Natural Language Processing

Info 159/259
Lecture 22: Social NLP (April 13, 2023)

David Bamman, UC Berkeley

Info 259
Project presentations

- 2-3:30pm Thursday 4/27 (last day of class) — on Zoom.
- Prepare a 5-minute presentation of your project to present to the class; be prepared to take questions from the audience.
- The project presentations won’t be recorded.

great

We are met on a great battle-field of that war.

λx.λy.meet(x,y) **λx.battlefield(x)** **λx.war(x)**

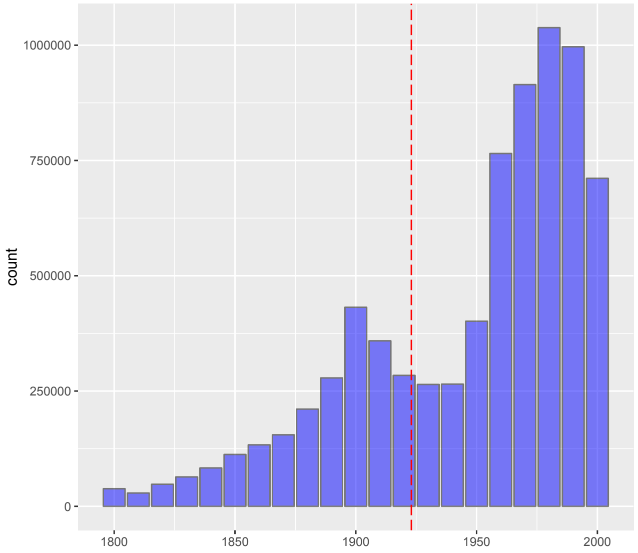
We are met on a great battle-field of that war.

Four score and seven years ago our fathers brought forth on this continent, a new nation, conceived in Liberty, and dedicated to the proposition that all men are created equal. Now we are engaged in a great civil war, testing whether that nation, or any nation so conceived and so dedicated, can long endure. We are met on a great battle-field of that war. We have come to dedicate a portion of that field, as a final resting place for those who here gave their lives that that nation might live. It is altogether fitting and proper that we should do this. But, in a larger sense, we can not dedicate -- we can not consecrate -- we can not hallow -- this ground. The brave men, living and dead, who struggled here, have consecrated it, far above our poor power to add or detract. The world will little note, nor long remember what we say here, but it can never forget what they did here. It is for us the living, rather, to be dedicated here to the unfinished work which they who fought here have thus far so nobly advanced. It is rather for us to be here dedicated to the great task remaining before us -- that from these honored dead we take increased devotion to that cause for which they gave the last full measure of devotion -- that we here highly resolve that these dead shall not have died in vain -- that this nation, under God, shall have a new birth of freedom -- and that government of the people, by the people, for the people, shall not perish from the earth.

Social NLP

- Social NLP covers a range of applications that analyze how language interacts with people in social settings.
- We leave behavioral traces in our interactions with others.
    - Social media
    - Books
    - Emails
    - Audio transcripts

Data

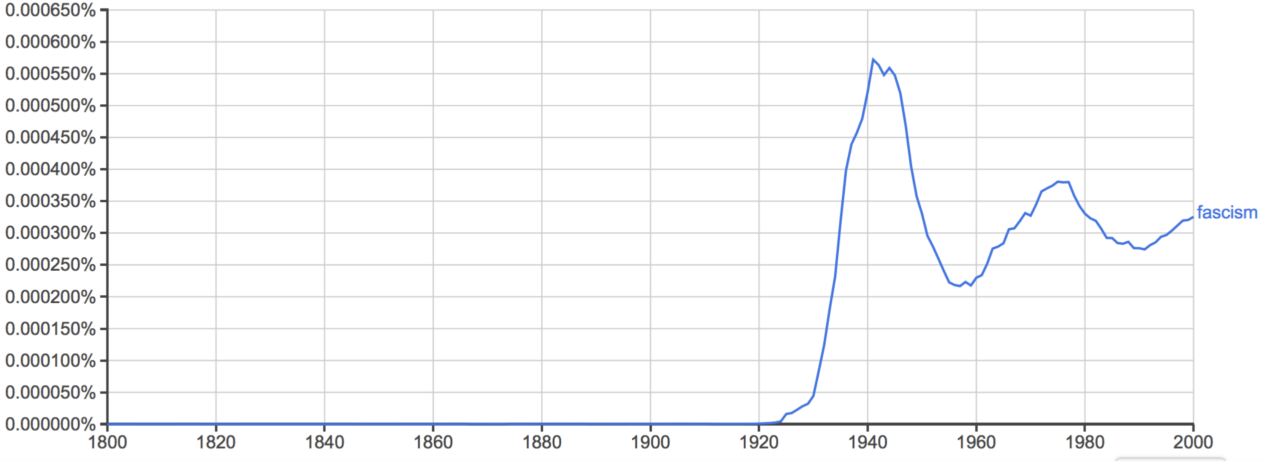


HathiTrust: 18.4M books

Affordances

Large-scale analysis of culture

(*as recorded in published books, with university library accession policies)



“Raw” data

- Social NLP often makes claims about the world using textual data.
- Data is not self-evident, neutral or objective
- Data is collected, stored, processed, mined, interpreted; each stage requires our participation.
- What is the process by which the data you have got to you?

Gitelman and Jackson (2013)

Data Collection

- Data → Research Question
    - “Opportunistic data”
    - Research questions are shaped by what data you can find
- Research Question → Data
    - Research is driven by questions, find data to support answering it.

Social NLP

- What are the research questions that we can ask when applying NLP to text to answer social and cultural questions?
- How do we answer those research questions using methods we've learned about?
    - data
    - algorithms
    - evaluation

Social NLP

- Manifestations of power in text
- Inferring social networks from text
- Measuring respect
- Examining gender bias in novels
- Asking for a favor
- Discovering structure in birth stories

Power

How is power manifested in language?

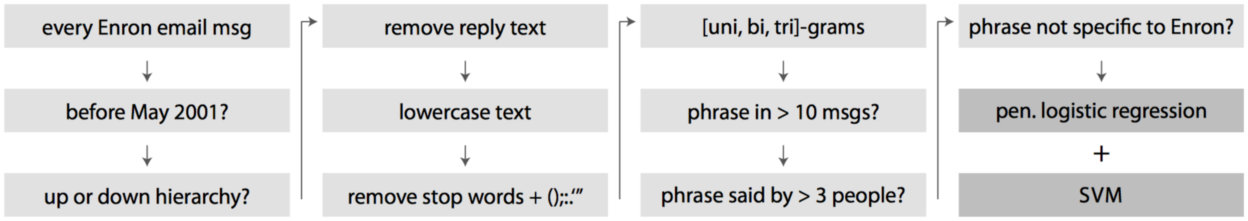
Power

- Text data: Enron emails
- Response: Enron org chart — for all pairs of entities in email (sender/recipients), who is higher on the org chart?

Gilbert 2012 (“Phrases that signal workplace hierarchy”)

Power

- Bag of words representation of text + binary classification.



Gilbert 2012 (“Phrases that signal workplace hierarchy”) these predict message going up

↑ phrases	β	↑ phrases	β
the ability to	6.76	attach	6.72
I took	6.57	that we might	6.54
are available	6.52	the calendar	6.06
kitchen	5.72	can you get	5.72
thought you would	5.65	driving	5.61
, I'll be	5.51	thoughts on	5.51
looks fine	5.50	shit	5.45
voicemail	5.43	we can talk	5.41
tremendous	5.27	it does	5.21
will you	5.17	involving	5.15
left a	5.07	the report	5.04
I put	4.90	please change	4.88
you ever	4.80	issues I	4.76
I'll give	4.69	is really	4.65
okay ,	4.60	your review	4.56
to send it	4.48	europe	4.45
communications	4.38	weekend .	4.35
a message	4.35	have our	4.33
one I	4.28	interviews	4.28
can I get	4.28	you mean	4.26
worksheet	4.15	haven't been	4.10
liked	4.07	me . 1	4.07
I gave you	3.95	tiger	3.94
credit will	3.88	change in	3.88
you make	3.86	item	3.84
together and	3.82	a decision	3.82
have presented	3.78	a discussion	3.74
think about	3.71	sounds good	3.65

not going up

↔↓ phrases	β	↔↓ phrases	β
have you been	-8.46	to manage the	-6.66
you gave	-6.64	let's discuss	-5.72
we are in	-5.44	publicly	-5.24
title	-5.05	promotion	-5.02
need in	-4.80	good one	-4.62
opened	-4.57	determine the	-4.47
initiatives	-4.38	is difficult	-4.36
. I would	-4.34	man	-4.26
we will probably	-4.12	number we	-4.11
any comments	-4.06	contact you	-4.05
you said	-3.99	the problem is	-3.97
I left	-3.88	you did	-3.78
can you help	-3.68	cool	-3.54
send this	-3.47	your attention	-3.44
whether we	-3.44	to think	-3.44
the trade	-3.40	addition to the	-3.30
and I thought	-3.28	great thanks	-3.24
should include	-3.19	selected	-3.16
please send	-3.14	ext	-3.13
existing	-3.06	and let me	-3.05
mondays	-3.02	security	-3.01
presentation on	-2.95	got the	-2.94
let's talk	-2.94	get your	-2.88
the items	-2.78	this week and	-2.77
i hope you	-2.77	team that	-2.75
did it	-2.75	a deal	-2.71
test	-2.69	yours .	-2.68
be sure	-2.65	briefing	-2.60

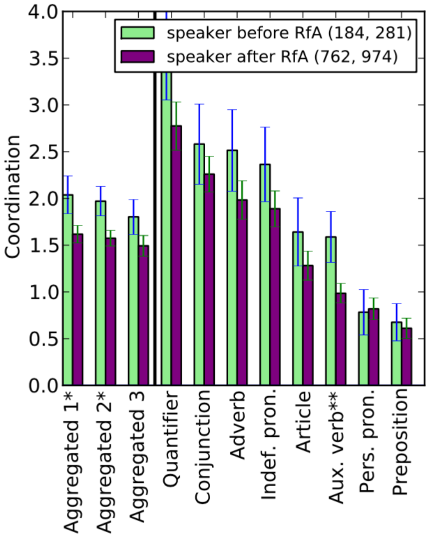
Power

- Text data: Wikipedia discussions, SCOTUS arguments
- Response: Wikipedia admins/non-admins; SCOTUS justices/laywers

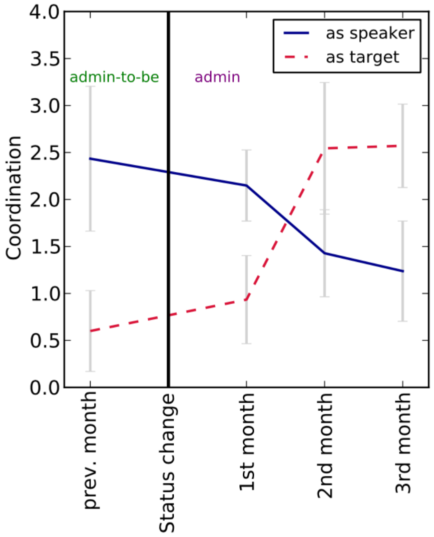
Danescu-Niculescu-Mizil et al. 2012 (“Echoes of Power”)

Power

- LIWC representation of text + measurements of accommodation (adapting your speech to the language of your interlocutor)



(a) Supporting $\mathcal{P}'_{speaker}$



(b) Timed effect of **status change** ($\mathcal{P}$)

Networks

Can we learn a social network from mentions in text?

Networks

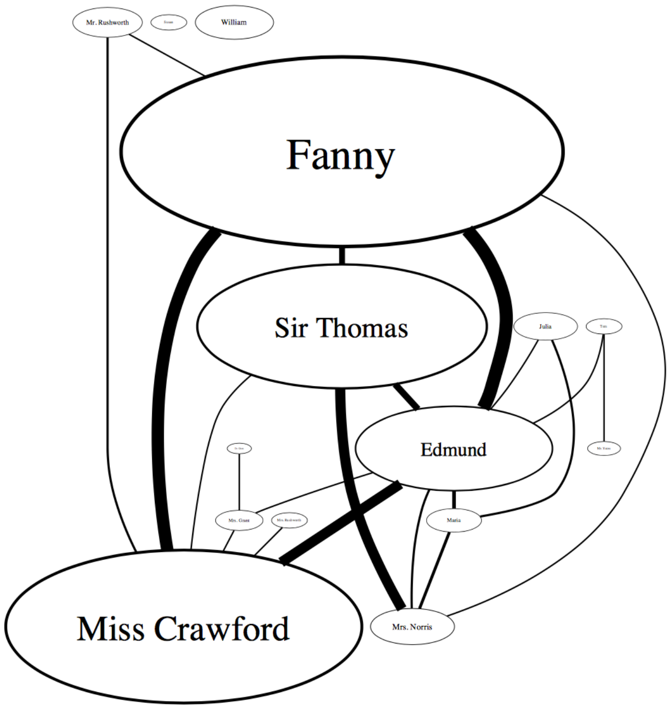
- Data: 60 novels from Project Gutenberg
- Conversational network:
    - Characters are in the same place at the same time
    - Characters take turns speaking
    - The characters are mutually aware of each other and each character's speech is mutually intended for the other to hear.

Elson et al. (2010), "Extracting Social Networks from Literary Fiction"

Networks

- Alias clustering: Tom, Tom Sawyer, Mr. Saywer = TOM SAWYER)
- Quoted speech attribution (“Yes,” said TOM SAWYER)
- Network construction
  - Divide book into 10-paragraph sections, count number of sections with two characters
  - Count occurrences of one character mentioning another in dialogue

Networks


Mr. Rushworth
William
Fanny
Sir Thomas
Julia
Edmund
Mrs. Grant
Maria
Miss Crawford
Mrs. Norris

Respect

- Data: transcripts of 981 OPD traffic stops (everyday interactions)
- Response: race

Voigt et al. 2017, “Language from police body camera footage shows racial disparities in officer respect”

Respect

- Present one dialogue turn (police/driver) to be rated by people for respect (4-point Likert scale). High IAA.

- Build a predictive model mapping text to respect.

Voigt et al. 2017, “Language from police body camera footage shows racial disparities in officer respect”

<table>
<tr><th>Feature Name</th><th>Implementation</th></tr>
<tr><td>Adverbial "Just"</td><td>"Just" occurs in a dependency arc as the head of an advmod relation</td></tr>
<tr><td>Apologizing</td><td>Lexicon: "sorry", "oops", "woops", "excuse me", "forgive me", "apologies", "apologize", "my bad", "my fault"</td></tr>
<tr><td>Ask for Agency</td><td>Lexicon: "do me a favor", "let me", "allow me", "can i", "should i", "may i", "might i", "could i"</td></tr>
<tr><td>Bald Command</td><td>The first word in a sentence is a bare verb with part-of-speech tag VB ("look", "give", "wait" etc.) but is not one of "be", "do", "have", "thank", "please", "hang".</td></tr>
<tr><td>Colloquialism</td><td>Regular expression capturing "y'all", "ain't" and words ending in "in'" such as "walkin'", "talkin'", etc., as marked by transcribers</td></tr>
<tr><td>Conditional</td><td>Lexicon: "if"</td></tr>
<tr><td>Disfluency</td><td>Word fragment ("Well I thi-") as indicated by transcribers</td></tr>
<tr><td>Filled Pauses</td><td>Lexicon: "um", "uh"</td></tr>
<tr><td>First Names</td><td>Top 1000 most common first names from the 1990 US Census, where first letter is capitalized in transcript</td></tr>
<tr><td>Formal Titles</td><td>Lexicon: "sir", "ma'am", "maam", "mister", "mr*", "ms*", "madam", "miss", "gentleman", "lady"</td></tr>
<tr><td>For Me</td><td>Lexicon: "for me"</td></tr>
<tr><td>For You</td><td>Lexicon: "for you"</td></tr>
<tr><td>Give Agency</td><td>Lexicon: "let you", "allow you", "you can", "you may", "you could"</td></tr>
<tr><td>Gratitude</td><td>Lexicon: "thank", "thanks", "appreciate"</td></tr>
<tr><td>Goodbye</td><td>Lexicon: "goodbye", "bye", "see you later"</td></tr>
<tr><td>Hands on the Wheel</td><td>Regular expression capturing cases like "keep your hands on the wheel" and "leave your hands where I can see them": "hands? ([.,?!:;]+)?(wheel|see)"</td></tr>
</table>

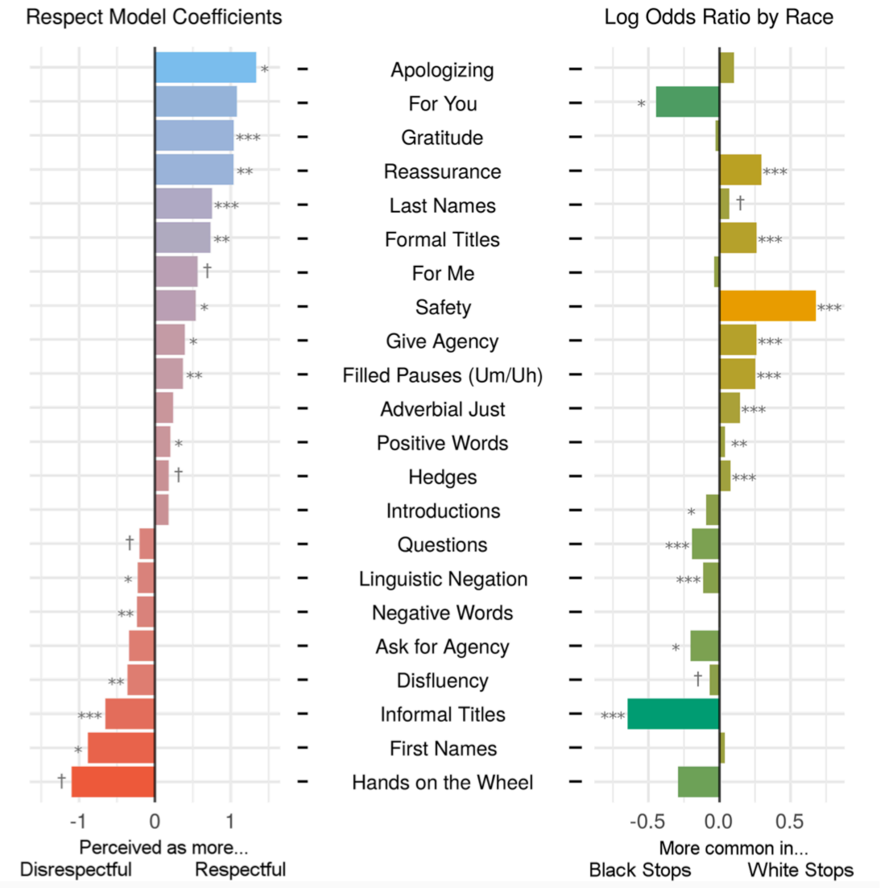
Voigt et al. 2017, "Language from police body camera footage shows racial disparities in officer respect"

<table>
<tr><td>Hedges</td><td>All words in the "Tentat" LIWC lexicon</td></tr>
<tr><td>Impersonal Pronoun</td><td>All words in the "Imppron" LIWC lexicon</td></tr>
<tr><td>Informal Titles</td><td>Lexicon: "dude*", "bro*", "boss", "bud", "buddy", "champ", "man", "guy*", "guy", "brotha", "sista", "son", "sonny", "chief"</td></tr>
<tr><td>Introductions</td><td>Regular expression capturing cases like "I'm Officer [name] from the OPD" and "How's it going?": "((i|my name).+officer | officer.+(oakland|opd))|((hi|hello|hey|good afternoon|good morning|good evening|how are you doing|how 's it going))"</td></tr>
<tr><td>Last Names</td><td>Top 5000 most common last names from the 1990 US Census, where first letter is capitalized in transcript</td></tr>
<tr><td>Linguistic Negation</td><td>All words in the "Negate" LIWC lexicon</td></tr>
<tr><td>Negative Words</td><td>All words in the "Negativ" category in the Harvard General Inquierer, matching on word lemmas</td></tr>
<tr><td>Positive Words</td><td>All words in the "Positiv" category in the Harvard General Inquierer, matching on word lemmas</td></tr>
</table>

<table>
<tr><td>Please</td><td>Lexicon: "please"</td></tr>
<tr><td>Questions</td><td>Occurrence of a question mark</td></tr>
<tr><td>Reassurance</td><td>Lexicon: "'s okay", "n't worry", "no big deal", "no problem", "no worries", "'s fine", "you 're good", "is fine", "is okay"</td></tr>
<tr><td>Safety</td><td>Regular expression for all words beginning with the prefix "safe", such as "safe", "safety", "safely"</td></tr>
<tr><td>Swear Words</td><td>All words in the "Swear" LIWC lexicon</td></tr>
<tr><td>Tag Question</td><td>Regular expression capturing cases like "..., right?" and "..., don't you?": ", (((all right|right|okay|yeah|please|you know)(sir| ma'am| miss| son)?)|((are|is|do|can|have|will|won't) (n't)?(i|me|she|us|we|you|he|they|them))) [?]"</td></tr>
<tr><td>The Reason for the Stop</td><td>Lexicon: "reason", "stop* you", "pull* you", "why i", "why we", "explain", "so you understand"</td></tr>
<tr><td>Time Minimizing</td><td>Regular expression capturing cases like "in a minute" and "let's get this done quick": "(a|one|a few) (minute|min|second|sec|moment)s?|this[.,?!]+quick|right back"</td></tr>
</table>

Respect

Example	Respect Score
First Name · Ask For Agency · Questions [name], can I see that driver's license again? It- it's showing suspended. Is that- that's you? Disfluency · Negative Word · Disfluency	-1.07
Informal Title · Ask For Agency · Adverbial "Just" All right, my man. Do me a favor. Just keep your hands on the steering wheel real quick. "Hands On The Wheel"	-0.51
Apology · Introduction · Last Name Sorry to stop you. My name's Officer [name] with the Police Department.	0.84
Formal Title · Safety · Please There you go, ma'am. Drive safe, please.	1.21
Adverbial "Just" · Filled Pause · Reassurance It just says that, uh, you've fixed it. No problem. Thank you very much, sir. Gratitude · Formal Title	2.07



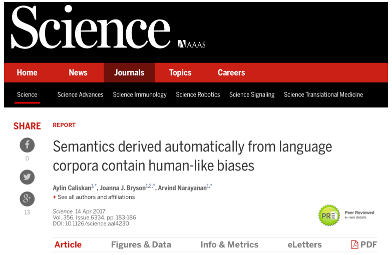
Respect Model Coefficients
Log Odds Ratio by Race
Apologizing
For You
Gratitude
Reassurance
Last Names
Formal Titles
For Me
Safety
Give Agency
Filled Pauses (Um/Uh)
Adverbial Just
Positive Words
Hedges
Introductions
Questions
Linguistic Negation
Negative Words
Ask for Agency
Disfluency
Informal Titles
First Names
Hands on the Wheel
-1
0
1
Perceived as more...
Disrespectful
Respectful
-0.5
0.0
0.5
More common in...
Black Stops
White Stops

Respect

- Higher respect to white drivers, older drivers, when a citation is issued.
- Lower respect when a search is conducted.

Are there differences in the depiction of men and women in fiction?

Gender bias


Science AAAS

Home News Journals Topics Careers

Science Science Advances Science Immunology Science Robotics Science Signaling Science Translational Medicine

SHARE

REPORT

Semantics derived automatically from language corpora contain human-like biases

Aylin Caliskan[1,*], Joanna J. Bryson[1,2,*], Arvind Narayanan[1,*]

+ See all authors and affiliations

Science 14 Apr 2017:
Vol. 356, Issue 6334, pp. 183-186
DOI: 10.1126/science.aal4230

Peer Reviewed

Article | Figures & Data | Info & Metrics | eLetters | PDF

- 15% of Wikipedia biographies are of women
- Women's biographies are 2.58x more likely to mention divorce, 1.57x more likely to mention marriage [Bamman and Smith 2015]
- Word embeddings encode cultural bias implicit in natural language usage [Caliskan et al. 2017; Bolukbasi et al. 2016]

Measurement

This is fundamentally a problem of measurement: how do we design an algorithmic instrument that can transform an entire novel into a quantity expressing the depiction of gender?

“TOM!” No answer. “TOM!” No answer. “What's gone with that boy, I wonder? You TOM!” No answer. The old lady pulled her spectacles down and looked over them about the room; then she put them up and looked out under them. She seldom or never looked *through* them for so small a thing as a boy; they were her state pair, the pride of her heart, and were built for “style,” not service--she could have seen through a pair of stove-lids just as well. She looked perplexed for a moment, and then said, not fiercely, but still loud enough for the furniture to hear: “Well, I lay if I get hold of you I'll--” She did not f time she was bending down and punching under the bed with 0.53 so she needed breath to punctuate the punches with. She hing but the cat. “I never did see the beat of that boy!” She went to the open door and stood in it and looked out among the tomato vines and “jimpson” weeds that constituted the garden. No Tom. So she lifted up her voice at an angle calculated for distance and shouted: “Y-o-u-u TOM!” There was a slight noise behind her and she turned just in time to seize a small boy by the slack of his roundabout and arrest his flight. “There! I might 'a' thought of that closet. What you been doing in there?” “Nothing.” “Nothing! Look at your hands. And look at your mouth. What *is* that truck?” “I don't

Measurement

Measurement here is dependent on character. We'll infer the characters in a novel and measure:

- the amount of "screen time" they get
- what those characters do.
- [dialogue, number of characters, etc.]

Becky Thatcher

Aunt Polly

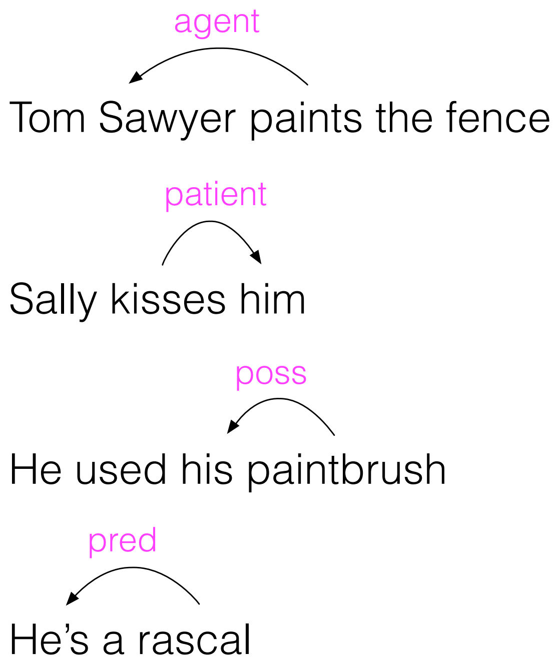
Huckleberry Finn

Tom Sawyer

Character depiction

Typed dependency information anchored on each character

Tom Sawyer		N
agent	paints, runs, …	137
patient	kisses, saw …	41
poss	paintbrush, …	30
pred	rascal, boy, …	8
total		216



Data

- 104,000 narratives from the HathiTrust Digital Library (www.hathitrust.org), published between 1703-2009.

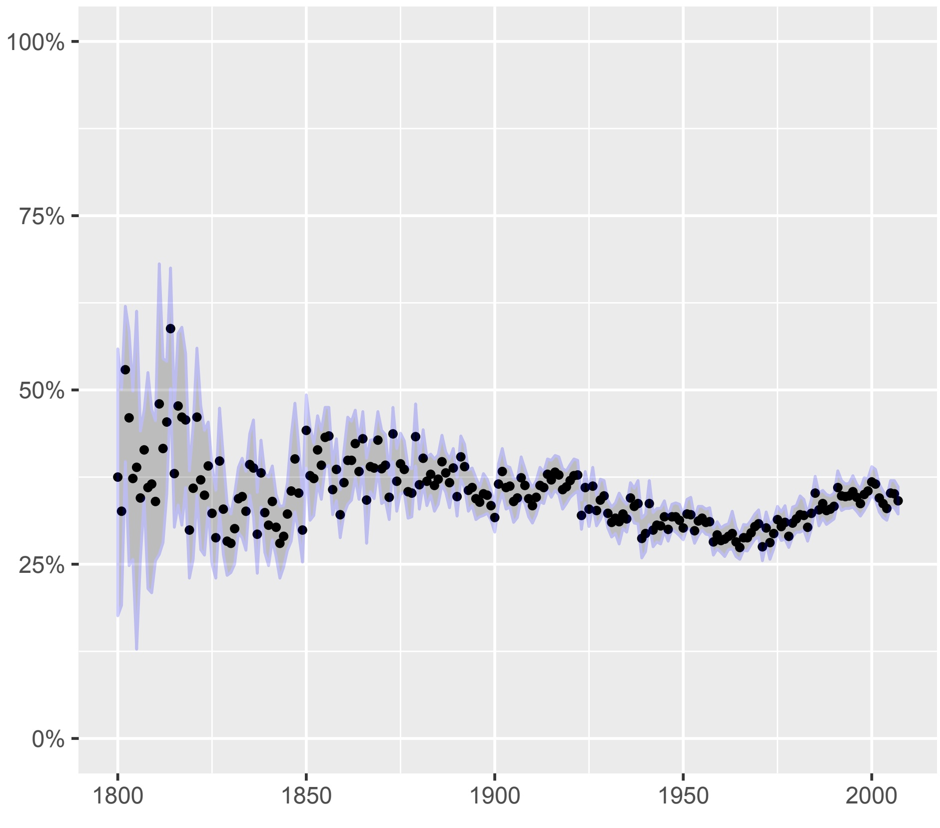
- OCR'd from page scans.

- Subject each narrative to an NLP pipeline to extract entity data.

- 10,000 narratives from the Chicago Corpus, published between 1880-1989

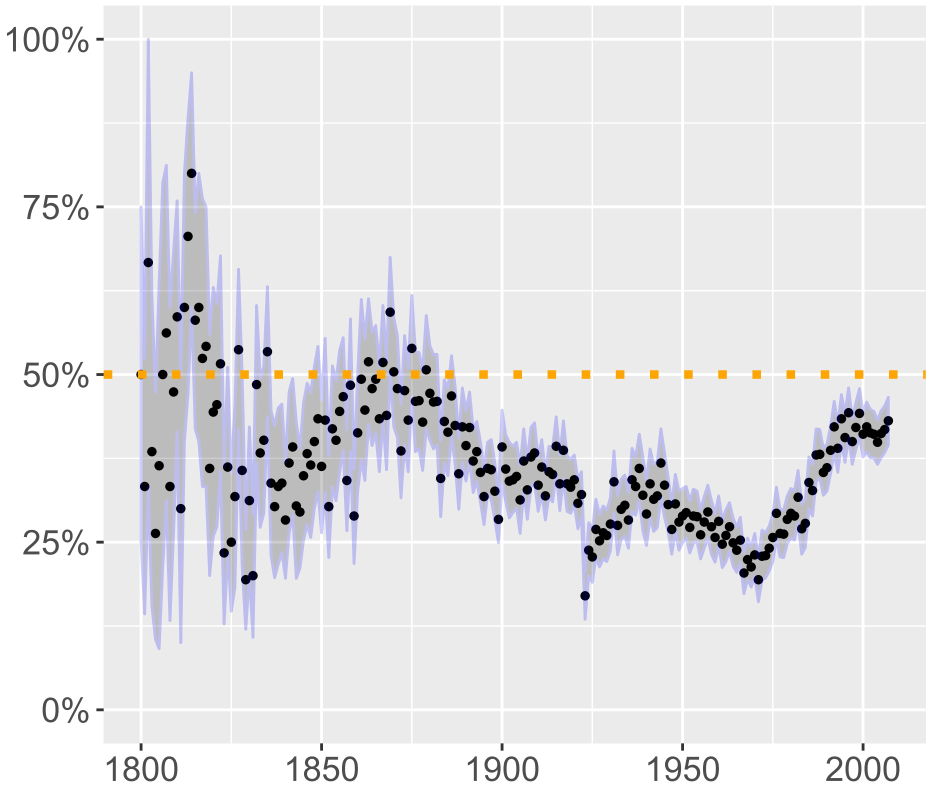
HATHI TRUST

Are there differences in the attention given to men and women in fiction?



Words about women, as a function of all characterization

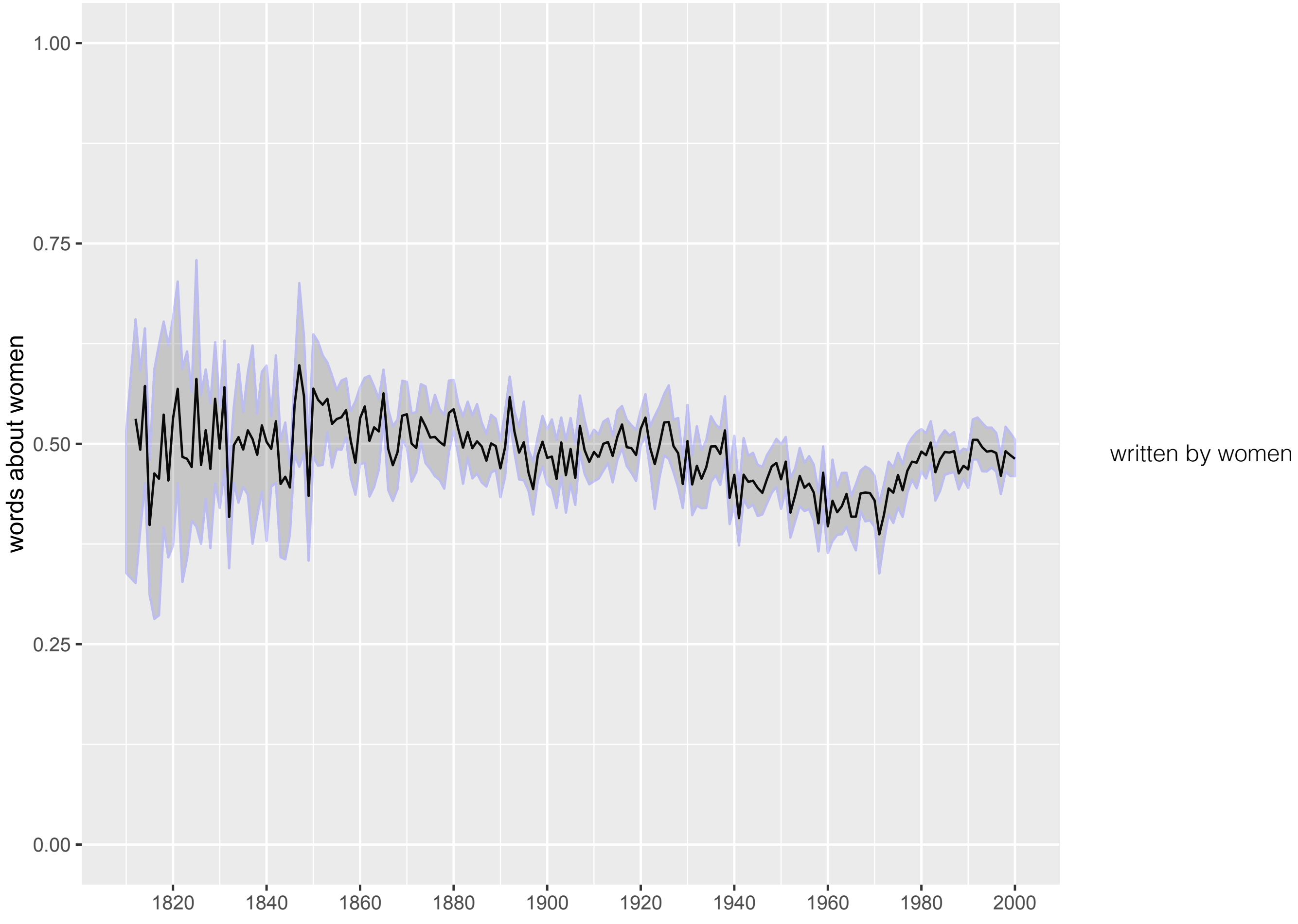
$$\frac{C(a, g = \text{female})}{C(a, g = \text{female}) + C(a, g = \text{male})}$$



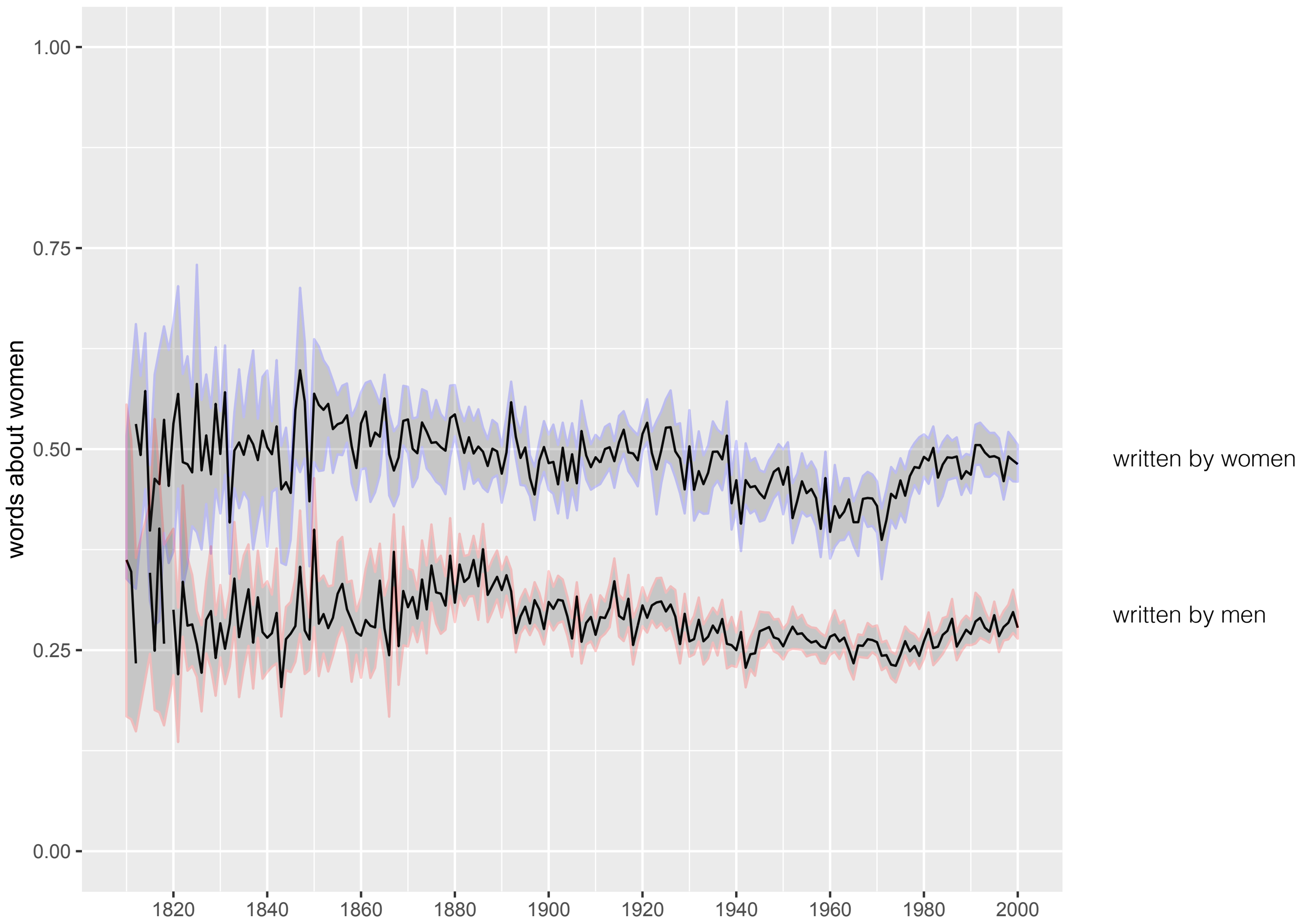
Books published by women

Tuchman and Fortin (2012), *Edging Women Out: Victorian Novelists and Social Change*

- "Before 1840 at least half of all novelists were women; by 1917 most high-culture novelists were men."
  - Early 19th c. fiction was not yet high-status career; after 1840, increasingly brought status
  - Terms of publisher's contracts became disadvantageous to women
  - Careers others than "novelist" were opening up to women



words about women
1.00
0.75
0.50
0.25
0.00
1820
1840
1860
1880
1900
1920
1940
1960
1980
2000
written by women



words about women
1.00
0.75
0.50
0.25
0.00
1820
1840
1860
1880
1900
1920
1940
1960
1980
2000
written by women
written by men

- Data: Random acts of pizza (subreddit)
- Response: Is a request successful in getting a pizza?

Althoff et al. (2014), "How to Ask for a Favor: A Case Study on the Success of Altruistic Requests

Favors

- Representation of text:
    - Temporal features of post/author
    - Politeness
    - Evidentiality (image link in text)
    - Reciprocity (“pay it forward”, “return the favor”)
    - Sentiment
    - Length
    - Status (reddit karma points)
    - “Money”, “Job” “Student”, “Family”, “Craving” topics

Coefficient	Estimate	SE
Intercept	-2.02^{***}	0.14
Community Age (Decile)	-0.13^{***}	0.01
First Half of Month (Binary)	0.22^{**}	0.08
Gratitude (Binary)	0.27^{**}	0.08
Including Image (Binary)	0.81^{***}	0.17
Reciprocity (Binary)	0.32^{**}	0.10
Strong Positive Sentiment (Binary)	0.14	0.08
Strong Negative Sentiment (Binary)	-0.07	0.08
Length (in 100 Words)	0.30^{***}	0.05
Karma (Decile)	0.13^{***}	0.02
Posted in RAOP before (Binary)	1.34^{***}	0.16
Narrative Craving (Binary)	-0.34^{***}	0.09
Narrative Family (Binary)	0.22^{*}	0.09
Narrative Job (Binary)	0.26^{**}	0.09
Narrative Money (Binary)	0.19^{**}	0.08
Narrative Student (Binary)	0.09	0.09

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Features	ROC AUC
Random Baseline	0.500
Unigram Baseline	0.621^{***}
Bigram Baseline	0.618^{***}
Trigram Baseline	0.618^{***}
Text Features	0.625^{***}
Social Features	0.576^{***}
Temporal Features	0.579^{***}
Temporal + Social	0.638^{***}
Temporal + Social + Text	$\mathbf{0.669}^{***}$
Temporal + Social + Text + Unigram	$\mathbf{0.672}^{***}$

Social NLP

- Many different methods from NLP
    - Text-based classification, regression, sequence labeling (e.g., NER)
- Representation is important:
    - Bag of words
    - Features derived from parts of speech, syntax
- When testing theories, interpretable models are important.

Social NLP

Text also provides a lens into exploratory analysis of social and cultural phenomena

What is the structure of birth stories?

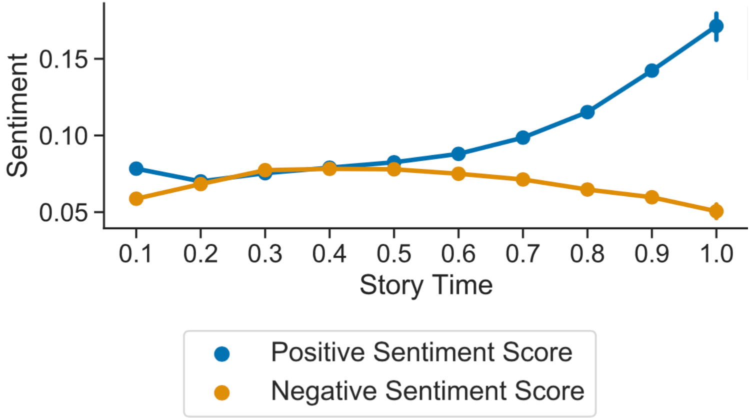
Birth stories

- Text data: 2,847 birth stories from r/BabyBumps — "narratives of individual experiences giving birth, often in great medical and emotional detail"
- Methods:
  - Topic modeling (clustering tokens in documents into coherent "topics")
  - Sentiment analysis
  - Connotation frames of power

Antoniak et al. (2019), "Narrative Paths and Negotiation of Power in Birth Stories," CSCW

Birth stories

- Dictionary-based sentiment analysis with VADER lexicon (Hutto and Gilbert 2014)



Topic Models

- A probabilistic model for discovering hidden “topics” or “themes” (groups of terms that tend to occur together) in documents.
- Unsupervised (find *interesting structure* in the data)

Blei et al. (2003), “Latent Dirichlet Allocation” (JMLR)

Topic Models

- **Input**: set of documents, number of clusters to learn.
- **Output**:
  - topics
  - topic ratio in each document
  - topic distribution for each word in doc

{album, band, music}	{government, party, election}	{game, team, player}
album	government	game
band	party	team
music	election	player
song	state	win
release	political	play

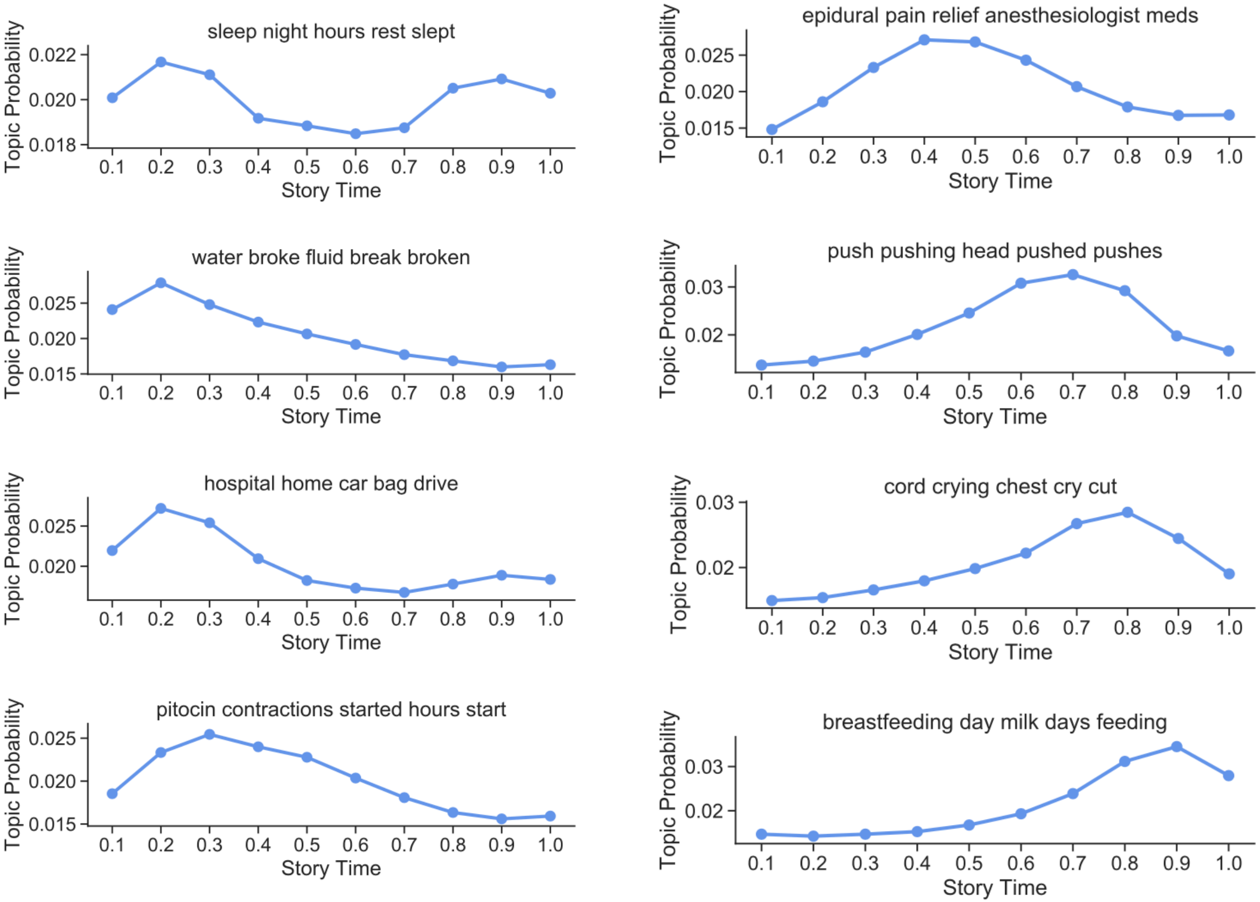
{god, call, give}	{company, market, business}	{math, number, function}
god	company	math
call	market	number
give	business	function
man	year	code
time	product	set

{city, large, area}	{math, energy, light}	{law, state, case}
city	math	law
large	energy	state
area	light	case
station	field	court
include	star	legal

Topic modeling

- Run Latent Dirichlet Allocation (LDA) on training birth stories, each divided into 100-word chunks
- 50 topics
- Divide each story into 10 chunks, plot aggregate topic distribution over narrative time.

Antoniak et al. (2019), “Narrative Paths and Negotiation of Power in Birth Stories,” CSCW



sleep night hours rest slept
Topic Probability
0.022
0.020
0.018
0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Story Time
epidural pain relief anesthesiologist meds
Topic Probability
0.025
0.020
0.015
0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Story Time
water broke fluid break broken
Topic Probability
0.025
0.020
0.015
0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Story Time
push pushing head pushed pushes
Topic Probability
0.03
0.02
0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Story Time
hospital home car bag drive
Topic Probability
0.025
0.020
0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Story Time
cord crying chest cry cut
Topic Probability
0.03
0.02
0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Story Time
pitocin contractions started hours start
Topic Probability
0.025
0.020
0.015
0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Story Time
breastfeeding day milk days feeding
Topic Probability
0.03
0.02
0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Story Time

Personas

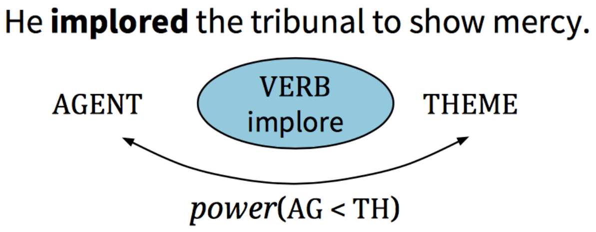
Persona	N-Grams	Total Mentions	Stories Containing Mentions	Average Mentions per Story
Author	I, me, myself	210,795	2,846	74.0
We	we, us, ourselves	24,757	2,764	8.7
Baby	baby, son, daughter	14,309	2,668	5.0
Doctor	doctor, dr, doc, ob, obgyn, gynecologist, physician	10,025	2,262	3.5
Partner	partner, husband, wife	8,998	2,006	3.2
Nurse	nurse	7,080	2,012	2.5
Midwife	midwife	4,069	886	1.4
Family	mom, dad, mother, father, brother, sister	3,490	1,365	1.2
Anesthesiologist	anesthesiologist	1,398	876	0.5
Doula	doula	896	256	0.3

Table 5. Personas identified in the birth stories collection and the n-grams used to classify the personas.

- Dictionary-based method to group word types into “personas” — e.g., partner, husband, wife → Partner

Power frames

- Verbs imply power differential between their agent/theme



$power(\text{AG}<\text{TH})$

$power(\text{AG}>\text{TH})$

Sap et al. (2017), "Connotation Frames of Power and Agency in Modern Films" (EMNLP)

Power frames

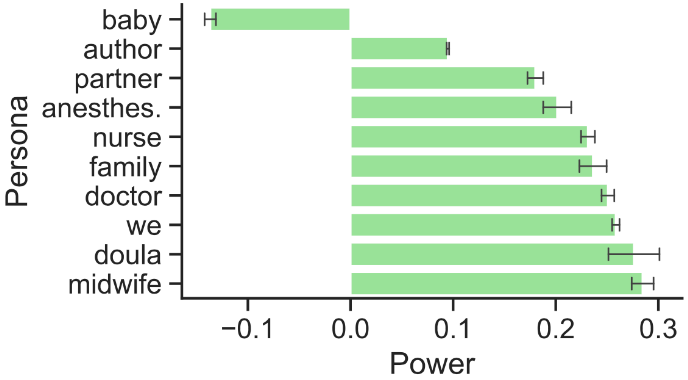
- The only time I got upset was when the nurse accused me of not feeding my child.
- The doctor broke my water.

Antoniak et al. (2019), “Narrative Paths and Negotiation of Power in Birth Stories,” CSCW

Birth stories

- The author is framed as having the least power (except for the baby).

- Clinicians are framed as having high power



Antoniak et al. (2019), "Narrative Paths and Negotiation of Power in Birth Stories," CSCW