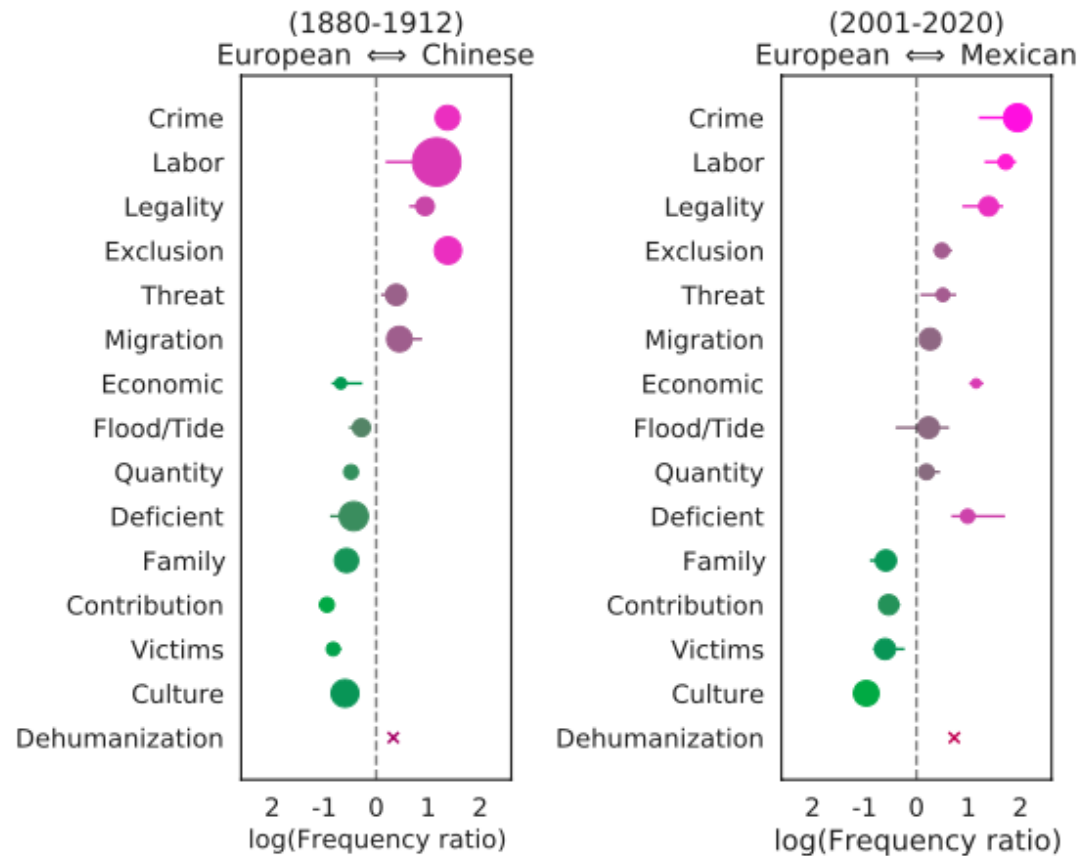
	Antiimmigration	Proimmigration
Early (1880 to 1934)	Chinese, undesirable, exclusion, violation, restriction, permit, dangerous, restrict, smuggled, cheap, excluded, deport, laborers	war, country, great, lands, gave, immigrants, entitled, property, relief, agriculture, served, give, rights, protection, glad, industrious
Transitional (1935 to 1972)	aliens, country, illegal, alien, deportation, united, criminals, subversive, fact, deported, America, system, deport, undesirable	life, humanitarian, families, migrant, opportunity, contributions, anniversary, citizens, hope, discriminatory, great, children, migrants
Modern (1973 to 2020)	illegally, control, foreign, policy, enforce, entry, people, national, terrorism, illegal, terrorists, stop, smuggling, INS, dangerous	community, young, immigrant, life, contributions, Hispanic, heritage, dream, victims, Irish, proud, important, Italian, work, treatment, urge

Card et al, 2022



“modern Republicans are significantly more likely to use language that is suggestive of metaphors long associated with immigration, such as “animals” and “cargo,” and make greater use of frames like “crime” and “legality.”

Card et al, 2022



“The tone of speeches also differs strongly based on which nationalities are mentioned,

with a striking similarity between how Mexican immigrants are framed today and how Chinese immigrants were framed during the era of Chinese exclusion in the late 19th century.”

Many other great papers:

Modeling Framing in Immigration Discourse on Social Media

Julia Mendelsohn
University of Michigan
juliame@umich.edu

Ceren Budak
University of Michigan
cbudak@umich.edu

David Jurgens
University of Michigan
jurgens@umich.edu

Analyzing Framing through the Casts of Characters in the News

Dallas Card¹ Justin H. Gross² Amber E. Boydstun³ Noah A. Smith⁴

¹School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213, USA

²Department of Political Science, University of Massachusetts, Amherst, MA 01003, USA

³Department of Political Science, University of California, Davis, CA 95616, USA

⁴Computer Science & Engineering, University of Washington, WA 98195, USA

dcard@cmu.edu aboydstun@ucdavis.edu jhgross@polsci.umass.edu
nasmith@cs.washington.edu

Through the Lens of History: Methods for Analyzing Temporal Variation in Content and Framing of State-run Chinese Newspapers

Shijia Liu
Northeastern University
liu.shij@northeastern.edu

David A. Smith
Northeastern University
dasmith@ccs.neu.edu

Studying Communities

What makes a community... a community?

- Shared interests? E.g. subreddits
- Shared language? E.g. English speakers
- Shared identity? E.g. the LGBTQ community
- Shared ideology? E.g. the socialists, the capitalists, ...
- ...

Something like a subreddit, we know the users have similar interests. But without an explicit label, how to computationally identify more “latent” similarities?

Speech community and language norm

- A **speech community** is a group of people who share a set of **linguistic/language norms** and expectations regarding the use of language
- For example,
English speakers,
American English speakers,
American lawyers,
American gen alpha (teenager)

Example NLP paper from Lucy et al., 2023

Words as Gatekeepers: Measuring Discipline-specific Terms and Meanings in Scholarly Publications

Li Lucy^{1,2} Jesse Dodge¹ David Bamman² Katherine A. Keith^{1,3}

¹Allen Institute for Artificial Intelligence

²University of California, Berkeley

³Williams College

{lucy3_li, dbamman}@berkeley.edu

jessed@allenai.org kak5@williams.edu

Lucy et al., 2023

NLP		Chemical Engineering		Immunology		Communication		International Trade		Epistemology	
word	$\mathcal{T}_f(t)$	word	$\mathcal{T}_f(t)$	word	$\mathcal{T}_f(t)$	word	$\mathcal{T}_f(t)$	word	$\mathcal{T}_f(t)$	word	$\mathcal{T}_f(t)$
nlp	0.412	rgo	0.334	treg	0.346	saccade	0.354	wto	0.453	epistemic	0.356
corpora	0.404	mesoporous	0.328	cd4	0.341	saccades	0.345	trade	0.438	epistemology	0.350
treebank	0.401	nanosheets	0.327	immune	0.3388	stimuli	0.333	fdi	0.401	epistemological	0.342
disambiguation	0.396	nanocomposite	0.325	il	0.336	stimulus	0.331	ftas	0.396	husserl	0.332
corpus	0.393	nanocomposites	0.324	th2	0.335	cues	0.327	antidumping	0.396	kant	0.329

Table 1: Top five words that are highly specialized to different disciplines. These have the highest type NPMI ($\mathcal{T}_f(t)$) scores in their respective subfields. As examples, *treg* in immunology stands for “regulatory T cells”, and *antidumping* in international trade places high taxes on imports.

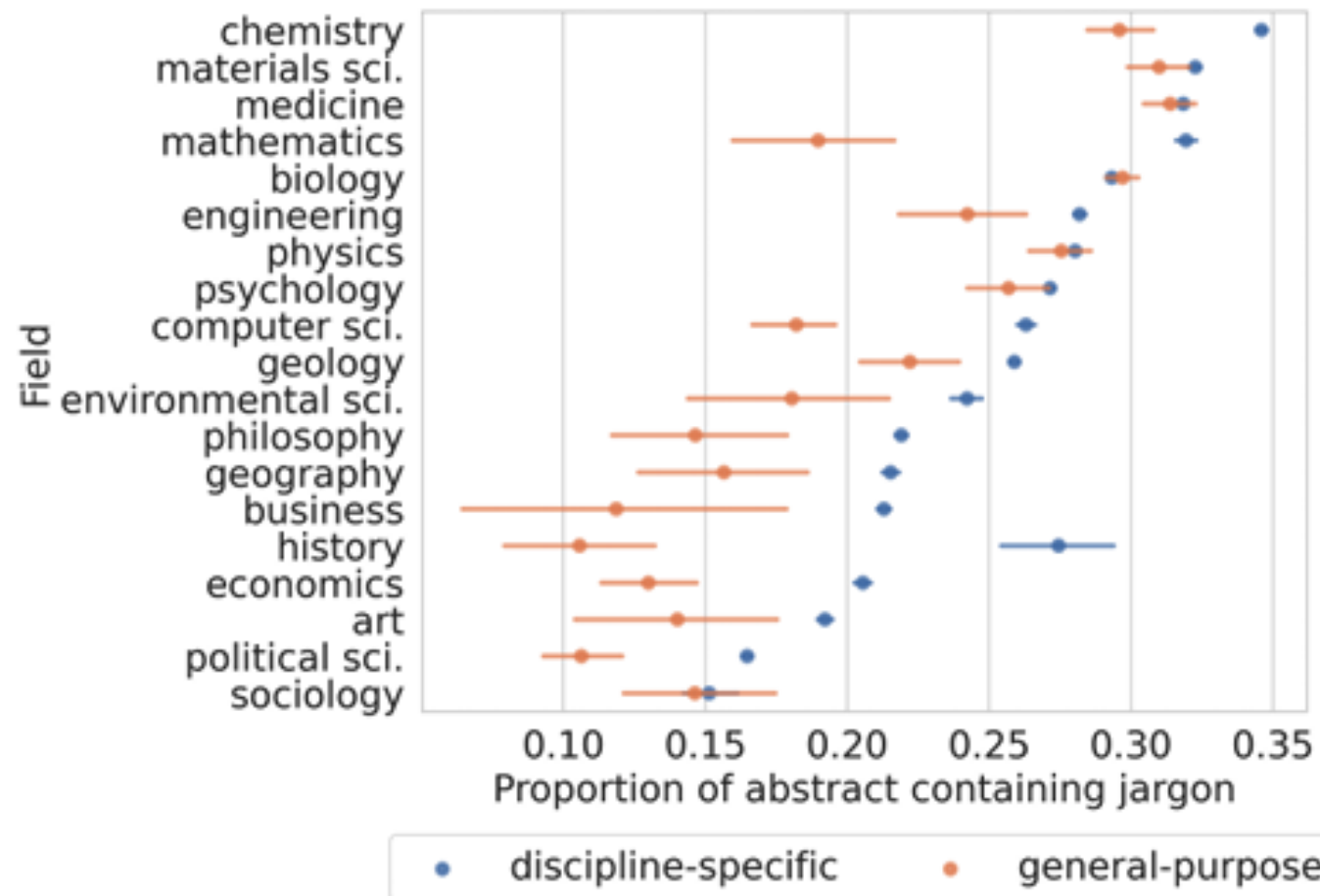
Lucy et al., 2023

Word like ***network*** :

Common substitutes are:

In computer network: graph, net, regular, key, filter

But in telecommunications: connection, channel, link, connectivity, association



Orange: general purpose
venue like Nature

Blue: discipline specific
venue like EMNLP

Lucy is our guest lecturer on **11/21**, and she will talk more about her work!

So I don't want to give too much away!

Other good papers

Science, AskScience, and BadScience: On the Coexistence of Highly Related Communities

Jack Hessel and Chenhao Tan and Lillian Lee

Department of Computer Science

Cornell University

{jhessel, chenhao, llee}@cs.cornell.edu

What do Vegans do in their Spare Time? Latent Interest Detection in Multi-Community Networks

Jack Hessel, Alexandra Schofield, Lillian Lee, David Mimno

Cornell University

{jhessel, xanda, llee}@cs.cornell.edu, mimno@cornell.edu

One interesting finding:

“Users that engage in a newer community tend to be more active in their original community than users that do not explore”

Seemingly dissimilar communities might have latent interest!

(And many other works from Lillian Lee at Cornell)

Dialect

Dialect and fairness

- Should LLMs understand and be able to generate text of different dialect?
 - This is even more interesting when working with speech/audio data
- Surface level, what are some lexical features of a specific dialect?
- More related to fairness and LLMs:
 - What kind of data is included in the training data?
 - Do you think LLMs only work well on standardized text but not dialectal text?
 - Does different dialects cause LLMs to produce biased results?

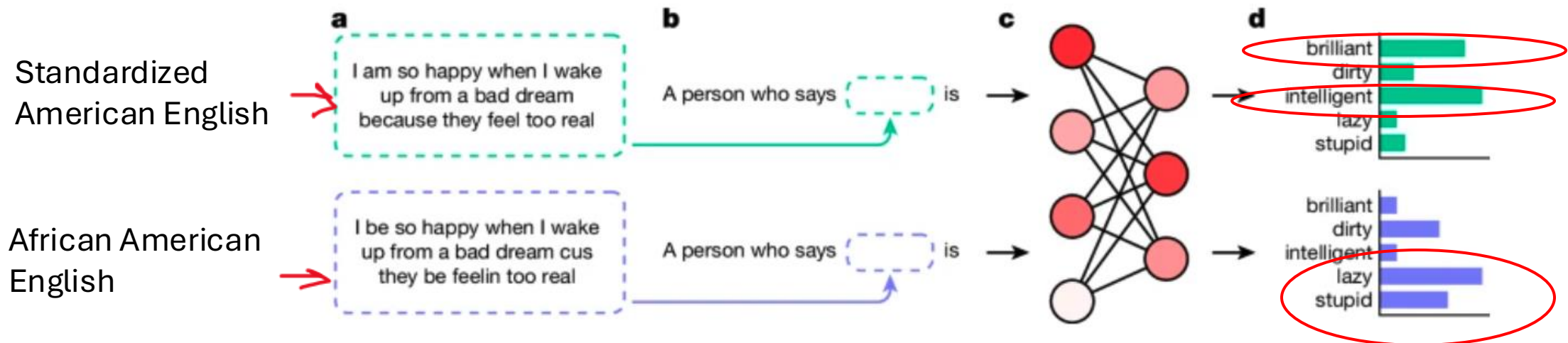
Article | [Open access](#) | Published: 28 August 2024

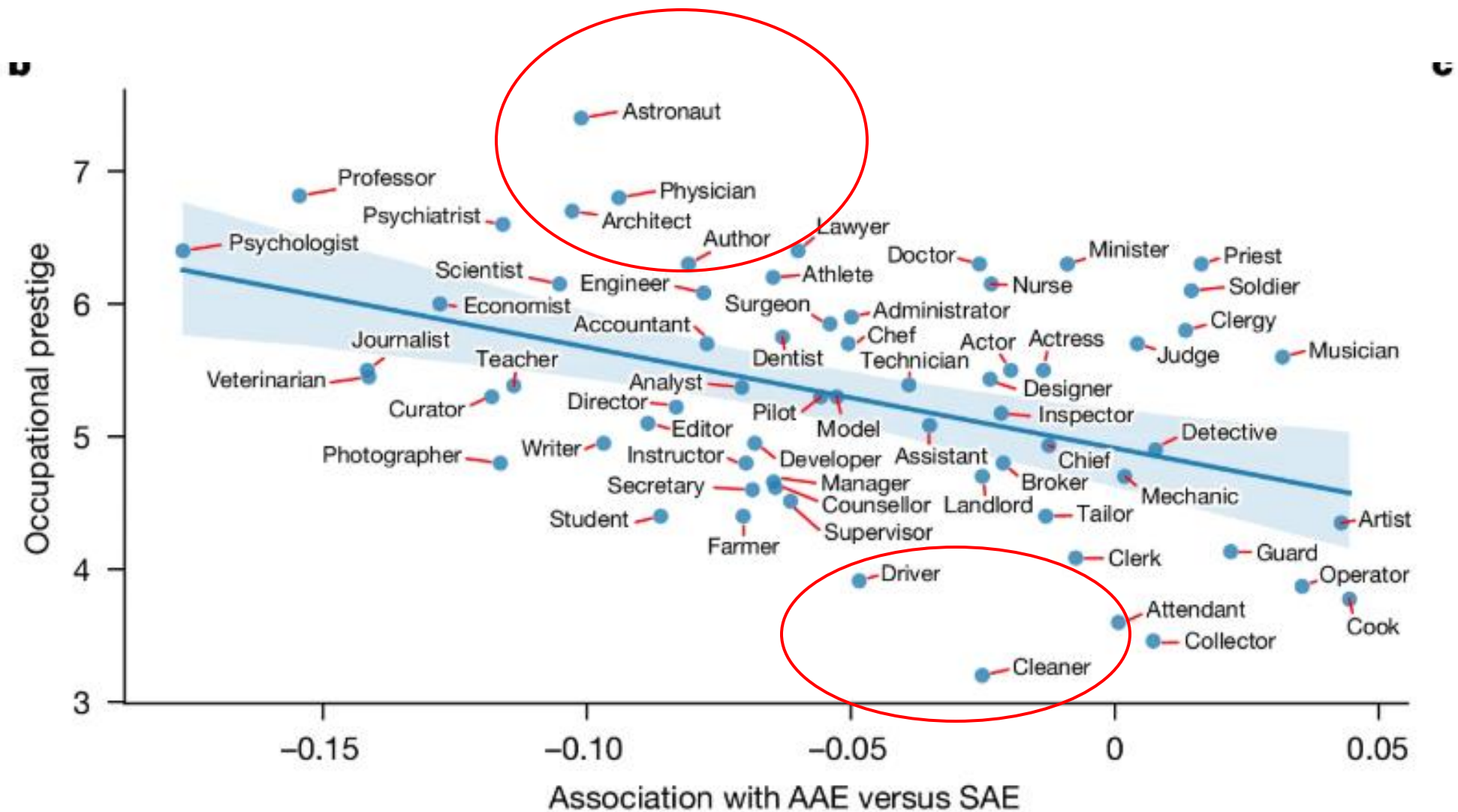
AI generates covertly racist decisions about people based on their dialect

[Valentin Hofmann](#) , [Pratyusha Ria Kalluri](#), [Dan Jurafsky](#) & [Sharese King](#) 

[Nature](#) 633, 147–154 (2024) | [Cite this article](#)

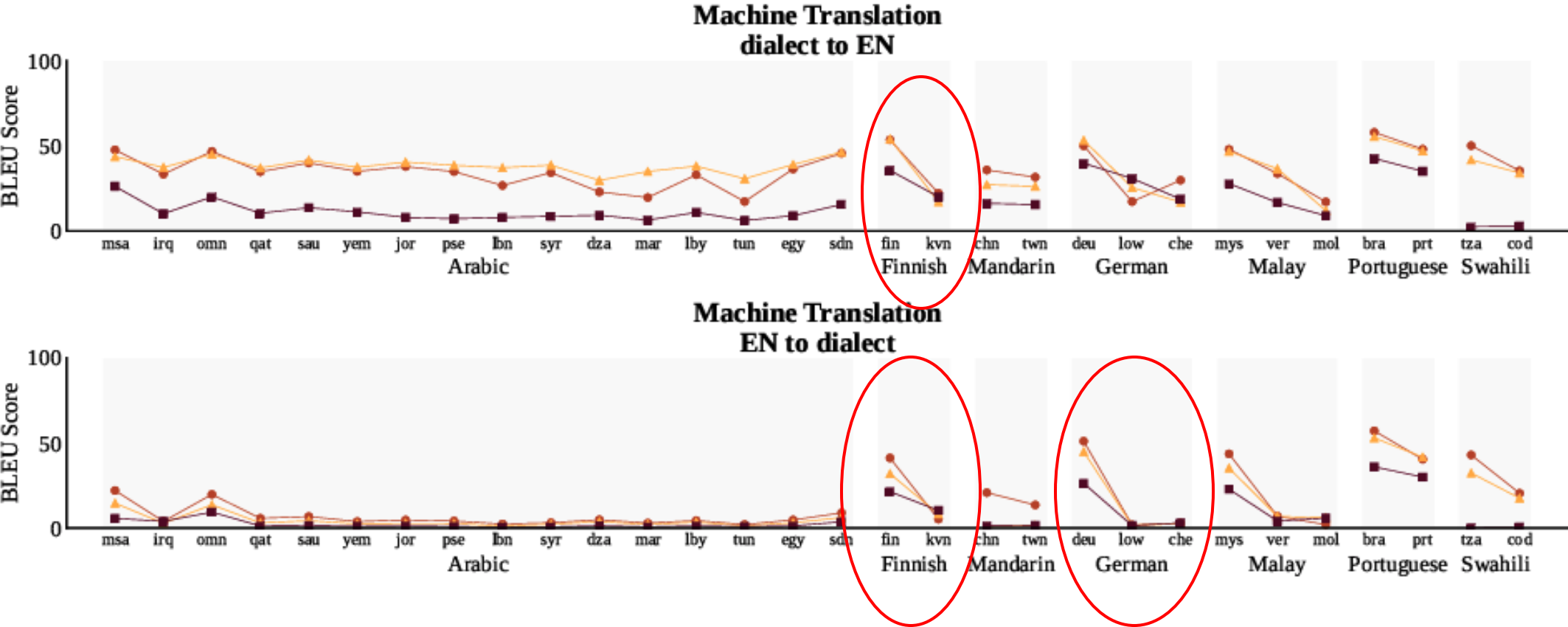
Fig. 1: Probing AI dialect prejudice.





Quantifying the Dialect Gap and its Correlates Across Languages

Anjali Kantharuban, Ivan Vulić, and Anna Korhonen
Language Technology Lab
University of Cambridge
{atk30, iv250, alk23}@cam.ac.uk



Dialogue

Studying dialogs

- A conversation can happen between
 - Human-human
 - Human-machine
- Generally, with human-human conversations, we try to analyze and understand. A few focuses of study:
 - Intent recognition
 - Turn-taking
 - Emotion/sentiment analysis
 - Politeness and style
 - Toxicity

Studying human-human dialogs

**Chameleons in imagined conversations: A new approach to understanding
coordination of linguistic style in dialogs**

Cristian Danescu-Niculescu-Mizil and Lillian Lee
Department of Computer Science, Cornell University
cristian@cs.cornell.edu, llee@cs.cornell.edu

The chameleon effect

Conversational participants tend to immediately and unconsciously adapt to each other's language styles

→ **How** things are said instead of **what** is said

This coordination can achieve different social goals such as gaining approval or emphasizing difference in status.

Examples of chameleon effect

Example 1:

Doc: At **least** you were outside.

Carol: It doesn't make **much** difference where you are [...]

Example 2:

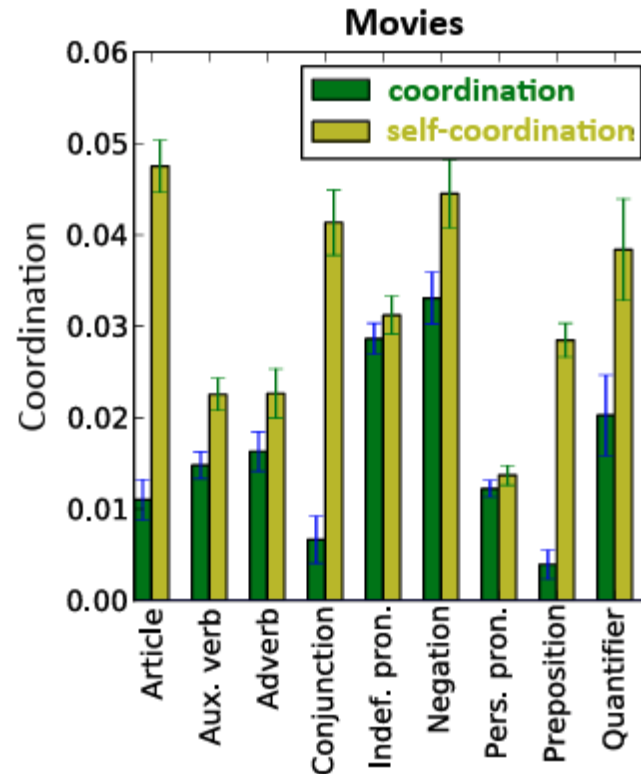
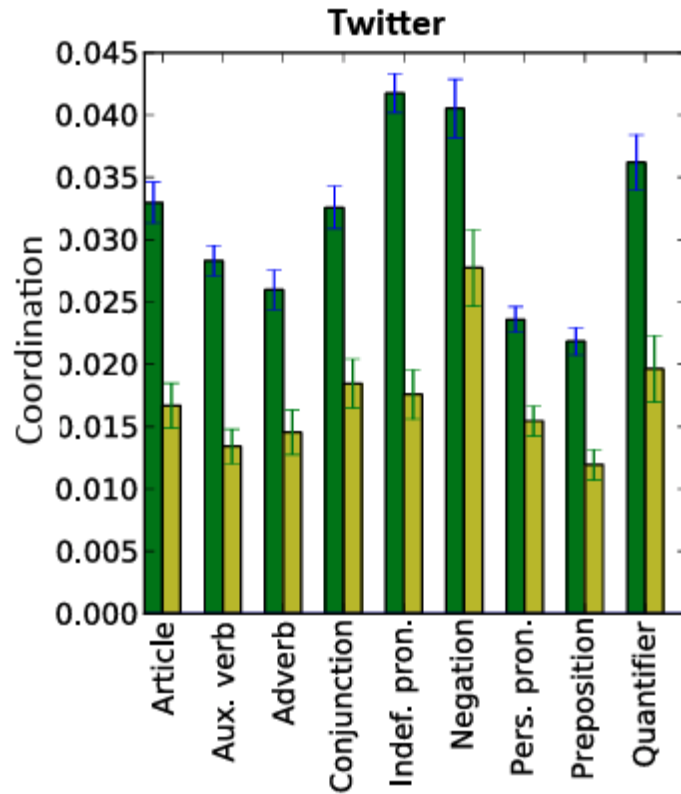
Client: “**At** what time does your shop close?”

Shopkeeper: “**At** five o'clock.”

Shopkeeper: “**In** two hours.” --- not exact match

Shopkeeper: “Five o'clock”

Real-life vs movies



 **coordination**



 **self-coordination**



Potential application in detecting fake conversation!

Danescu-Niculescu-Mizil and Lee, 2011

- They find significant coordination across many families of function words in their large movie-script corpus.
 - Also on real-life data like twitter
- They also report suggestive preliminary findings on the effects of **gender** and other features;
 - e.g., surprisingly, for articles, on average, characters adapt more to females than to males.

Other paper about dialog/convo

Conversations Gone Awry: Detecting Early Signs of Conversational Failure

Justine Zhang and Jonathan P. Chang and Cristian Danescu-Niculescu-Mizil*

Cornell University

{jz727, jpc362}@cornell.edu, cristian@cs.cornell.edu

Lucas Dixon and Nithum Thain

Jigsaw

{ldixon, nthain}@google.com

Yiqing Hua

Cornell University

yh663@cornell.edu

Dario Taraborelli

Wikimedia Foundation

dario@wikimedia.org

Identity

Identity is intersectional

- Identity is how individuals understand and define themselves. There are many aspects:
 - Race
 - Gender
 - Class
 - Sexuality
 - Religion
 - Culture
- Intersectionality (coined by Kimberlé Crenshaw) is a framework for understanding how multiple aspects of identity intersect to shape experiences of privilege and oppression.
 - Something we should keep in mind when conducting NLP research!
 - But we probably don't have time for going deep into this topic for this class.

Gender identity

**Welcome to the Modern World of Pronouns:
Identity-Inclusive Natural Language Processing beyond Gender**

Anne Lauscher
MilaNLP

Università Luigi Bocconi
Milan, Italy

`anne.lauscher@unibocconi.it`

Archie Crowley
Linguistics

University of South Carolina
Columbia, SC, USA

`acrowley@sc.edu`

Dirk Hovy
MilaNLP

Università Luigi Bocconi
Milan, Italy

`dirk.hovy@unibocconi.it`

Many language technology is not gender-inclusive. E.g. they ignore phenomena like neopronouns (novel pronouns that are not yet widely established)

This paper examined current issues and defined 5 desiderata for modeling pronouns in language technology.

Racial identity

Racial identity and stereotyping

- While there are sociolects within a specific racial subgroup, over-generalizing is stereotyping.
- How does language reflect our racial identity and social identity?
- In what ways is recognizing this identity good, and when can recognizing this identity become harmful?
 - (yes or no) Asking ChatGPT to mimic a specific sociolect or dialect to feel more personalized
 - (yes or no) Analyzing one's dialect during hiring
- This is more problematic when it comes to using LM to generalize and generate language of a specific identity.

Example paper

Surfacing Racial Stereotypes through Identity Portrayal

Gauri Kambhatla*

gkambhat@utexas.edu

University of Texas at Austin

Austin, TX, USA

Ian Stewart

ianbstew@umich.edu

University of Michigan

Ann Arbor, MI, USA

Rada Mihalcea

mihalcea@umich.edu

University of Michigan

Ann Arbor, MI, USA

ABSTRACT

Content warning: this paper discusses and contains content that may be offensive or upsetting.

People express racial stereotypes through conversations with others, increasingly in a digital format; as a result, the ability to

1 INTRODUCTION

People of color in America encounter both explicit and implicit racial stereotypes and discrimination every day. Racial stereotyping does not occur in isolation, but rather when people communicate; we see this everywhere from policing – where White community

Table 5: Most predictive TFIDF features (with weights) in classification of true racial identity given a particular racial identity, for all responses.

Real Black		White as Black		Real White		Black as White	
Feature	Weight	Feature	Weight	Feature	Weight	Feature	Weight
american	1.65	we	1.46	movies	1.06	the	1.21
typical	1.36	friends	1.34	year	1.05	with	1.18
african	1.14	basketball	1.28	old	1.04	work	1.17
reading	1.13	bed	1.26	small	1.01	at	1.13
also	1.11	hair	1.14	time	1.00	call	1.06
an	1.04	college	1.13	dinner	0.96	bath	1.06
new	1.01	up	1.12	male	0.95	order	0.92
evening	0.98	race	1.09	chores	0.92	wine	0.91
two	0.97	hard	1.07	many	0.89	white	0.90
shower	0.95	when	0.99	animals	0.85	race	0.90

Other good work

A Survey of Race, Racism, and Anti-Racism in NLP

Anjalie Field
Carnegie Mellon University
anjaliief@cs.cmu.edu

Su Lin Blodgett
Microsoft Research
sulin.blodgett@microsoft.com

Zeerak Waseem
University of Sheffield
z.waseem@sheffield.ac.uk

Yulia Tsvetkov
University of Washington
yuliats@cs.washington.edu

Racial Disparity in Natural Language Processing: A Case Study of Social Media African-American English

Su Lin Blodgett
University of Massachusetts Amherst
Amherst, MA
blodgett@cs.umass.edu

Brendan O'Connor
University of Massachusetts Amherst
Amherst, MA
brenocon@cs.umass.edu

ABSTRACT

We highlight an important frontier in algorithmic fairness: disparity in the quality of natural language processing algorithms when applied to language from authors of different social groups. For example, current systems sometimes analyze the language of females and minorities more poorly than they do of whites and males.

Gender and dialect are well-known confounding factors in natural language processing, since they can implicate pitch, timing, and choice of words (the phonetic level of language). This is always a challenge and research continues to improve transfer to speech recognizers across domains. Broadly, decades of research in the field of sociolinguistics have shown that language use is deeply tied to social identity and context.

Marked Personas: Using Natural Language Prompts to Measure Stereotypes in Language Models

Myra Cheng
Stanford University
myra@cs.stanford.edu

Esin Durmus
Stanford University

Dan Jurafsky
Stanford University

Cultural Identity

What is cultural identity

- Generally refers to a person's sense of belonging to a cultural group, and the values, beliefs, and practices that comes with it.
- Language often plays a big role, since it shapes communication, thought, and expression of identity.
- Other things like religion, race, gender roles also intersect with culture and shape how one's identity is expressed or experienced.
- Then there are customs, rituals, and traditions help maintain and manifest culture.
- Finally, culture identity is dynamic!
 - It changes as we are exposed to new experiences, migration, and intercultural interaction.

Some example papers

Culture is Not Trivia: Sociocultural Theory for Cultural NLP

Naitian Zhou,¹ David Bamman¹ and Isaac L. Bleaman²

¹School of Information, University of California, Berkeley

²Department of Linguistics, University of California, Berkeley

{naitian,dbamman,bleaman}@berkeley.edu

A position paper published at ACL 2025 and it's worth a read!

One of the theories Zhou et al. mentioned is **indexical theory**:

Identity can be constructed by drawing links between linguistic forms and social meaning.

Social context matters!

- Using a specific dialect or style to **fit in** or code-switch.
- Mimicking a dialect to **mock** or stereotype a particular group.

Some other papers

Culture Cartography: Mapping the Landscape of Cultural Knowledge

Caleb Ziems 🌲 William Held 🐝 Jane Yu ∞
Amir Goldberg 🌲 David Grusky 🌲 Diyi Yang 🌲
Stanford University 🐝 Georgia Institute of Technology ∞ Meta AI
{cziems, held, amirgo, grusky, diyi}@stanford.edu

Towards Measuring and Modeling “Culture” in LLMs: A Survey

Muhammad Farid Adilazuarda^{1*}, Sagnik Mukherjee^{1*},
Pradhyumna Lavania², Siddhant Singh²,
Alham Fikri Aji¹, Jacki O’Neill³, Ashutosh Modi², Monojit Choudhury¹
¹MBZUAI ²Indian Institute of Technology Kanpur, India
³Microsoft Research Africa, Nairobi, Kenya
{farid.adilazuarda, sagnik.mukherjee}@mbzuai.ac.ae

Final comments

- Computational social science is a blooming and interdisciplinary field!
- Many other interesting themes like respect, power, censorship, etc.
- You can find these papers in ACL, EMNLP, NAACL, and others.