Applied NLP: Static word embeddings in Computational Social Science

CS 1671/2071 guest lecture

Neha Kennard

February 21 2024

A little about me

- PhD student at IESL
- UMass Amherst
- NLP for Computational Social Science
- End users: sociologists

A little about you?

Please go to:

menti.com

1383 3875

Course Schedule

Subject to change. Last revised 2024-02-12. All due dates are at 11:59pm ET except when indicated.

All

Schedule

Date	Topics	Slides (update after class)	Readings	Assignments & Project
Monday, Jan 8	Introduction	01.pdf		
Wednesday, Jan 10	Text Processing	Basic Text Processing	Jurafsky and Martin Chapter 2 (2.1-2.4)	HW1 (Basic Text Processing) out
Wednesday, Jan 17	Text Classification	Sparse word representation	Jurafsky and Martin Chapter 6 (6.3-6.7)	
Monday, Jan 22	Text Classification	Naive Bayes	Jurafsky and Martin Chapter 4 (4-4.5)	
Wednesday, Jan 24	Text Classification	Naive Bayes cont & Classification Evaluation	Jurafsky and Martin Chapter 4 (4.7-4.10) Bender & Friedman 2018 (data statements) Mitchell et al. 2019 (model cards)	HW1 (Basic Text Processing) due
Monday, Jan 29	Text Classification	Logistic Regression	Jurafsky and Martin Chapter 5 (5-5.3)	HW2 (Text Classification) out
Wednesday, Jan 31	Text Classification	Logistic Regression 2	Jurafsky and Martin Chapter 5 (5.4-5.6, 5.11)	
Monday, Feb 5	Representation Learning	Logistic Regression 3 & Static Word Embeddings	Jurafsky and Martin Chapter 5 (5.6-5.9) Jurafsky and Martin Chapter 6 (6-6.2, 6.8-6.13) Arora et al.2020 Blodgett et al. 2020	
Wednesday, Feb 7	Neural Networks	Static Word Embeddings & Feedforward Neural Networks	Jurafsky and Martin Chapter 7 (7-7.1, 7.3-7.4, 7.6, 7.8)	HW2 (Text Classification) due
Monday, Feb 12	Language Modeling	FFNN & N-gram language models 1	Jurafsky and Martin Chapter 3 (3.2)	Project proposal due
Wednesday, Feb 14	Language Modeling	N-gram language models 2	Jurafsky and Martin Chapter 3 (3.3-3.6, 3.9)	HW3 (N-gram Language Model) out
Monday, Feb 19	Language Modeling Neural Networks	Recurrent Neural Networks	Jurafsky and Martin Chapter 9 (9-9.2, 9.6-9.9)	
Wednesday, Feb 21	Guest Lecture	Neha Kennard — NLP for Social Problems (Remote)		
Monday, Feb 26	Language Modeling Neural Networks	Transformers	Jurafsky and Martin Chapter 10 (10-10.2, 10.4)	HW3 (N-gram Language Model) due
Wednesday, Feb 28	Language Modeling Neural Networks	Transformers 2	Jurafsky and Martin Chapter 10 (10.7) Chapter 11 (11-11.3.2)	HW4 (Sentiment with Neural Nets) out
Monday, Mar 4	Coding Walk Through	Mid-term review & Assignment 4 Help Session		
Wednesday, Mar 6	First Exam			
Monday, Mar 18	Language Modeling	Pre-training, GPT		Project mid-term report due
Wednesday, Mar 20	Language Modeling	LLM & Prompt Tuning		HW4 (Sentiment with Neural Nets) due
Monday, Mar 25	Sequence Labeling	HMMs, forward algorithm.		HW5 (Prompting) out
Wednesday, Mar 27	Sequence Labeling	HMM, Viterbi		
Monday, Apr 1	Parsing	Constituency Parsing		
Wednesday, Apr 3	Parsing	Dependency Parsing		
Monday, Apr 8	NLP Applications	Commonsense Knowledge		HW5 (Prompting) due
Wednesday, Apr 10	NLP Applications	Summarization		
Monday, Apr 15	NLP Applications	Question Answering		
Wednesday, Apr 17	Social Factors in NLP	Fairness and Bias		Project presentation; Project final report due on 4/19
TBD	Second Exam			

COUNT 28

Language Modeling Neural Networks	Transformers
Language Modeling Neural Networks	Transformers 2
Coding Walk Through	Mid-term review & Assignment 4 Help Session
First Exam	
Language Modeling	Pre-training, GPT
Language Modeling	LLM & Prompt Tuning

The Brilliance and Weirdness of ChatGPT

...penAI is inspiring awe, fear, stunts and ...t its guardrails.

The Year Chatbots Were Tamed

A year ago, a rogue A.I. tried to break ... backlash help make chatbots too borin...

OpenAI Gives ChatGPT a Better 'Memory'

The A.I. start-up is releasing a new version of ChatGPT that stores what users say and applies it to future chats.

Course Schedule

Subject to change. Last revised 2024-02-12. All due dates are at 11:59pm ET except when indicated.

All

Schedule

Date	Topics	Slides (update after class)	Readings	Assignments & Project
Monday, Jan 8	Introduction	01.pdf		
Wednesday, Jan 10	Text Processing	Basic Text Processing	Jurafsky and Martin Chapter 2 (2.1-2.4)	HW1 (Basic Text Processing) out
Wednesday, Jan 17	Text Classification	Sparse word representation	Jurafsky and Martin Chapter 6 (6.3-6.7)	
Monday, Jan 22	Text Classification	Naive Bayes	Jurafsky and Martin Chapter 4 (4-4.5)	
Wednesday, Jan 24	Text Classification	Naive Bayes cont & Classification Evaluation	Jurafsky and Martin Chapter 4 (4.7-4.10) Bender & Friedman 2018 (data statements) Mitchell et al. 2019 (model cards)	HW1 (Basic Text Processing) due
Monday, Jan 29	Text Classification	Logistic Regression	Jurafsky and Martin Chapter 5 (5-5.3)	HW2 (Text Classification) out
Wednesday, Jan 31	Text Classification	Logistic Regression 2	Jurafsky and Martin Chapter 5 (5.4-5.6, 5.11)	
Monday, Feb 5	Representation Learning	Logistic Regression 3 & Static Word Embeddings	Jurafsky and Martin Chapter 5 (5.6-5.9) Jurafsky and Martin Chapter 6 (6-6.2, 6.8-6.13) Arora et al.2020 Blodgett et al. 2020	
Wednesday, Feb 7	Neural Networks	Static Word Embeddings & Feedforward Neural Networks	Jurafsky and Martin Chapter 7 (7-7.1, 7.3-7.4, 7.6, 7.8)	HW2 (Text Classification) due
Monday, Feb 12	Language Modeling	FFNN & N-gram language models 1	Jurafsky and Martin Chapter 3 (3.2)	Project proposal due
Wednesday, Feb 14	Language Modeling	N-gram language models 2	Jurafsky and Martin Chapter 3 (3.3-3.6, 3.9)	HW3 (N-gram Language Model) out
Monday, Feb 19	Language Modeling Neural Networks	Recurrent Neural Networks	Jurafsky and Martin Chapter 9 (9-9.2, 9.6-9.9)	
Wednesday, Feb 21	Guest Lecture	Neha Kennard — NLP for Social Problems (Remote)		
Monday, Feb 26	Language Modeling Neural Networks	Transformers	Jurafsky and Martin Chapter 10 (10-10.2, 10.4)	HW3 (N-gram Language Model) due
Wednesday, Feb 28	Language Modeling Neural Networks	Transformers 2	Jurafsky and Martin Chapter 10 (10.7) Chapter 11 (11-11.3.2)	HW4 (Sentiment with Neural Nets) out
Monday, Mar 4	Coding Walk Through	Mid-term review & Assignment 4 Help Session		
📌 Wednesday, Mar 6	First Exam			
Monday, Mar 18	Language Modeling	Pre-training, GPT		Project mid-term report due
Wednesday, Mar 20	Language Modeling	LLM & Prompt Tuning		HW4 (Sentiment with Neural Nets) due
Monday, Mar 25	Sequence Labeling	HMMs, forward algorithm.		HW5 (Prompting) out
Wednesday, Mar 27	Sequence Labeling	HMM, Viterbi		
Monday, Apr 1	Parsing	Constituency Parsing		
Wednesday, Apr 3	Parsing	Dependency Parsing		
Monday, Apr 8	NLP Applications	Commonsense Knowledge		HW5 (Prompting) due
Wednesday, Apr 10	NLP Applications	Summarization		
Monday, Apr 15	NLP Applications	Question Answering		
Wednesday, Apr 17	Social Factors in NLP	Fairness and Bias		Project presentation; Project final report due on 4/19
📌 TBD	Second Exam			

COUNT 28

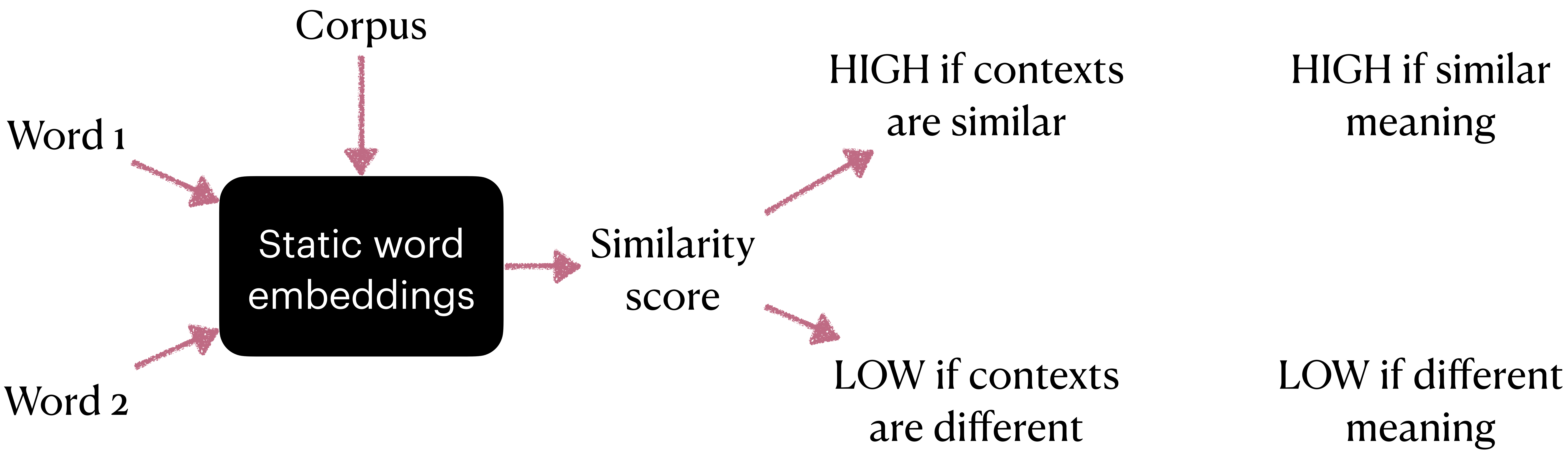
Language Modeling Neural Networks	Transformers
Language Modeling Neural Networks	Transformers 2
Coding Walk Through	Mid-term review & Assignment 4 Help Session
First Exam	
Language Modeling	Pre-training, GPT
Language Modeling	LLM & Prompt Tuning

Representation Learning	Logistic Regression 3 & Static Word Embeddings
Neural Networks	Static Word Embeddings & Feedforward Neural Networks

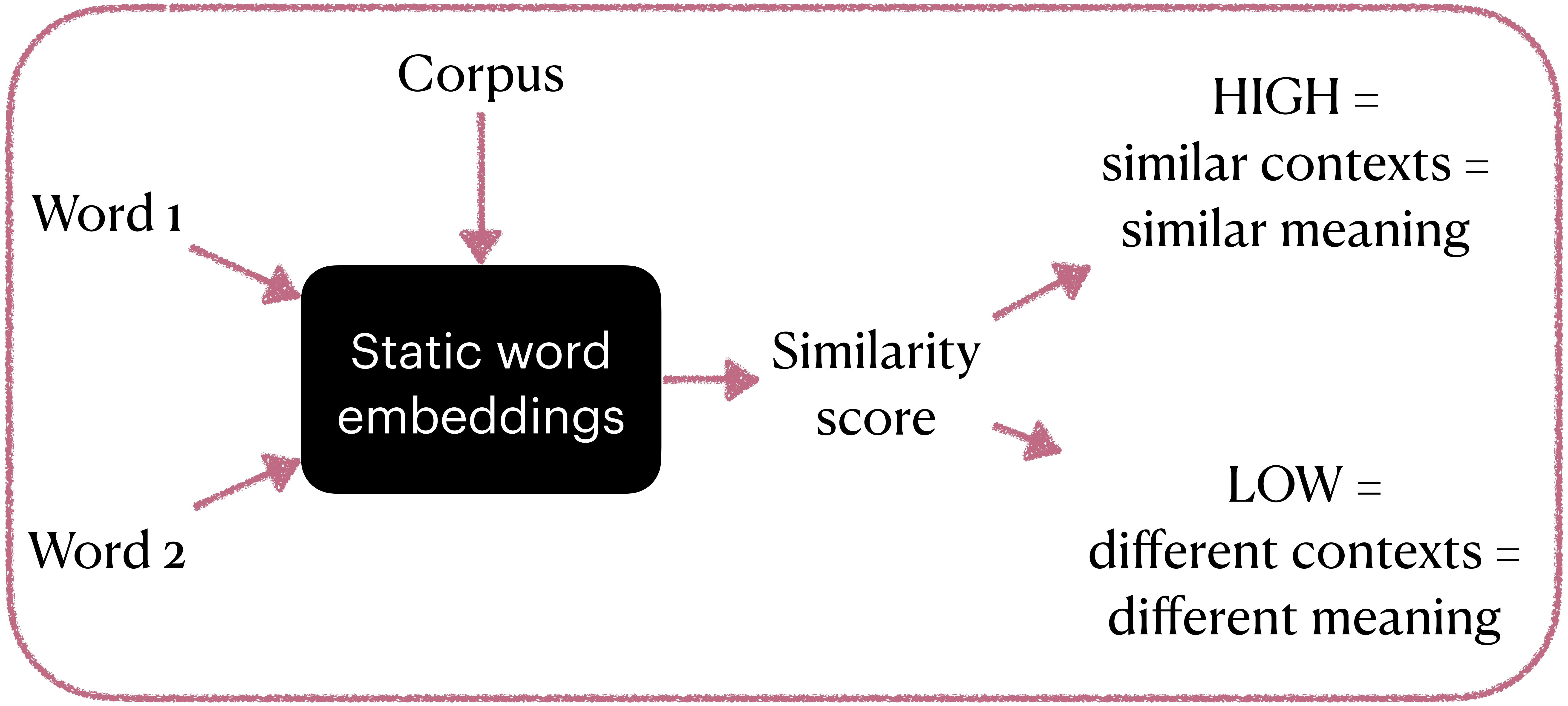
What is a static word embedding?

ICYMI

- Set of vectors — points in a vector space
- Trained on a corpus
- One for each word or ‘type’
- If two words have similar contexts in the corpus, their vectors are close together
- Corpus -> vectors conversion: word2vec, GloVe, fasttext…



"You shall know a word by the company it keeps" - [Firth 1957]


Corpus
Word 1
Word 2
Static word embeddings
Similarity score
HIGH = similar contexts = similar meaning
LOW = different contexts = different meaning

Questions for today

- Who is still using these, and for what?
- What can I, a computer scientist, contribute?
- Why not just use LLMs instead?

Research questions

In NLP and in other fields

NLP:
(From CS3730)

- Can models learn language without embodiment?
- Should knowledge be neuralized or indexed?
- How will the understanding of language benefit multi-modal applications and embodied agents?

- Questions *about* NLP models
- Answered *with*… various techniques

🤔

Research questions

In NLP and in other fields

Political Science:

- How, and by whom, is emotional language employed in US Congress debates?

- Questions *about* people and their interactions
- Answered *with*… NLP techniques we have already learned about! 🥳

How, and by whom, is emotional language employed in US Congress debates?

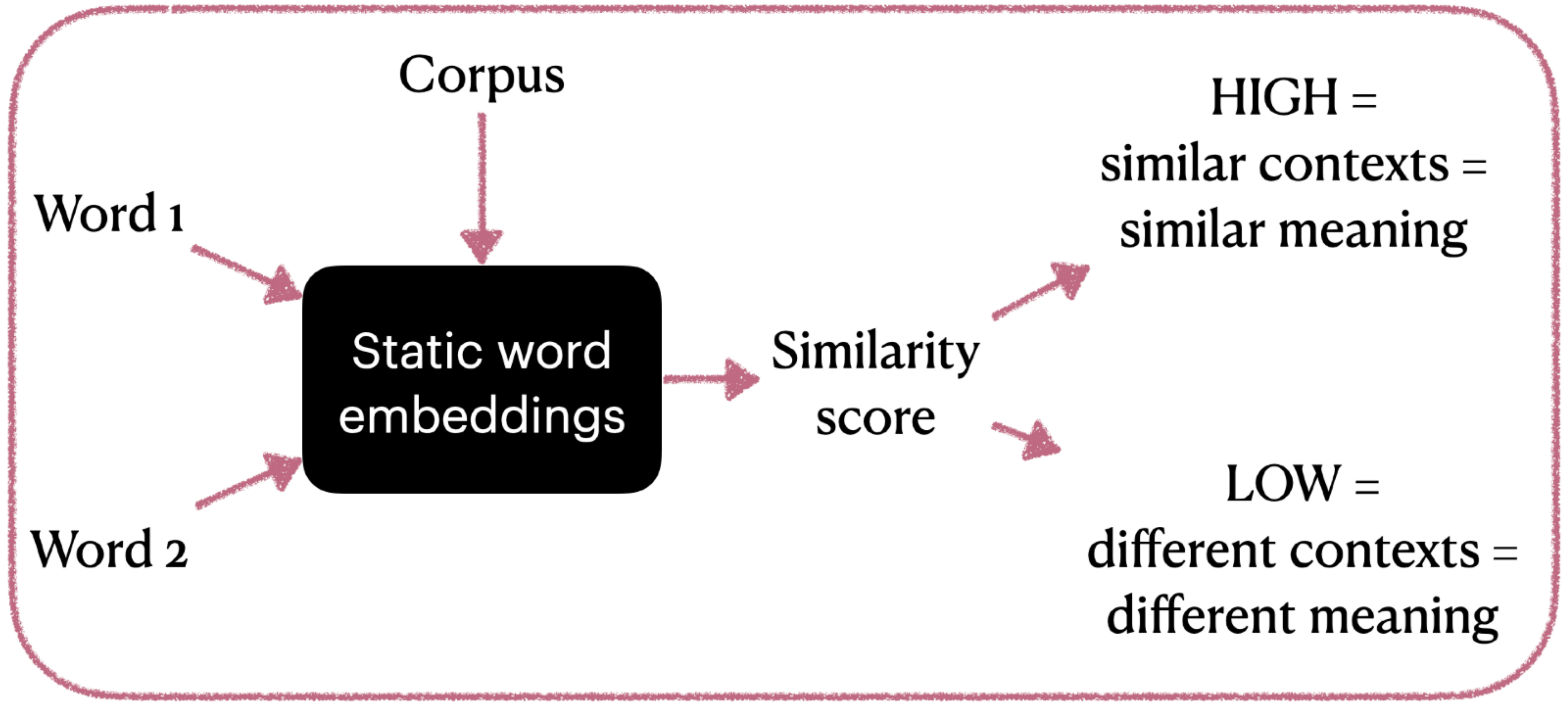
Gloria Gennaro, Elliott Ash, Emotion and Reason in Political Language, *The Economic Journal*, Volume 132, Issue 643, April 2022, Pages 1037–1059, https://doi.org/10.1093/ej/ueab104

How, and by whom, is emotional language employed in US Congress debates?

- In his treatise on Rhetoric, Aristotle suggested that persuasion can be achieved through **either logical argumentation or emotional arousal** in the audience; success depends on selecting the most appropriate strategy for the given context.
- The extent to which politicians engage with this trade-off … **is largely unknown**.
- **Providing empirical evidence on these questions has been difficult** due to the lack of a reproducible, validated and scalable measure of emotionality in political language.

How, and by whom, is emotional language employed in US Congress debates?

Theory	
Emotional arousal	Logical argumentation
Pathos	Logos
Emotion and Affect	Rationality and Cognition
Politicians trade off — When? Why?	



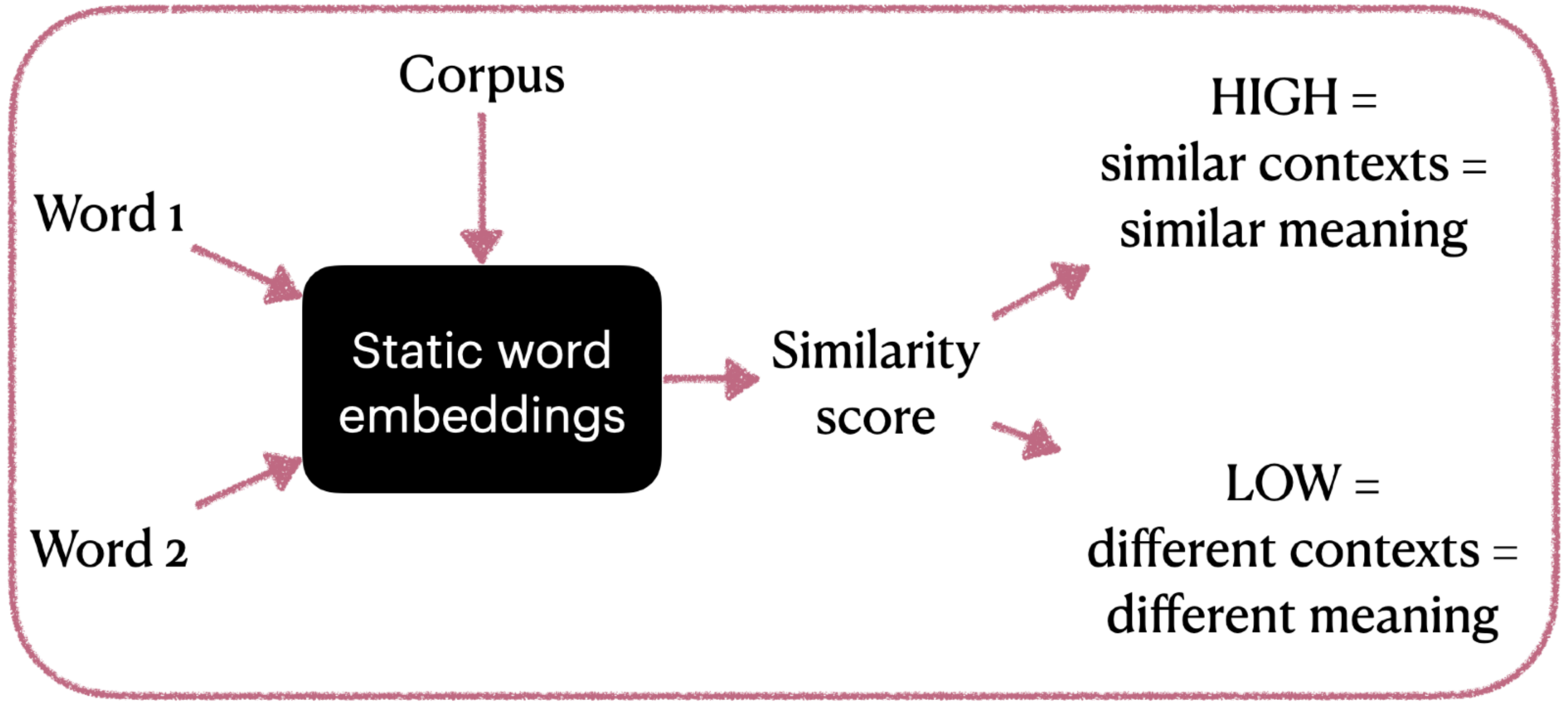
- Answer using static embeddings

Your toolkit

How, and by whom, is emotional language employed in US Congress debates?

- Digitized transcripts of speeches in the U.S. House and Senate between 1858 and 2014
- For each speech:
  - Full text
  - Date of speech
    - Speaker's political party

Theory	
Emotional arousal	Logical argumentation
Pathos	Logos
Emotion and Affect	Rationality and Cognition
Politicians trade off — When? Why?	

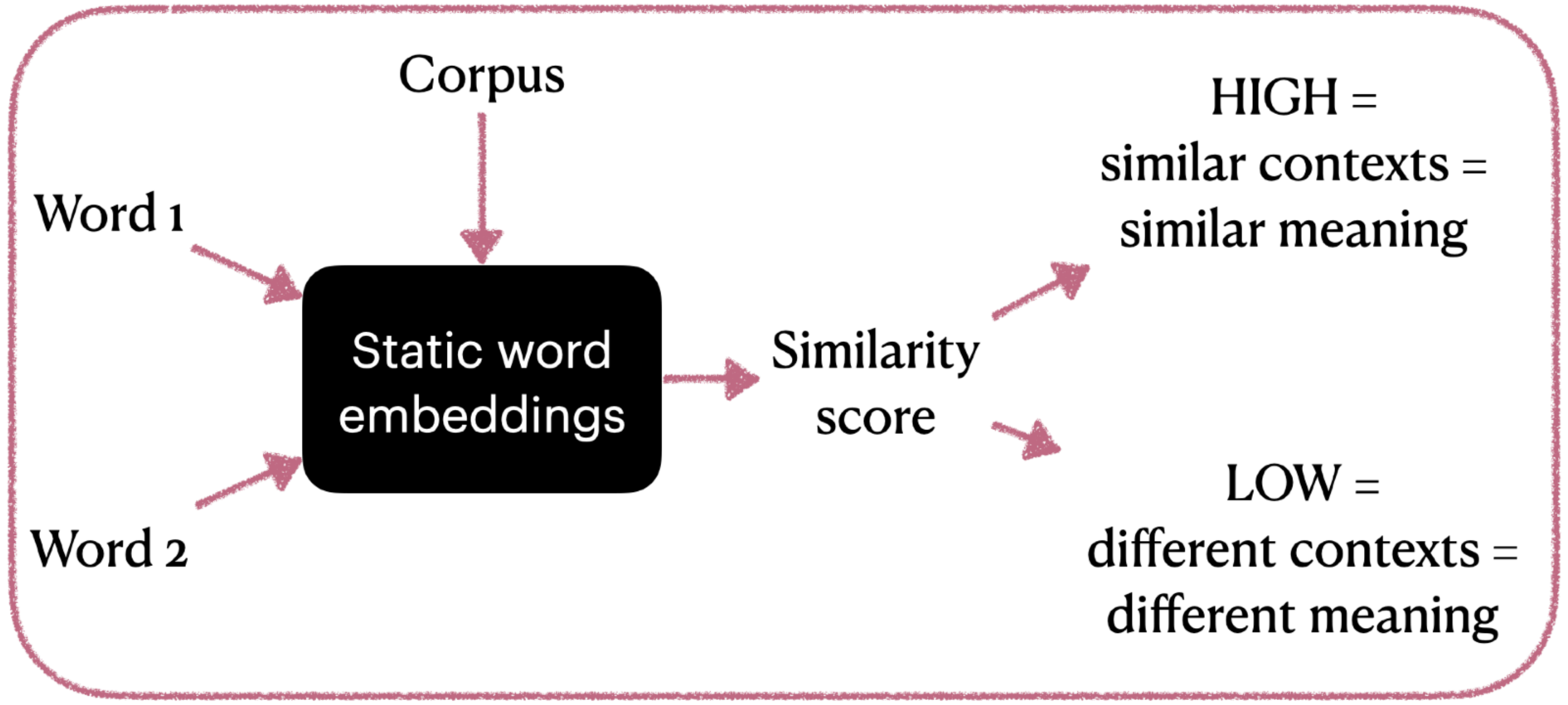


Preparation

How, and by whom, is emotional language employed in US Congress debates?

- Concatenate all speeches from 1858 - 2014 into one corpus
- Clean data (part-of-speech tagging, removing stopwords, etc.)
- Train a word2vec model

Theory	
Emotional arousal	Logical argumentation
Pathos	Logos
Emotion and Affect	Rationality and Cognition
Politicians trade off — When? Why?	

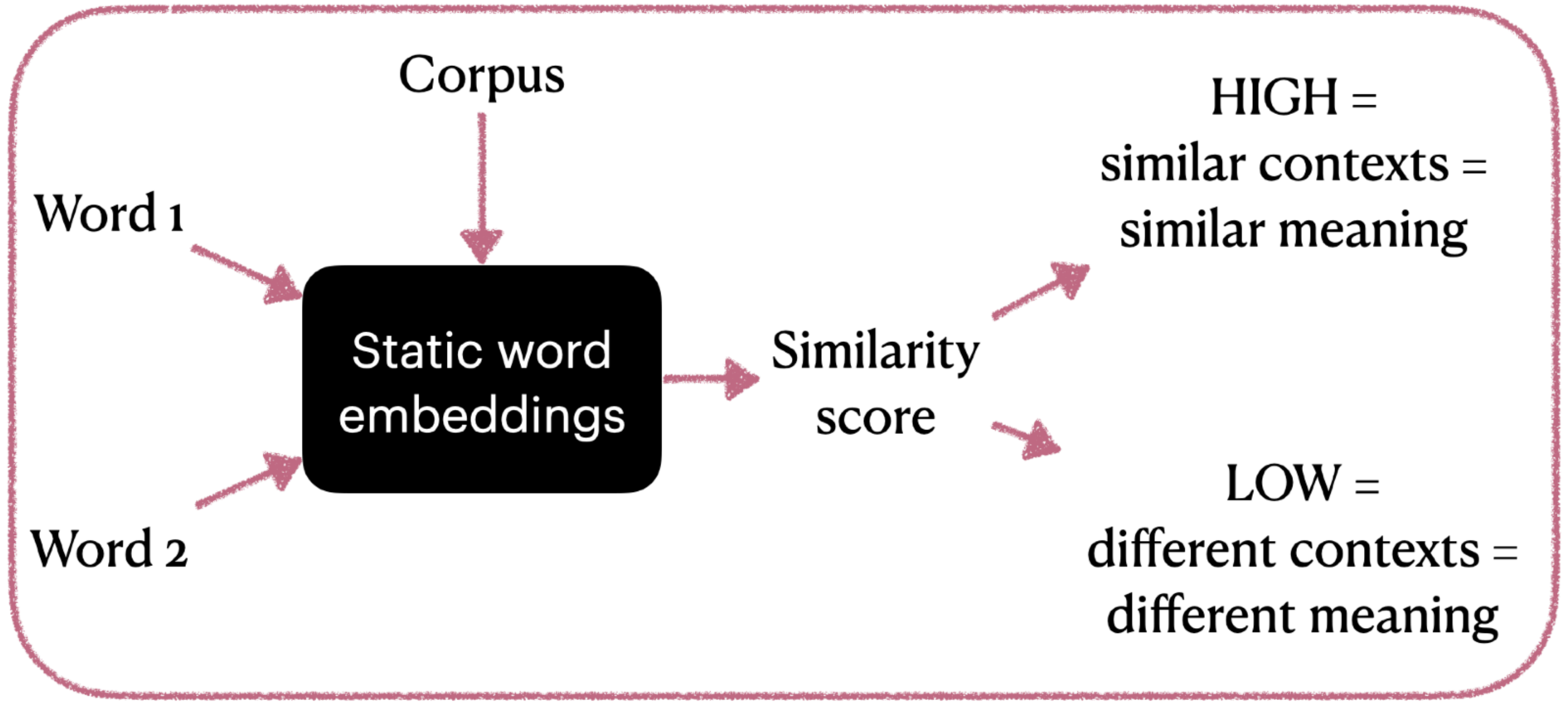


Method

How, and by whom, is emotional language employed in US Congress debates?

- We want to compare individual *speeches* to *concepts* (**Emotion** or **Cognition**)
- Speeches can be represented as a set of words
  - Average of vectors in the set
- Concepts can be represented as sets of words e.g. *{thrill, serene, frighten, ...}*
  - Similarly, average of vectors in the set

Theory	
Emotional arousal	Logical argumentation
Pathos	Logos
Emotion and Affect	Rationality and Cognition
Politicians trade off — When? Why?	



We report there the affect dictionary words with their count in the corpus:

support (1765047), import (1421018), like (1327182), great (1195251), agre (1147658), care (1018579), help (945406), concern (834363), thank (746428), opportun (662106), defens (647623), polit (560160), interest (511530), critic (358826), credit (355314), favor (344079), open (330082), give (312834), person (297694), valu (295900), fight (273278), encourag (255137), fail (254356), relief (244541), argument (234996), attack (231244),

We report there the cognition dictionary words with their count in the corpus:

think (2222390), want (1933090), need (1858735), question (1765467), know (1761052), believ (1294547), fact (1278946), resolut (1204296), reason (870024), understand (860049), effect (829068), consid (802972), chang (800344), purpos (794236), make (755361), allow (741097), product (738070), recogn (722642), result (685842), control (675044), distinguish (672218), respons (669281), statement (649465), inform (628884), differ (616581), refer

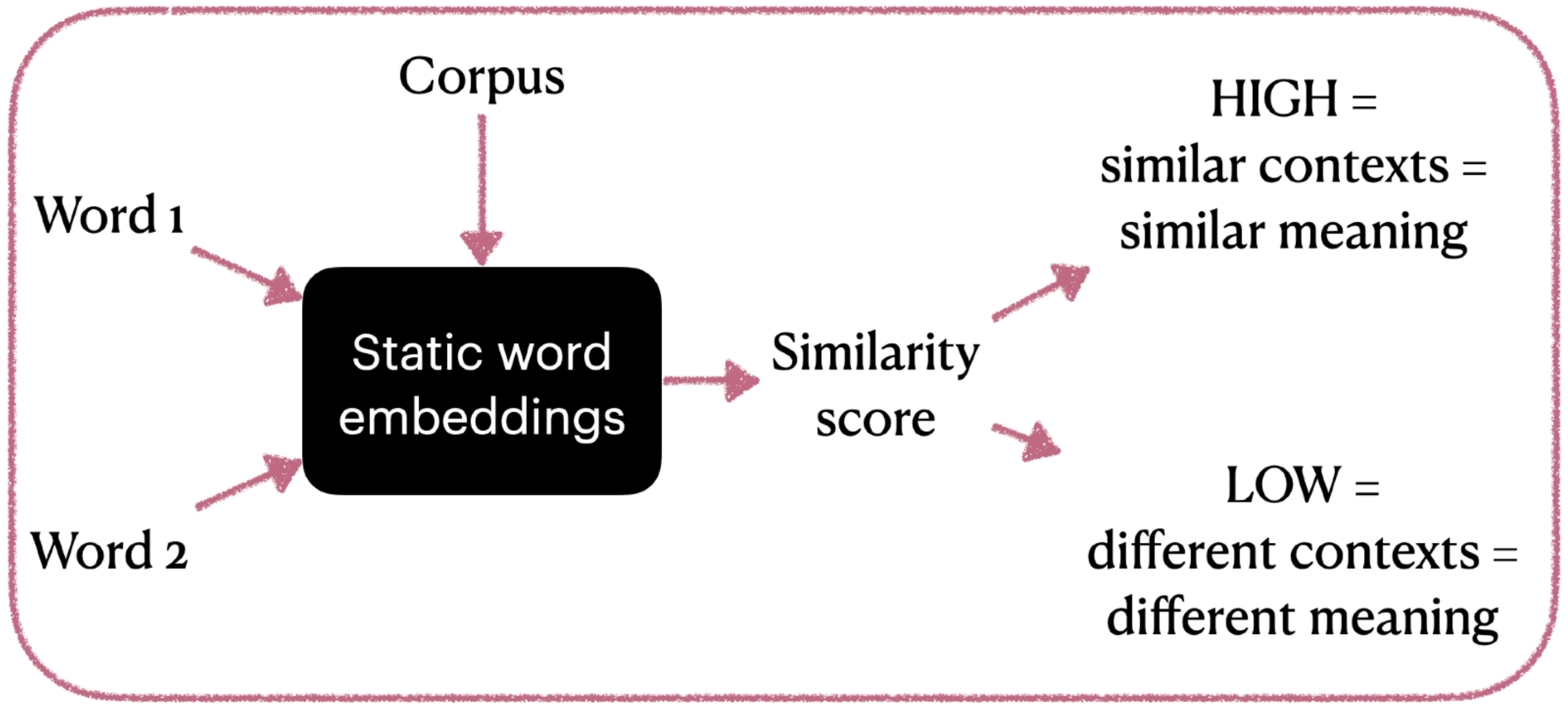
Method

How, and by whom, is emotional language employed in US Congress debates?

- Emotion is represented by vector **E** — mean of emotion words' vectors
- Cognition is represented by vector **C** — mean of cognition words' vectors
- Speech *i* is represented by vector $\mathbf{d}_i$ — mean of vectors of all words in the speech
- *Emotionality* Y_i of speech *i*

$$Y_i = \frac{sim(\mathbf{d}_i, \mathbf{E}) + b}{sim(\mathbf{d}_i, \mathbf{C}) + b}$$

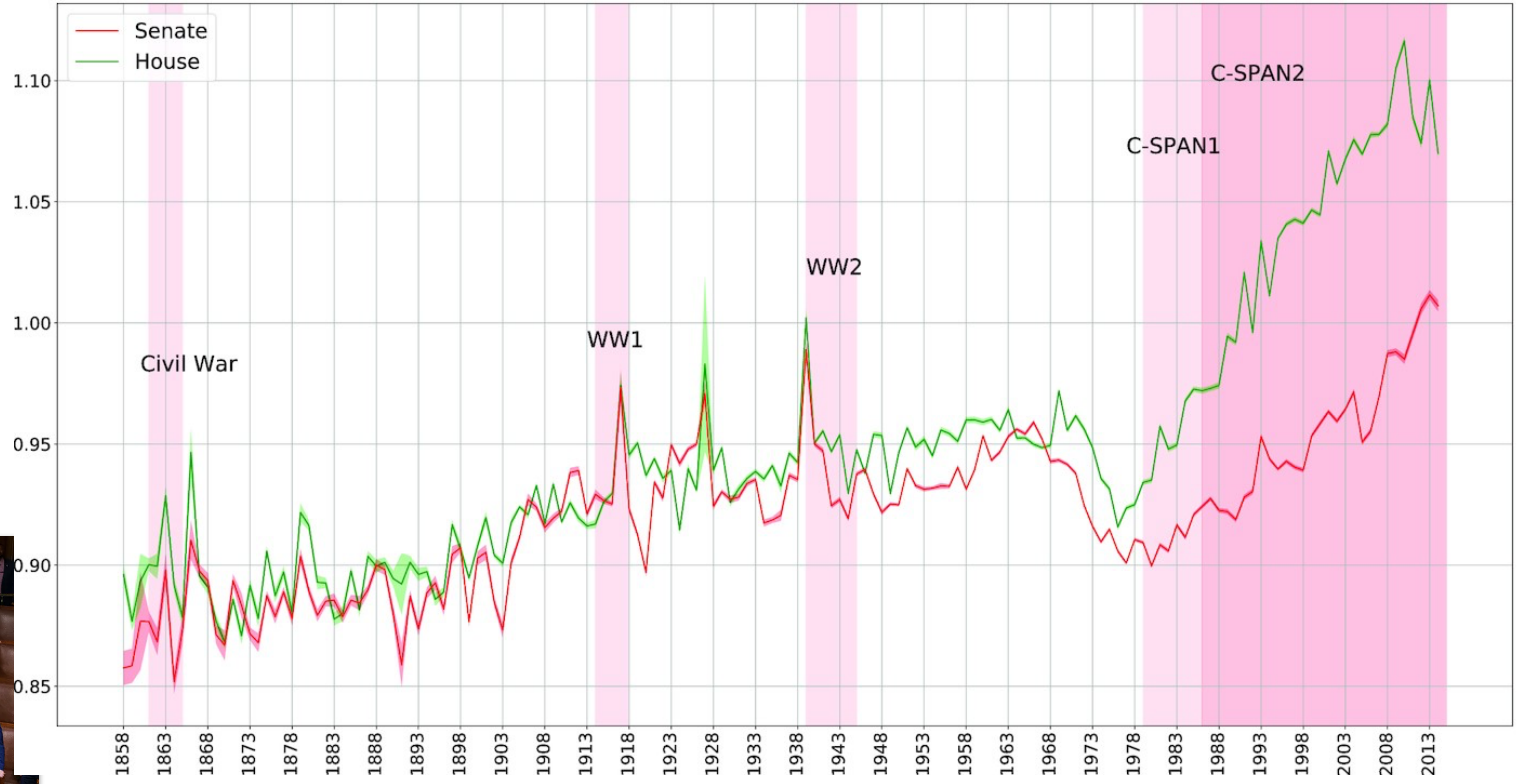
Theory	
Emotional arousal	Logical argumentation
Pathos	Logos
Emotion and Affect	Rationality and Cognition
Politicians trade off — When? Why?	



Findings

How, and by whom, is emotional language employed in US Congress debates?

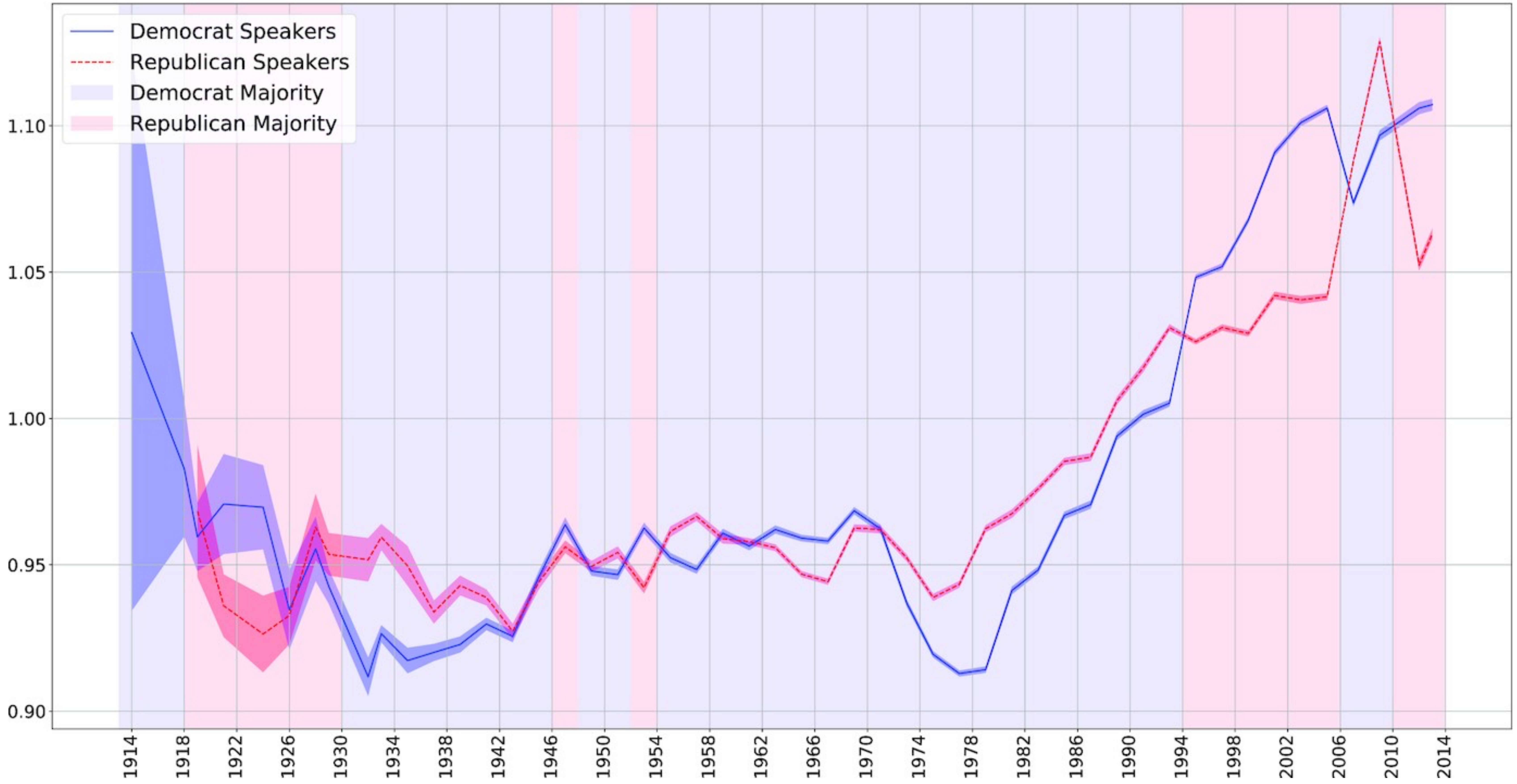
- Emotionality over time



Findings

How, and by whom, is emotional language employed in US Congress debates?

- Emotionality by party and party majority



Double checking

How, and by whom, is emotional language employed in US Congress debates?

- Could this measure accidentally be measuring positive v/s negative sentiment?
  - No! They run the same experiment with positive and negative words, and find that they are *not correlated*

(a) Cognitive Positive Language (b) Emotional Positive Language

(c) Cognitive Negative Language (d) Emotional Negative Language

Double checking

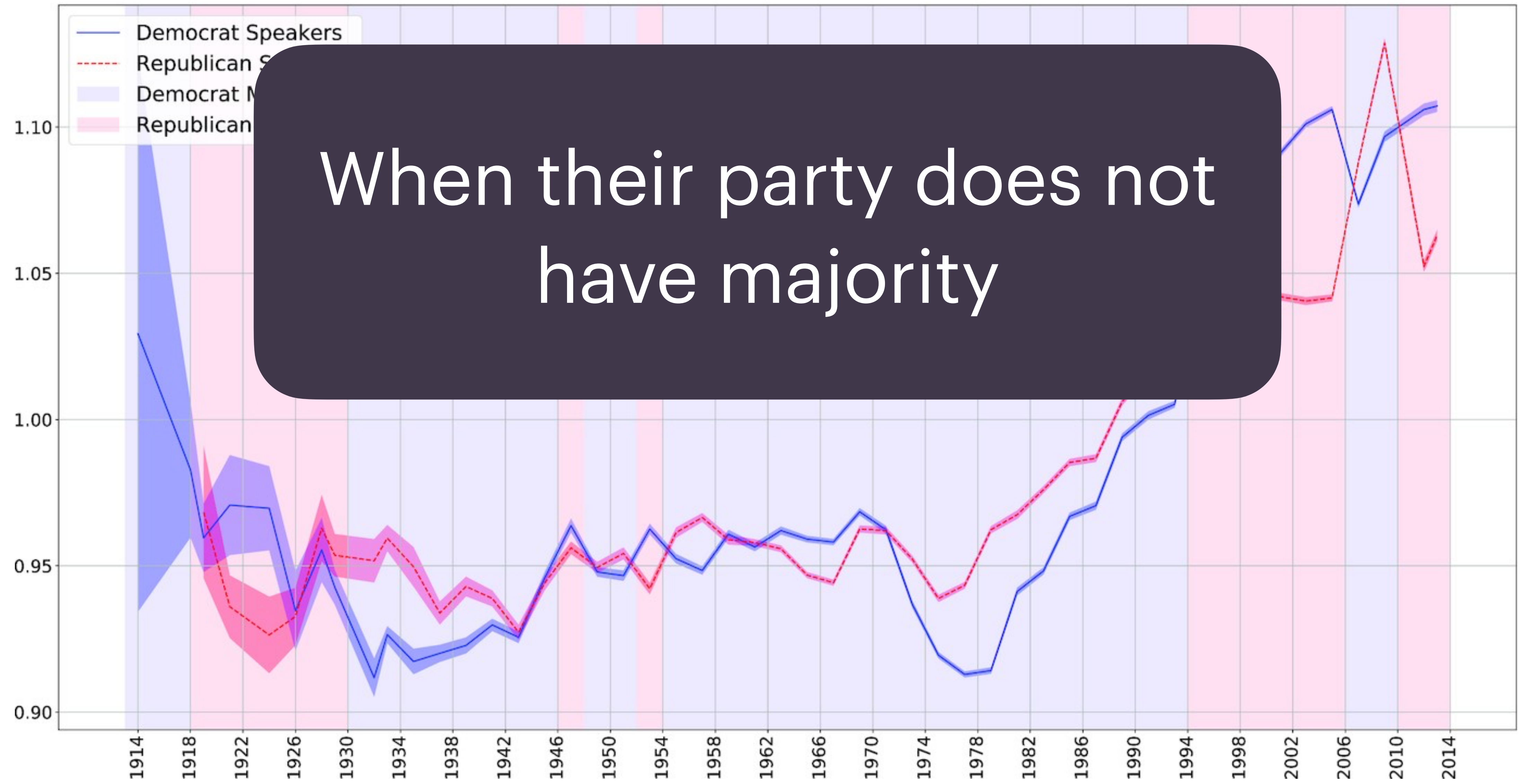
How, and by whom, is emotional language employed in US Congress debates?

- Is this general language change, rather than something specific occurring in politics?
  - No! They run the same experiment for Google Books and find emotionality *decreasing*
- Is this the same as polarization? (Different parties gravitating to different topics)
  - No! Prior work has found polarization, but *starting in the 1990s*

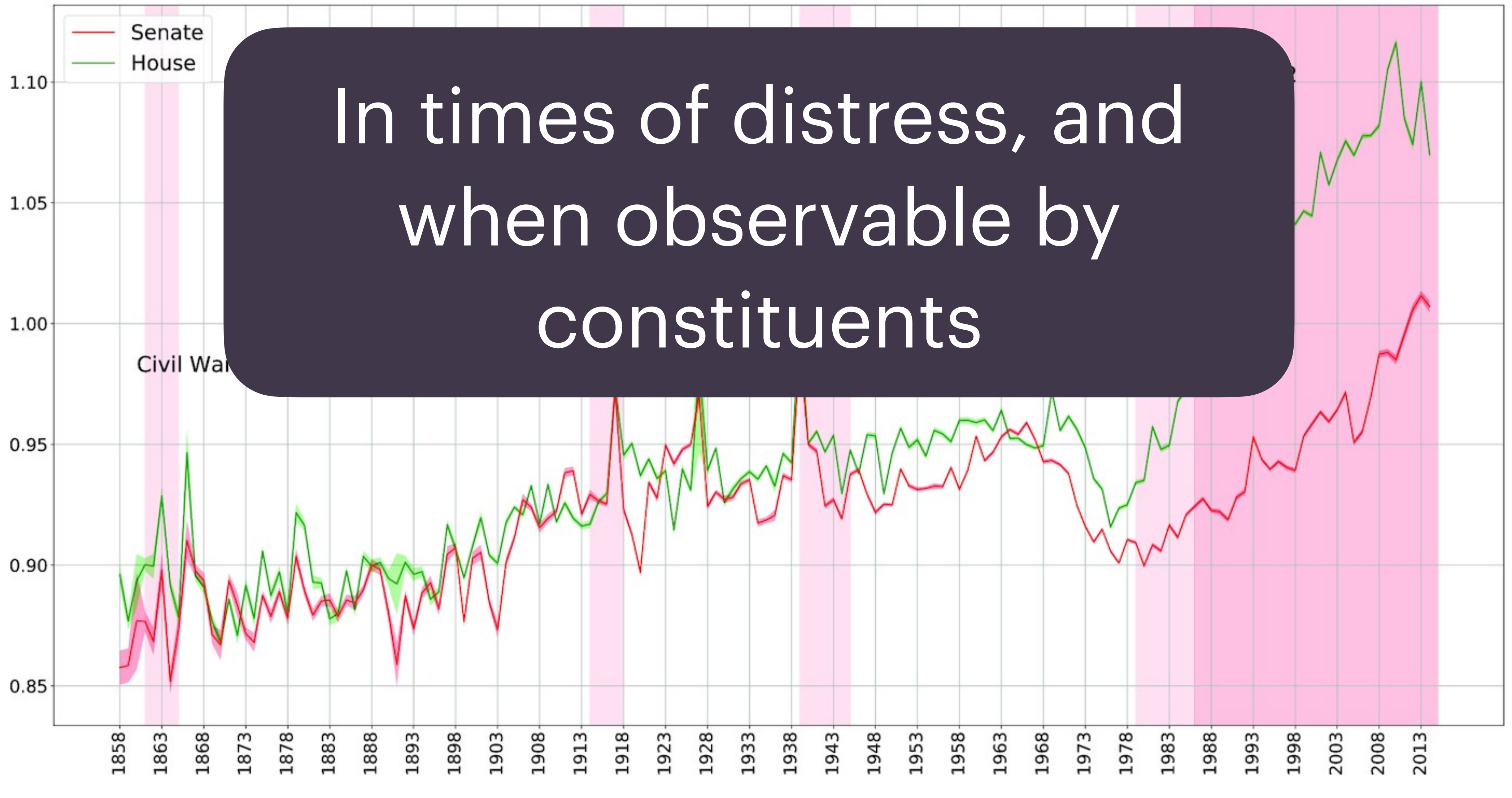
Theory	
Emotional arousal	Logical argumentation
Pathos	Logos
Emotion and Affect	Rationality and **Cognition**
Politicians trade off — When? Why?	

Conclusion

- The extent to which politicians engage with this trade-off … **is largely unknown**.
- But, in the US Congress, we can say quantitatively using **static word embeddings** that politicians employ emotional language…



When their party does not have majority



In times of distress, and when observable by constituents

Gloria Gennaro, Elliott Ash, Emotion and Reason in Political Language, *The Economic Journal*, Volume 132, Issue 643, April 2022, Pages 1037–1059, https://doi.org/10.1093/ej/ueab104

Once there's a working GitHub repo out there, are computer scientists of any use?

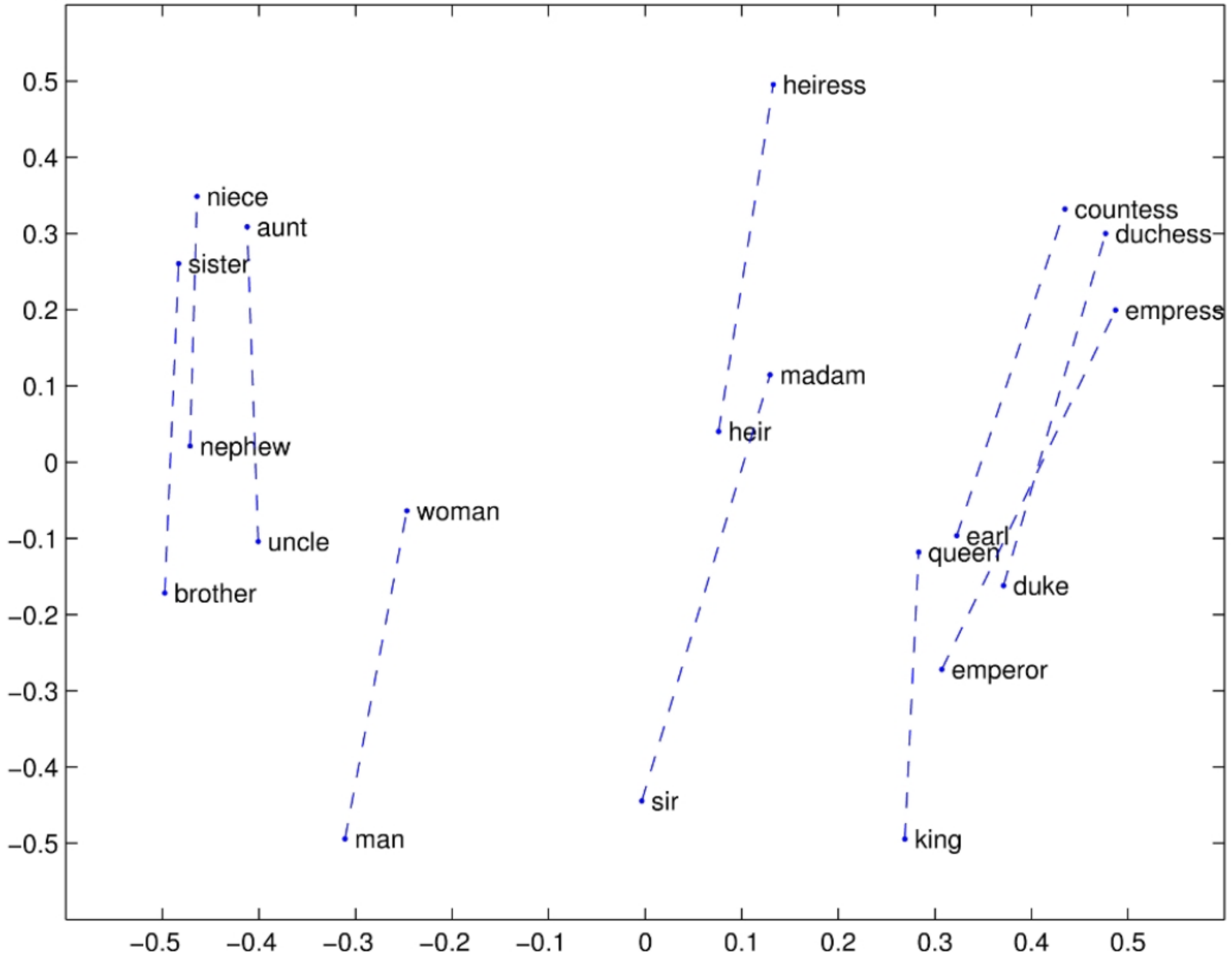
Scenario

- Observation: $$\overrightarrow{CocaCola} - \overrightarrow{rich} + \overrightarrow{poor} \approx \overrightarrow{Pepsi}$$
- Claim: Pepsi is the ‘poor people version’ of Coca Cola.

Any other possible explanations?

Semantic properties on embeddings

Caveats: only seems to work for frequent words, small distances and certain relations, like relating countries to capitals, or parts of speech. [Linzen 2016, Gladkova et al. 2016, Ethayarajh et al. 2019a]



Scenario

Pepsi is the 'poor people version' of Coca Cola?

$$\overrightarrow{CocaCola} - \overrightarrow{rich} + \overrightarrow{poor} \approx \overrightarrow{Pepsi}$$

- Linzen 2016:
  - Rich and poor are antonyms:
    - They also occur in similar contexts
    - Their vectors are very similar
    - Their vector difference is small and noisy $\overrightarrow{poor} - \overrightarrow{rich} \approx \overrightarrow{\epsilon}$
  - Pepsi and CocaCola are practically synonyms
    - They occur in very similar contexts
    - You might find that Pepsi is CocaCola's nearest neighbor
    - It's also the nearest neighbor of $\overrightarrow{CocaCola} + \overrightarrow{\epsilon}$

Distributed representations of words and phrases and their compositionality 42655 2013
T Mikolov, I Sutskever, K Chen, GS Corrado, J Dean
Neural information processing systems

Glove: Global vectors for word representation 39354 2014
J Pennington, R Socher, CD Manning
Proceedings of the 2014 conference on empirical methods in natural language …

Issues in evaluating semantic spaces using word analogies 178 2016
T Linzen
Proceedings of the First Workshop on Evaluating Vector Space Representations …

Analogy-based detection of morphological and semantic relations with word embeddings: what works and what doesn't. 260 2016
A Gladkova, A Drozd, S Matsuoka
Proceedings of the NAACL Student Research Workshop, 8-15

Caveats: only seems to work for frequent words, small distances and certain relations, like relating countries to capitals, or parts of speech. [Linzen 2016, Gladkova et al. 2016, Ethayarajh et al. 2019a]

Conclusion

- Best practices evolve quickly
- Computer scientists are well positioned to keep updated on these things
- Researchers aren’t incentivized to be explicit about shortcomings

- Our community is presenting technologies with certain promises
- Some parts of our community should help responsibly contextualize those

Isn’t it easier to ask ChatGPT?

ChatGPT baffles users by speaking 'Spanglish' as AI goes rogue

Reports of OpenAI-owned chatbot talking gibberish emerge on social media

Matthew Field

21 February 2024 • 11:43am

ChatGPT Has Gone Berserk, Giving Nonsensical Responses Right Now

OpenAI's ChatGPT is experiencing an unusual bug, spouting gibberish responses to user queries all night.

By **Maxwell Zeff** Published 36 minutes ago | Comments (2)

- Providing empirical evidence on these questions has been difficult due to the lack of a **reproducible, validated and scalable** measure of emotionality in political language.

Even word embeddings don't work out of the box

- Different types of data: much smaller corpora, words of interest might be rare
- Different end uses: Rather than just describing, make complex statistical statements

American Political Science Review (2023) 117, 4, 1255–1274
doi:10.1017/S0003055422001228

Embedding Regression: Models for Context-Specific Description and Inference

PEDRO L. RODRIGUEZ *New York University, United States*
ARTHUR SPIRLING *New York University, United States*
BRANDON M. STEWART *Princeton University, United States*

Why not LLMs?

- Different types of data: much smaller corpora, obscure and new

Automated Annotation with Generative AI Requires Validation

Nicholas Pangakis,[*] Samuel Wolken,[†] and Neil Fasching[‡]

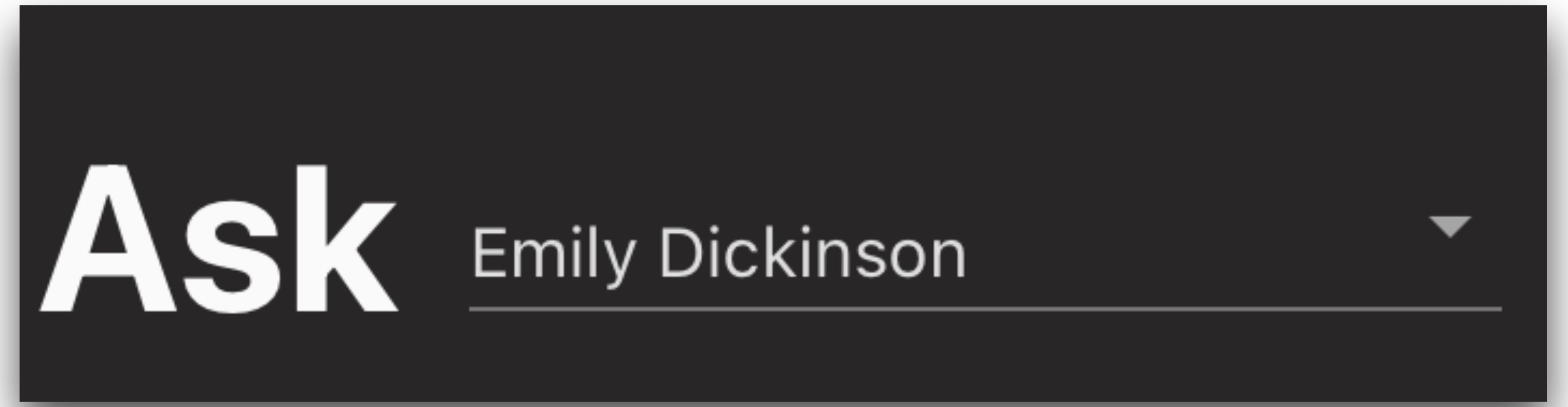
June 2, 2023

Metric	Minimum	25th percentile	Mean	Median	75th percentile	Maximum
Accuracy	0.674	0.808	0.855	0.85	0.905	0.981
Precision	0.033	0.472	0.615	0.650	0.809	0.957
Recall	0.25	0.631	0.749	0.829	0.899	0.982
F1	0.059	0.557	0.660	0.707	0.830	0.969

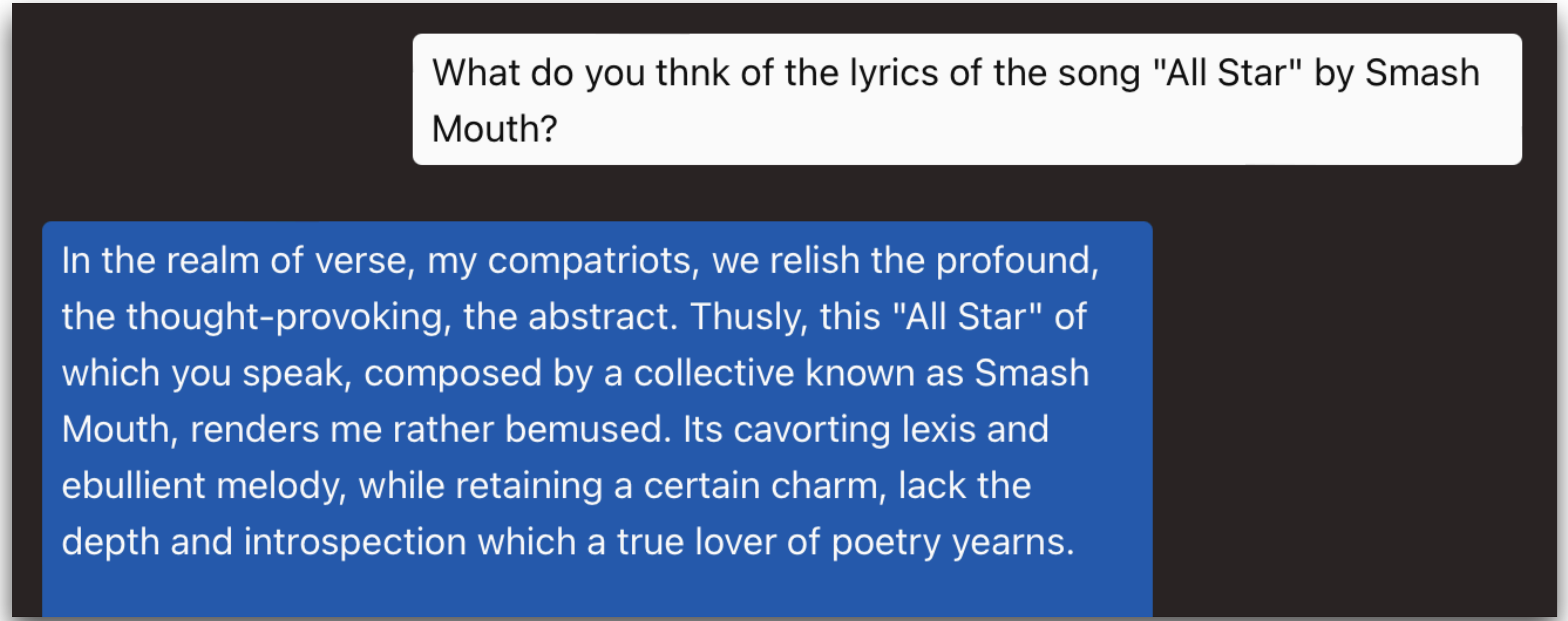
Table 1: LLM classification performance across 27 tasks from 11 datasets.

Why not LLMs?

Source criticism



- Something they are known to have said
- Expert reconstruction of what they might have said
- A fictional account?
- A reconstruction by a non-expert with very particular biases?



Laura K. Nelson
@LauraNelson

Dec 22, 2022

@alex @TedUnderwood @emilymbender I guess I don't see the leap from "we don't know enough about them to know the perspective" to "we need to assume a hegemonic view from nowhere." The hegemonic view *is* a view from somewhere. And that can tell us a lot about society. Maybe we start there?

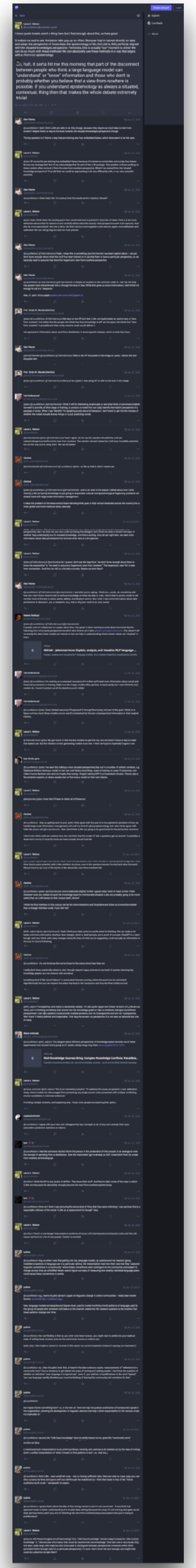
1

Alex Hanna
@alex@dair-community.social

Dec 22, 2022

@LauraNelson @TedUnderwood @emilymbender I see what you're saying. I think one _could_ do something with that, but I don't know what it'd tell us without knowledge of what the data is. Like, I don't need to prod a model to tell me that most of the text is racist, sexist, ableist, and Western-centric. But I wish I had more information about data provenance to discuss it _as_ a viewpoint. (e.g. this is why your work is so cool, Laura)

1



There's a lot of work to be done!

- Social science research imposes different and interesting constraints on NLP algorithms
- These require additional work on and around the tools we present to social scientists
  - These are questions about NLP *models* that help *others* answer questions about people and their interactions

Thank you!