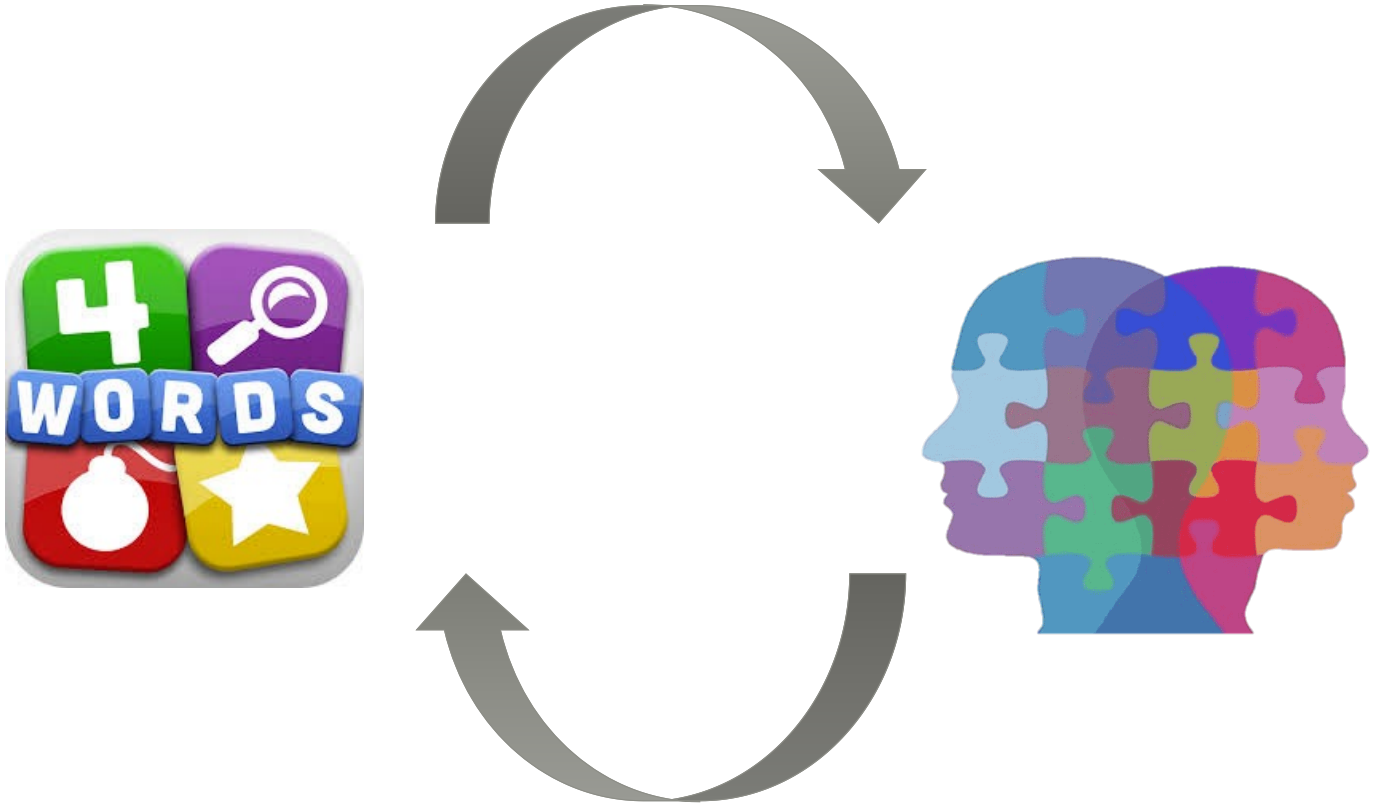
Language and social cognition



SERIAL TRANSMISSION PARADIGM (LYONS & KASHIMA, 2003)

describe more behaviors that are consistent, versus inconsistent, with stereotypes

CATE IS ITALIAN. SHE IS LOVES SHARING FOOD WITH FRIENDS, SHE DOES NOT LIKE NOISE.

CATE IS ITALIAN. SHE COOKS FOR HER FRIENDS, TYPICALLY PIZZA OR SPAGHETI. SHE IS VERY WARM AND AFFECTIONATE. SHE IS OUTGOING AND EXPRESSIVE.

THREE METAPHORS OF LANGUAGE

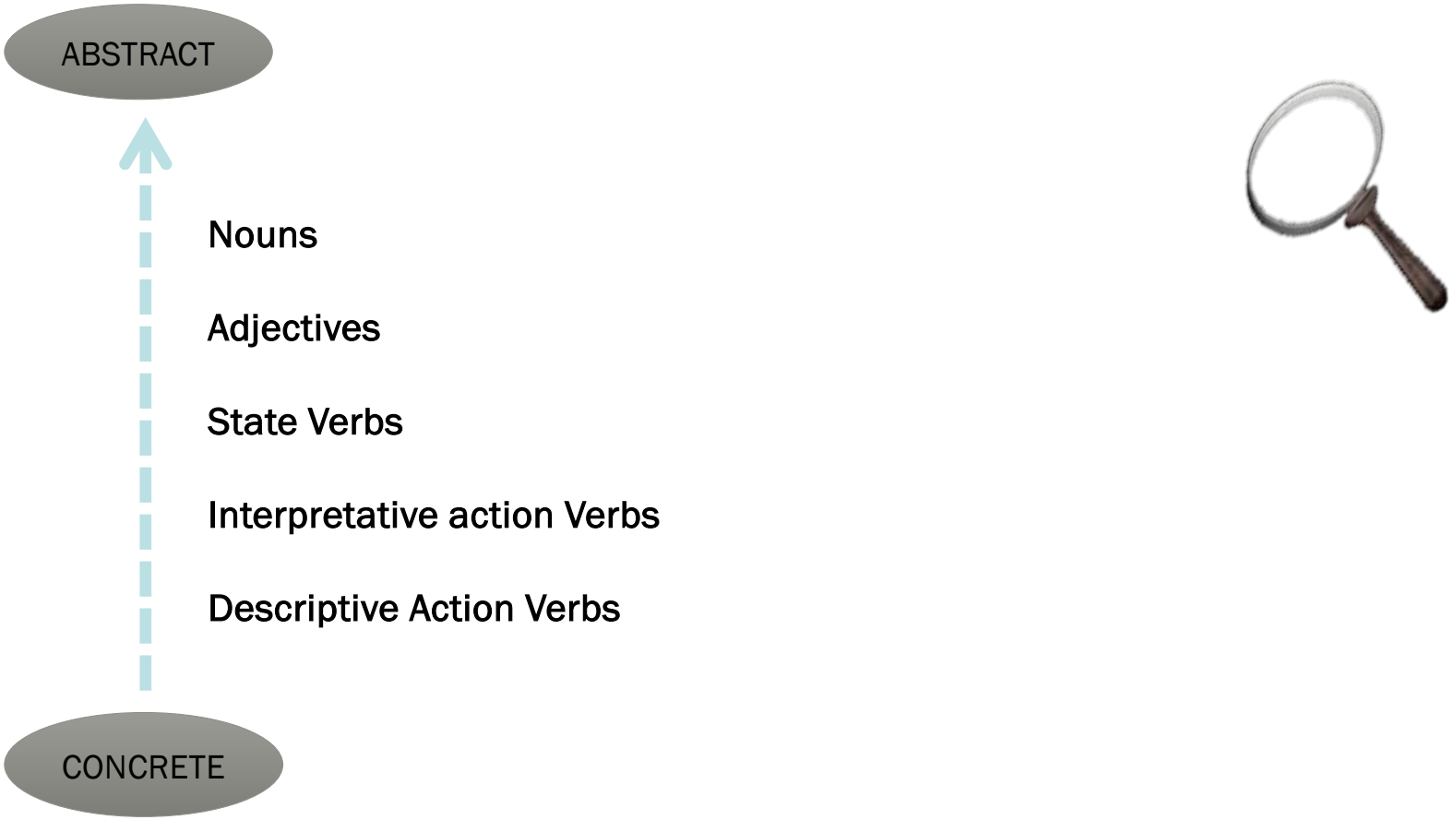
- VESSEL in which thoughts are encapsulated and transmitted from one mind to another

- LENS which focuses cognition on certain aspects of the world and away from others

- BAROMETER reveals a communicator's cognition to the audience

LINGUISTIC ABSTRACTION

Linguistic Category Model (Semin & Fiedler, 1988); Nouns (Carnaghi et al. 2008)



LINGUISTIC ABSTRACTION

Linguistic Category Model (Semin & Fiedler, 1988); Nouns (Carnaghi et al. 2008)

Nouns → Category a person belong to → Eg: a genious, a criminal

Adjectives → characteristics of a person → Eg: intelligent, aggressive

State Verbs → cognitive or emotional states, perduring in time, without specific beginning and end → Eg: admire, hate, appreaciate, ecc..

Interpretative action Verbs → verbs referring to a set of actions with a specific beginning and end → Eg: help, provoce, avoid, ecc..

Descriptive Action Verbs → verbs referring to a single action with a specific beginning and end → Eg: hit, scream, walk, ecc..

LINGUISTIC INTERGROUP BIAS (Maass, 1999)

Abstraction -> focus to stability

+ve ingroup behaviors & –ve outgroup behaviors

If communicators use abstract language to describe a person‟s negative behavior and concrete language for their positive behavior, they are also seen as having negative attitudes and intentions (Douglas & Sutton, 2006)

generalization of positive behaviors to the enitire ingroup

generalization of negative behaviors to the enitire outgroup

Nouns (vs. adjectives)

Carnaghi et al. 2008

- Promote essentialization
- facilitate stereotype-congruent inferences
- inhibit incongruent ones
- Inihibit alternative classifications.
- Transmit prejudice

Reynaert and Gelman (2007)

he has baxtermia", "he is baxtermic", "he is a baxtermic"

Illness permanence----------→

Verbs as carriers of AGENCY

Formanowicz et al. 2017, 2021

- Pseudo-Verbs are perceived as more agentic

- Verbs appears more often in association to agentic social targets (male, young etc)

- Verbs enhance persuasiveness of a message

Generic masculine

- linguistic convention in English has long had it that masculine terms such as “man”, “his”, and the collective noun “Man”, can be used without reference to gender.
- *fireman*
- *native language* rather than *mother tongue*, *police officers* rather than *policemen, humans* rather than *men* to refer to human beings)

Generic masculine

- masculine generic inhibits the availabilty of female examplars (Stahlberg et al., 2007).

- the ratio of male to female pronouns reflected the status of women in the United States (1.2 million U.S. books, 1900–2008; Google Books database; Twenge et al., 2012)
- Countries with grammatical gender languages had lower levels of social gender equality than countries with natural gender languages or genderless languages (Prewitt-Freilino et al., 2012)

- participants with modern sexist beliefs were found to use more traditional, gender-unfair language (Swim et al., 2004).

DEROGATORY LABELS: *FAG IS NOT A SYNONYMOUS OF GAY*

"the overhearing of derogatory labels would automatically activate negative feelings and beliefs associated with the group in question" (Greenberg and Pyszczynski, 1985, p.156)

people may infer that prejudice is normative when they hear others using hate terms.

-> self-perpetuating cycle of prejudice

REAPPROPRIATION OF DEROGATORY LABELS

Galinsky, Hugenberg, Groom, & Bodenhausen, 2003
a stigmatized group has the possibility to renegotiate the connotation of that word, transforming it from a negative expression to an empowering one.

Order and comparison asymmetry

Primacy effect: first mentioned target is more likely to capture the attention, is better remembered, is more likely to be perceived as the cause (e.g., Bettinsoli et al.)

the partner possessing more stereotypically masculine traits is mentioned first (Hegarty et al. 2001)

when men are presented as referent group (e.g., compared to males, females are …), gender differences in status were perceived as larger and more legitimate (Bruckmüller et al. 2012)

Semantic Networks: a definition

- WHAT — graphical representations of knowledge based on meaningful relationships of written text, structured as a network of labeled nodes cognitively related to one another
- WHY — GOAL: extract meanings
- HOW — semantic networks connect words to words/hashtags/phrases, based on their co-occurrence
- WHO — human and computerized methods, dealing with challenges such as co-reference resolution, synonym resolution, and ambiguity

How good are the retrieved docs?

- *Precision* : “purity” Fraction of retrieved docs that are relevant to the user’s information need (reject irrelevant)

- *Recall* : “completeness” Fraction of relevant docs in collection that are retrieved (select relevant)

CLEAN DATA

Pre-processing starts the text preparation into a more structured representation.

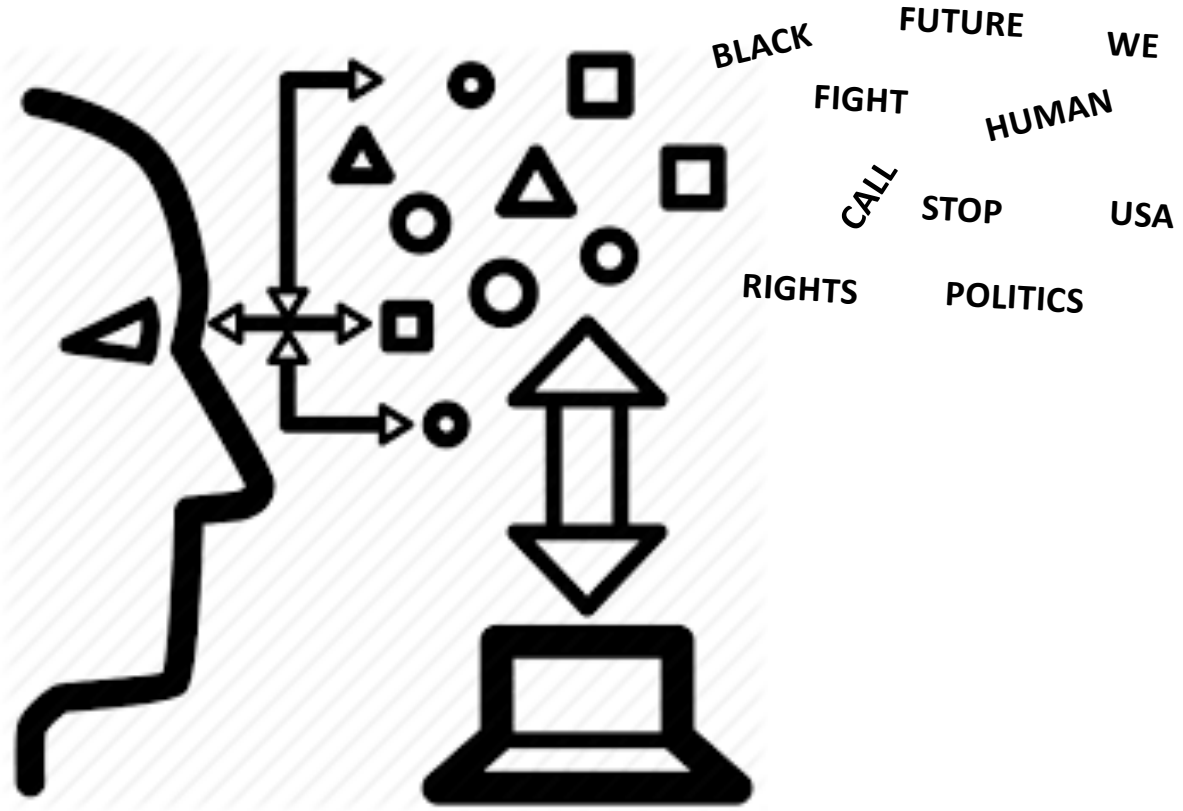
1) **Tokenization:** Tokenization is used to identify all words in a given text.
2) **Data Filtering:** People use a lot of casual language on twitter. To improve this and make words more similar to generic words, such sets of repeated letters are replaced by two occurrences.
haaaaappy -> haappy.

3) **Stop Word Removal**: Is used to eliminate that words that occurs frequently such as article, prepositions, conjunction and adverbs. These stop words depends on language of the text in questions. For example, words like the, and, before, while, and so on do not contribute to the sentiment.

4) **Stemming:** In information retrieval, stemming is the process of reducing a word to its root form.
walking, walker, walked ->walk

PROCESS DATA

Dealing with textual data: from text to numbers



Words or Hashtags

- Top down semantic/sentiment classification: bag of words
- Bottom up semantic/sentiment classification: human coding
- Meta-semantic classification: pronouns, nouns, verbs, adjectives
- Meta-semantic structural properties: word order, dropping
- Semantic & grammar: future/past/present tense

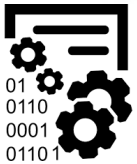
- topical signifier : shared conversation marker,
- can also represent the context of a tweet
- flag an individual's community membership
- indicate shared interests

Dealing with textual data: from text to numbers

Theory Driven

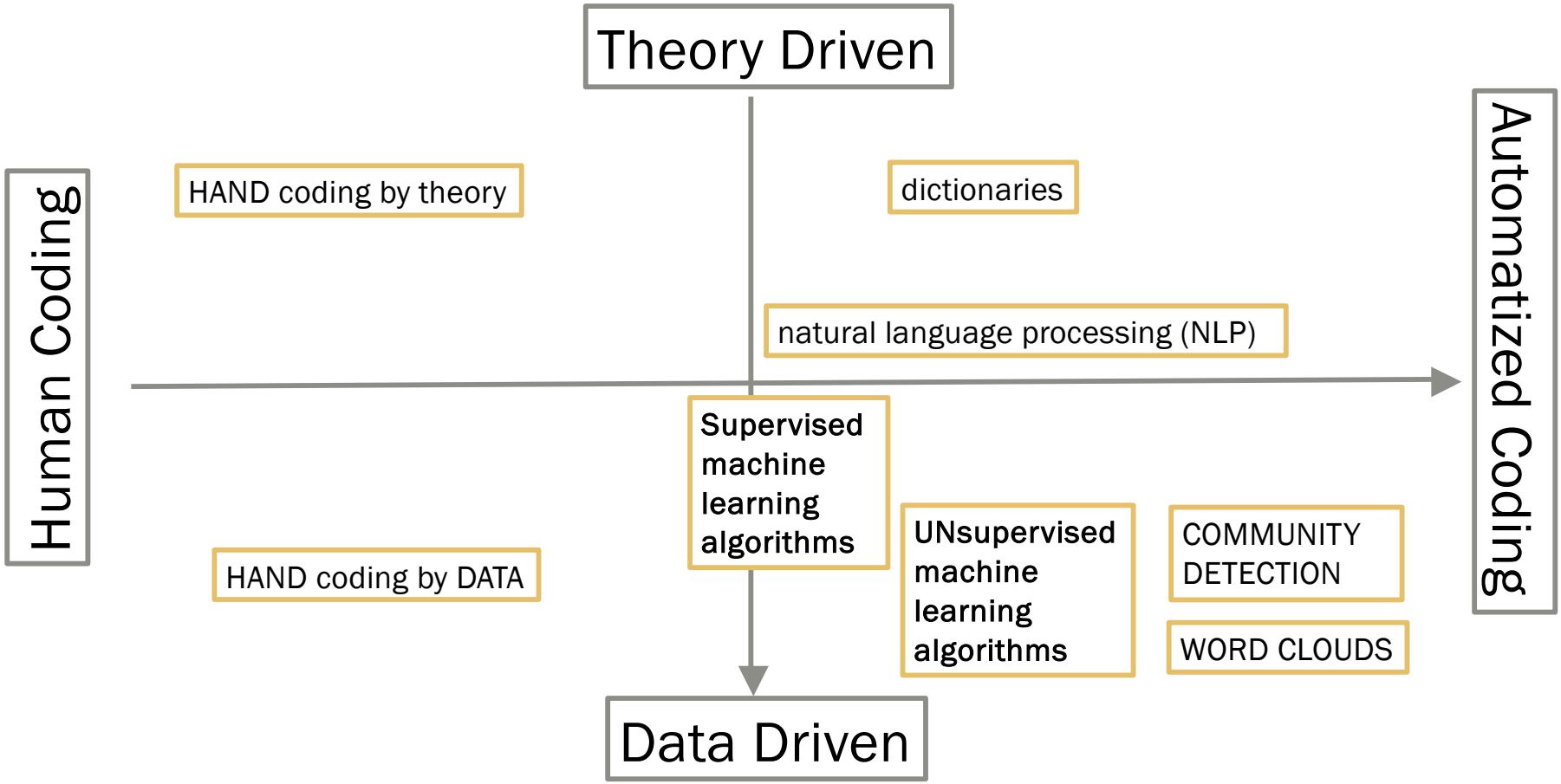
Human Coding

Automatized Coding



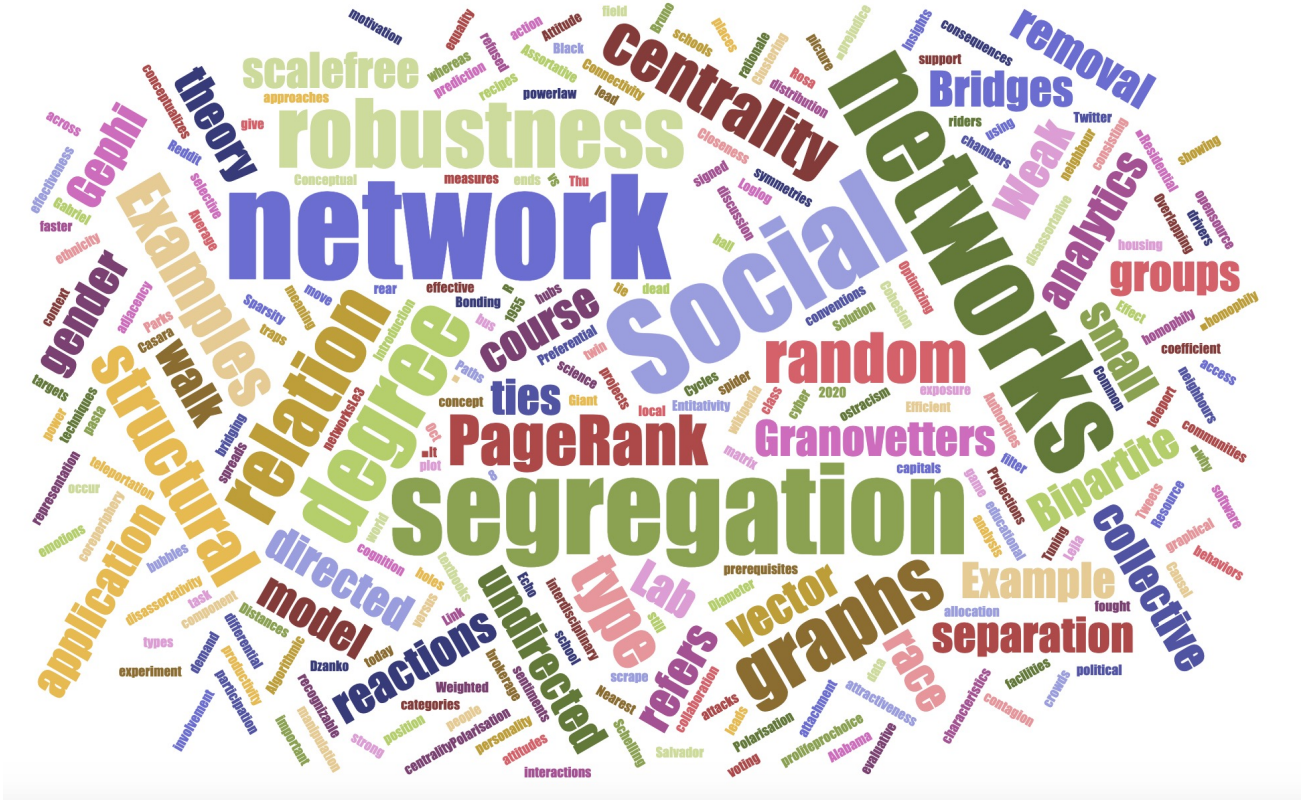
Data Driven

Dealing with textual data: from text to numbers



Bag of words: word cloud

- Based of word count
- Bigger words= more frequent
- bottom

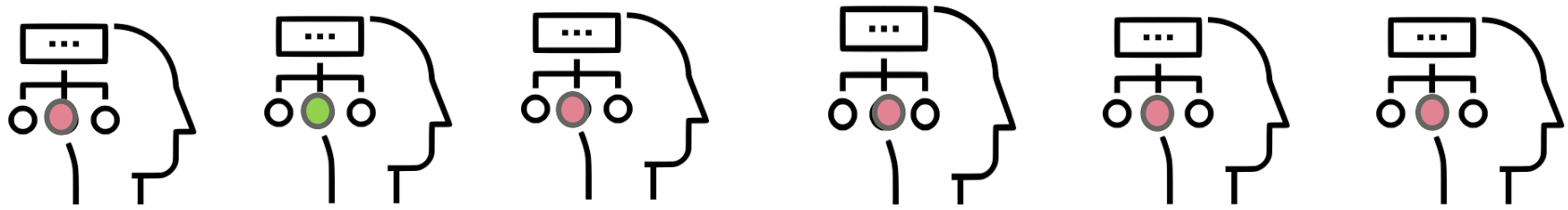


Human Coding

- *top down (coding by theory):* initial coding scheme developed from the from pre-existing theory or assumptions

- *bottom up (grounded theory):* initial coding scheme developed from the data

- *THE SUBJECTIVITY ISSUE: intercoder & intracoder reliability*
  - a classification procedure is reliable when it is consistent: Different people should code the same text in the same way



Dictionaries

- A **sentiment analysis dictionary** contains information about the emotions or polarity expressed by words, phrases, or concepts. In practice, a **dictionary** usually provides one or more scores for each word. We can then use them to compute the overall **sentiment** of an input sentence based on individual words.
- top down
- create you own dictionary
- Use a dictionary developed by other scientists
- LIWC, bing (in R), WordNet (Miller, 1990)
- Word Association nets: https://wordassociations.net/en

LIWC... *Psychometrics of Word Usage*

$109.74

Journal of Language and Social Psychology
29(1) 24–54

DOI: 10.1177/0261927X09351676
http://jls.sagepub.com
SAGE

The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods

Yla R. Tausczik[1] and James W. Pennebaker[1]

Abstract

We are in the midst of a technological revolution whereby, for the first time, researchers can link daily word use to a broad array of real-world behaviors. This article reviews several computerized text analysis methods and describes how Linguistic Inquiry and Word Count (LIWC) was created and validated. LIWC is a transparent text analysis program that counts words in psychologically meaningful categories. Empirical results using LIWC demonstrate its ability to detect meaning in a wide variety of experimental settings, including to show attentional focus, emotionality, social relationships, thinking styles, and individual differences.

https://s3-us-west-2.amazonaws.com/downloads.liwc.net/LIWC2015_OperatorManual.pdf

LIWC

Summary Variable
Analytical Thinking
Clout
Authentic
Emotional Tone

Informal Speech	informal
Swear words	swear
Netspeak	netspeak
Assent	assent
Nonfluencies	nonfl
Fillers	filler

With the exception of the summary variables and words per sentence, all LIWC2015 output variables are expressed as percentage of total words.

All Punctuation[5]	Allpunc
Periods	Period
Commas	Comma
Colons	Colon
Semicolons	SemiC
Question marks	QMark
Exclamation marks	Exclam
Dashes	Dash
Quotation marks	Quote
Apostrophes	Apostro
Parentheses (pairs)	Parenth
Other punctuation	OtherP

Language Metrics	
Words per sentence[1]	WPS
Words>6 letters	Sixltr
Dictionary words	Dic
Function Words	function
Total pronouns	pronoun
Personal pronouns	ppron
1st pers singular	i
1st pers plural	we
2nd person	you
3rd pers singular	shehe
3rd pers plural	they
Impersonal pronouns	ipron
Articles	article
Prepositions	prep
Auxiliary verbs	auxverb
Common adverbs	adverb
Conjunctions	conj
Negations	negate

Grammar Other	
Regular verbs	verb
Adjectives	adj
Comparatives	compare
Interrogatives	interrog
Numbers	number
Quantifiers	quant

Word count: people who is lying use more words!!!
Hancock, Curry, Goorha, and Woodworth (2008)
Extrovert people use more words (Pennebaker & King, 1999

LIWC

Summary Variable
Analytical Thinking
Clout
Authentic
Emotional Tone

Informal Speech	informal
Swear words	swear
Netspeak	netspeak
Assent	assent
Nonfl	nonfl
Fillers	fil

With the exception of the summary va
sentence, all LIWC2015 output variabl
percentage of total words.

Fong, A., Roozenbeek, J., Goldwert, D., Rathje, S., & van der Linden, S. (2021). The language of conspiracy: A psychological analysis of speech used by conspiracy theorists and their followers on Twitter. *Group Processes & Intergroup Relations*, *24*(4), 606-623.

Language Metrics	
Words per sentence[1]	WPS
Words>6 letters	Sixltr
Dictionary words	Dic
Function Words	function
Total pronouns	pronoun
Personal pronouns	ppron
1st pers singular	i
1st pers plural	we
	you
	shehe
	they
	ipron
	article
	prep
	auxverb
	adverb
	conj
Negations	negate

All Punctuation[5]	Allpunc
Periods	Period
Commas	Comma
Colons	Colon
Semicolons	SemiC
Question marks	QMark
Exclamation marks	Exclam
Dashes	Dash
Quotation marks	Quote
Apostrophes	Apostro
Parentheses (pairs)	Parenth
Other punctuation	OtherP

Grammar Other	
Regular verbs	verb
Adjectives	adj
Comparatives	compare
Interrogatives	interrog
Numbers	number
Quantifiers	quant

LIWC

Summary Variable
Analytical Thinking
Clout
Authentic
Emotional Tone

Informal Speech	informal
Swear words	swear
Netspeak	netspeak
Assent	assent
Nonfluencies	nonfl
Fillers	filler

Language Metrics	
Words per sentence[1]	WPS
Words>6 letters	Sixltr
Dictionary words	Dic
Function Words	function
Total pronouns	pronoun
Personal pronouns	ppron
1st pers singular	i
1st pers plural	we
2nd person	you
3rd pers singular	shehe
3rd pers plural	they
Impersonal pronouns	ipron
Articles	article
Prepositions	prep
Auxiliary verbs	auxverb
Common adverbs	adverb
Conjunctions	conj
Negations	negate

With the exception of the summary variables [illegible] per sentence, al[illegible] percentage

People who are experiencing physical or emotional pain tend to have their attention drawn to themselves and subsequently use more first-person singular pronouns (e.g., Rude, Gortner, & Pennebaker, 2004).
When people sit in front of a mirror and complete a questionnaire, they use more words such as “I” and “me” than when the mirror is not present (Davis & Brock, 1975)

All Punctuation[5]
Periods
Commas
Colons
Semicolons
Question marks
Exclamation marks
Dashes
Quotation marks
Apostrophes
Parentheses (pairs)
Other punctuation

Grammar Other	
Regular verbs	verb
Adjectives	adj
Comparatives	compare
Interrogatives	interrog
Numbers	number
Quantifiers	quant

LIWC

Summary Variable
Analytical Thinking
Clout
Authentic
Emotional Tone

Informal Speech	informal
Swear words	swear
Netspeak	netspeak
Assent	assent
Nonfluencies	nonfl
Fillers	filler

Language Metrics	
Words per sentence[1]	WPS
Words>6 letters	Sixltr
Dictionary words	Dic
Function Words	function
Total pronouns	pronoun
Personal pronouns	ppron
1st pers singular	i
1st pers plural	we
2nd person	you
3rd pers singular	shehe
3rd pers plural	they
Impersonal pronouns	ipron
Articles	article
Prepositions	prep
Auxiliary verbs	auxverb
Common adverbs	adverb
Conjunctions	conj
Negations	negate

With the exception of the summary variables a per sentence, all LIWC2015 output variables are
percentag

All Punctuation[5]
Periods
Commas
Colons
Semicolons
Question marks
Exclamation marks
Dashes
Quotation marks
Apostrophes
Parentheses (pairs)
Other punctuation

STATUS Across five studies in which status was either experimentally manipulated, determined by partner ratings, or based on existing titles, increased use of first-person plural was a good predictor of higher status, and in four of the studies increased use of first-person singular was a good predictor of lower status (Kacewicz, Pennebaker, Davis, Jeon, & Graesser, 2009)

Grammar Other	
Regular verbs	verb
Adjectives	adj
Comparatives	compare
Interrogatives	interrog
Numbers	number
Quantifiers	quant

LIWC

Summary Variable
Analytical Thinking
Clout
Authentic
Emotional Tone

Informal Speech	informal
Swear words	swear
Netspeak	netspeak
Assent	assent
Nonfluencies	nonfl
Fillers	filler

Language Metrics	
Words per sentence[1]	WPS
Words>6 letters	Sixltr
Dictionary words	Dic
Function Words	function
Total pronouns	pronoun
Personal pronouns	ppron
1st pers singular	i
1st pers plural	we
2nd person	you
3rd pers singular	shehe
3rd pers plural	they
Impersonal pronouns	ipron
Articles	article
Prepositions	prep
Auxiliary verbs	auxverb
Common adverbs	adverb
Conjunctions	conj
Negations	negate

With the exception of the summary variables and [illegible] per sentence, all LIWC2015 output variables [illegible] as percentag[illegible]

All Punctuation[5]
Periods
Commas
Colons
Semicolons
Question marks
Exclamation marks
Dashes
Quotation marks
Apostrophes
Parentheses (pairs)
Other punctuation

relationship quality

first-person plural (“we”) has not been found to be related to higher relationship quality, instead use of second person (“you”) is more important in predicting lower-quality relationships. Simmons, Chambless, and Gordon (2008

Grammar Other	
Regular verbs	verb
Adjectives	adj
Comparatives	compare
Interrogatives	interrog
Numbers	number
Quantifiers	quant

LIWC

Summary Variable
Analytical Thinking
Clout
Authentic
Emotional Tone

Informal Speech	informal
Swear words	swear
Netspeak	netspeak
Assent	assent
Nonfluencies	nonfl
Fillers	filler

With the exception of the summary variables and words per sentence, all LIWC2015 output variables are expressed as percentag

All Punctuation[5]
Periods
Commas
Colons
Semicolons
Question marks
Exclamation marks
Dashes
Quotation marks
Apostrophes
Parentheses (pairs)
Other punctuation

COHERENCE

Conjunctions (e.g., and, also, although) join multiple thoughts together and are important for creating a coherent narrative (Graesser, McNamara, Louwerse, & Cai, 2004).

Language Metrics	
Words per sentence[1]	WPS
Words>6 letters	Sixltr
Dictionary words	Dic
Function Words	function
Total pronouns	pronoun
Personal pronouns	ppron
1st pers singular	i
1st pers plural	we
2nd person	you
3rd pers singular	shehe
3rd pers plural	they
Impersonal pronouns	ipron
Articles	article
Prepositions	prep
Auxiliary verbs	auxverb
Common adverbs	adverb
Conjunctions	conj
Negations	negate

Grammar Other	
Regular verbs	verb
Adjectives	adj
Comparatives	compare
Interrogatives	interrog
Numbers	number
Quantifiers	quant

LIWC

Summary Variable
Analytical Thinking
Clout
Authentic
Emotional Tone

Informal Speech	informal
Swear words	swear
Netspeak	netspeak
Assent	assent
Nonfluencies	nonfl
Fillers	filler

With the exception of the summary variables and words per sentence, all LIWC2015 output variables are expressed as percentage of total words.

All Punctuation[5]	Allpunc
Periods	Period
Commas	Comma
Colons	Colon
Semicolons	SemiC
Question marks	QMark
Exclamation marks	Exclam
Dashes	Dash
Quotation marks	Quote
Apostrophes	Apostro
Parentheses (pairs)	Parenth
Other punctuation	OtherP

Language Metrics	
Words per sentence[1]	WPS
Words>6 letters	Sixltr
Dictionary words	Dic
Function Words	function
Total pronouns	pronoun
Personal pronouns	ppron
1st pers singular	i
1st pers plural	we
2nd person	you
3rd pers singular	shehe
3rd pers plural	they
Impersonal pronouns	ipron
Articles	article
Prepositions	prep
Auxiliary verbs	auxverb
Common adverbs	adverb
Conjunctions	conj
Negations	negate

Grammar Other	
Regular verbs	verb
Adjectives	adj
Comparatives	compare
Interrogatives	interrog
Numbers	number
Quantifiers	quant

LIWC

People experiencing physical or emotional pain tend to use more first-person singular pronouns (Rude, Gortner, & Pennebaker, 2004).

Depressed patients are more likely to use more first-person singular and more negative emotion words than participants who have never been depressed in emotional writings (Rude et al., 2004)

Language Metrics	
Words per sentence[1]	WPS
Words>6 letters	Sixltr
Dictionary words	Dic
Function Words	function
Total pronouns	pronoun
Personal pronouns	ppron
1st pers singular	i
1st pers plural	we
2nd person	you
3rd pers singular	shehe
3rd pers plural	they
Impersonal pronouns	ipron
Articles	article
Prepositions	prep
Auxiliary verbs	auxverb
Common adverbs	adverb
Conjunctions	conj

When people sit in front of a mirror use more words such as "I" and "me" than when the mirror is not present (Davis & Brock, 1975)

Semicolons	SemiC
Question marks	QMark
Exclamation marks	Exclam
Dashes	Dash
Quotation marks	Quote

Comparatives	compare
Interrogatives	interrog
Numbers	number
Quantifiers	quant

"we" can signal a sense of group identity, such as when couples are asked to evaluate their marriages to an interviewer, the more the participants use "we," the better their marriage (Simmons, Gordon, & Chambless, 2005)

www.secretlifeofpronouns.com)

Psycho-social index

Affect Words	affect
Positive emotion	posemo
Negative emotion	negemo
Anxiety	anx
Anger	anger
Sadness	sad

Social Words	social
Family	family
Friends	friend
Female referents	female
Male referents	male

Positive political ads used more present and future tense verbs, and negative ads used more past tense verbs (Gunsch et al., 2000). From the tense of the verbs and the personal pronouns used, we can infer that negative ads focus on past actions of the opponent, and positive ads focus on the present and future acts of the candidate.

Core Drives and Needs	drives
Affiliation	affiliation
Achievement	achieve
Power	power
Reward focus	reward
Risk/prevention focus	risk
Time Orientation[4]	
Past focus	focuspast
Present focus	focuspresent
Future focus	focusfuture
Relativity	relativ
Motion	motion
Space	space
Time	time

Personal Concerns	
Work	work
Leisure	leisure
Home	home
Money	money
Religion	relig
Death	death

Psycho-social index

Social Words	social
Family	family
Friends	
Female	
Male re	

Depressed and suicidal individuals are more self-focused, express more negative emotion and sometime use more death-related words. . Depressed patients are more likely to use more first-person singular and more negative emotion words than participants who have never been depressed in emotional writings (Rude et al., 2004

Affect Words	affect
Positive emotion	posemo
Negative emotion	negemo
Anxiety	anx
Anger	anger
Sadness	sad

Core Drives an us	drives
Affiliation	affiliation
Achiev	achieve
Po	power
cus	reward
revention focus	risk
e Orientation[4]	
ast focus	focuspast
Present focus	focuspresent
Future focus	focusfuture
Relativity	relativ
Motion	motion
Space	space
Time	time

Personal Concerns	
Work	work
Leisure	leisure
Home	home
Money	money
Religion	relig
Death	death

Psycho-social index

Social Words	social
Family	family
Friends	friend
Female referents	female
Male referents	male

Pasupathi, 2007
Participants were asked to either recall an event that they had discussed with someone else, or an undisclosed event
past tense in discussing a disclosed event and greater present tense in discussing an undisclosed event.

Affect Words	affect
Positive emotion	posemo
Negative emotion	negemo
Anxiety	anx
Anger	anger
Sadness	sad

Core Drives and Needs	drives
Affiliation	affiliation
Achievement	achieve
Power	power
Reward focus	reward
Risk/prevention focus	risk
Time Orientation[4]	
Past focus	focuspast
Present focus	focuspresent
Future focus	focusfuture
Relativity	relativ
Motion	motion
Space	space
Time	time

Personal Concerns	
Work	work
Leisure	leisure
Home	home
Money	money
Religion	relig
Death	death

Cognition & perceptio

LANGUEAGE AMBIGUITY (insight, tentat, Roos et al.'s (2020) is related to dogmatism (Fast & Horvitz, 2016) and politeness (Li et al., 2020).

Cognitive Processes[2]	cogproc
Insight	insight
Cause	cause
Discrepancies	discrep
Tentativeness	tentat
Certainty	certain
Differentiation[3]	differ
Perpetual Processes	percept
Seeing	see
Hearing	hear
Feeling	feel
Biological Processes	bio
Body	body
Health/illness	health
Sexuality	sexual
Ingesting	ingest

C esses[2]	cogproc
	insight
Cause	cause
Discrepancies	discrep
Tentativeness	tentat
Certainty	certain
Differentiation[3]	differ

Prepositions (e.g., to, with, above), cognitive mechanisms (e.g., cause, know, ought), and words greater than six letters are all also indicative of more complex language. Cognitive complexity can be thought of as a richness of two components of reasoning: the extent to which someone differentiates between multiple competing solutions and the extent to which someone integrates among solutions (Tetlock, 1981)

Incivility score in LIWC

- Addition of Swear, Anger, and Negative Emotions (based on previous research, see Ksiazek et al., 2015; Stoll et al., 2020

Sentiment /emotion tools

- vader_df function of the VADER package (version 0.2.1, Roehrick, 2020). VADER Sentiment Analysis. VADER (Valence Aware Dictionary and sEntiment Reasoner) is a lexicon and rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media, and works well on texts from other domains. https://github.com/cjhutto/vaderSentiment
- EmoLex, ANEW, SentiWordNet are designed to analyze larger sets of emotional categories
- General Inquirer (GI) human curated dictionary that operates over a broader set of topics (e.g., power, weakness)
- Empath allows researchers to performtext analyses over a broader set of topical and emotional categories than existing tools, and also to create and validate newcategories on demand
  (PDF) Empath: Understanding Topic Signals in Large-Scale Text. Available from: https://www.researchgate.net/publication/301872654_Empath_Understanding_Topic_Signals_in_Large-Scale_Text [accessed Nov 08 2023]. deceptive reviews convey strongersentiment across both positively and negatively charged cat-egories, and tend towards exaggerated language

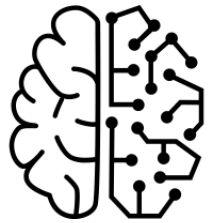
Natural language processing (NLP)

Natural language processing (**NLP**) is a subfield of linguistics, computer science, and artificial intelligence concerned with the interactions between computers and human language, in particular how to program computers to process and analyze large amounts of natural language data.

- tokenization
- grammatical role POS (part of speech) tagging (sbj, obj..)
- stemming
- thesauri
- shallow parsing : identifies constituent parts of sentences (nouns, verbs, adjectives, etc.)

the hand-coding of a set of rules, coupled with a dictionary lookup

Machine learning



Supervised machine learning algorithms *apply what has been learned in the past to new data using labeled examples to predict future events. Starting from the analysis of a known training dataset, the learning algorithm produces an inferred function to make predictions about the output values. The system is able to provide targets for any new input after sufficient training. The learning algorithm can also compare its output with the correct, intended output and find errors in order to modify the model accordingly.*

Unsupervised machine learning algorithms *are used when the information used to train is neither classified nor labeled. Unsupervised learning studies how systems can infer a function to describe a hidden structure from unlabeled data. The system doesn't figure out the right output, but it explores the data and can draw inferences from datasets to describe hidden structures from unlabeled data.*

Content Analysis

- Detect systematic patterns in communication
  - -> *topic identification*

→ opinions

Sentiment Analysis

- extract, quantify, and study **affective states** and subjective information

→ attitudes

refers to the use of natural language processing, text analysis, computational linguistics, and biometrics to systematically identify

ANALYSE DATA

- -> frequency
- -> correlations
- -> source comparison
- -> networks: centrality measures, community detection etc