Language Patterns in International Relations

Leah Windsor, Ph.D.
Associate Professor of Applied Linguistics
Institute for Intelligent Systems
Department of English
PI, Languages Across Cultures Lab

Institute for Intelligent Systems
THE UNIVERSITY OF MEMPHIS

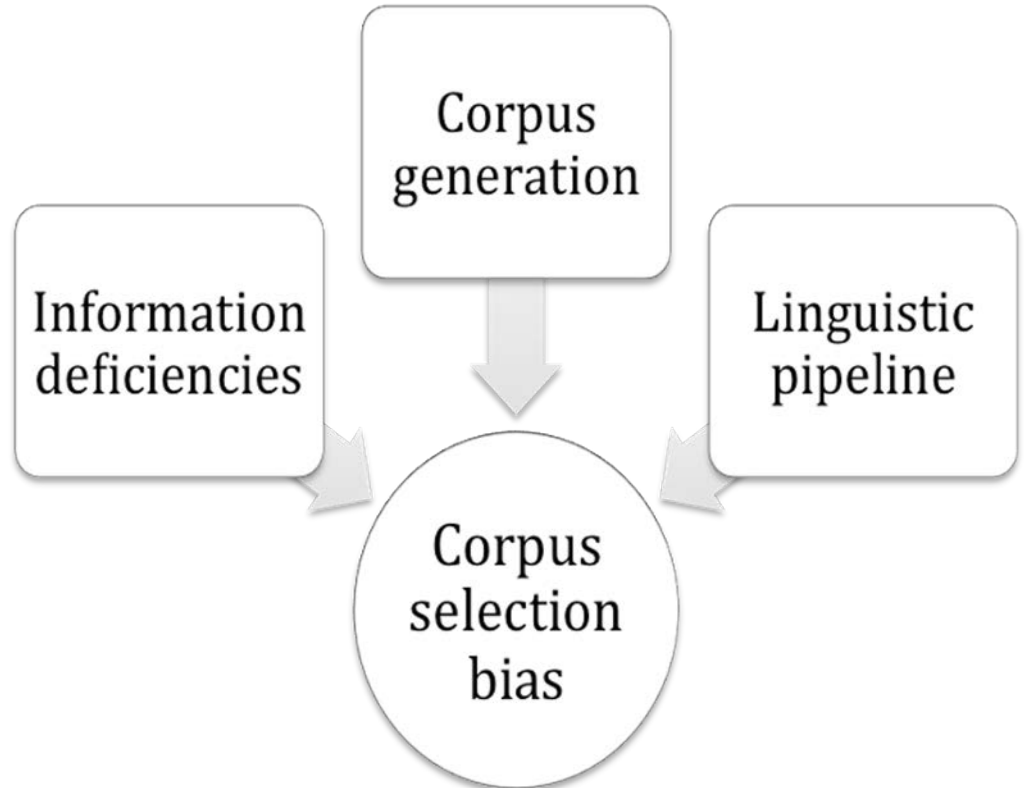THE UNIVERSITY OF MEMPHIS®

Overview

THE BIG PICTURE

How does what we say, reveal who we are?

- Is it what we say,
  or how we say it?

- What biases influence our understanding of international relations (and life), and how do we fix this?

How we learn about the world through text analysis

Research question: What's puzzling you?

"Words in the Wild"

Data Generating Process/Corpus Identification

Theoretical framework: What informs your research question?

Feature selection based on theoretical expectations
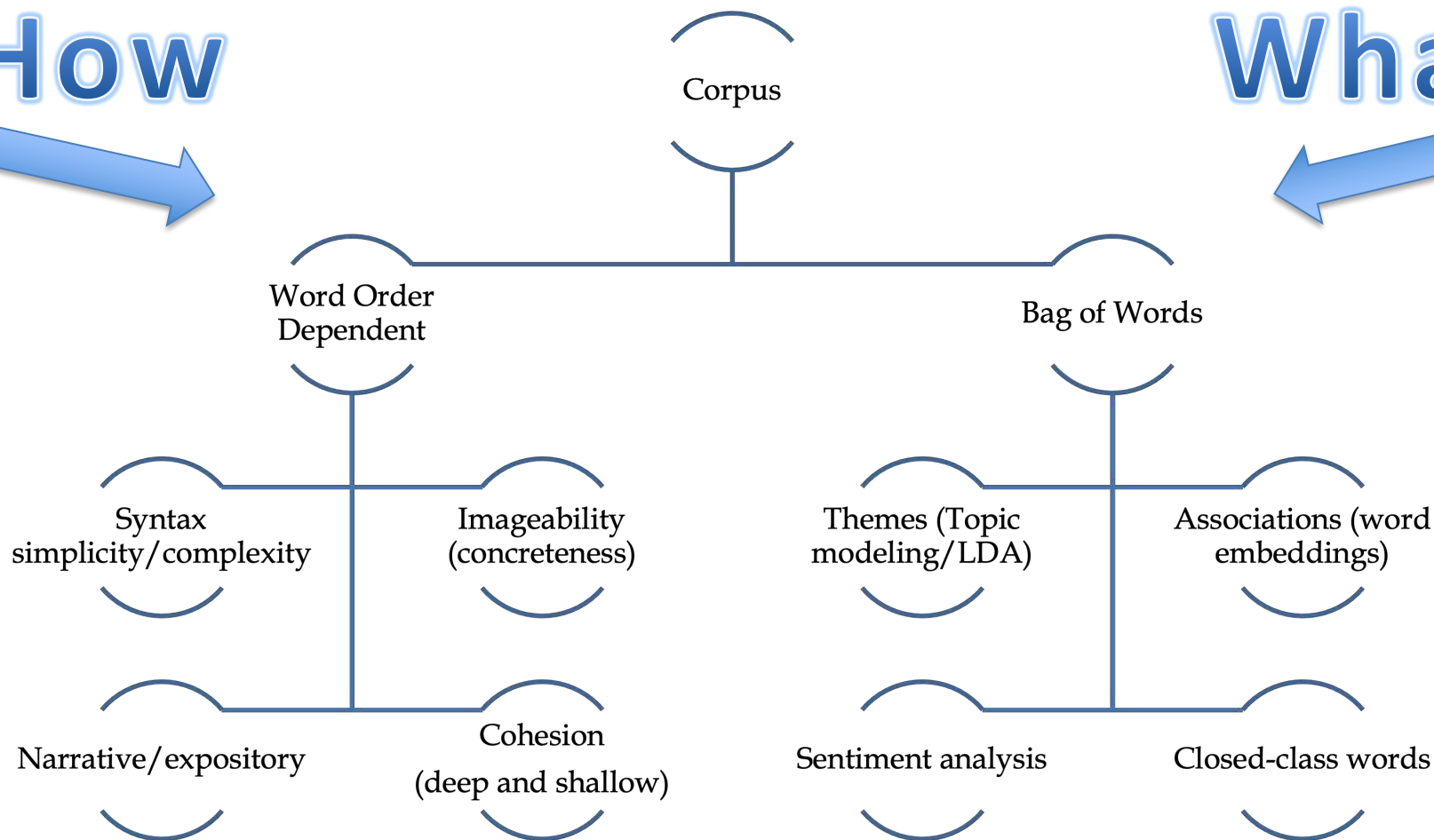
In what language? With what tools?
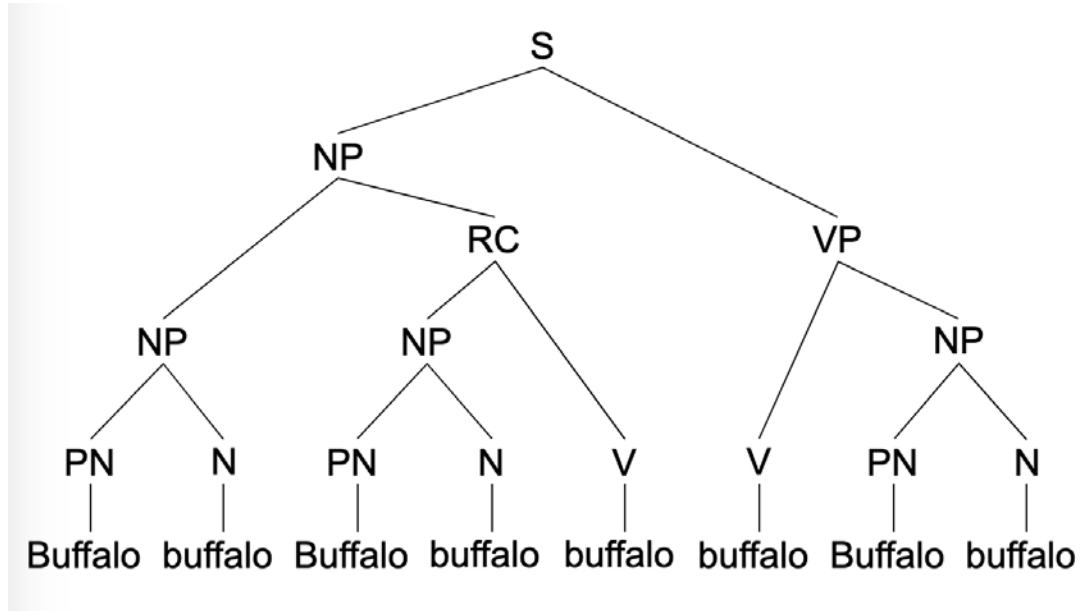
What sources inform theory?

Institute for Intelligent Systems
THE UNIVERSITY OF MEMPHIS

THE UNIVERSITY OF MEMPHIS.

It's what you say *and* how you say it

Overview

BAG OF WORDS + WORD ORDER DEPENDENT

Function words (closed class) words comprise much (30-50%) of spoken and written language

TABLE 12.2 Frequency of the 20 Most Commonly-Used Words as a Function of Genre (from our text archive)

	Descriptions	Experiments	Internet	Published	Personal	Spoken	Mean
I	2.63	5.75	2.57	1.04	5.35	4.47	3.64
The	3.99	3.18	3.00	4.93	2.98	2.77	3.48
And	2.48	3.28	1.90	3.14	3.25	3.46	2.92
To	2.99	3.57	2.31	2.54	3.20	2.83	2.91
A	1.99	1.95	1.76	1.84	2.08	2.02	1.94
Of	1.96	1.57	1.33	3.02	1.65	1.47	1.83
That	1.29	1.67	1.06	0.90	1.92	2.06	1.48
In	1.33	1.20	1.09	1.83	1.24	1.06	1.29
It	1.07	1.26	0.97	0.71	1.39	1.75	1.19
My	0.64	2.28	0.65	0.37	1.53	0.99	1.08
Is	1.06	0.91	1.29	0.64	1.30	1.15	1.06
You	1.45	0.32	1.06	0.70	0.84	1.93	1.05
Was	0.72	1.40	0.56	0.67	1.25	1.45	1.01
For	0.87	0.91	0.79	0.89	0.74	0.61	0.80
Have	0.63	0.86	0.70	0.39	0.88	0.77	0.70
With	0.68	0.75	0.55	0.71	0.69	0.63	0.67
He	0.56	0.64	0.36	0.60	0.80	1.03	0.66
Me	0.55	1.03	0.44	0.31	0.82	0.70	0.64
On	0.77	0.67	0.60	0.65	0.55	0.56	0.63
But	0.50	0.71	0.48	0.38	0.83	0.80	0.62
Top 10 words	21.18	25.91	17.37	20.86	24.65	24.21	21.76
Top 20 words	28.16	33.91	23.47	26.26	33.29	32.51	29.60
Top 50 words	39.85	47.26	34.55	34.95	47.71	47.95	41.82

Note: Numbers reflect percentage of total words within any given text. For example, in any given text from the Description archive, 2.64% of all words are the word "I" (this includes I'm, I'd, I'll, I've).
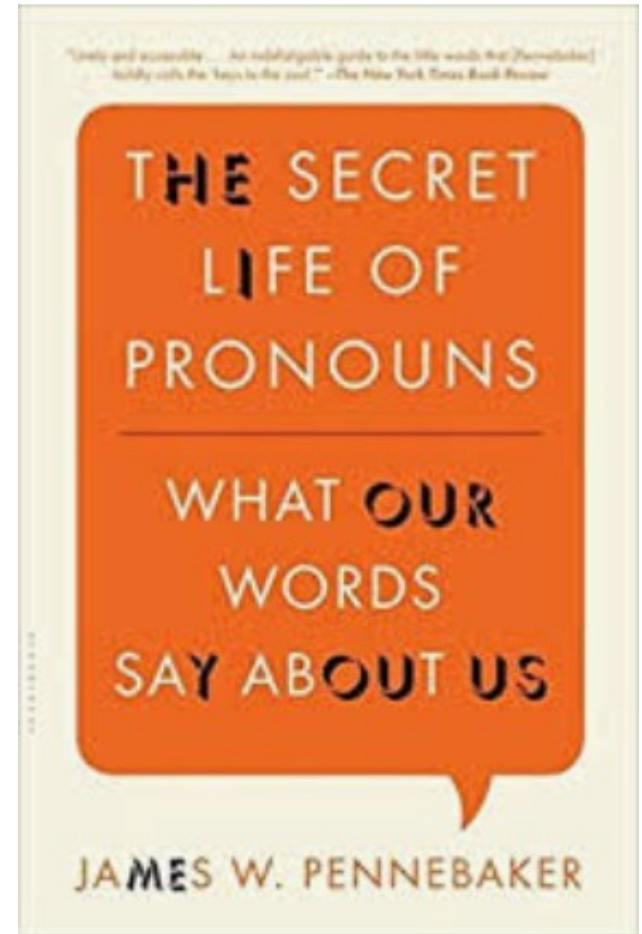
Institute for Intelligent Systems
THE UNIVERSITY OF MEMPHIS

THE UNIVERSITY OF MEMPHIS.

Sentiment Analysis:
Pronouns as language truth-o-meter

"Hey very sorry. Meant no harm. And I apologize for whatever I may have said or what was said. And my overtures are strictly professional. Again I apologize didn't mean to be inappropriate at all. Sorry that impression was left."

-email by Harold Ford, Jr., in response to sexual harassment allegations

THE SECRET LIFE OF PRONOUNS

WHAT OUR WORDS SAY ABOUT US

JAMES W. PENNEBAKER

Institute for Intelligent Systems
THE UNIVERSITY OF MEMPHIS

THE UNIVERSITY OF MEMPHIS.

Syntactic simplicity/complexity

connor
@connorhannigan4

text: i can't
email: i cannot
essay of 2000 words: i am unable
to can

4:06 · 26 Feb 19 · Twitter for iPhone

41.1K Retweets   **191K** Likes

THE UNIVERSITY OF
MEMPHIS.

Word concreteness/abstractness

*"China will enhance **friendly relations** and **cooperation** with all other **countries** on the basis of the Five **Principles** of **Peaceful Coexistence** and strive to build a new type of **relations** between major countries based on mutual **respect** and win-win **cooperation**. Following a policy of building **good-neighborly relationships** and **partnerships** with **neighboring** countries, China has actively expanded exchanges with countries in its **neighborhood**."*

https://gadebate.un.org/en/67/china

Narrativity vs. Expository

- Lists are harder to remember
- Stories are easier to remember
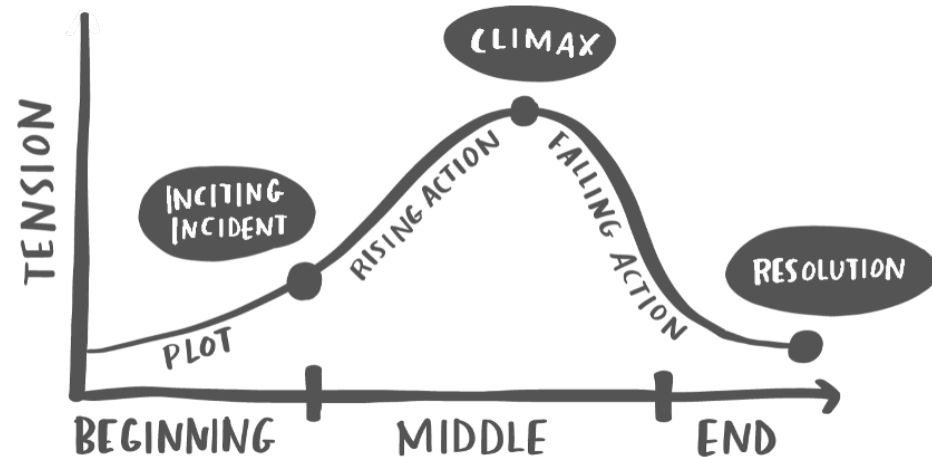
Deep cohesion

But everybody else, the 170 million of us who get health insurance through our employees got big benefits. *Number one*, insurance companies can't deny you coverage because of a pre-existing condition. *Number two*, no lifetime limits, which is a big deal if you have serious health problems. *Number three*, women can't be charged more than men for our health insurance, which is the way it used to be before the Affordable Care Act. *Number four*, if you're under 26, and your parents have a policy, you can be on that policy until the age of 26, something that didn't happen before.

Connectives (causal, temporal, and logical) help the reader follow along with the speaker, text, and message.

Shallow cohesion

Donald Trump, 2ⁿᵈ Presidential Debate 2016

Obamacare will never *work*. It's *very bad*, *very bad* health insurance. Far too *expensive*. And not only *expensive* for the person that has it, unbelievab**ly** *expensive* for our country. It's going to be one of the biggest line items very short**ly**. We have to repeal it and replace it with something absolute**ly** much less *expensive* and something that *works*, where your plan can actual**ly** be tailored. We have to get rid of the *lines* around the state, artificial *lines*, where we stop insurance *companies* from coming in and *competing*, because they want -- and President Obama and whoever was working on it -- they want to leave those *lines*, because that gives the insurance *companies* essentially monopolies. We want *competition*. You will have the finest health care plan there is.

UNGA General Debate, Venezuala 2004 (Ambassador Jesus Arnaldo Perez, Minister of Foreign Affairs)

"There are **moments** we can describe as historical turning points, when **nations** and **peoples** must decide where they **stand**. This is one of these **moments**, when history will judge us as **leaders**, and examine if we were democratic **leaders** that represented the will of our **peoples**. It is clear that the **people** of the **world** are taking a **stand, against** neoliberal economics and **war**. **They** are **fighting against** those who would impose their will by **military** and economic **force**. **They** are resisting those who would undermine and even overthrow the basic principles that founded the United Nations itself."

Repetition enhances memorability

Types of textual data + questions they can help answer

Open-ended survey responses

Speeches

Interviews

Conversations

Books

Songs/lyrics

Transcripts

Social media posts

Radicalization

In-group/ out-group membership

Bluffs/ threats

Deception/lying

**Group cohesion and dispersion**

**Changes in status or power**

Current projects

LANGUAGE IN ACTION

Time and space: Contextualizing your data

- Leaders
  - Leader assumes power
  - Leader removed from power
- Social movements
  - Onset/trajectory of social movement
  - Repression of social movements
  - Leaders/followers
- War
  - Onset of war
  - Duration of war
  - Escalation of war
  - Termination of war
  - Casualties
- Demographics
  - Gender
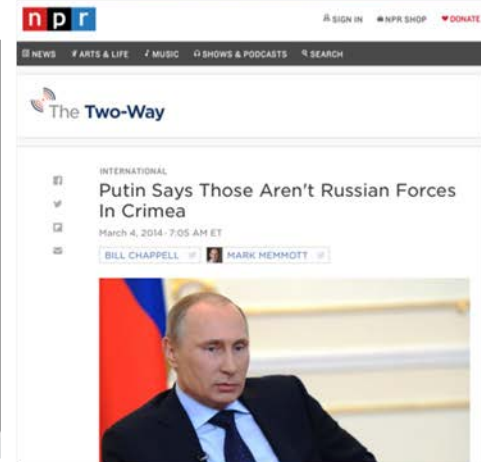 - Age
  - Identity
- Events/event data
- Speech acts

North Korea

Polyarchy / Formality vs Year

Legend: Polyarchy, Formality, Formality
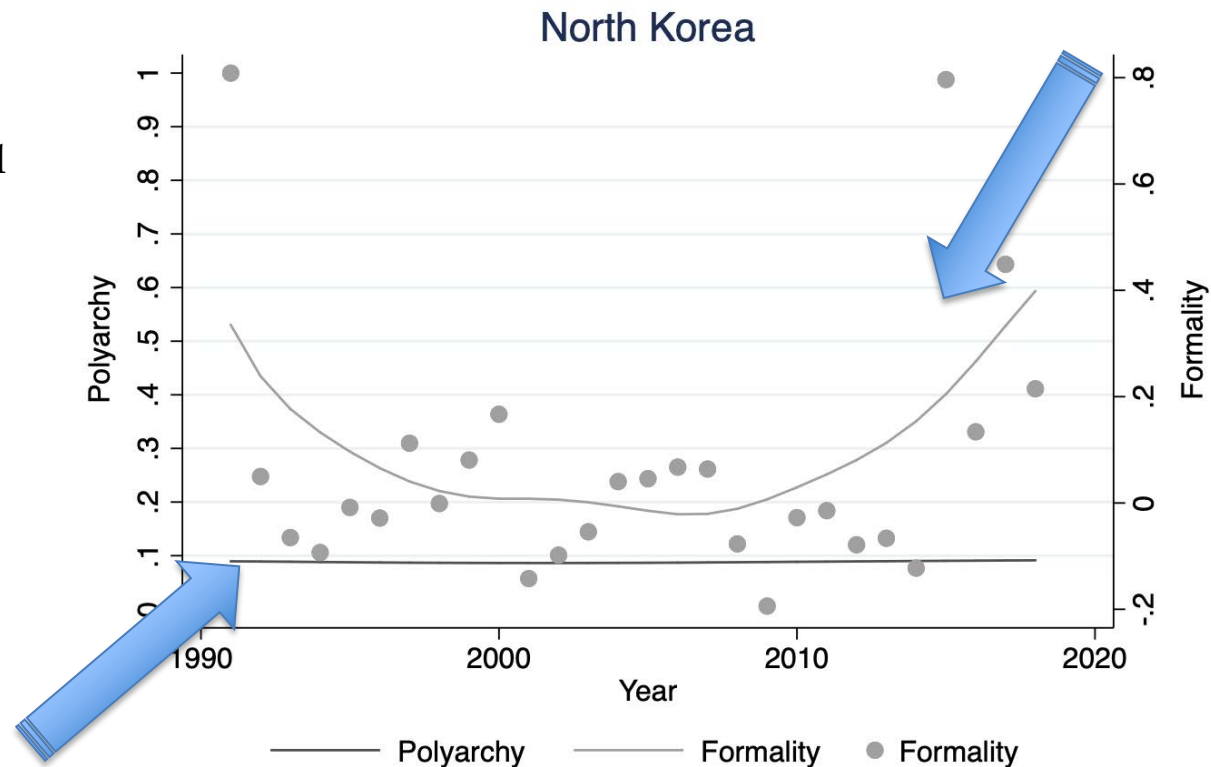
Language and culture shape how we see the world (What we say reveals who we are)
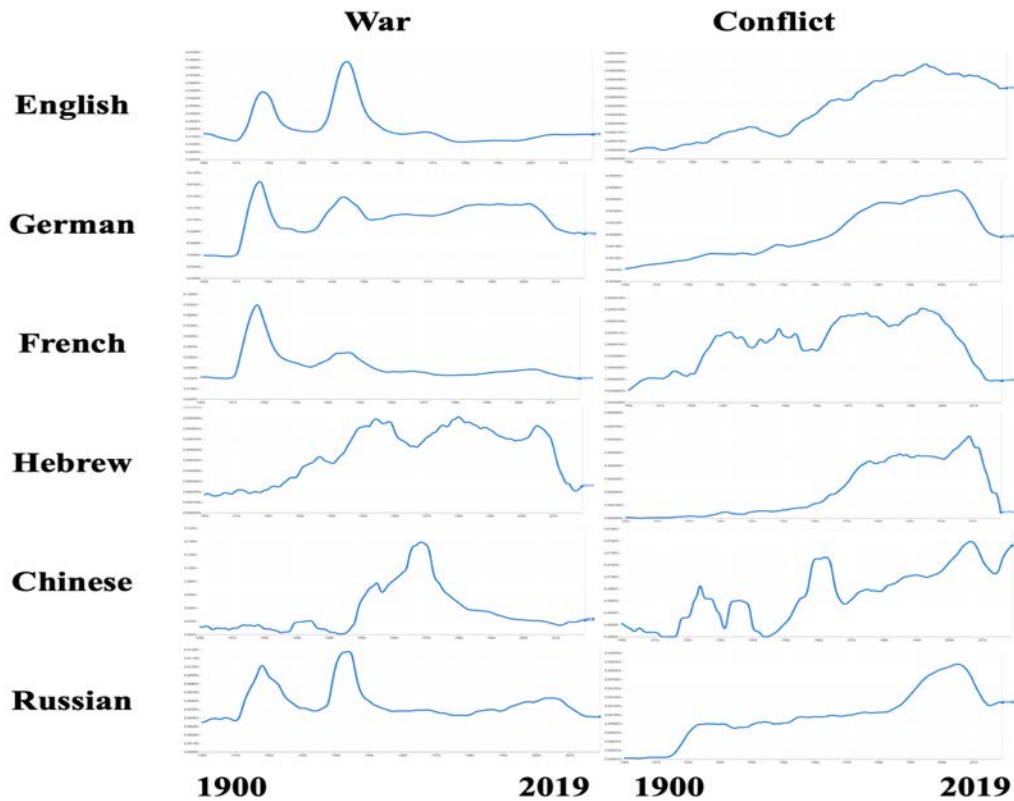
- War and conflict are semantically similar terms (interchangeable, even!)
- The patterns of usage (in a Google n-gram search) differ by:
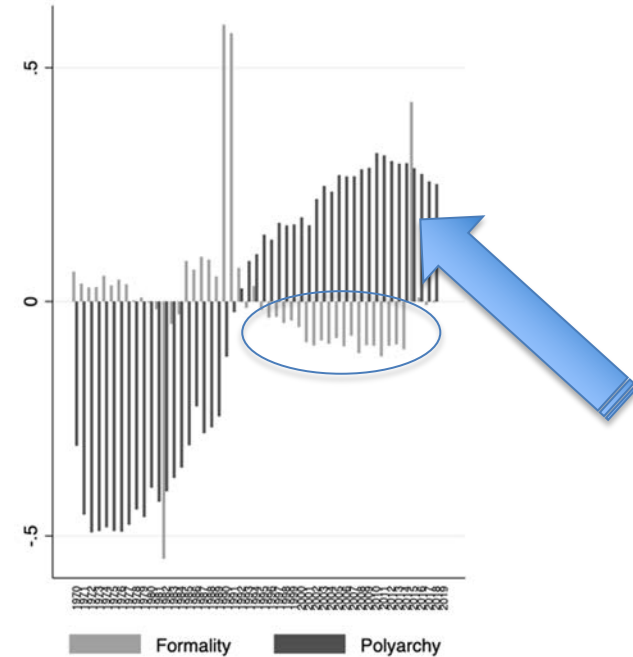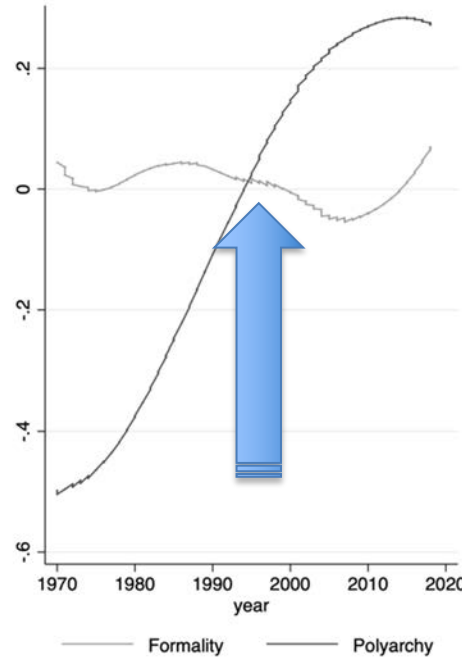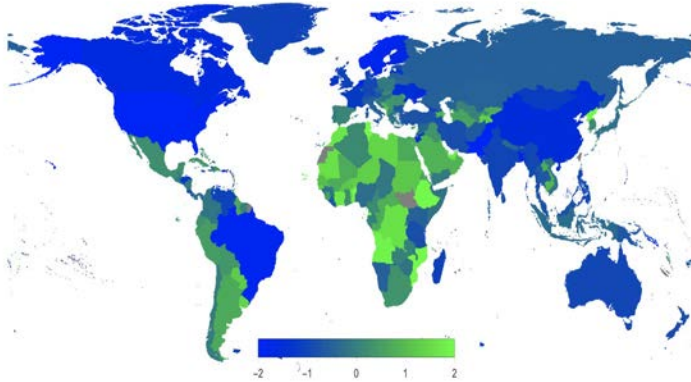 - Language
  - Time
  - Term

Windsor, L. (2021) "Seeing the Forest and the Trees: AI Bias, Political Relativity, and the Language of International Relations." Psychology of AI Ethics, edited by Andrew J. Hampton and Jeanine DeFalco.

Culturally Implicit Associations: Differing Perspectives on the Same Phenomenon

Arthur Graesser and Leah Windsor. (2022) "Text and Discourse with Humans and Computers." To appear in Atlas of Language Analysis in Psychology, edited by Ryan Boyd, Guilford Press

Language patterns can tell us about big picture systemic changes…

|	Cold War	Post-Cold War	Post-2013
Syntax simplicity	More complex	More simple	More complex
Word concreteness	More concrete	More abstract	More concrete
Narrativity	More expository	More narrative	More expository
Deep cohesion	More deep cohesion	Less deep cohesion	More deep cohesion
Shallow cohesion	More shallow cohesion	Less shallow cohesion	More shallow cohesion
Formality	More formal	Less formal	More formal
Systemic changes	**More stress**	**More cooperation**	**More stress**

The Linguistic Foundations of Populist Rhetoric: How Language Conditions World Politics

Windsor, L. and Thies, C. The Linguistic Basis of Populist Rhetoric: How Language Conditions World Politics. (2022) Under contract with Brill.
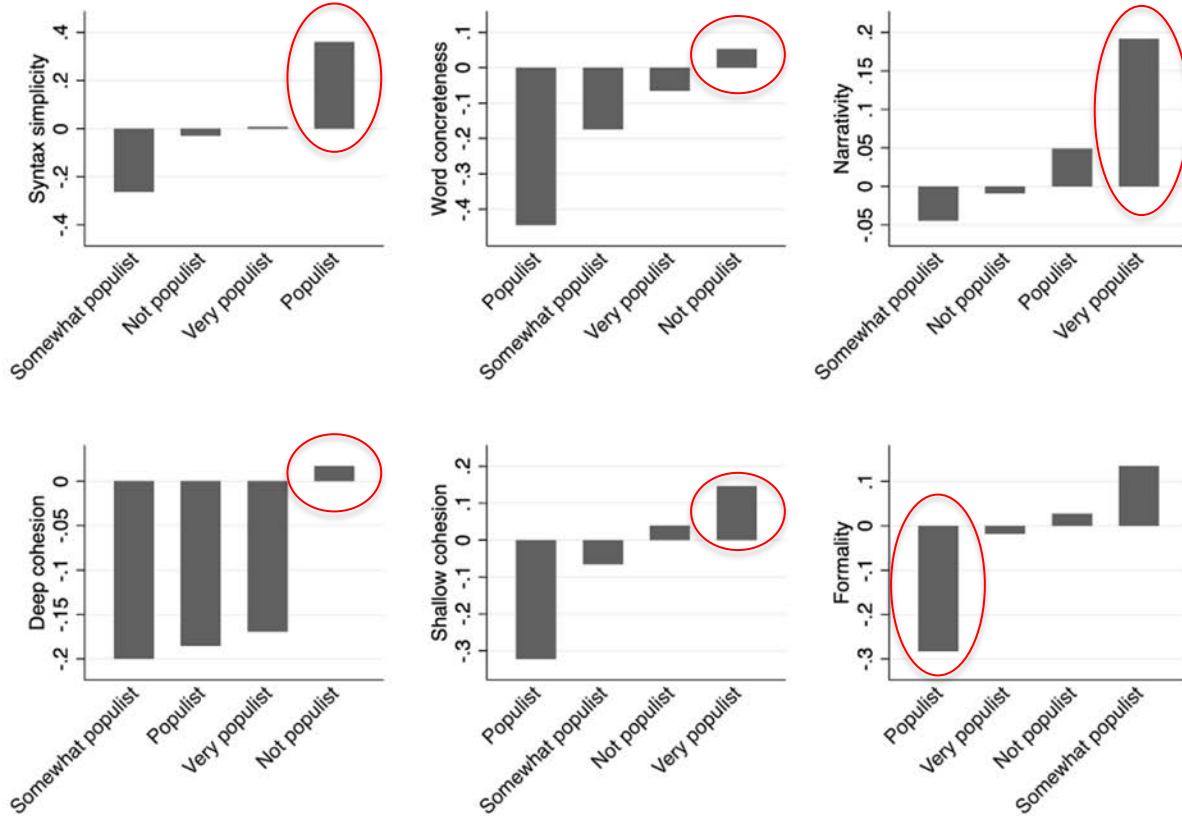
Multimodal communication
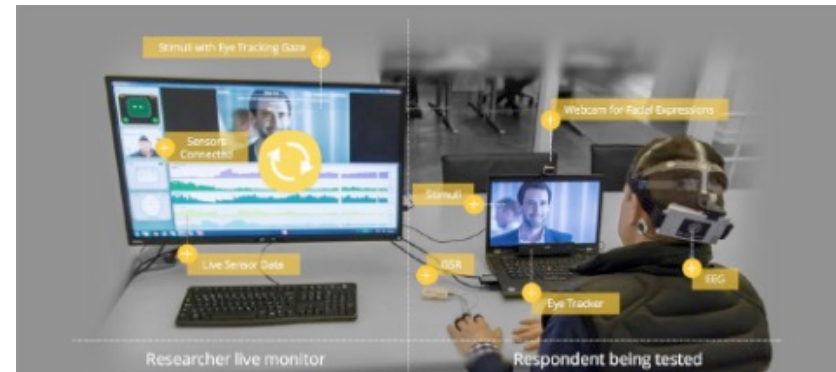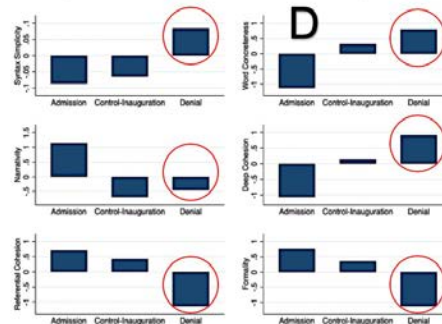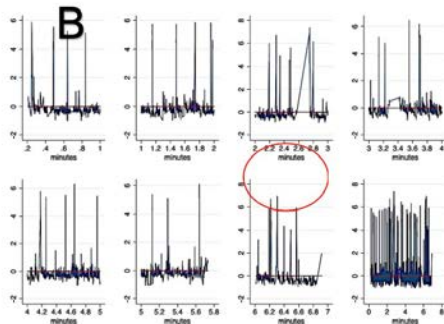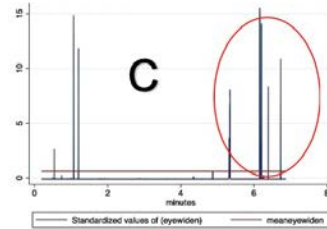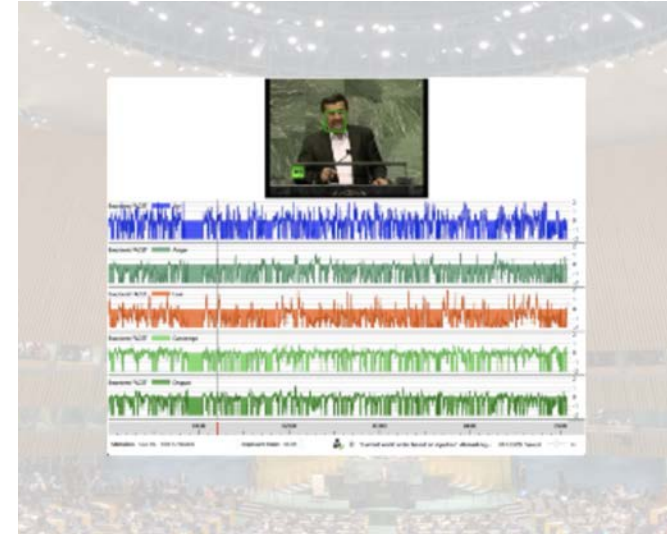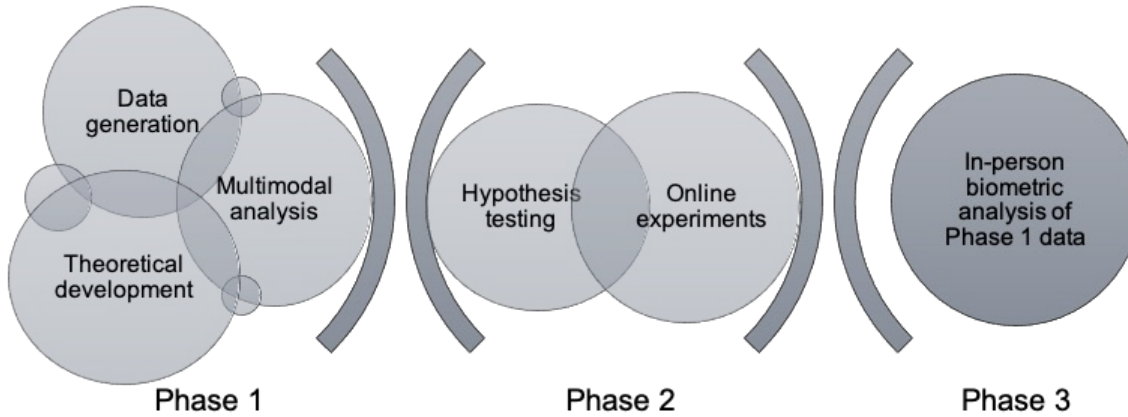
Widening the practice

BRIDGING TECHNICAL DETAILS AND GAPS

From text to data (abc → 123)

Why convert words to numbers?
  Capture trends that may be harder to see in unprocessed natural language
  De-bias analysis
  Contextualize with other numerical data
Hand coding
  Benefits: up close and personal with the data; uncover unexpected relationships
  Drawbacks: time consuming; potential bias
Automated analysis
  Benefits: quick (most of the time); potentially less biased
  Drawbacks: data at a distance; numbers less intuitive (sometimes) than words

Challenges

Where to start?

   You don't know what you don't know!

Picking the right approach

   What do you want to do with your data?

   How much data do you have?

   What language is your data in?

   What tools do you have?

   What theory do you want to explore?

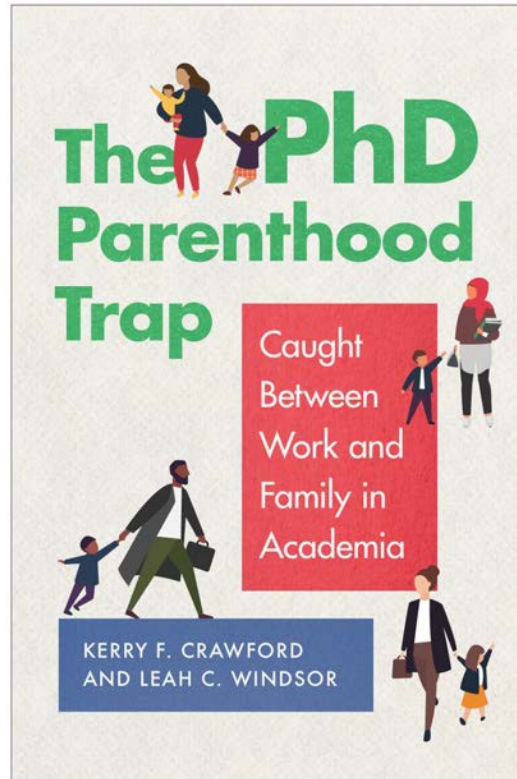   What hypotheses do you want to test?

Getting tech'd up

   Connecting theory-data-hypotheses

   Interpreting results

Connecting multiple modalities (acoustic, linguistic, gestural, neurological)

Interdisciplinary collaboration

Institute for Intelligent Systems
THE UNIVERSITY OF MEMPHIS

Chutes and Ladders:
The Hidden Curriculum
that Makes or Breaks
Academic Careers

Stay tuned!

MENTORSHIP FOR DIVERSITY, EQUITY + INCLUSION

THE UNIVERSITY OF MEMPHIS.

Mentorship + the hidden curriculum

Trends in Genetics

Science & Society

Women and Minorities Encouraged to Apply (Not Stay)

Leah C. Windsor[1,*] and Kerry F. Crawford[2,*]

The Coronavirus 2019 (COVID-19) pandemic has deepened gender and racial diversity problems in academia. Mentorship shows women and other under-represented groups where the ladders to success are, and helps them avoid the chutes, a revised leaky pipeline metaphor. Here, we identify tangible strategies that will improve gender equity, including increasing active mentorship by male academics.

MENTORSHIP: "MEN IN THE MIDDLE" AND THEIR ROLE AS ALLIES IN ADDRESSING GENDER BIAS

Leah C. Windsor, *The University of Memphis*
Cameron G. Thies, *Arizona State University*

NOT A LEAKY PIPELINE! ACADEMIC SUCCESS IS A GAME OF CHUTES AND LADDERS

Leah Cathryn Windsor, *University of Memphis*
Kerry F. Crawford, *James Madison University*
Marijke Breuning, *University of North Texas*

The Profession

Grant Writing and the Hidden Curriculum: Mentoring and Collaborating Across Disciplines

Leah C. Windsor, *University of Memphis, USA*
Christian Kronsted, *University of Memphis, USA*

ABSTRACT Submitting grant proposals is becoming an increasingly common expectation— and, in some cases, a requirement—in the discipline of political science as well as other social sciences and the humanities. However, writing a grant with a good chance of success at getting funded is not part of standard mentorship or pedagogy in our discipline. It is a part of the hidden curriculum, where grant-writing skills often are taught informally in working with a principal investigator. This article describes the process and structure of writing a grant to provide a roadmap for scholars to follow in submitting externally funded projects. The article describes an Institutional Review Board–approved survey about mentorship and grant writing and discusses the importance of socialization, professionalization, and administration in supporting scholars in writing and obtaining grants.

Conditional accept with International Studies Perspectives (July 2022)

Call Me, Maybe? Gender, Bias, and Virtual Mentorship in Post-Covid Academia (with Kerry F. Crawford)
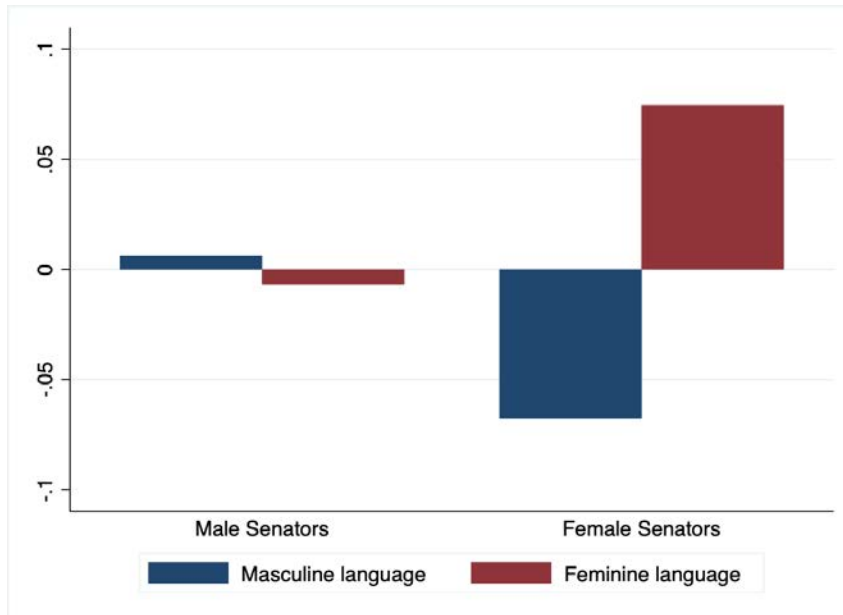
SOME FUN PROJECTS

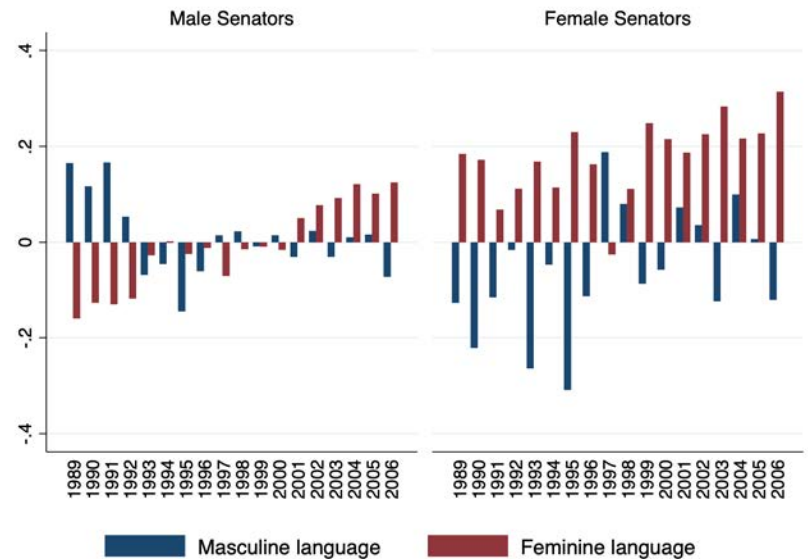"It's a full-bodied wine with hints of acrimony, partisanship, and moral outrage."

Gender and language

Women Senators speak like women (population norm-referenced)

Men Senators adopt more feminine linguistic styles as women's representation in the Senate increases
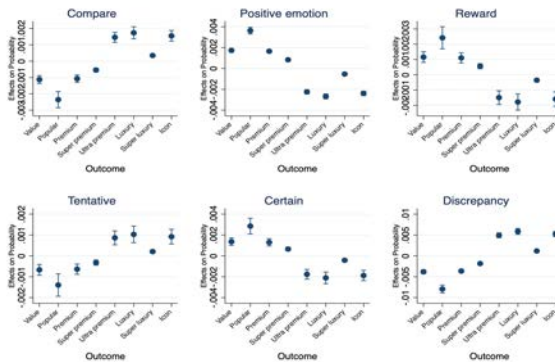

Windsor, L. (2022) Gender, Language, and Representation in the U.S. Senate. With Sara Mitchell, Tracy Osborn, Bryce Dietrich, and Andrew J. Hampton. Journal of Language and Politics

THE UNIVERSITY OF MEMPHIS.

Wine reviews:
Snooty language sells wine

Table 2. Summary of price variable trends

	Less expensive wines	**More expensive wines**
Use more	Verbs, adjectives, prepositions, positive emotion, reward, certainty, family affiliation, perception, body, hear, bio,	Total words, big words, impersonal pronouns, auxiliary verbs, articles, comparisons, tentative, discrepancy, social, leisure, see, health, ingest,
Use less/fewer	Total words, big words, impersonal pronouns, auxiliary verbs, articles, comparisons, tentative, discrepancy, social, leisure, see, health, ingest,	Verbs, adjectives, prepositions, positive emotion, reward, certainty, family affiliation, perception, body, hear, bio,

Windsor, L. (2022) Linguistic Correlates of Wine Ratings and Prices. With Jennifer Tatara, Christian Kronsted, Alistair Windsor, and Courtney Peters.
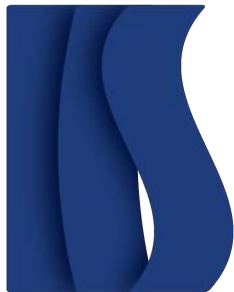
The Voynich Manuscript: Linguistic analysis when you don't know the language

- What can we learn about language patterns when we don't know anything about the language itself?
  - Does it follow common language patterns, like Zipf's, Heaps', and Benford's Laws?
  - How does this help us investigate questions and check our own biases?

Thank you!

Leah.Windsor@memphis.edu
https://www.thewordwitch.dev/home

@leahcwindsor

@lecawe

Institute for Intelligent Systems
THE UNIVERSITY OF MEMPHIS

THE UNIVERSITY OF
MEMPHIS.

Picking the right approach

- Bag of words
  - Word counting
  - Frequency
- Nearest neighbors
  - Word embeddings
- Topic modeling
  - Identifying themes
- Syntax/semantics
  - Word order dependent

QuantText
https://sites.google.com/view/quanttext/home

Organizing + pre-processing data

- **File-naming conventions**
- **Bag of Words**
  - Pre-processing less important
  - Things to consider:
    - Html tags
    - Page numbers
    - Different speakers
    - Spreadsheet or text files?
- **Word Order Dependent**
  - Pre-processing important
  - Hard line breaks
  - Non-ascii characters

| year	president	despc	dessc	deswc	despl	despld	dessl	dessld	deswlsy	deswlsyd	deswllt	deswlltd	narrativity	pcnarp
1790	washington	17	41	1401	2.4119999	1.1210001	34.243999	19.367001	1.6210001	0.97600001	4.8930001	2.931	-0.494	31.209999
1790	washington	20	27	1085	1.35	0.58700001	40.333	27.509001	1.711	1.0599999	5.0700002	3.1259999	-0.682	24.83
1791	washington	41	63	2303	1.5369999	0.83999997	36.651001	19.216	1.682	1.035	5.0370002	3.0280001	-0.929	17.879999
1792	washington	32	62	2099	1.938	1.045	33.903	19.319	1.6569999	1.013	4.9380002	2.924	-0.798	21.48
1793	washington	24	57	1964	2.375	1.377	34.578999	18.839001	1.6160001	0.94800001	4.823	2.825	-0.625	26.76
1794	washington	30	81	2914	2.7	1.745	36.049	20.804001	1.652	0.97299999	4.921	2.8840001	-0.446	33
1795	washington	25	53	1985	2.1199999	1.166	37.472	22.427999	1.714	1.018	5.072	2.973	-0.503	30.85
1796	washington	36	79	2869	2.194	1.215	36.443001	19.995001	1.6619999	0.99199998	4.9390001	2.947	-0.726	23.58
1797	jadams	23	61	2060	2.652	1.799	33.77	24.204	1.648	0.96399999	4.9429998	2.825	-0.733	23.27
1798	jadams	27	57	2218	2.1110001	1.1210001	38.911999	25.013	1.659	0.98400003	4.9460001	2.8970001	-0.48	31.559999
1799	jadams	15	38	1505	2.533	1.4069999	39.605	28.438	1.704	1.03	5.026	3.0079999	-0.761	22.360001
1800	jadams	20	43	1372	2.1500001	0.98799998	31.952999	17.299999	1.691	1.034	4.9879999	2.9400001	-0.623	26.76
1801	tjefferson	32	91	3224	2.8440001	1.5470001	35.451	25.924999	1.645	0.99199998	4.8909998	2.8870001	-0.45	32.639999
1802	tjefferson	23	64	2196	2.783	2.0220001	34.422001	16.736	1.591	0.91900003	4.7859998	2.734	-0.455	32.639999
1803	tjefferson	28	54	2267	1.929	1.12	42.056	22.764999	1.658	0.98000002	5.0339999	2.9100001	-0.519	30.15
1804	tjefferson	20	49	2095	2.45	1.432	42.877998	27.731001	1.623	0.96700001	4.8899999	2.848	-0.607	27.43
1805	tjefferson	25	77	2924	3.0799999	1.681	38.077999	23.447001	1.5650001	0.898	4.7259998	2.6900001	-0.306	38.209999
1806	tjefferson	25	72	2859	2.8800001	2.108	39.75	21.493999	1.614	0.972	4.8569999	2.796	-0.267	39.740002
1807	tjefferson	21	61	2385	2.905	1.48	39.18	17.405001	1.615	0.92199999	4.8660002	2.823	-0.367	35.939999
1808	tjefferson	21	71	2676	3.381	1.596	37.761002	19.645	1.64	0.97500002	4.9180002	2.8710001	-0.585	28.1
1809	jmadison	17	41	1828	2.4119999	1.502	44.633999	26.759001	1.669	1.0140001	4.941	2.9990001	-0.545	29.459999
1810	jmadison	26	62	2443	2.385	1.329	39.5	19.671	1.696	1.029	5.0440001	3.0079999	-1.068	14.46
1811	jmadison	27	47	2271	1.7410001	0.81300002	48.362	24.506001	1.686	1.012	5.026	2.9489999	-0.637	26.43
1812	jmadison	35	88	3238	2.5139999	1.541	36.919998	22.177999	1.664	1	4.9439998	2.9119999	-0.834	20.33
1813	jmadison	38	68	3254	1.789	1.143	47.956001	22.580999	1.6440001	0.98500001	4.8790002	2.842	-0.741	22.959999
1814	jmadison	23	51	2109	2.217	1.38	41.451	32.713001	1.638	0.95300001	4.862	2.7279999	-0.755	22.66
1815	jmadison	29	59	3142	2.0339999	0.98100001	53.321999	28.629	1.712	1.033	5.0469999	3	-0.841	20.049999
1816	jmadison	29	73	3363	2.517	1.4299999	46.137001	38.134998	1.676	1.012	4.9380002	2.9260001	-0.729	23.27
1817	jmonroe	43	123	4416	2.8599999	1.641	36.033001	19.237	1.603	0.94	4.8080001	2.777	-0.78	21.77
1818	jmonroe	42	115	4371	2.7379999	1.345	38.051998	22.17	1.6440001	0.98900002	4.941	2.8310001	-0.692	24.51
1819	jmonroe	46	133	4699	2.891	1.337	35.405998	18.450001	1.594	0.94499999	4.8210001	2.7839999	-0.466	32.279999
1820	jmonroe	33	85	3444	2.576	1.3	40.729	27.013	1.597	0.94499999	4.8249998	2.7820001	-0.596	27.76
1821	jmonroe	51	151	5817	2.961	1.442	38.780998	23.545	1.592	0.92199999	4.7950001	2.744	-0.713	23.889999
1822	jmonroe	66	140	4735	2.1210001	1.705	34	20.620001	1.598	0.94	4.822	2.7869999	-0.625	26.76
1823	jmonroe	77	191	6360	2.4809999	1.382	33.627998	20.514	1.6210001	0.96499997	4.9089999	2.8570001	-0.651	25.780001
1824	jmonroe	91	254	8403	2.7909999	1.546	33.201	17.028	1.596	0.95200002	4.816	2.7920001	-0.717	23.889999
1825	jqadams	63	212	9007	3.365	1.609	42.873001	24.528	1.642	0.991	4.8839998	2.891	-0.933	17.620001
1826	jqadams	52	175	7739	3.365	1.521	44.777	26.275999	1.627	0.98799999	4.8759999	2.908	-0.839	20.049999
1827	jqadams	72	167	6965	2.319	1.341	42.334999	29.471001	1.638	0.986	4.973	2.9300001	-0.964	16.85
1828	jqadams	61	196	7316	3.2130001	1.55	37.723999	21.035	1.641	0.97600001	4.9229999	2.8900001	-0.929	17.620001
1829	ajackson	84	300	10528	3.5710001	1.971	35.237	16.058001	1.6390001	0.96799999	4.8779998	2.8610001	-0.529	29.809999
1830	ajackson	129	392	15102	3.039	1.783	38.569	21.556	1.642	0.98500001	4.9089999	2.9070001	-0.57	28.43
1831	ajackson	56	166	7185	2.964	1.401	43.433998	24.940001	1.6289999	0.96799999	4.8759999	2.8340001	-0.491	31.209999
1832	ajackson	66	200	7879	3.03	1.569	39.535	19.191999	1.632	0.95899999	4.8559999	2.8380001	-0.554	29.120001
1833	ajackson	68	182	7888	2.6760001	1.511	43.582001	22.881001	1.614	0.95700002	4.8579998	2.846	-0.48	31.559999
1834	ajackson	119	323	13442	2.714	1.341	41.845001	24.514	1.633	0.97100002	4.882	2.8740001	-0.535	29.809999

How do people think about problems? Language styles reveal preferences for top-down and bottom-up processing
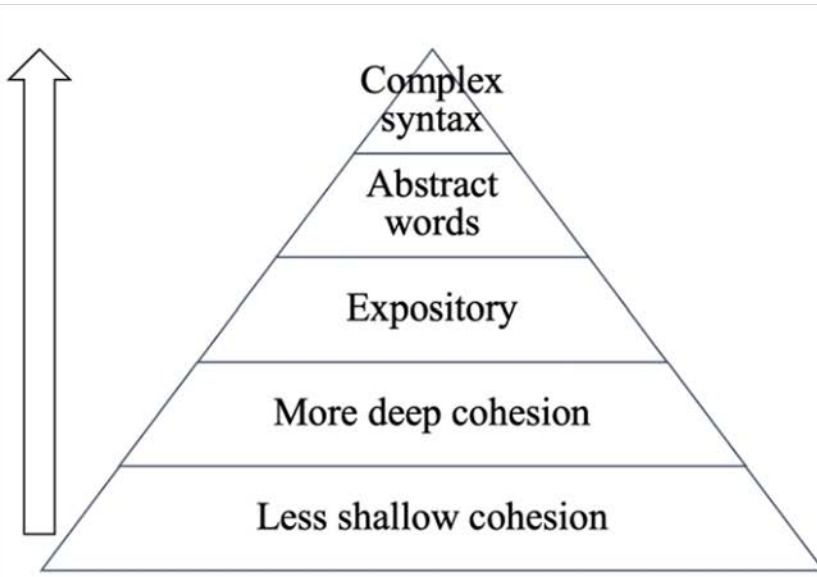
Bottom-up   **Top-down**
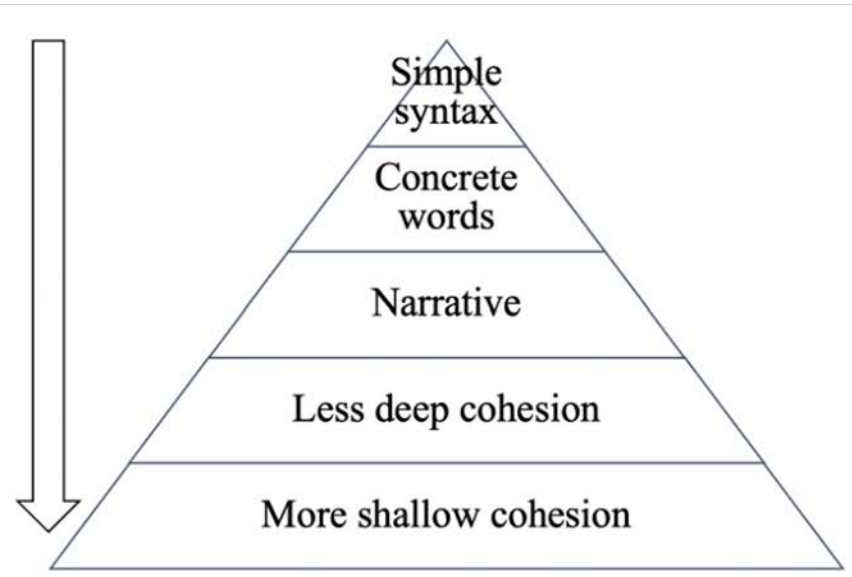
Some final thoughts on being and working happy

- **Academia is a game of chutes and ladders, and a progression of deliberate choices along a path**
  - Although the "hidden curriculum" obscures the path along the way
  - Academic progress is the result of a rigorous, but generous and flexible, process
- **Success by a thousand cuts**
  - Chip away at projects
  - Reclaim tidbits of time
  - Don't wait until you have a chunk of time
  - Be relentlessly efficient
- **Don't believe the hype**
  - No FOMO
  - Ignore humblebrags
  - Don't underplay your achievements
  - Don't be less than as amazing as you can be, just to fit in.
- **Work-life BOUNDARIES (not balance)**
  - Clock out, go home
  - Emergency self-care is a gimmick designed to sell products; real self-care is an ongoing commitment
  - A pedicure is not the cure