

J. Sytsma

Corpus Methods & Experimental Philosophy



Broad Conception:

Experimental Philosophy is the practice of systematically collecting and analyzing empirical data in attempting to answer philosophical questions... and doing so in a way that seriously engages with the surrounding philosophical literature.

Corpus Methods & Experimental Philosophy

Have employed many methods...

But the dominant tool has been
questionnaire methods.

Often give participants a vignette (story),
and ask a question involving the application
of a term of interest (“concept”).

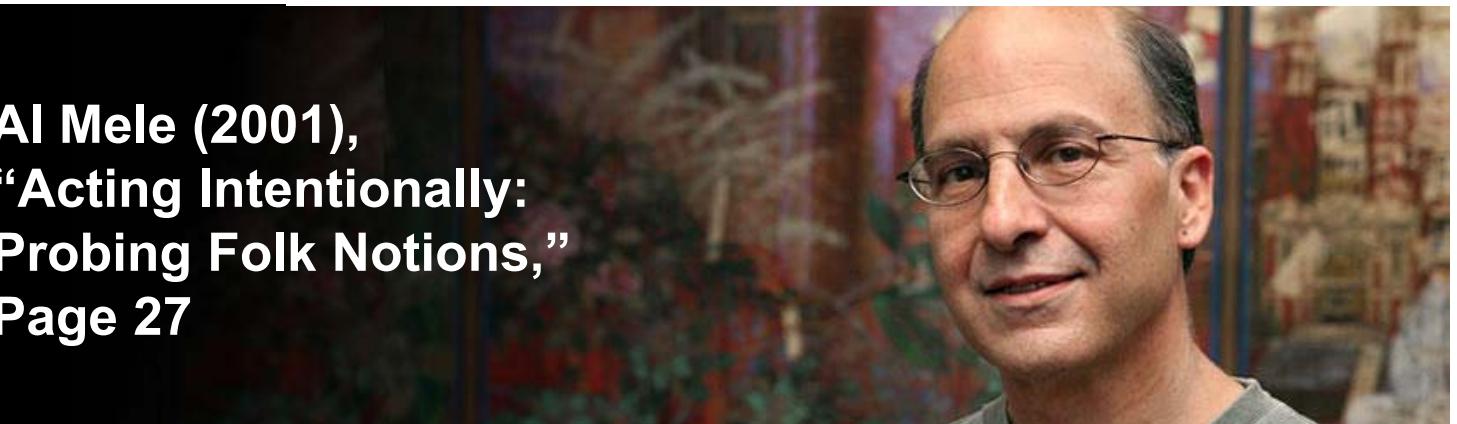
For example, whether the concept of
intentionality applies to a given action...

Corpus Methods & Experimental Philosophy

“What is it to do something *intentionally*?

“Philosophical work on this question is motivated by a variety of interrelated concerns. In trying to understand and explain human action, a project that is as old as Plato and Aristotle, philosophers of action are concerned primarily with intentional actions. In discussions of freedom of action, intentional action also naturally occupies center stage. And although people are morally accountable for some unintentional actions, as in cases of negligence, moral assessment of actions is focused primarily upon intentional actions.”

Al Mele (2001),
“Acting Intentionally:
Probing Folk Notions,”
Page 27



Corpus Methods & Experimental Philosophy

“In my opinion, any adequate answer... will be **anchored by common-sense judgments about particular hypothetical or actual actions**. One can test attempted philosophical analyses of intentional action partly by ascertaining whether what these analyses entail about particular actions is in line with what the majority of non-specialists would say....

“It is also worth noting that if there is a widely shared concept of intentional action, such judgments provide evidence about what that concept is, and a philosophical analysis of intentional action that is wholly unconstrained by that concept runs the **risk of having nothing more than a philosophical fiction as its subject matter.**”

Al Mele (2001),
“Acting Intentionally:
Probing Folk Notions,”
Page 27



Corpus Methods & Experimental Philosophy

One question (among many): Does the ordinary concept of intentionality allow that a side effect could be brought about intentionally?

Philosophers have given different answers, generally calling on their own sense of the concept or what is meant by the term.

A classic piece of experimental philosophy complicated the answer...

Corpus Methods & Experimental Philosophy



**Knobe (2003),
“Intentional Action and
Side Effects in Ordinary
language,” *Analysis*,
63(3): 190-4**

78 Participants
Manhattan Park
One of two conditions

Harm Condition

The vice-president of a company went to the chairman of the board and said, “We are thinking of starting a new program. It will help us increase profits, but it will also harm the environment.”

The chairman of the board answered, “I don’t care at all about harming the environment. I just want to make as much profit as I can. Let’s start the new program.”

They started the new program. Sure enough, the environment was harmed.

- (1) How much blame does the chairman deserves (0 to 6)
- (2) Did he intentionally harm the environment?

82% = Intentionally Harmed

Help Condition

The vice-president of a company went to the chairman of the board and said, “We are thinking of starting a new program. It will help us increase profits, but it will also help the environment.”

The chairman of the board answered, “I don’t care at all about helping the environment. I just want to make as much profit as I can. Let’s start the new program.”

They started the new program. Sure enough, the environment was helped.

- (1) How much praise does the chairman deserves (0 to 6)
- (2) Did he intentionally help the environment?

23% = Intentionally Helped

82% vs 23%!

General type of worry: Maybe such results are driven by something about the experimental context.

Perhaps the vignette suggests to p's that they're expected to answer in this way, or...



Corpus Methods & Experimental Philosophy

Especially when we're investigating the ordinary use of terms ("conceptual analysis"), relevant to also get information about the use of these terms *in the wild*, outside of the experimental context.

Address worries

Consilience of evidence

General methodological advice is to bring multiple channels of evidence to bear on a question of interest.



Corpus Methods & Experimental Philosophy

Corpus Analysis is especially well-suited to investigating the ordinary meaning of terms.

Corpus = collection of texts

Analysis = looking at linguistic data in a corpus and assessing it for a research purpose

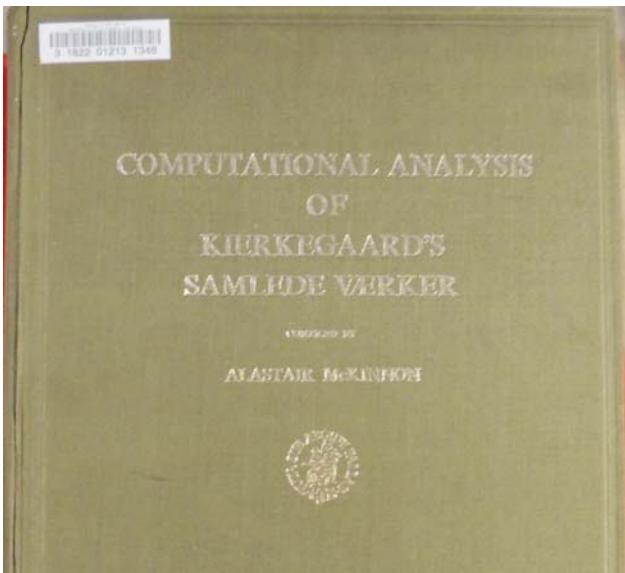
Lots of corpora available (and you can build your own)...

And a lot of things that can be done with them...

Corpus Methods & Experimental Philosophy

A decent amount of work has been done in philosophy discussing or using corpus methods: [corpus_refs.pdf](#) ★

In fact, a reasonably long history...



Computers and the Humanities, Vol. 10, pp. 281-286. PERGAMON PRESS, 1976. Printed in the U.S.A.

A System for Text and Content Analysis

JEAN-GUY MEUNIER, STANISLAS ROLLAND, and FRANÇOIS DAOUST

THE INTRODUCTION of the computer in the last ten years as a tool for the exploration and analysis of literary data has not been substitution for the intelligence of the researcher but an extension of his memory and organizing capabilities. In the investigation of many disciplines based on textual data (literature, history, sociology, philosophy, theology), the researcher often has to manage a vast quantity of information. As he must frequently parse, recall or organize his text and make inferences from it, he may write cards, underline, annotate with numbers, copy references, generate new texts, produce concept associations, build concordances and so on. In fact, many of his procedures treat the text as nothing more than a set of elements held together by an organizing structure. At this physical level, a computer can be of great assistance, for it can offer many systematically organized data structures for a text and recall hierarchically organized information that varies according to the chosen level of interpretation, be it morphological, lexical or even conceptual.

A textual data processing system for literary and social sciences, called SATO (Système d'analyse des

content analysis, such as the General Inquirer.¹ In the first of these, many textpack systems oriented towards the manipulation of large textual data bases have provided indexes, concordances and basic frequency counts. Used mainly for stylistic, philological and lexical studies, as well as the statistical analysis of authorship, these textpacks are usually programmed to present a simple interface to the literary scholar who has no special knowledge of computer programming. The second of these areas, content analysis, has been more concerned with methodology. Of secondary importance is the occasionally large volume of textual data, which may come from newspapers, interviews, etc. The analysis techniques are more sophisticated, for example cluster analysis, factorial analysis, categorial and correlation analysis. The SATO system has been conceived with these two functions in mind, i.e., classical literary research and content analysis either on standard or transformed texts. Like the other textpacks, it does not require user programming.

The Text. The first step in textual data processing is usually pre-editing for a computer. In the wide variety of texts to be manipulated, one encounters

Includes use of the Web as a corpus (WaC).

Phenom Cogn Sci
DOI 10.1007/s11097-007-9066-y

REGULAR ARTICLE

Intuitions about consciousness: Experimental studies

Joshua Knobe · J. Prinz

© Springer Science

Abstract When having certain kinds of thoughts, we often do not standpoints or frames of reference to determine how we experience them. The results point to the importance of phenomenological descriptions of states of mind, and the sensitivity to particular kinds of thoughts that they can have on our conscious experience.

Keywords Consciousness · Experimental philosophy · Phenomenal states · Non-phenomenal states · Intuitions · Microsoft

Here are the phrases ascribing non-phenomenal states, along with the number of hits that each phrase received:

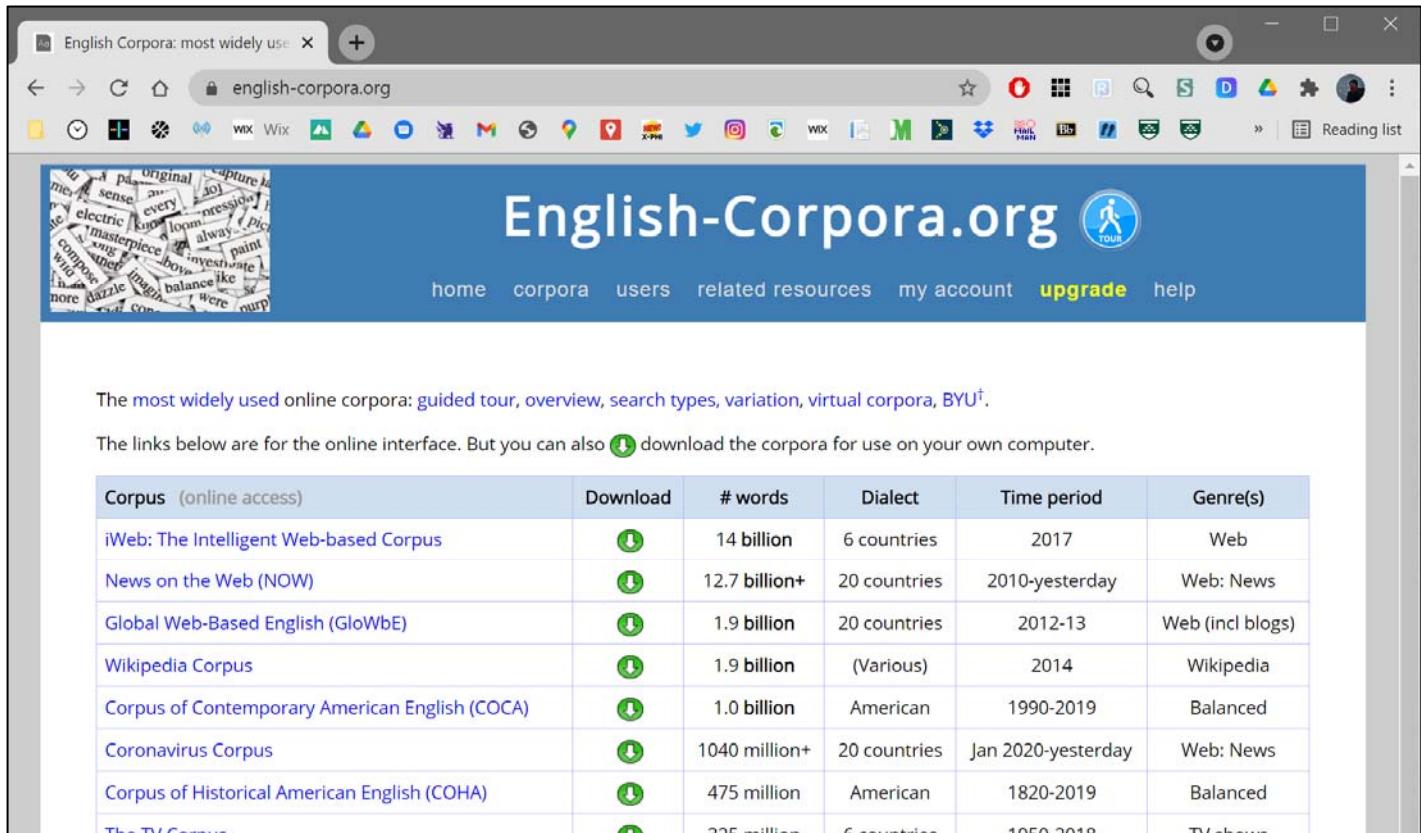
'Microsoft intends'	25,700
'Microsoft decides'	11,400
'Microsoft tries'	52,600
'Microsoft wants'	135,000
'Microsoft believes'	31,100
'Microsoft hopes'	56,600
'Microsoft loves'	4,100
'Microsoft hates'	970

And here are the phrases ascribing phenomenal states:

'Microsoft feels depressed'	0
'Microsoft experiences joy'	0
'Microsoft feels happy'	0
'Microsoft feels pain'	2
'Microsoft feels angry'	0
'Microsoft feels scared'	0

Corpus Methods & Experimental Philosophy

And the Web for building a corpus (*WfC*),
as well as accessing them.

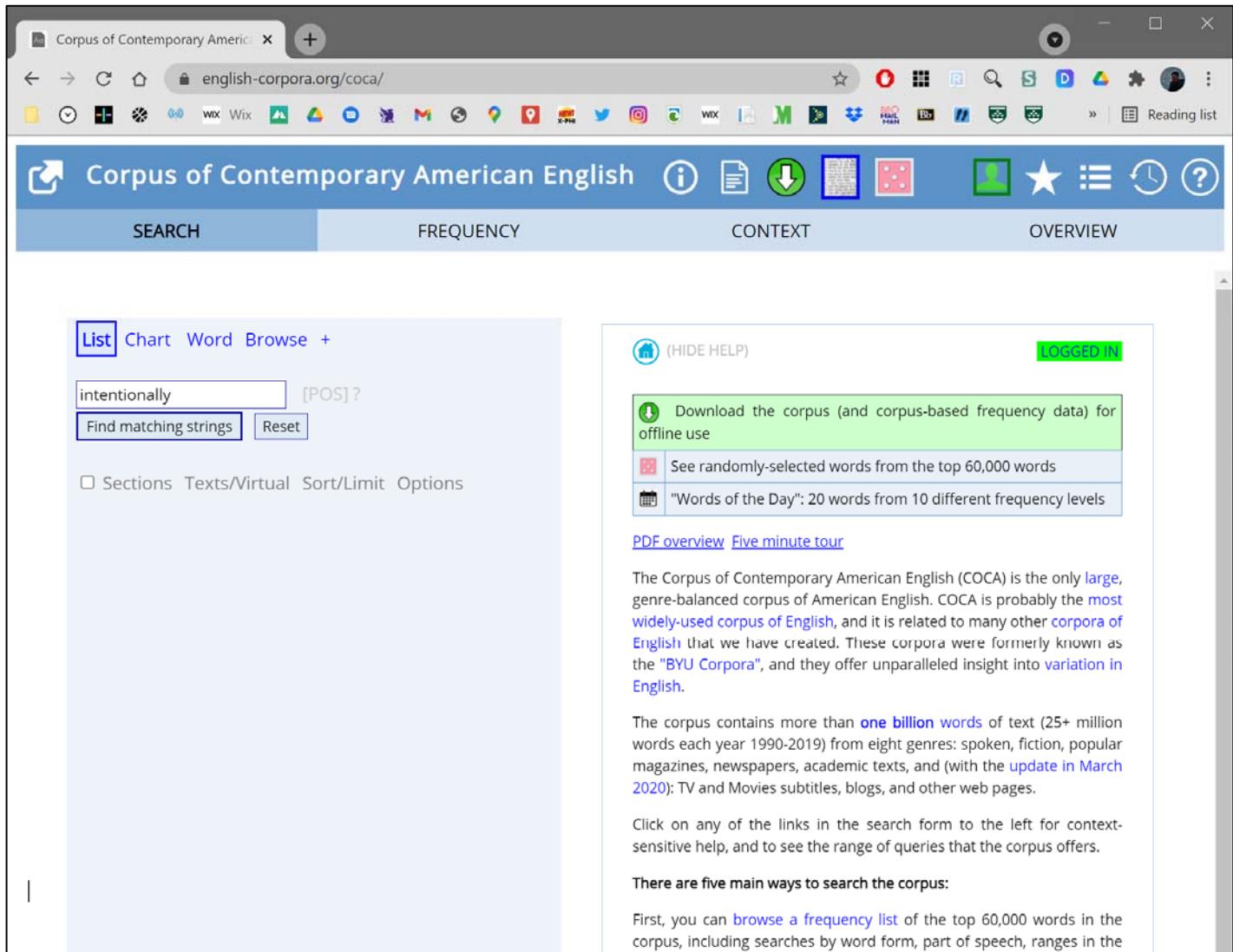


The most widely used online corpora: guided tour, overview, search types, variation, virtual corpora, BYU[†].

The links below are for the online interface. But you can also  download the corpora for use on your own computer.

Corpus (online access)	Download	# words	Dialect	Time period	Genre(s)
iWeb: The Intelligent Web-based Corpus		14 billion	6 countries	2017	Web
News on the Web (NOW)		12.7 billion+	20 countries	2010-yesterday	Web: News
Global Web-Based English (GloWbE)		1.9 billion	20 countries	2012-13	Web (incl blogs)
Wikipedia Corpus		1.9 billion	(Various)	2014	Wikipedia
Corpus of Contemporary American English (COCA)		1.0 billion	American	1990-2019	Balanced
Coronavirus Corpus		1040 million+	20 countries	Jan 2020-yesterday	Web: News
Corpus of Historical American English (COHA)		475 million	American	1820-2019	Balanced
The TV Corpus		325 million	6 countries	1950-2010	TV shows

Corpus Methods & Experimental Philosophy



The screenshot shows a web browser window displaying the COCA website at english-corpora.org/coca/. The interface includes a search bar with the query "intentionally", a frequency chart, and sections for context and overview. On the right, there's a sidebar with user information and links to download the corpus or see words of the day.

Corpus of Contemporary American English

SEARCH FREQUENCY CONTEXT OVERVIEW

List Chart Word Browse +

intentionally [POS] ?

Find matching strings Reset

Sections Texts/Virtual Sort/Limit Options

(HIDE HELP) LOGGED IN

Download the corpus (and corpus-based frequency data) for offline use

See randomly-selected words from the top 60,000 words

"Words of the Day": 20 words from 10 different frequency levels

PDF overview Five minute tour

The Corpus of Contemporary American English (COCA) is the only large, genre-balanced corpus of American English. COCA is probably the most widely-used corpus of English, and it is related to many other corpora of English that we have created. These corpora were formerly known as the "BYU Corpora", and they offer unparalleled insight into variation in English.

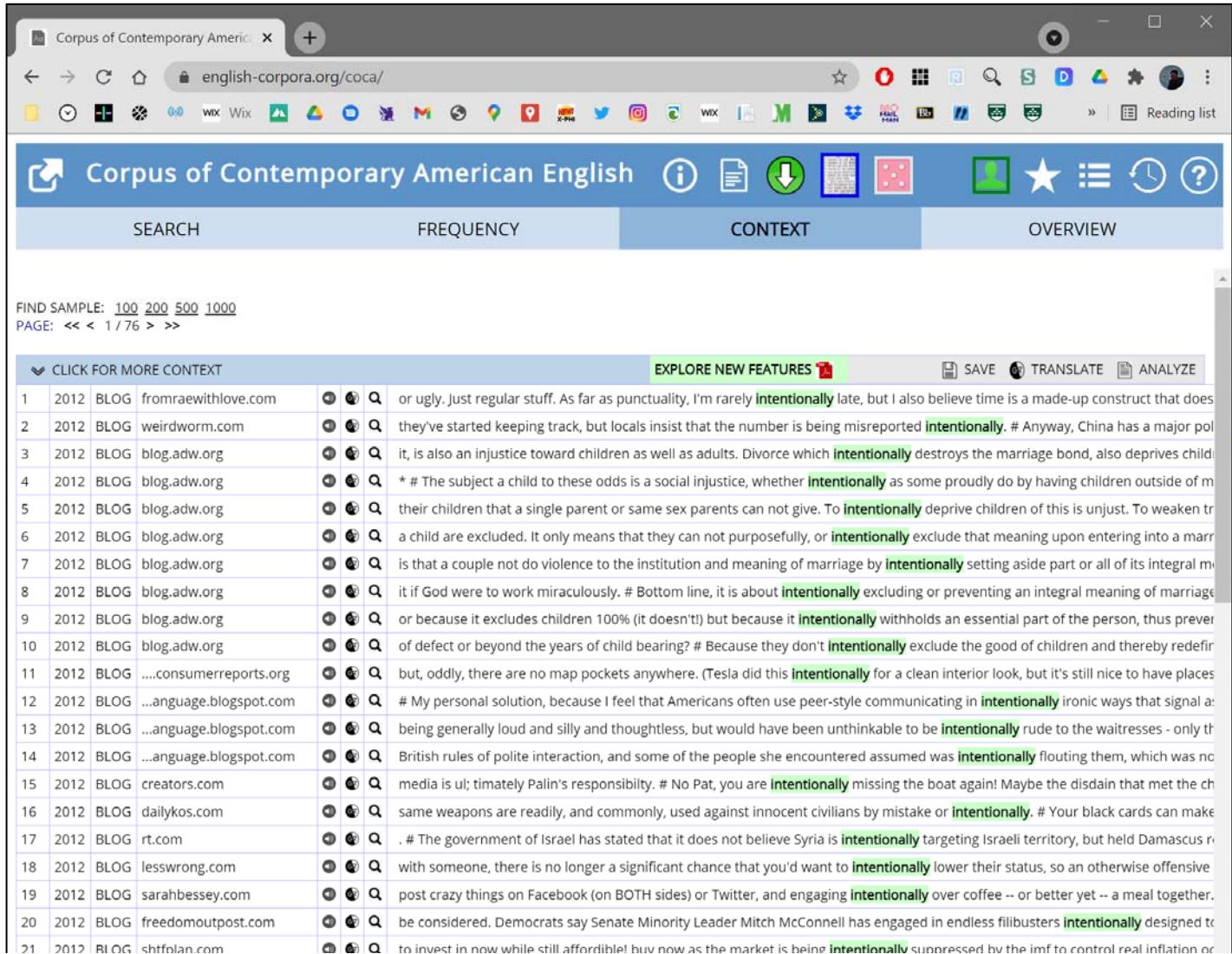
The corpus contains more than one billion words of text (25+ million words each year 1990-2019) from eight genres: spoken, fiction, popular magazines, newspapers, academic texts, and (with the update in March 2020): TV and Movies subtitles, blogs, and other web pages.

Click on any of the links in the search form to the left for context-sensitive help, and to see the range of queries that the corpus offers.

There are five main ways to search the corpus:

First, you can browse a frequency list of the top 60,000 words in the corpus, including searches by word form, part of speech, ranges in the

Corpus Methods & Experimental Philosophy



The screenshot shows a web browser displaying the COCA interface. The URL in the address bar is english-corpora.org/coca/. The interface has a blue header with the title "Corpus of Contemporary American English" and various navigation icons. Below the header, there are four tabs: "SEARCH" (selected), "FREQUENCY", "CONTEXT", and "OVERVIEW". Under the "SEARCH" tab, there are buttons for "CLICK FOR MORE CONTEXT" and "EXPLORE NEW FEATURES". The main content area displays a list of 21 search results, each with a number, date, source type, source URL, and a snippet of text containing the word "intentionally". The results are from various sources in 2012, including blogs like fromraewithlove.com, weirdworm.com, blog.adw.org, and creators.com.

1	2012	BLOG	fromraewithlove.com	or ugly. Just regular stuff. As far as punctuality, I'm rarely intentionally late, but I also believe time is a made-up construct that does
2	2012	BLOG	weirdworm.com	they've started keeping track, but locals insist that the number is being misreported intentionally . # Anyway, China has a major pol
3	2012	BLOG	blog.adw.org	it, is also an injustice toward children as well as adults. Divorce which intentionally destroys the marriage bond, also deprives child
4	2012	BLOG	blog.adw.org	* # The subject a child to these odds is a social injustice, whether intentionally as some proudly do by having children outside of m
5	2012	BLOG	blog.adw.org	their children that a single parent or same sex parents can not give. To intentionally deprive children of this is unjust. To weaken tr
6	2012	BLOG	blog.adw.org	a child are excluded. It only means that they can not purposefully, or intentionally exclude that meaning upon entering into a marr
7	2012	BLOG	blog.adw.org	is that a couple not do violence to the institution and meaning of marriage by intentionally setting aside part or all of its integral m
8	2012	BLOG	blog.adw.org	it if God were to work miraculously. # Bottom line, it is about intentionally excluding or preventing an integral meaning of marriage
9	2012	BLOG	blog.adw.org	or because it excludes children 100% (it doesn't!) but because it intentionally withholds an essential part of the person, thus preve
10	2012	BLOG	blog.adw.org	of defect or beyond the years of child bearing? # Because they don't intentionally exclude the good of children and thereby redefin
11	2012	BLOG	...consumerreports.org	but, oddly, there are no map pockets anywhere. (Tesla did this intentionally for a clean interior look, but it's still nice to have places
12	2012	BLOG	...anguage.blogspot.com	# My personal solution, because I feel that Americans often use peer-style communicating in intentionally ironic ways that signal a
13	2012	BLOG	...anguage.blogspot.com	being generally loud and silly and thoughtless, but would have been unthinkable to be intentionally rude to the waitresses - only th
14	2012	BLOG	...anguage.blogspot.com	British rules of polite interaction, and some of the people she encountered assumed was intentionally flouting them, which was no
15	2012	BLOG	creators.com	media is ul; timately Palin's responsibility. # No Pat, you are intentionally missing the boat again! Maybe the disdain that met the ch
16	2012	BLOG	dailykos.com	same weapons are readily, and commonly, used against innocent civilians by mistake or intentionally . # Your black cards can make
17	2012	BLOG	rt.com	. # The government of Israel has stated that it does not believe Syria is intentionally targeting Israeli territory, but held Damascus r
18	2012	BLOG	lesswrong.com	with someone, there is no longer a significant chance that you'd want to intentionally lower their status, so an otherwise offensive
19	2012	BLOG	sarahbessey.com	post crazy things on Facebook (on BOTH sides) or Twitter, and engaging intentionally over coffee -- or better yet -- a meal together.
20	2012	BLOG	freedomoutpost.com	be considered. Democrats say Senate Minority Leader Mitch McConnell has engaged in endless filibusters intentionally designed to
21	2012	BLOG	shtfplan.com	to invest in now while still affordable! buy now as the market is being intentionally suppressed by the imf to control real inflation or

Corpus Methods & Experimental Philosophy

Can also (purchase and) download COCA to do various things with it.

Distributional Semantic Model (DSM)
using word2vec on non-academic COCA...

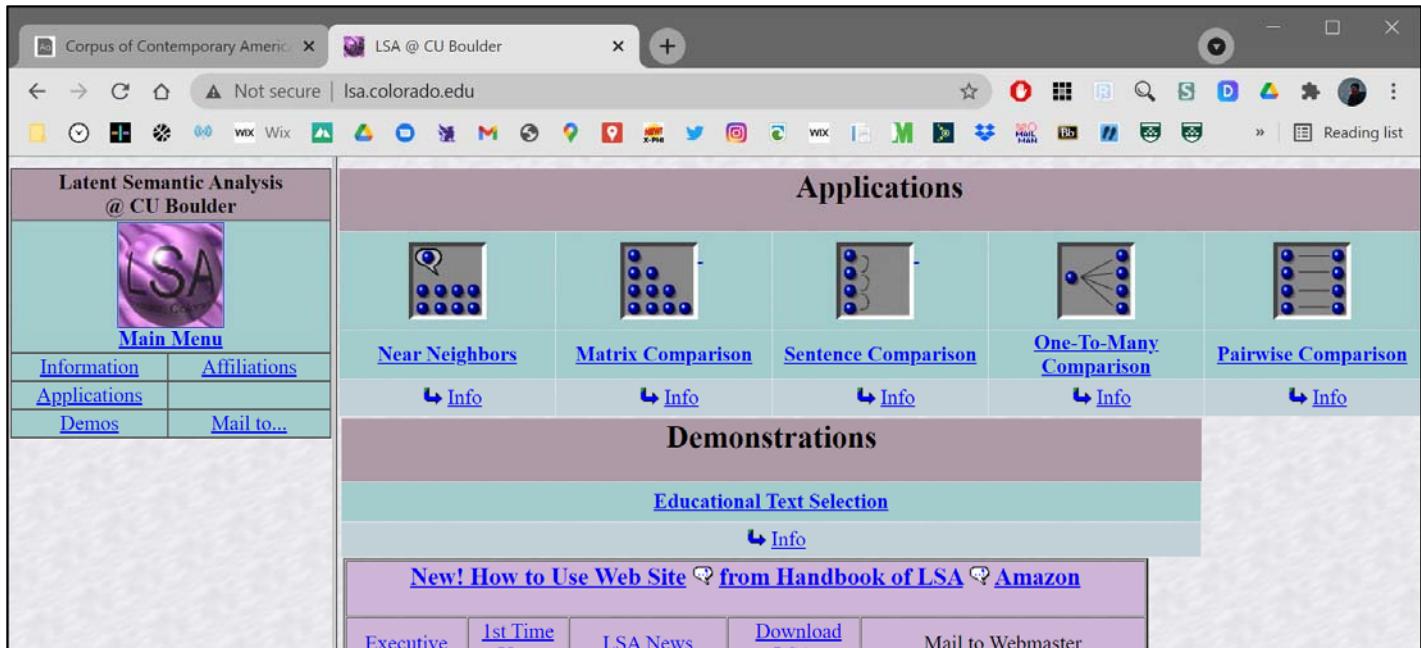
Nearest neighbors to “intentionally”:

'deliberately'	0.8082486391067505
'purposely'	0.7684776186943054
'willfully'	0.7106383442878723
'maliciously'	0.6992570161819458
'inadvertently'	0.6933962106704712
'unfairly'	0.615240216255188

Corpus Methods & Experimental Philosophy

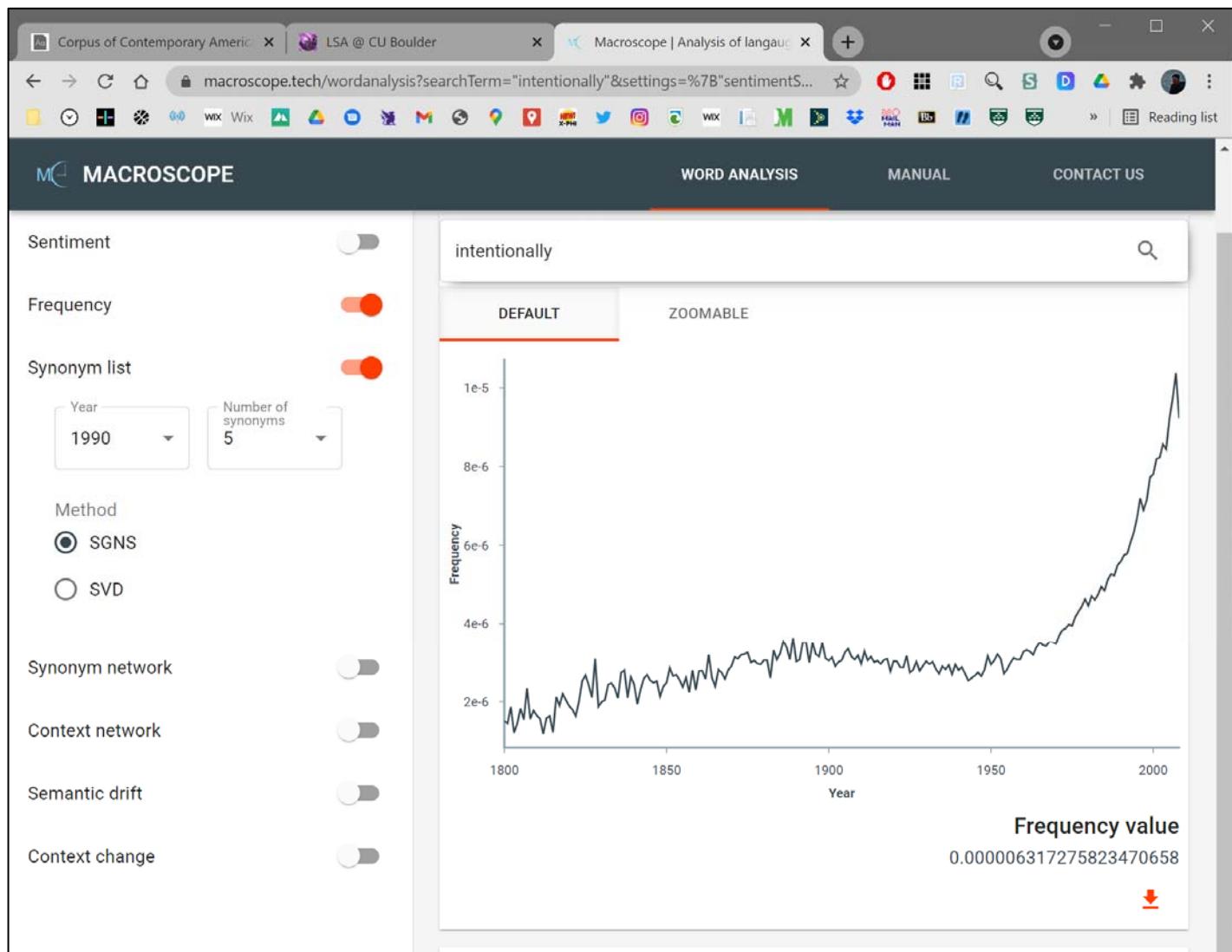
Will look more at COCA and DSM
in later sessions...

including DSM tools available on the web



The screenshot shows a web browser window with two tabs: "Corpus of Contemporary American English" and "LSA @ CU Boulder". The "LSA @ CU Boulder" tab is active, displaying the LSA homepage. The page has a sidebar on the left titled "Latent Semantic Analysis @ CU Boulder" with links for "Main Menu", "Information", "Affiliations", "Applications", and "Demos". The main content area is titled "Applications" and features five grid-based icons representing different comparison methods: "Near Neighbors", "Matrix Comparison", "Sentence Comparison", "One-To-Many Comparison", and "Pairwise Comparison". Below each icon is a "Info" link. The page also includes sections for "Demonstrations", "Educational Text Selection", and a purple banner at the bottom with the text "New! How to Use Web Site" and links to "from Handbook of LSA" and "Amazon". Navigation links at the bottom include "Executive", "1st Time", "LSA News", "Download", and "Mail to Webmaster".

Corpus Methods & Experimental Philosophy



J. Sytsma

C
E

MACROSCOPE

WORD ANALYSIS

MANUAL

CONTACT US

Sentiment

Frequency

Synonym list

Year Number of synonyms

Method

SGNS

SVD

Synonym network

Context network

Semantic drift

Context change

intentionally



DEFAULT

ZOOMABLE



Frequency value

0.000006317275823470658



Synonym list

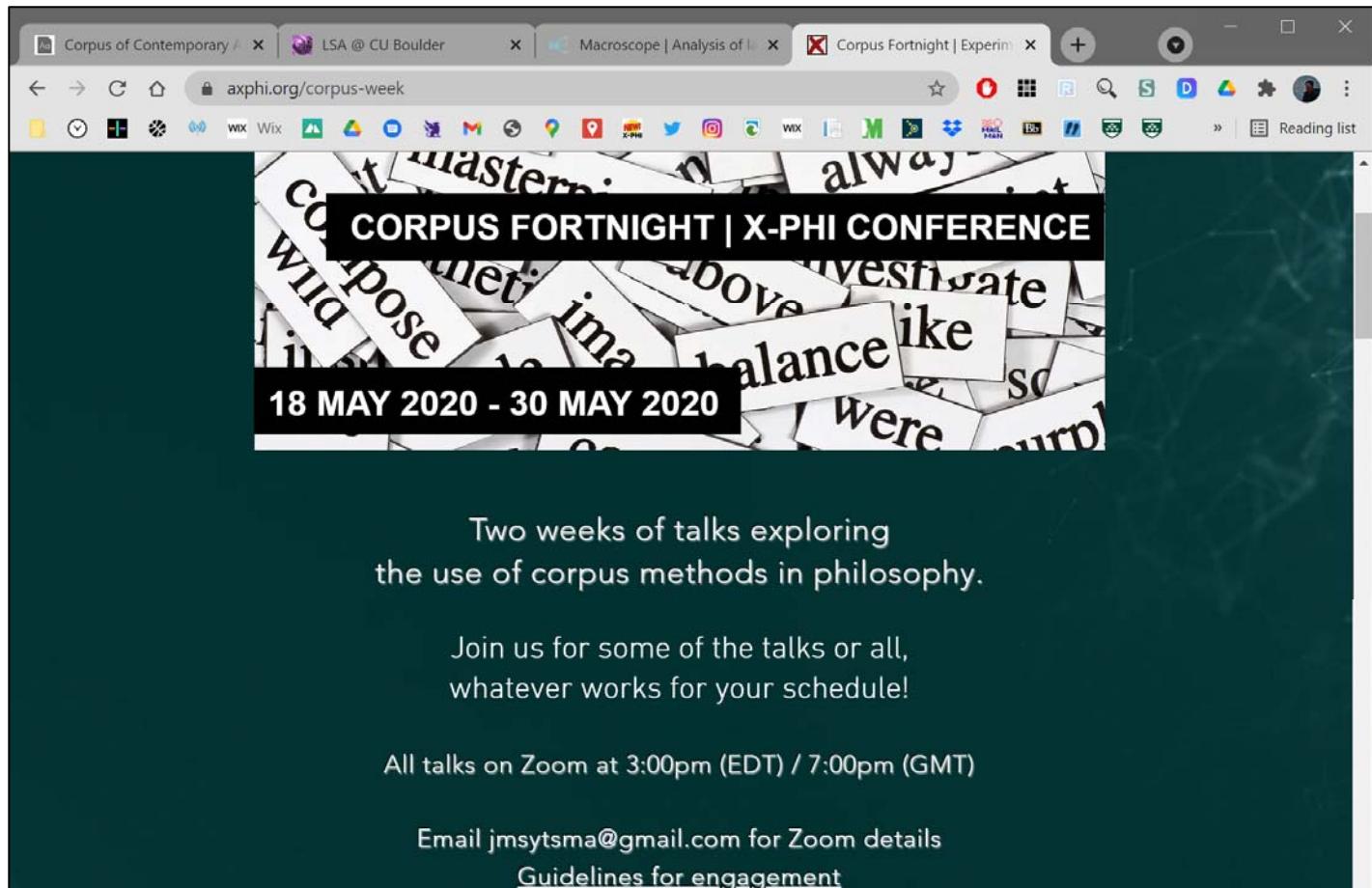
Synonyms of intentionally (1990)

Relative score ↓

knowingly	0.434855179019946
recklessly	0.4289375670632001
willfully	0.4169229154313807
purposely	0.3857500039284977
negligently	0.37632531595259316

Corpus Methods & Experimental Philosophy

Some further examples...



The screenshot shows a web browser window with several tabs open at the top. The active tab displays the website for the 'CORPUS FORTNIGHT | X-PHI CONFERENCE' held from '18 MAY 2020 - 30 MAY 2020'. The background of the page features a collage of various words and phrases. Below the main banner, text reads: 'Two weeks of talks exploring the use of corpus methods in philosophy.' A message encourages visitors to 'Join us for some of the talks or all, whatever works for your schedule!'. At the bottom, it states 'All talks on Zoom at 3:00pm (EDT) / 7:00pm (GMT)' and provides an email address 'Email jmsytsma@gmail.com for Zoom details' and a link to 'Guidelines for engagement'.

CORPUS FORTNIGHT | X-PHI CONFERENCE

18 MAY 2020 - 30 MAY 2020

Two weeks of talks exploring the use of corpus methods in philosophy.

Join us for some of the talks or all, whatever works for your schedule!

All talks on Zoom at 3:00pm (EDT) / 7:00pm (GMT)

Email jmsytsma@gmail.com for Zoom details

[Guidelines for engagement](#)

Corpus tools can be used for **testing** claims.
And they can be used for **exploration**.

Episteme (2019), 1–27
doi:10.1017/epi.2019.15

episteme

ARTICLE

A Corpus Study of “Know”: On The Verification of Philosophers’ Frequency Claims about Language

Nat Hansen^{1*}, J.D. Porter² and Kathryn Francis¹

¹Department of Philosophy, University of Reading and ²Stanford University Literary Lab

*Corresponding author. Email: n.d.hansen@reading.ac.uk

(Received 29 May 2018; revised 9 January 2019; accepted 4 April 2019)

Abstract

We investigate claims about the frequency of “know” made by philosophers. Our investigation has several overlapping aims. First, we aim to show what is required to confirm or disconfirm philosophers’ claims about the comparative frequency of different uses of philosophically interesting expressions. Second, we aim to show how using linguistic corpora as tools for investigating meaning is a productive methodology, in the sense that it yields discoveries about the use of language that philosophers would have overlooked if they remained in their “armchairs of an afternoon”, to use J.L. Austin’s phrase. Third, we discuss facts about the meaning of “know” that so far have been ignored in philosophy, with the aim of reorienting discussions of the relevance of ordinary language for philosophical theorizing.

Corpus Methods & Experimental Philosophy



“Observations about what we ‘ordinarily’ or ‘typically’ or ‘most commonly’ mean by some expression of interest have been a staple of philosophical practice at least since the heyday of ordinary language philosophy in the 1950s.” (1)

Includes philosophy claims about “know” talk:

“most of the time, outside of epistemology, when **we** consider whether somebody knows something, **we** are mainly interested in.... Similarly, when **we** say that someone does not know something, typically **we** mean....”
(Bach 2005, 62-3)

Corpus Methods & Experimental Philosophy

What would justify philosophers' frequency claims about how people ordinarily talk?

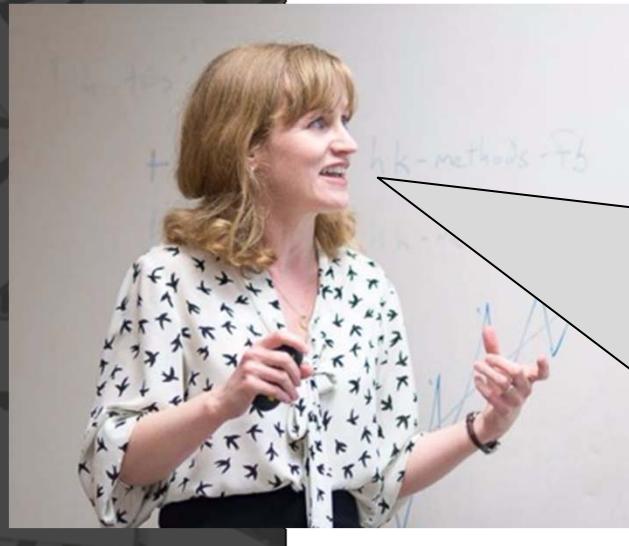


Fortunately for the philosopher of language, however, there now exist resources to investigate objective frequencies of occurrences of linguistic expressions, namely *linguistic corpora*, which are organized bodies of text, purpose-built for answering linguistic questions, such as what the comparative frequencies of various expressions are. (2)

Corpus Methods & Experimental Philosophy

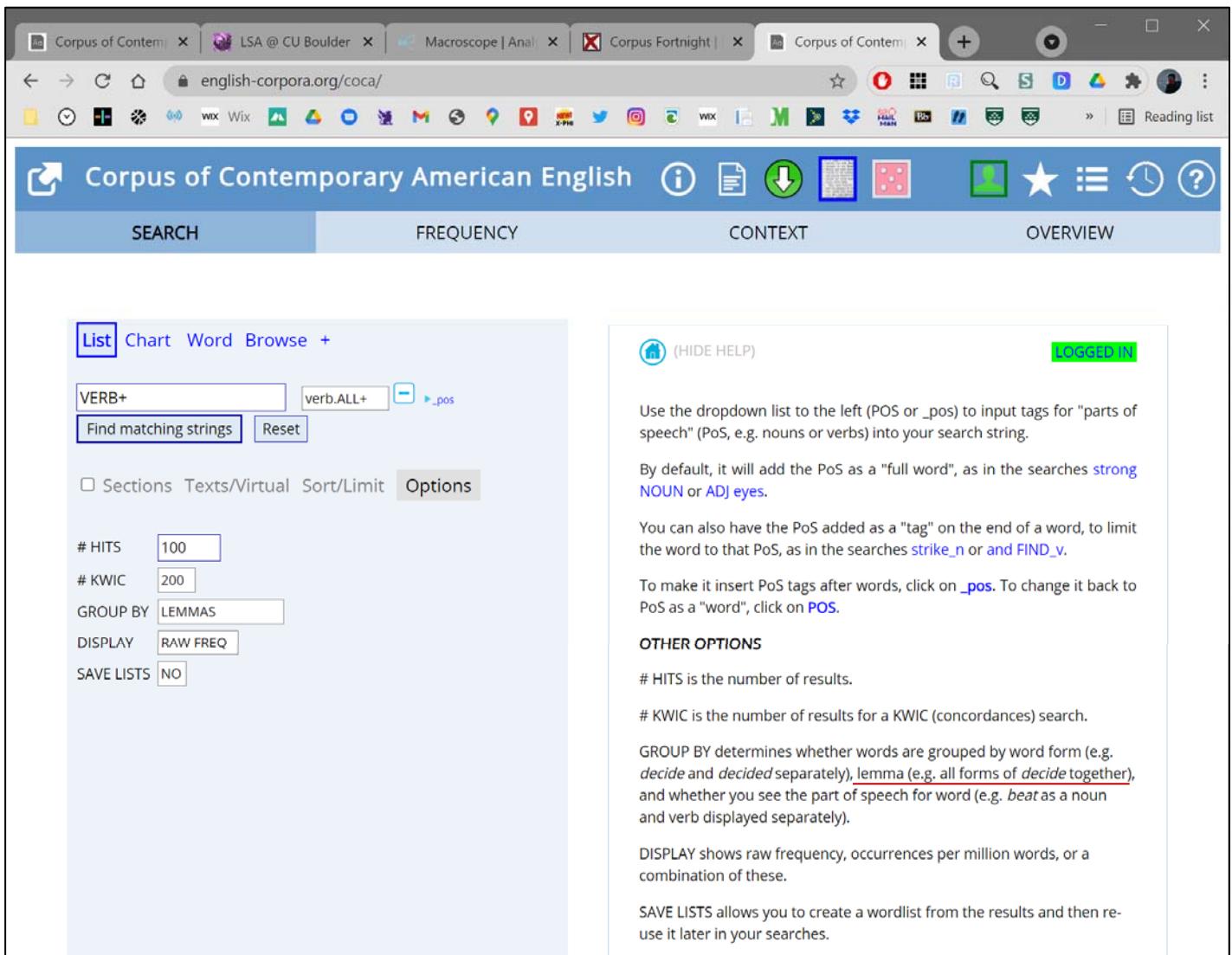


Note use of COCA by Nagel (2014) to justify that investigating “know” (*knowledge*) is of central importance.



In spoken English [in COCA] ‘know’... [figures] as the sixth... most commonly used [verb].... Spoken English is deeply invested in knowing.... To be fair, **some** of the conversational traffic in ‘know’ is coming from fixed phrases, like – *you know* – invitations to conversational partners to make some inference, or – *I know* – indications that you are accepting what conversational partners are saying. But even after we strip out those formulaic uses, the database’s randomly sampled conversations **remain thickly larded** with genuine references to knowing....

Corpus Methods & Experimental Philosophy



Corpus of Contemporary American English

SEARCH **FREQUENCY** **CONTEXT** **OVERVIEW**

List [Chart](#) [Word](#) [Browse](#) [+](#)

VERB+ [verb.ALL+](#) [▶_pos](#)

[Find matching strings](#) [Reset](#)

Sections [Texts/Virtual](#) [Sort/Limit](#) [Options](#)

HITS # KWIC GROUP BY [LEMMAS](#)

DISPLAY [RAW FREQ](#) SAVE LISTS [NO](#)

(HIDE HELP) **LOGGED IN**

Use the dropdown list to the left (POS or _pos) to input tags for "parts of speech" (PoS, e.g. nouns or verbs) into your search string.

By default, it will add the PoS as a "full word", as in the searches [strong](#) [NOUN](#) or [ADJ eyes](#).

You can also have the PoS added as a "tag" on the end of a word, to limit the word to that PoS, as in the searches [strike_n](#) or [and FIND_v](#).

To make it insert PoS tags after words, click on [_pos](#). To change it back to PoS as a "word", click on [POS](#).

OTHER OPTIONS

HITS is the number of results.

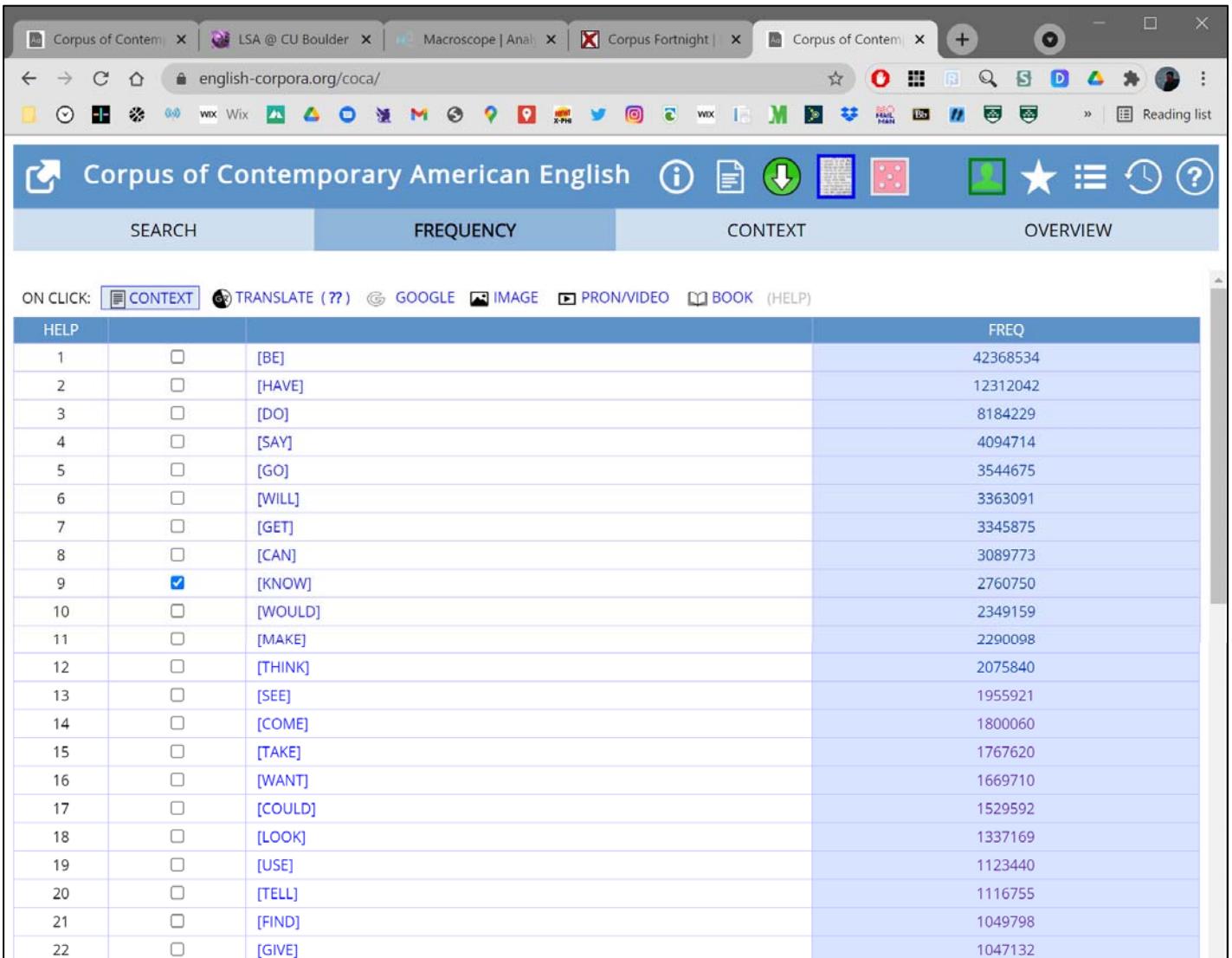
KWIC is the number of results for a KWIC (concordances) search.

GROUP BY determines whether words are grouped by word form (e.g. *decide* and *decided* separately), lemma (e.g. all forms of *decide* together), and whether you see the part of speech for word (e.g. *bear* as a noun and verb displayed separately).

DISPLAY shows raw frequency, occurrences per million words, or a combination of these.

SAVE LISTS allows you to create a wordlist from the results and then reuse it later in your searches.

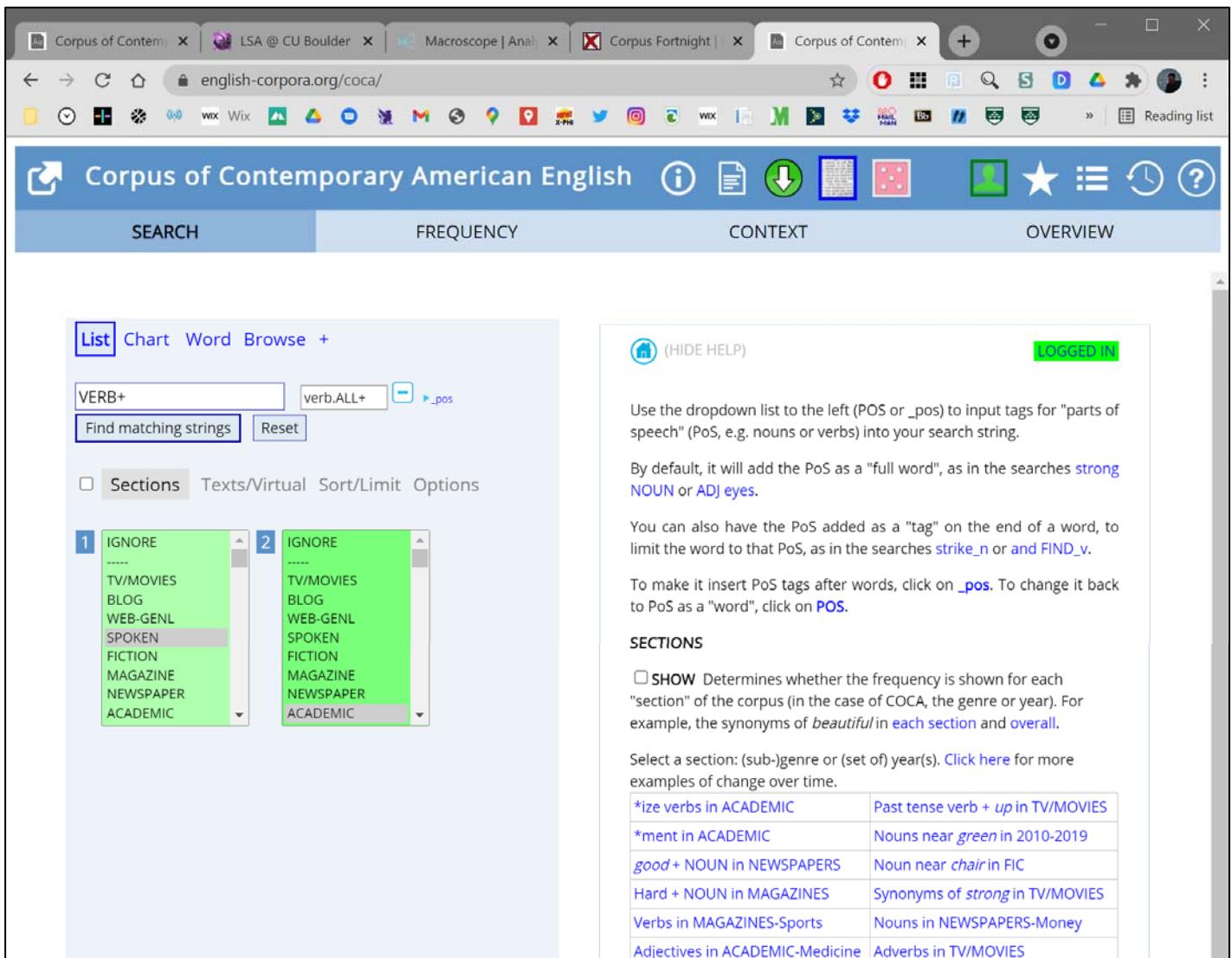
Corpus Methods & Experimental Philosophy



The screenshot shows a web browser window with multiple tabs open, including "Corpus of Contem...", "LSA @ CU Boulder", "Macroscope | Anal...", "Corpus Fortnight |", and "Corpus of Contem...". The main content is the COCA homepage. At the top, there are tabs for "SEARCH", "FREQUENCY" (which is selected), "CONTEXT", and "OVERVIEW". Below this is a toolbar with icons for "CONTEXT", "TRANSLATE (??)", "GOOGLE", "IMAGE", "PRON/VIDEO", "BOOK", and "(HELP)". A "ON CLICK:" dropdown menu is also present. The main area is a table titled "HELP" with columns for "HELP" (verb forms) and "FREQ" (frequency). The table lists 22 entries, starting with [BE] at 42368534 and ending with [GIVE] at 1047132. The verb "KNOW" is highlighted with a checked checkbox in the "HELP" column.

HELP	FREQ
[BE]	42368534
[HAVE]	12312042
[DO]	8184229
[SAY]	4094714
[GO]	3544675
[WILL]	3363091
[GET]	3345875
[CAN]	3089773
<input checked="" type="checkbox"/> [KNOW]	2760750
[WOULD]	2349159
[MAKE]	2290098
[THINK]	2075840
[SEE]	1955921
[COME]	1800060
[TAKE]	1767620
[WANT]	1669710
[COULD]	1529592
[LOOK]	1337169
[USE]	1123440
[TELL]	1116755
[FIND]	1049798
[GIVE]	1047132

Corpus Methods & Experimental Philosophy



Corpus of Contemporary American English

SEARCH FREQUENCY CONTEXT OVERVIEW

List Chart Word Browse +

VERB+ verb.ALL+ ▶_pos

Find matching strings Reset

Sections Texts/Virtual Sort/Limit Options

1 IGNORE ----
TV/MOVIES
BLOG
WEB-GENL
SPOKEN
FICTION
MAGAZINE
NEWSPAPER
ACADEMIC

2 IGNORE ----
TV/MOVIES
BLOG
WEB-GENL
SPOKEN
FICTION
MAGAZINE
NEWSPAPER
ACADEMIC

(HIDE HELP) LOGGED IN

Use the dropdown list to the left (POS or _pos) to input tags for "parts of speech" (PoS, e.g. nouns or verbs) into your search string.

By default, it will add the PoS as a "full word", as in the searches [strong](#) [NOUN](#) or [ADJ eyes](#).

You can also have the PoS added as a "tag" on the end of a word, to limit the word to that PoS, as in the searches [strike_n](#) or [and FIND_v](#).

To make it insert PoS tags after words, click on [_pos](#). To change it back to PoS as a "word", click on [POS](#).

SECTIONS

SHOW Determines whether the frequency is shown for each "section" of the corpus (in the case of COCA, the genre or year). For example, the synonyms of *beautiful* in [each section](#) and [overall](#).

Select a section: (sub-)genre or (set of) year(s). [Click here](#) for more examples of change over time.

*ize verbs in ACADEMIC	Past tense verb + <i>up</i> in TV/MOVIES
*ment in ACADEMIC	Nouns near <i>green</i> in 2010-2019
good + NOUN in NEWSPAPERS	Noun near <i>chair</i> in FIC
Hard + NOUN in MAGAZINES	Synonyms of <i>strong</i> in TV/MOVIES
Verbs in MAGAZINES-Sports	Nouns in NEWSPAPERS-Money
Adjectives in ACADEMIC-Medicine	Adverbs in TV/MOVIES

Corpus Methods & Experimental Philosophy

Corpus of Contem... LSA @ CU Boulder Macroscope | Anal... Corpus Fortnight | Corpus of Contem...

english-corpora.org/coca/

SEARCH FREQUENCY CONTEXT OVERVIEW

SEE CONTEXT: CLICK ON WORD (ALL SECTIONS) OR NUMBER (SPECIFIED SECTION) [HELP...]

SEC 1 (SPOKEN): 126,135,576 WORDS SEC 2 (ACADEMIC): 119,790,456 WORDS

	WORD/PHRASE	TOKENS 1	TOKENS 2	PM 1	PM 2	RATIO		WORD/PHRASE	TOKENS 2	TOKENS 1	PM 2	PM 1	RATIO
1	[BE]	7025832	3688728	55,700.6	30,793.2	1.8	1	[BE]	3688728	7025832	30,793.2	55,700.6	0.6
2	[HAVE]	1902743	937078	15,084.9	7,822.6	1.9	2	[HAVE]	937078	1902743	7,822.6	15,084.9	0.5
3	[DO]	1472220	320486	11,671.7	2,675.4	4.4	3	[DO]	320486	1472220	2,675.4	11,671.7	0.2
4	[GO]	777526	65456	6,164.2	546.4	11.3	4	[CAN]	269171	391778	2,247.0	3,106.0	0.7
5	[SAY]	668765	100967	5,302.0	842.9	6.3	5	[USE]	224192	95366	1,871.5	756.1	2.5
6	[KNOW]	647159	75084	5,130.7	626.8	8.2	6	[WOULD]	187694	345627	1,566.9	2,740.1	0.6
7	[GET]	561944	50321	4,455.1	420.1	10.6	7	[MAKE]	178483	288629	1,490.0	2,288.2	0.7
8	[THINK]	538853	52271	4,272.0	436.4	9.8	8	[WILL]	174037	452360	1,452.8	3,586.3	0.4
9	[WILL]	452360	174037	3,586.3	1,452.8	2.5	9	[MAY]	169976	78411	1,418.9	621.6	2.3
10	[CAN]	391778	269171	3,106.0	2,247.0	1.4	10	[SEE]	153677	298935	1,282.9	2,369.9	0.5
11	[WOULD]	345627	187694	2,740.1	1,566.9	1.7	11	[PROVIDE]	117142	18211	977.9	144.4	6.8
12	[SEE]	298935	153677	2,369.9	1,282.9	1.8	12	[FIND]	114395	110941	955.0	879.5	1.1
13	[COME]	295153	67221	2,340.0	561.2	4.2	13	[TAKE]	113904	250379	950.9	1,985.0	0.5
14	[WANT]	294463	37875	2,334.5	316.2	7.4	14	[COULD]	111981	180300	934.8	1,429.4	0.7
15	[MAKE]	288629	178483	2,288.2	1,490.0	1.5	15	[SHOULD]	103379	118191	863.0	937.0	0.9
16	[TAKE]	250379	113904	1,985.0	950.9	2.1	16	[SAY]	100967	668765	842.9	5,302.0	0.2
17	[LOOK]	198534	35907	1,574.0	299.7	5.3	17	[BECOME]	92328	57094	770.7	452.6	1.7
18	[MEAN]	184453	39404	1,462.3	328.9	4.4	18	[INCLUDE]	86276	14060	720.2	111.5	6.5
19	[COULD]	180300	111981	1,429.4	934.8	1.5	19	[GIVE]	86257	134394	720.1	1,065.5	0.7
20	[TALK]	180096	13399	1,427.8	111.9	12.8	20	[SHOW]	83825	68133	699.8	540.2	1.3
21	[TELL]	173377	29995	1,374.5	250.4	5.5	21	[KNOW]	75084	647159	626.8	5,130.7	0.1
22	[LET]	152081	12879	1,205.7	107.5	11.2	22	[NEED]	72026	106784	601.3	846.6	0.7
23	[ITDV1]	127604	27896	1,091.6	222.0	4.7	23	[MAINT]	70406	26044	599.5	306.5	2.0

Striking difference!

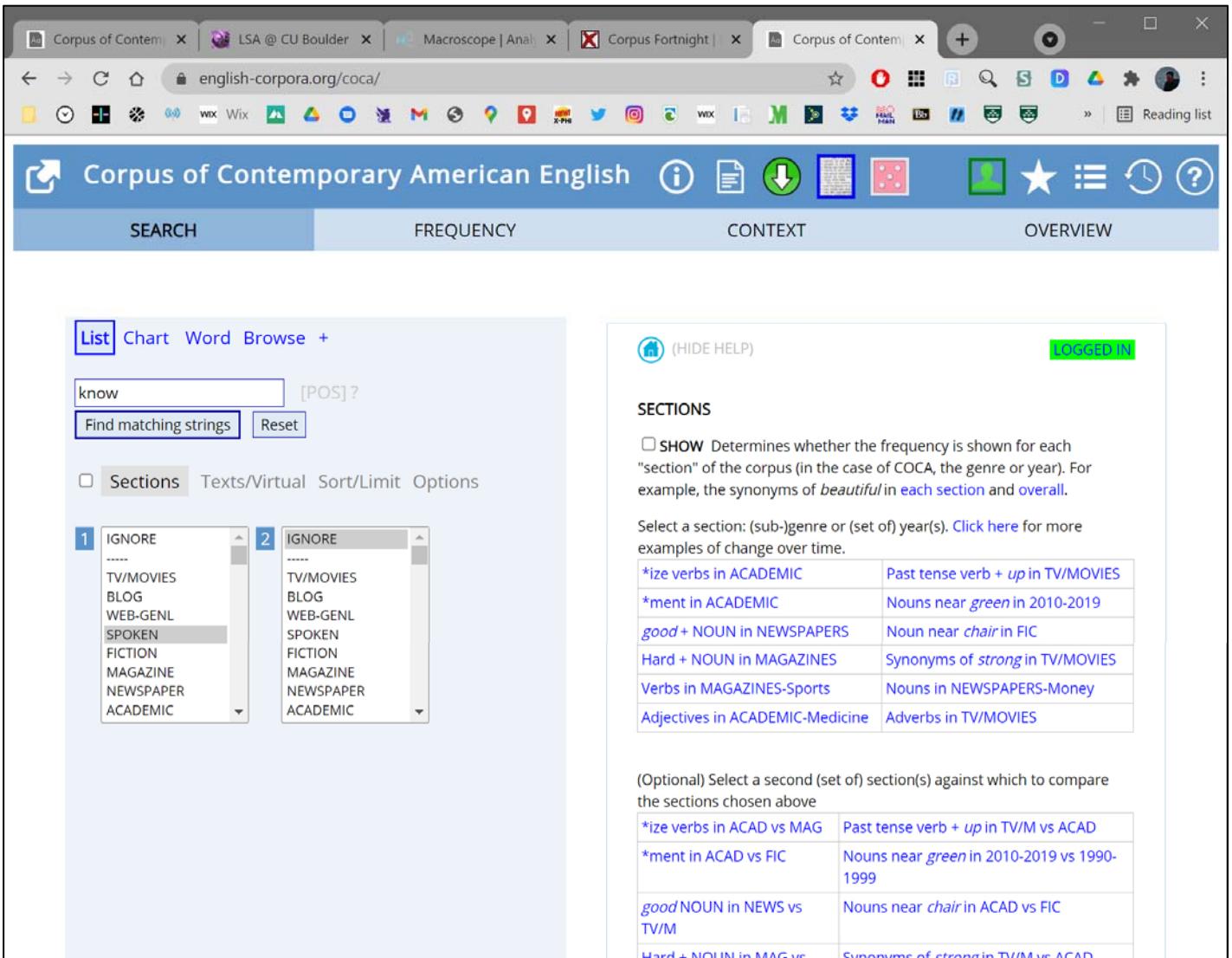
How much is due to Nagel's "fixed phrases" – i.e., "**discourse markers**, which, you know, probably don't occur as frequently in academic writing" (4)?

Random samples from COCA and code **concordance lines**:

Spoken: 254/500 = discourse markers

Academic: 8/500 = discourse markers
(all direct quotes)

Corpus Methods & Experimental Philosophy



The screenshot shows the COCA interface with the search term "know" entered. The results are displayed in two columns:

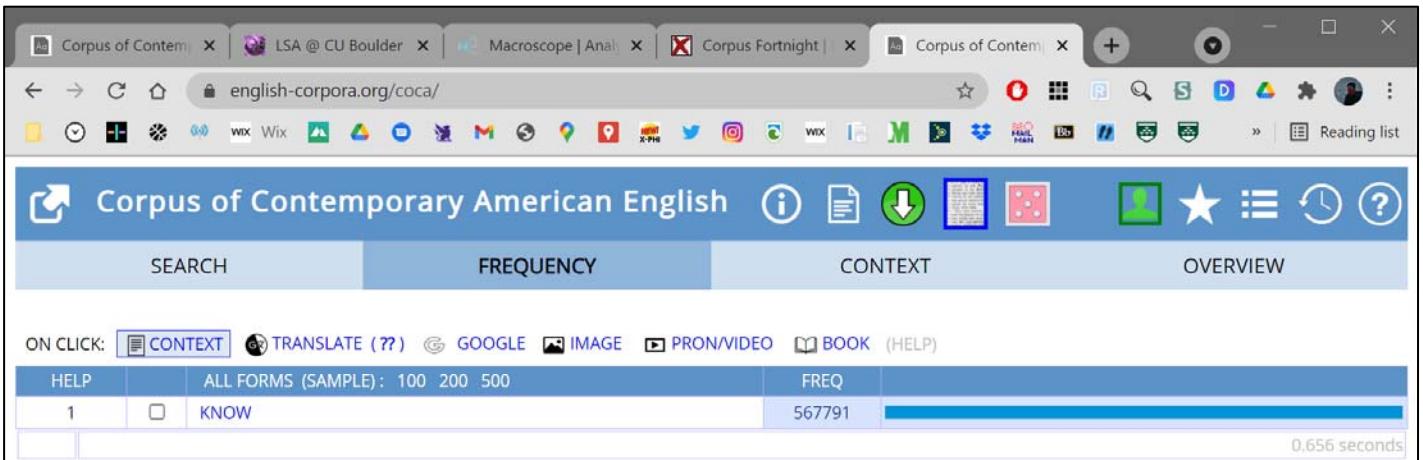
Section	Example
TV/MOVIES	Past tense verb + <i>up</i> in TV/MOVIES
BLOG	Nouns near <i>green</i> in 2010-2019
WEB-GENL	Noun near <i>chair</i> in FIC
SPOKEN	Synonyms of <i>strong</i> in TV/MOVIES
FICTION	Nouns in NEWSPAPERS-Money
MAGAZINE	Adverbs in TV/MOVIES
NEWSPAPER	
ACADEMIC	

Below the table, there is a note about optional section comparisons:

(Optional) Select a second (set of) section(s) against which to compare the sections chosen above

*ize verbs in ACAD vs MAG	Past tense verb + <i>up</i> in TV/M vs ACAD
*ment in ACAD vs FIC	Nouns near <i>green</i> in 2010-2019 vs 1990-1999
good+NOUN in NEWS vs TV/M	Nouns near <i>chair</i> in ACAD vs FIC
Hard+NOUN in MAG	Synonyms of <i>strong</i> in TV/M vs ACAD

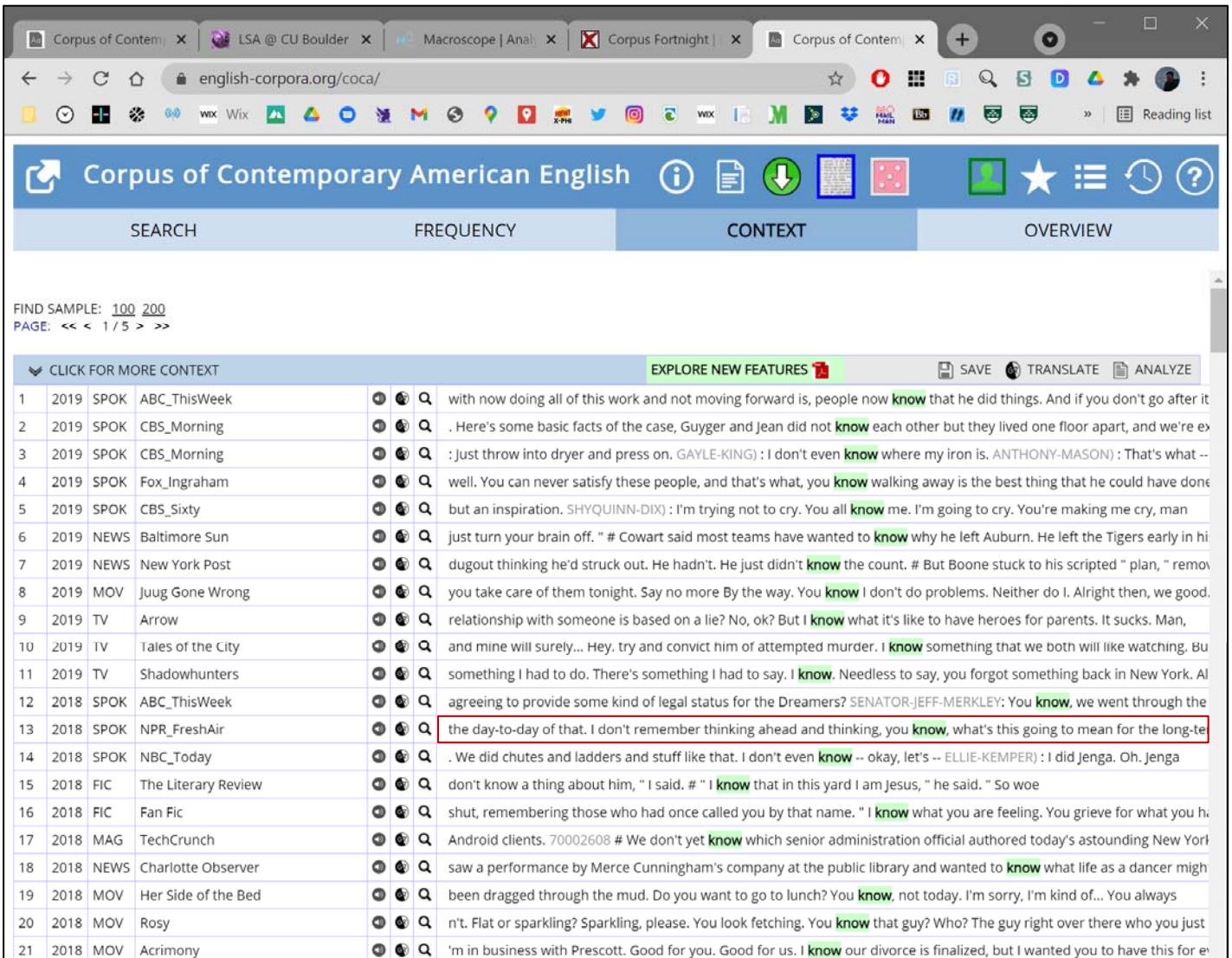
Corpus Methods & Experimental Philosophy



The screenshot shows a web browser window with multiple tabs open. The active tab is for the Corpus of Contemporary American English (COCA) at english-corpora.org/coca/. The interface has a blue header with the COCA logo and navigation icons. Below the header, there are four tabs: SEARCH, FREQUENCY, CONTEXT, and OVERVIEW. The FREQUENCY tab is selected. At the top of the main content area, there is a toolbar with various icons. Below that, a row of links includes "CONTEXT", "TRANSLATE (??)", "GOOGLE", "IMAGE", "PRON/VIDEO", "BOOK", and "(HELP)". Underneath these links is a table with three columns: "HELP", "ALL FORMS (SAMPLE): 100 200 500", and "FREQ". The first row of the table shows "1" in the HELP column, "KNOW" in the ALL FORMS column, and "567791" in the FREQ column. A progress bar indicates "0.656 seconds".

HELP	ALL FORMS (SAMPLE): 100 200 500	FREQ
1	KNOW	567791

Corpus Methods & Experimental Philosophy

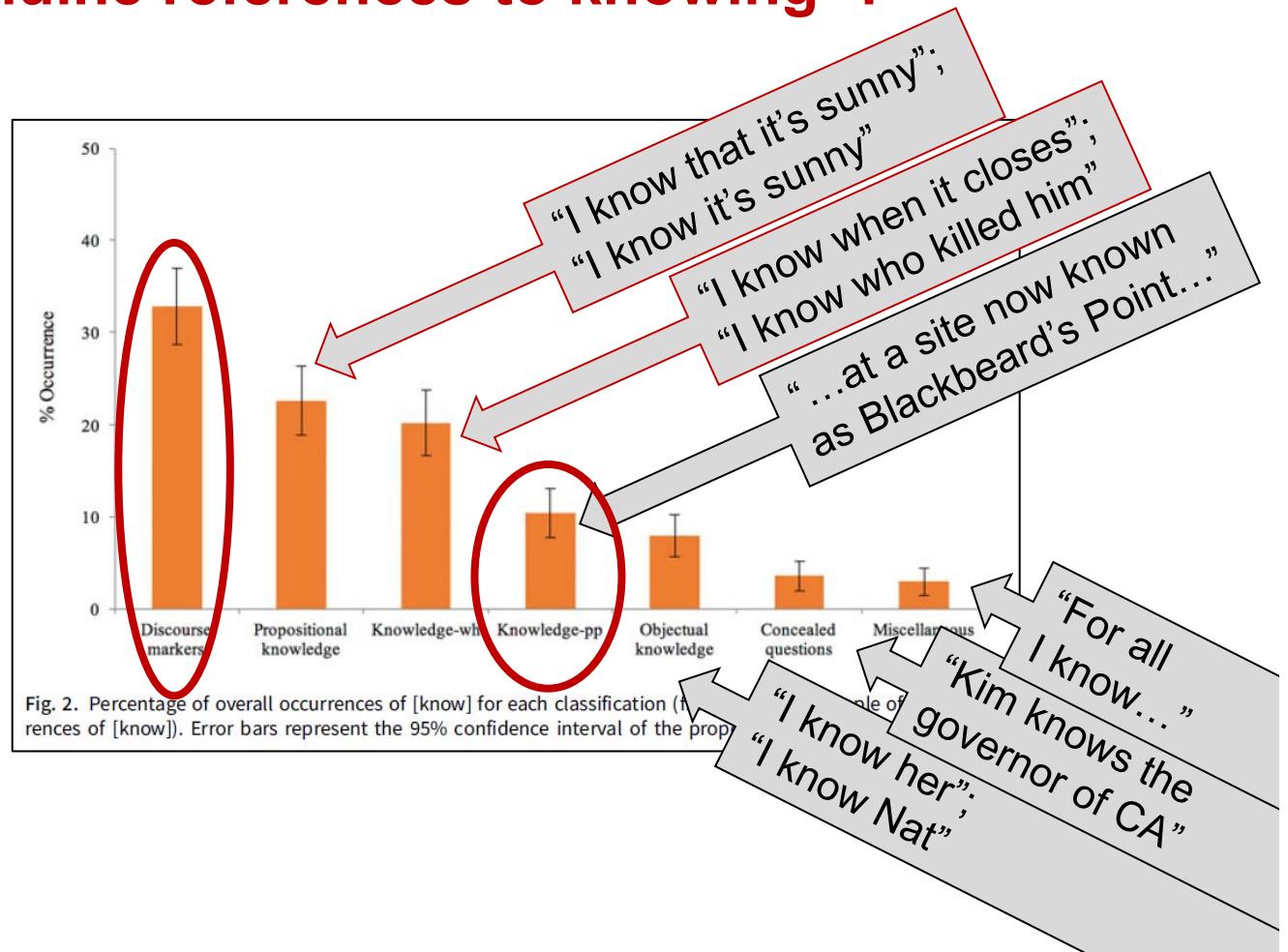


The screenshot shows a web browser window with multiple tabs open, including "Corpus of Contem...", "LSA @ CU Boulder", "Macroscope | Anal...", "Corpus Fortnight |", and "Corpus of Contem...". The main content is the COCA homepage. At the top, there are tabs for "SEARCH", "FREQUENCY", "CONTEXT", and "OVERVIEW". Below these tabs, there are buttons for "CLICK FOR MORE CONTEXT", "EXPLORE NEW FEATURES", "SAVE", "TRANSLATE", and "ANALYZE". The search results table lists 21 entries, each with a date, source, and snippet of text containing the word "know". The 13th entry, from 2018, is highlighted with a red border.

DATE	SOURCE	SNIPPET		
1	2019	SPOK	ABC_ThisWeek	with now doing all of this work and not moving forward is, people now know that he did things. And if you don't go after it
2	2019	SPOK	CBS_Morning	. Here's some basic facts of the case, Guyger and Jean did not know each other but they lived one floor apart, and we're ex
3	2019	SPOK	CBS_Morning	: Just throw into dryer and press on. GAYLE-KING : I don't even know where my iron is. ANTHONY-MASON) : That's what --
4	2019	SPOK	Fox_Ingraham	well. You can never satisfy these people, and that's what, you know walking away is the best thing that he could have done
5	2019	SPOK	CBS_Sixty	but an inspiration. SHYQUINN-DIX) : I'm trying not to cry. You all know me. I'm going to cry. You're making me cry, man
6	2019	NEWS	Baltimore Sun	just turn your brain off. " # Cowart said most teams have wanted to know why he left Auburn. He left the Tigers early in hi
7	2019	NEWS	New York Post	dugout thinking he'd struck out. He hadn't. He just didn't know the count. # But Boone stuck to his scripted " plan, " remov
8	2019	MOV	Juug Gone Wrong	you take care of them tonight. Say no more By the way. You know I don't do problems. Neither do I. Alright then, we good.
9	2019	TV	Arrow	relationship with someone is based on a lie? No, ok? But I know what it's like to have heroes for parents. It sucks. Man,
10	2019	TV	Tales of the City	and mine will surely... Hey. try and convict him of attempted murder. I know something that we both will like watching. Bu
11	2019	TV	Shadowhunters	something I had to do. There's something I had to say. I know . Needless to say, you forgot something back in New York. Al
12	2018	SPOK	ABC_ThisWeek	agreeing to provide some kind of legal status for the Dreamers? SENATOR-JEFF-MERKLEY: You know , we went through the
13	2018	SPOK	NPR_FreshAir	the day-to-day of that. I don't remember thinking ahead and thinking, you know , what's this going to mean for the long-te
14	2018	SPOK	NBC_Today	. We did chutes and ladders and stuff like that. I don't even know -- okay, let's -- ELLIE-KEMPER) : I did Jenga. Oh. Jenga
15	2018	FIC	The Literary Review	don't know a thing about him, " I said. # " I know that in this yard I am Jesus, " he said. " So woe
16	2018	FIC	Fan Fic	shut, remembering those who had once called you by that name. " I know what you are feeling. You grieve for what you ha
17	2018	MAG	TechCrunch	Android clients. 70002608 # We don't yet know which senior administration official authored today's astounding New York
18	2018	NEWS	Charlotte Observer	saw a performance by Merce Cunningham's company at the public library and wanted to know what life as a dancer migh
19	2018	MOV	Her Side of the Bed	been dragged through the mud. Do you want to go to lunch? You know , not today. I'm sorry, I'm kind of... You always
20	2018	MOV	Rosy	n't. Flat or sparkling? Sparkling, please. You look fetching. You know that guy? Who? The guy right over there who you just
21	2018	MOV	Acrimony	'm in business with Prescott. Good for you. Good for us. I know our divorce is finalized, but I wanted you to have this for ev

Corpus Methods & Experimental Philosophy

“genuine references to knowing”?



Corpus Methods & Experimental Philosophy



In addition to addressing general worries about the reliability and verifiability of armchair frequency judgments, looking at samples drawn from **linguistic corpora can reveal new facts** about the way linguistic expressions are used that have so far flown below philosophers' radar. For example, while philosophers have devoted a great deal of attention to knowledge-that, knowledge-wh, and objectual knowledge, so far they have **not discussed cases of knowledge + prepositional phrase complements**. A cursory look at occurrences of [know] in COCA reveals that such occurrences are roughly as frequent as occurrences of objectual knowledge. And most striking of all, **philosophers have so far not discussed discourse marking uses of [know]**, even though such uses are the most frequent occurrences of [know] in COCA. (22)

Another example looking at “know,” this time using a specialty corpus (CHILDES).

DOI: 10.1111/mila.12179

WILEY

ORIGINAL ARTICLE

Skepticism and the acquisition of “knowledge”

Shaun Nichols | N. Ángel Pinillos

Department of Philosophy, University of Arizona,
Tucson, Arizona

Correspondence

Shaun Nichols, Department of Philosophy,
University of Arizona, P.O. Box 210027 1145
E. South Campus Drive, Tucson, AZ 85721.
Email: shaunbnichols@gmail.com;
napinillos@gmail.com

Do you know you are not being massively deceived by an evil demon? That is a familiar skeptical challenge. Less familiar is this question: How do you have a conception of knowledge on which the evil demon constitutes a *prima facie* challenge? Recently several philosophers have suggested that our responses to skeptical scenarios can be explained in terms of heuristics and biases. We offer an alternative explanation, based in learning theory. We argue that, given the evidence available to the learner, it would be rational for the learner to infer an infallibilist conception of knowledge.

KEYWORDS

Corpus Methods & Experimental Philosophy

Rough idea: skepticism is naturally compelling because we tend to have an **infallibilist concept** of knowledge, such that if **S** knows **P**, then **S**'s epistemic state guarantees **P**.

Call on learning theory and corpus evidence to provide support for **Folk Infallibilism**:

“If we can determine the conception of knowledge that a child *should* learn, given the evidence, that will provide some reason to think that this is the conception he/she does learn.” (404)

The Size Principle

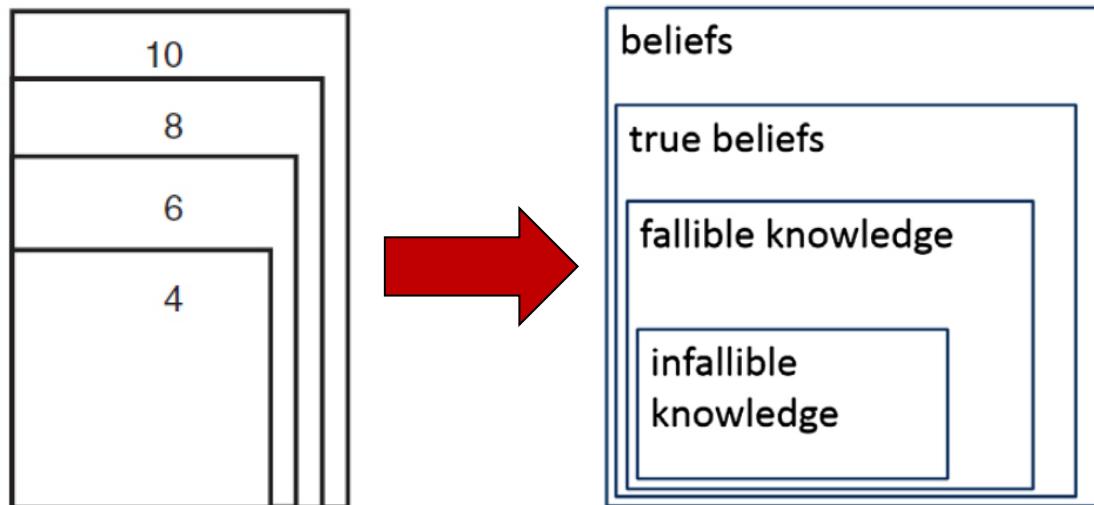
Xu and Tenenbaum (2007): Children only need a few positive examples to learn the word “dog.”



Compatible with *animal*, *living being*, etc., but don’t need negative instances. **Why?!**

Corpus Methods & Experimental Philosophy

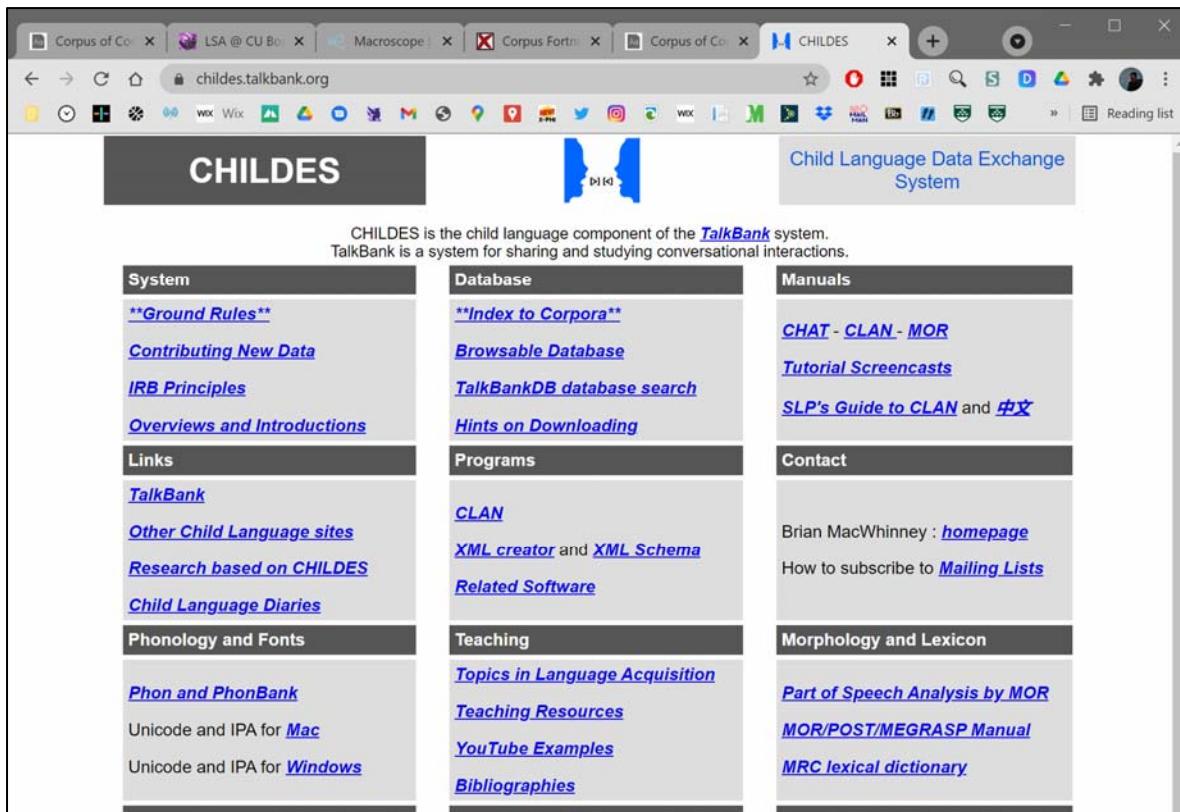
Smaller hypotheses that are consistent with the data are significantly preferred to larger hypotheses.



What evidence do children get?

Corpus Methods & Experimental Philosophy

Occurrences of “know” in **CHILDES**:
corpus of speech involving children
(directed at, made in the presence of).



The screenshot shows a web browser window with multiple tabs open. The active tab is for the CHILDES website at childestalkbank.org. The page title is "CHILDES" and it features a logo of two blue stylized figures. To the right, there is a sidebar with the text "Child Language Data Exchange System". The main content area is organized into several columns of links:

System	Database	Manuals
Ground Rules	**Index to Corpora**	CHAT - CLAN - MOR
Contributing New Data	Browsable Database	Tutorial Screencasts
IRB Principles	TalkBankDB database search	SLP's Guide to CLAN and 中文
Overviews and Introductions	Hints on Downloading	
Links	Programs	Contact
TalkBank	CLAN	Brian MacWhinney : homepage
Other Child Language sites	XML creator and XML Schema	How to subscribe to Mailing Lists
Research based on CHILDES	Related Software	
Child Language Diaries	Teaching	Morphology and Lexicon
Phonology and Fonts	Topics in Language Acquisition	Part of Speech Analysis by MOR
Phon and PhonBank	Teaching Resources	MOR/POST/MEGRASP Manual
Unicode and IPA for Mac	YouTube Examples	MRC lexical dictionary
Unicode and IPA for Windows	Bibliographies	

Corpus Methods & Experimental Philosophy

Three independent coders identified **802** instances of positive knowledge attributions:

Adult (Abe's father): I know you can't go away I'm not going away and you can't go away.

Adult (Nathan's mother): I know you want Leah to stay home honey

Adult (Lois, the investigator, talking to Peter): I don't see it but I know it was here yesterday.

Adult (Sarah's mother after Sarah says "once I hurt myself"): Yeah I know you did.

0 instances where the attribution of knowledge was accompanied by an expression of fallibilism about the proposition!

Comparison: **2,290** instances for "think," **91** accompanied by expression of fallibilism.

Adult (Lois, experimenter, speaking to Peter): I think that was right Peter. I'm not sure.

Adult (Naomi's mother): I don't know I think that might be a matzo crumb too.

One more example

While I've defined x-phi broadly, it is often treated as the empirical study of philosophical intuitions.

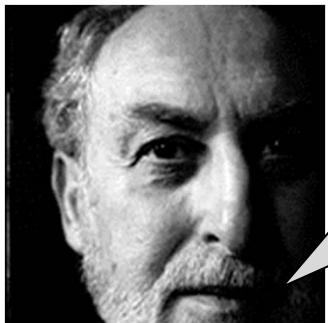
"Experimental philosophy is the name for a recent movement whose participants use the methods of experimental psychology to probe the way people make judgments that bear on debates in philosophy.... All of this work shares a commitment to using controlled and systematic experiments to explore people's intuitions and conceptual usage and to examine how the results of such experiments bear on traditional philosophical debates." (Nadelhoffer & Nahmias 2007, 123)

Corpus Methods & Experimental Philosophy

Makes sense given general assumptions that philosophy is steeped in and methodologically tied to the use of intuitions as evidence.



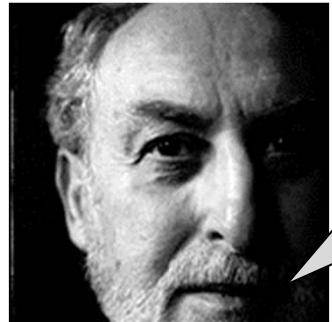
In contemporary analytic discussions... “intuition” has become the name for whatever it is that might provide philosophy with a distinctive method and hence preserve it as a separate (in principle) intellectual domain. (Gutting 1998, 7)



One thing that distinguishes philosophical methodology from the methodology of the sciences is its extensive and avowed reliance on intuition. (Goldman 2007, 1)

Corpus Methods & Experimental Philosophy

Interestingly, Goldman takes intuition talk to be a recent development.



As a historical matter, philosophers haven't always described their methodology in the language of intuitions. In fact, this seems to be a fairly recent bit of usage. (2)

New language or new method?

And is this historical claim even true?

If so, what explains the change?

Corpus Methods & Experimental Philosophy

Explored more systematically by James Andow using corpus methods.

METAPHILOSOPHY

© 2015 The Author Metaphilosophy published by Metaphilosophy LLC and John Wiley & Sons Ltd
METAPHILOSOPHY
Vol. 46, No. 2, April 2015
0026-1068

HOW “INTUITION” EXPLODED

JAMES ANDOW

Abstract: Recent decades have seen a surge in interest in metaphilosophy. In particular there has been an interest in philosophical methodology. Various questions have been asked about philosophical methods. Are our methods any good? Can we improve upon them? Prior to such evaluative and ameliorative concerns, however, is the matter of what methods philosophers actually use. Worryingly, our understanding of philosophical methodology is impoverished in various respects. This article considers one particular respect in which we seem to be missing an important part of the picture. While it is a received wisdom that the word “intuition” has exploded across analytic philosophy in recent decades, the article presents evidence that the explosion is apparent across a broad swathe of academia (and perhaps beyond). It notes various implications for current methodological debates about the role of intuitions in philosophy.

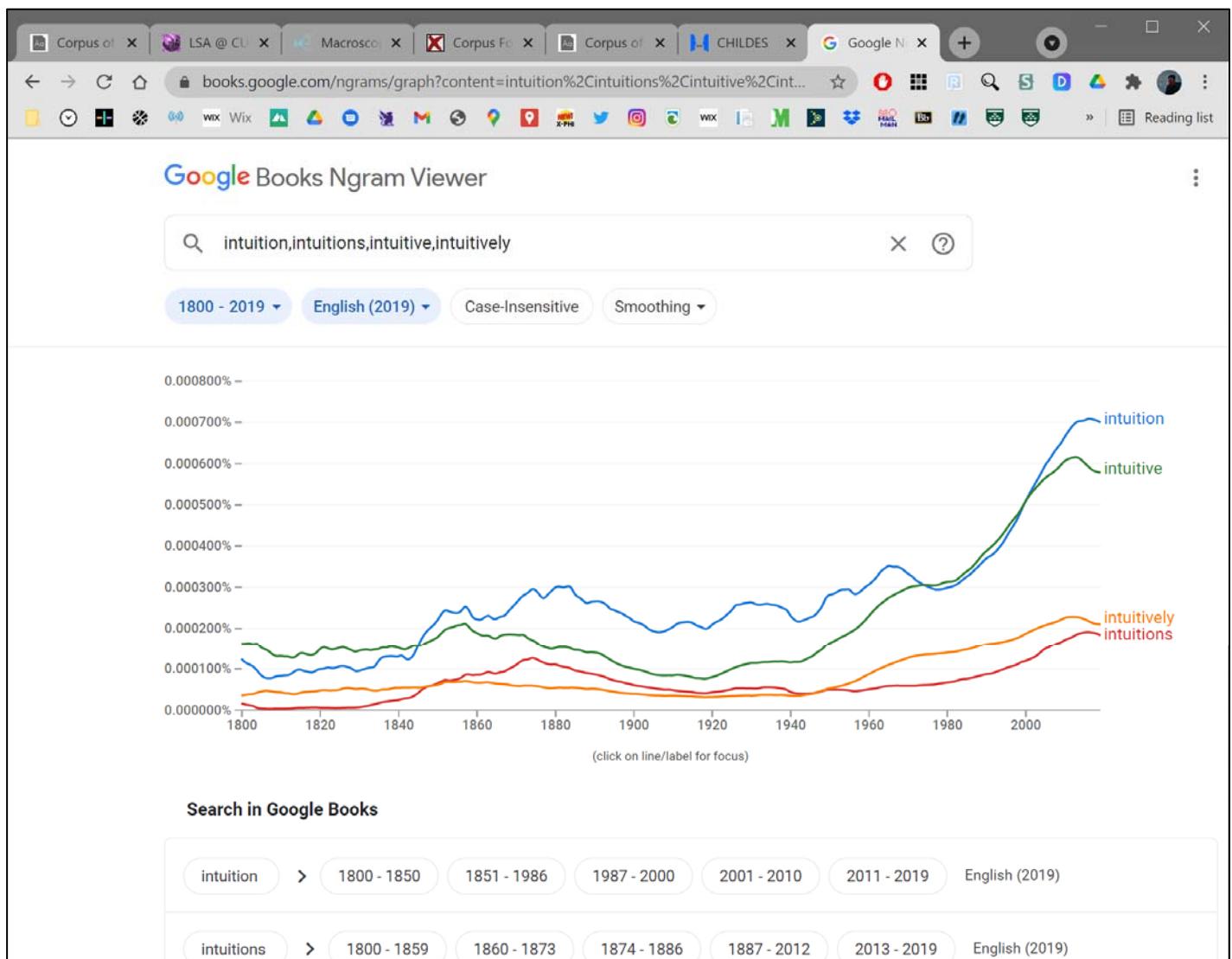
Corpus Methods & Experimental Philosophy

Tallant (2013): Use of “intuition” / “intuitive” / “intuitively” / “counter-intuitive” / “counter-intuitively” in five physics journals steadily increased from **7.44%** of articles in 2001 to **9.25%** of articles in 2011.

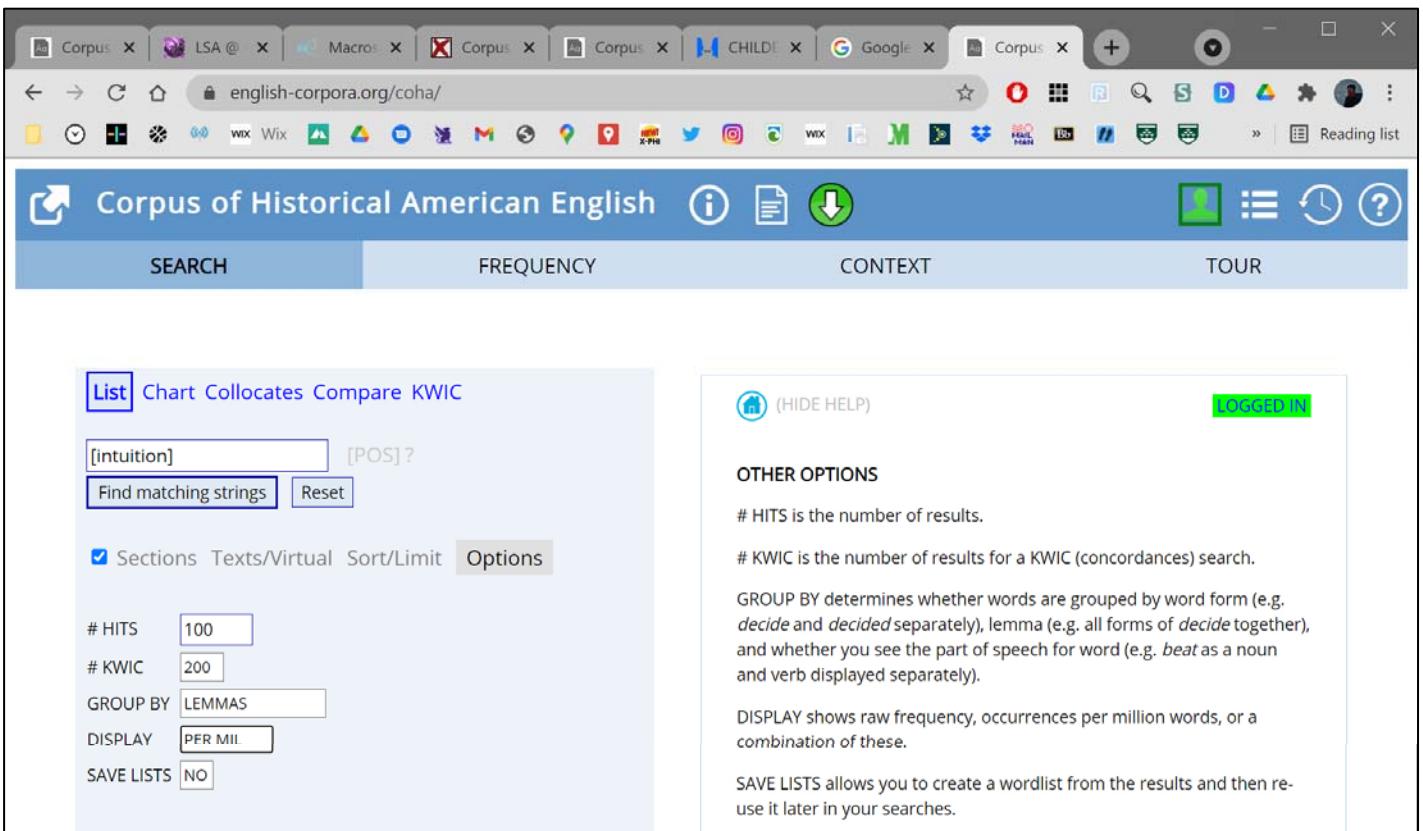
Andow: **1.95%** in the 1960s to **7.89%** in 2000s.

Google’s NGram corpus:
4x increase from 1800 to 2000.

Corpus Methods & Experimental Philosophy



Corpus Methods & Experimental Philosophy



The screenshot shows a web browser window with multiple tabs open, including "Corpus", "LSA @", "Macros", "Corpus", "Corpus", "CHILD!", "Google", and "Corpus". The main content is the COHA search interface at english-corpora.org/coha/. The interface has a blue header with the COHA logo, a search bar containing "[intuition]", and tabs for "SEARCH", "FREQUENCY", "CONTEXT", and "TOUR". On the left, there are buttons for "List", "Chart", "Collocates", and "Compare KWIC". Below the search bar are input fields for "# HITS" (100), "# KWIC" (200), "GROUP BY" (LEMMAS), "DISPLAY" (PER MIL.), and "SAVE LISTS" (NO). To the right, a sidebar titled "OTHER OPTIONS" explains terms like "# HITS", "# KWIC", "GROUP BY", "DISPLAY", and "SAVE LISTS". It also indicates the user is "LOGGED IN".

(HIDE HELP)

LOGGED IN

OTHER OPTIONS

HITS is the number of results.

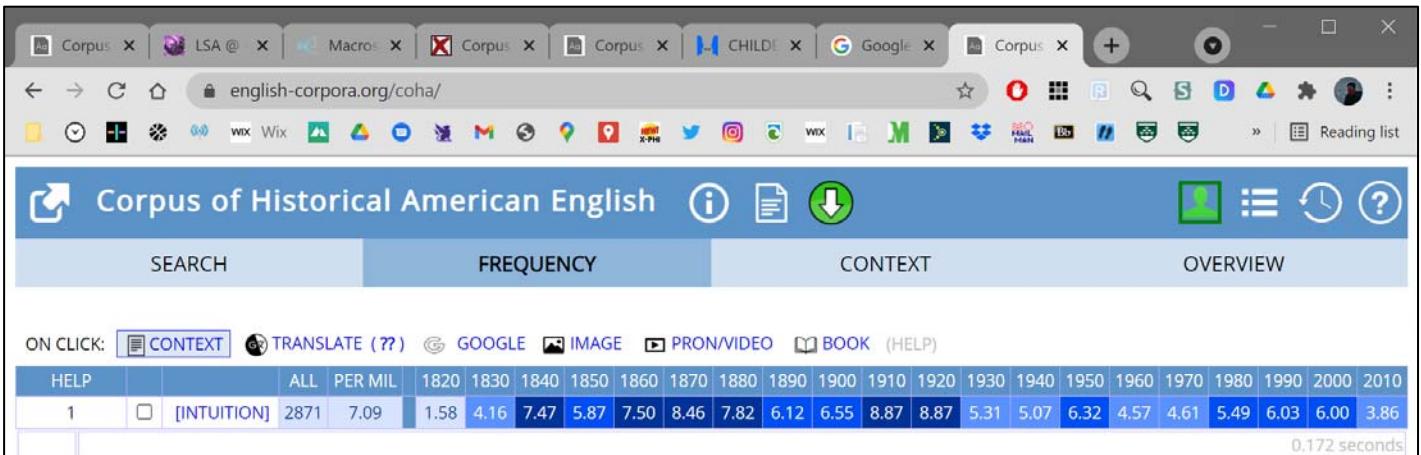
KWIC is the number of results for a KWIC (concordances) search.

GROUP BY determines whether words are grouped by word form (e.g., *decide* and *decided* separately), lemma (e.g., all forms of *decide* together), and whether you see the part of speech for word (e.g., *beat* as a noun and verb displayed separately).

DISPLAY shows raw frequency, occurrences per million words, or a combination of these.

SAVE LISTS allows you to create a wordlist from the results and then reuse it later in your searches.

Corpus Methods & Experimental Philosophy



A screenshot of a web browser displaying the Corpus of Historical American English (COHA) website at english-corpora.org/coha/. The interface includes a navigation bar with tabs for SEARCH, FREQUENCY, CONTEXT, and OVERVIEW. Below the navigation bar is a search bar with the query "[INTUITION]" and a dropdown menu showing results for "ALL", "PER MIL", and various years from 1820 to 2010. The main content area shows a table of frequency data for the word "compose". The table has columns for year (1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970, 1980, 1990, 2000, 2010) and frequency values (e.g., 2871, 7.09, 1.58, 4.16, 7.47, 5.87, 7.50, 8.46, 7.82, 6.12, 6.55, 8.87, 8.87, 5.31, 5.07, 6.32, 4.57, 4.61, 5.49, 6.03, 6.00, 3.86). The total time taken for the search is 0.172 seconds.

HELP	ALL	PER MIL	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010	
1	[INTUITION]	2871	7.09	1.58	4.16	7.47	5.87	7.50	8.46	7.82	6.12	6.55	8.87	8.87	5.31	5.07	6.32	4.57	4.61	5.49	6.03	6.00	3.86



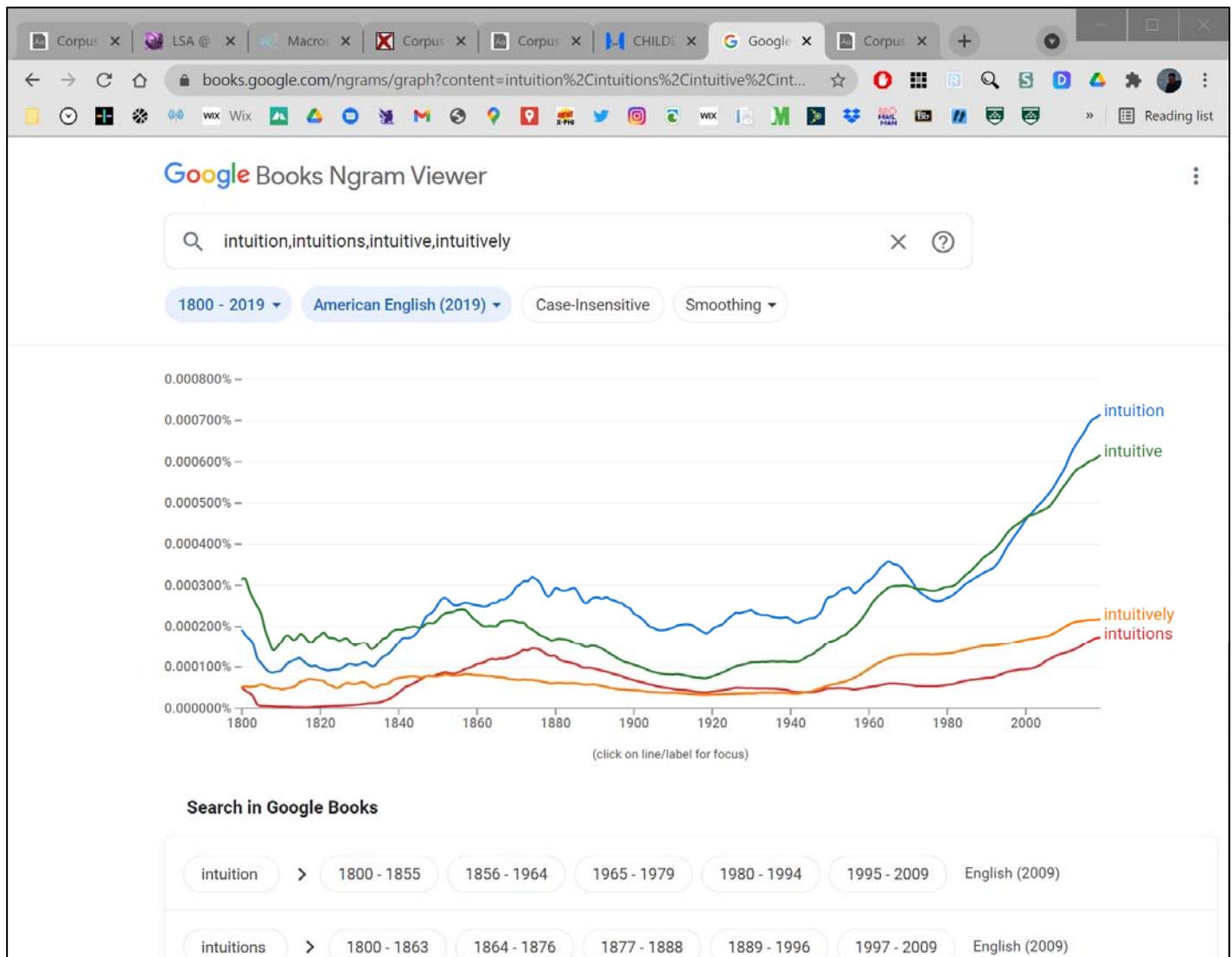
COHA contains more than [475 million words](#) of text from the 1820s-2010s (which makes it [50-100 times as large](#) as other comparable historical corpora of English) and the corpus is [balanced](#) by genre decade by decade. The creation of the corpus results from a grant from the [National Endowment for the Humanities](#) (NEH) from 2008-2010.

Corpus Methods & Experimental Philosophy

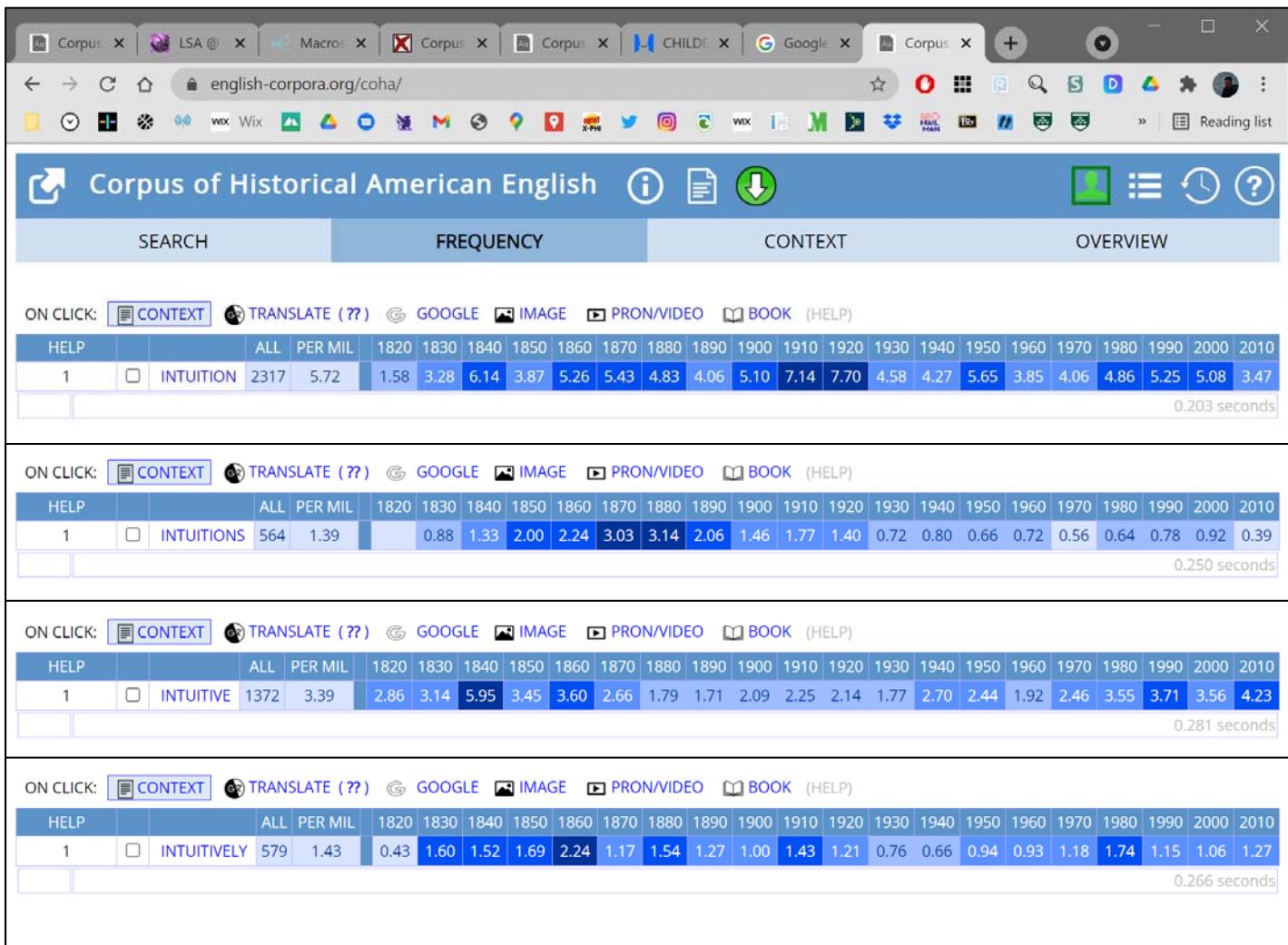
Below are descriptions of the corpora that can be searched with the Google Books Ngram Viewer. All corpora were generated in July 2009, July 2012, and February 2020; we will update these corpora as our book scanning continues, and the updated versions will have distinct persistent identifiers. Books with low OCR quality and serials were excluded.

Informal corpus name	Shorthand	Persistent identifier	Description
American English 2019	eng_us_2019	googlebooks-eng-us-20200217	Books predominantly in the English language that were published in the United States.
American English 2012	eng_us_2012	googlebooks-eng-us-all-20120701	
American English 2009	eng_us_2009	googlebooks-eng-us-all-20090715	
British English 2019	eng_gb_2019	googlebooks-eng-gb-20200217	Books predominantly in the English language that were published in Great Britain.
British English 2012	eng_gb_2012	googlebooks-eng-gb-all-20120701	
British English 2009	eng_gb_2009	googlebooks-eng-gb-all-20090715	
English 2019	eng_2019	googlebooks-eng-20200217	Books predominantly in the English language published in any country.
English 2012	eng_2012	googlebooks-eng-all-20120701	
English 2009	eng_2009	googlebooks-eng-all-20090715	
English Fiction 2019	eng_fiction_2019	googlebooks-eng-fiction-20200217	Books predominantly in the English language that a library or publisher identified as fiction.
English Fiction 2012	eng_fiction_2012	googlebooks-eng-fiction-all-20120701	
English Fiction 2009	eng_fiction_2009	googlebooks-eng-fiction-all-20090715	
			The "Google Million". All are in English with dates ranging from 1500 to 2008. No more than about 6000 books were

Corpus Methods & Experimental Philosophy



Corpus Methods & Experimental Philosophy



The screenshot shows a web browser window with multiple tabs open, including 'Corpus', 'LSA @', 'Macros', 'Corpus', 'Corpus', 'CHILD!', 'Google', and 'Corpus'. The main content is the COHA interface.

Corpus of Historical American English

ON CLICK: CONTEXT, TRANSLATE (??), GOOGLE, IMAGE, PRON/VIDEO, BOOK (HELP)

SEARCH FREQUENCY CONTEXT OVERVIEW

ON CLICK: CONTEXT, TRANSLATE (??), GOOGLE, IMAGE, PRON/VIDEO, BOOK (HELP)

HELP	INTUITION	ALL	PER MIL	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010
1	INTUITION	2317	5.72	1.58	3.28	6.14	3.87	5.26	5.43	4.83	4.06	5.10	7.14	7.70	4.58	4.27	5.65	3.85	4.06	4.86	5.25	5.08	3.47

0.203 seconds

ON CLICK: CONTEXT, TRANSLATE (??), GOOGLE, IMAGE, PRON/VIDEO, BOOK (HELP)

HELP	INTUITIONS	ALL	PER MIL	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010
1	INTUITIONS	564	1.39	0.88	1.33	2.00	2.24	3.03	3.14	2.06	1.46	1.77	1.40	0.72	0.80	0.66	0.72	0.56	0.64	0.78	0.92	0.39	

0.250 seconds

ON CLICK: CONTEXT, TRANSLATE (??), GOOGLE, IMAGE, PRON/VIDEO, BOOK (HELP)

HELP	INTUITIVE	ALL	PER MIL	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010
1	INTUITIVE	1372	3.39	2.86	3.14	5.95	3.45	3.60	2.66	1.79	1.71	2.09	2.25	2.14	1.77	2.70	2.44	1.92	2.46	3.55	3.71	3.56	4.23

0.281 seconds

ON CLICK: CONTEXT, TRANSLATE (??), GOOGLE, IMAGE, PRON/VIDEO, BOOK (HELP)

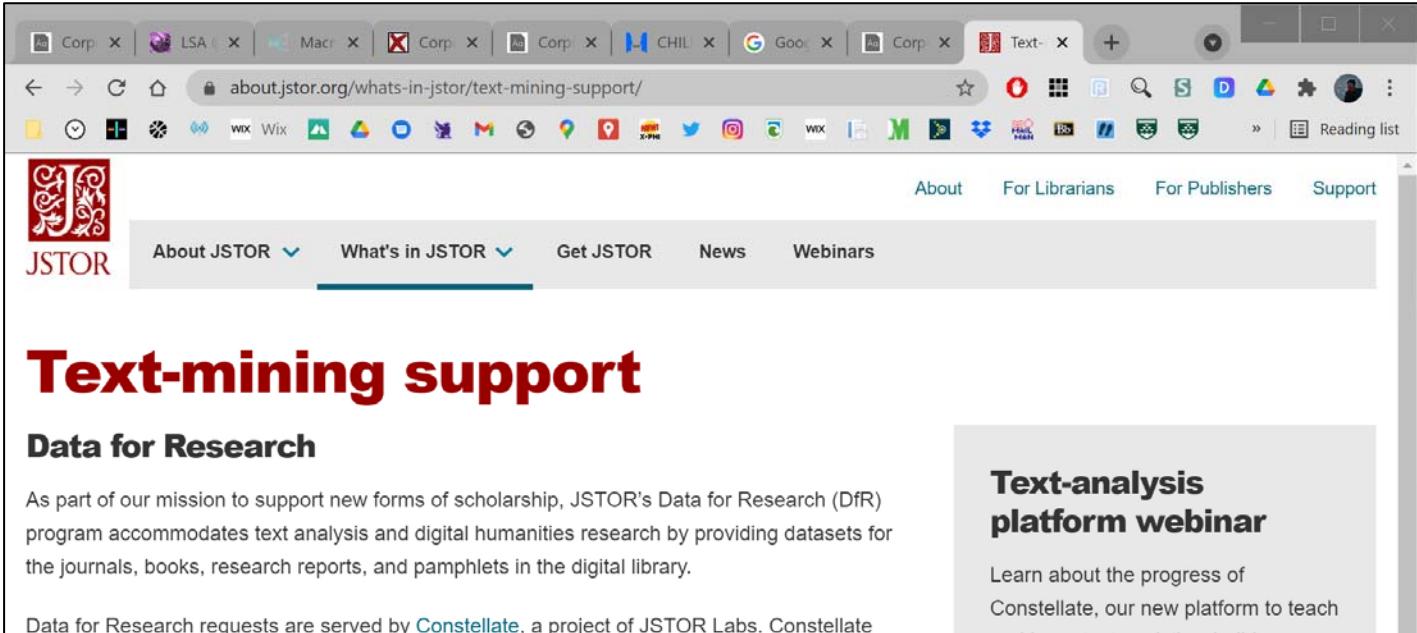
HELP	INTUITIVELY	ALL	PER MIL	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010
1	INTUITIVELY	579	1.43	0.43	1.60	1.52	1.69	2.24	1.17	1.54	1.27	1.00	1.43	1.21	0.76	0.66	0.94	0.93	1.18	1.74	1.15	1.06	1.27

0.266 seconds

Corpus Methods & Experimental Philosophy

Main Study

“intuition” / “intuitions” / “intuit” / “intuitive” /
“counter-intuitive” / “intuitively” / “counter-
intuitively” in journal research articles in
English in JSTOR



The screenshot shows a web browser window with multiple tabs open at the top. The active tab is for about.jstor.org/whats-in-jstor/text-mining-support/. The page content is from the JSTOR website, featuring the JSTOR logo and navigation links for "About", "For Librarians", "For Publishers", and "Support". The main heading is "Text-mining support" in red. Below it, there's a section titled "Data for Research" with a description of the DfR program and a note about Constellate. To the right, there's a box for "Text-analysis platform webinar" with a link to learn more about Constellate.

Text-mining support

Data for Research

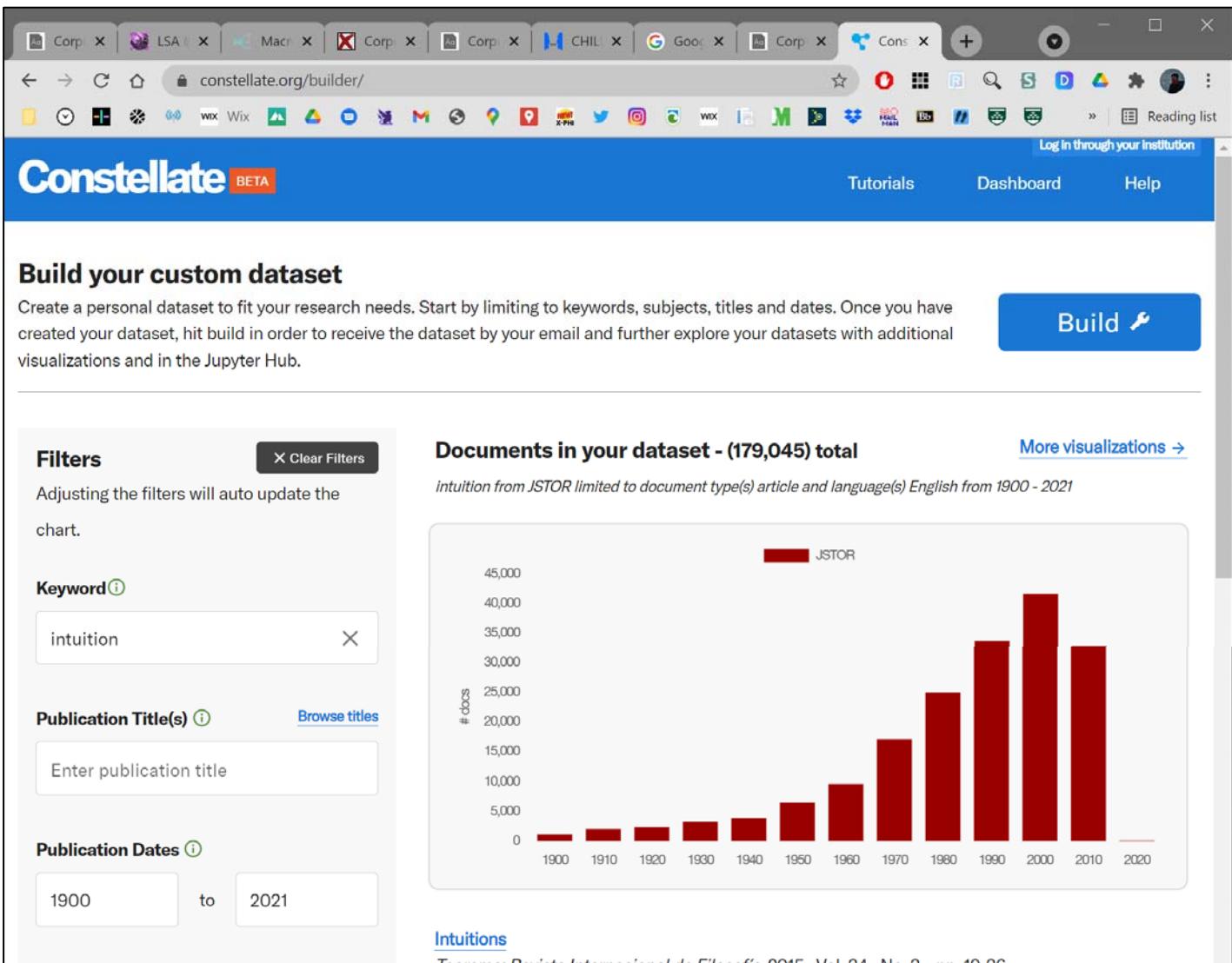
As part of our mission to support new forms of scholarship, JSTOR's Data for Research (DfR) program accommodates text analysis and digital humanities research by providing datasets for the journals, books, research reports, and pamphlets in the digital library.

Data for Research requests are served by [Constellate](#), a project of JSTOR Labs. Constellate

Text-analysis platform webinar

Learn about the progress of Constellate, our new platform to teach

Corpus Methods & Experimental Philosophy



The screenshot shows the Constellate.org builder interface. At the top, there's a navigation bar with tabs for Tutorials, Dashboard, and Help. Below that, a blue header bar says "Build your custom dataset". A sub-header below it reads: "Create a personal dataset to fit your research needs. Start by limiting to keywords, subjects, titles and dates. Once you have created your dataset, hit build in order to receive the dataset by your email and further explore your datasets with additional visualizations and in the Jupyter Hub." To the right of this text is a large blue "Build" button with a wrench icon.

On the left side, there's a "Filters" section with a "Clear Filters" button. It includes three input fields: "Keyword" containing "intuition", "Publication Title(s)" with a placeholder "Enter publication title", and "Publication Dates" set from 1900 to 2021. To the right of the filters is a chart titled "Documents in your dataset - (179,045) total". The chart is a bar chart showing the number of documents (# docs) from 1900 to 2020. The y-axis ranges from 0 to 45,000. The x-axis shows years from 1900 to 2020 in 10-year increments. The bars are red and labeled "JSTOR". The chart shows a steady increase in the number of documents over time, with a significant peak around 2000.

Build your custom dataset

Create a personal dataset to fit your research needs. Start by limiting to keywords, subjects, titles and dates. Once you have created your dataset, hit build in order to receive the dataset by your email and further explore your datasets with additional visualizations and in the Jupyter Hub.

Build 

Filters X Clear Filters

Adjusting the filters will auto update the chart.

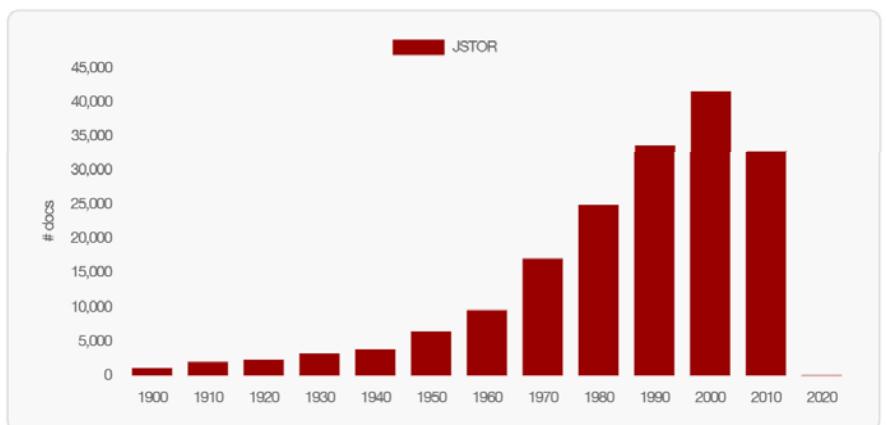
Keyword  X

Publication Title(s)  [Browse titles](#)

Publication Dates  to

Documents in your dataset - (179,045) total [More visualizations →](#)

intuition from JSTOR limited to document type(s) article and language(s) English from 1900 - 2021

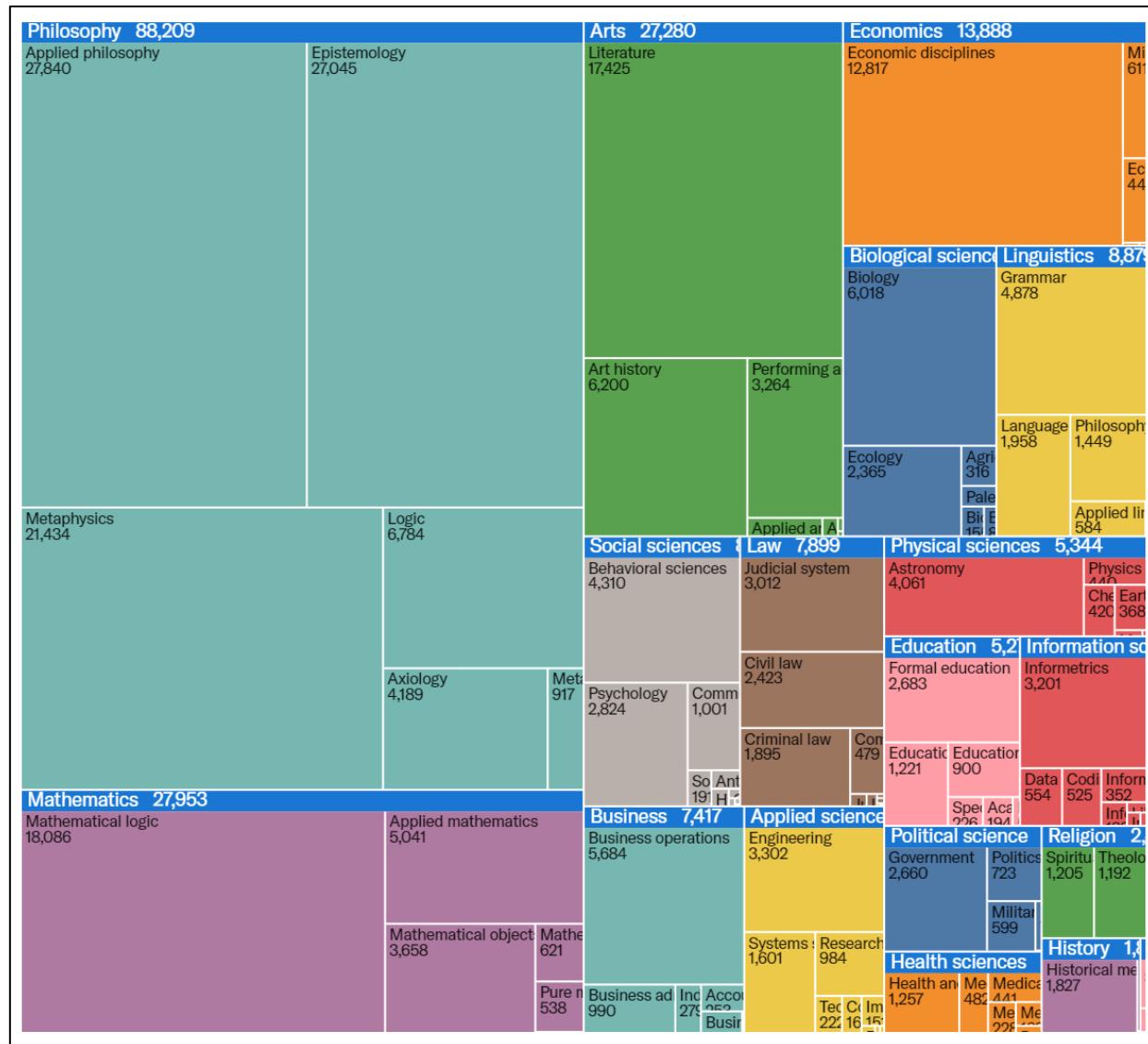


Year	# docs (JSTOR)
1900	~1,000
1910	~1,500
1920	~2,000
1930	~3,000
1940	~4,000
1950	~6,000
1960	~10,000
1970	~17,000
1980	~22,000
1990	~32,000
2000	~38,000
2010	~32,000
2020	~1,000

Intuitions

T... D... Intuitions / J. Philos. 2016, Vol. 94, No. 2, pp. 40–68

Corpus Methods & Experimental Philosophy



Corpus Methods & Experimental Philosophy



The headline finding is that... an increase in the proportion of articles using at least one of these terms was apparent across the broad swath of academia indexed (196)

TABLE 1. Indulgence in intuition talk by decade and discipline

Table Information: The first row gives the decade-by-decade proportions (in percentages) for all disciplines, including those not exhibited. The rest of the disciplines are ranked in decreasing order by the proportion of articles mentioning the terms for the 2000s. Dashes indicate that no articles are indexed for a given discipline and decade. The searches for the data in this data set were conducted on 4 September 2012.

Discipline	1900s	10s	20s	30s	40s	50s	60s	70s	80s	90s	2000s
All disciplines	2.6	3.6	4.2	5.2	6.2	8.3	10.5	13.1	14.6	16.3	18.3
Philosophy	21.7	29.6	34.5	33.5	4.2	32.7	34.9	44.1	47.5	50.5	53.6
Finance	7.2	5	7	6	12.4	10.1	16	18.7	25.8	36.7	47
Marketing & Advertising	—	—	—	11.1	11.4	14.6	17.3	25.6	31.8	36.8	39.9
Business	7.1	6.4	6.2	6.5	10.1	13.4	16.3	20.1	25.7	32	39.3
Linguistics	3.9	5.1	4.3	5.1	5.9	5.7	15.2	22.8	24.6	32.4	39.2
Economics	6.7	6	6.1	7.1	9.9	13.2	14	16.5	21.6	27	34.5
British Studies	8.1	6.4	8.9	12	13.3	22.9	19.5	24.8	27.8	30.1	33.5
Slavic Studies	—	—	18.8	17	19.2	20.7	19.4	24.5	23.7	27.6	33.2
Management & Org. Behav.	—	—	—	—	8	18.1	21.1	19.8	23	29.2	32.9
Hist. of Science & Technology	8.8	10.9	15.9	13	12.7	17.2	21.7	30.4	31.1	33.4	32.5

Corpus Methods & Experimental Philosophy

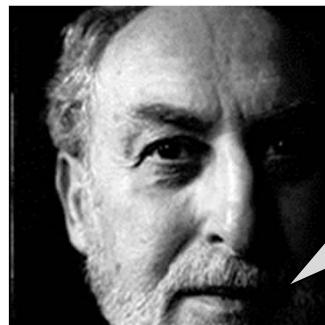
TABLE 2. Model 1

Model	B	Wald χ^2	p
Philosophy	2.233	9347.372	< 0.0005
Decade by philosophy	-.062	450.758	< 0.0005
Decade	.165	12817.243	< 0.0005
Constant	-2.324	40517.458	< 0.0005

TABLE 5. Model 2 (after 1950)

Model	Wald	χ^2	p
Philosophy	1.606	1613.772	< 0.0005
Decade by philosophy	.010	4.595	.032
Decade	.166	4798.034	< 0.0005
Constant	-2.327	13551.247	< 0.0005

Corpus Methods & Experimental Philosophy



in the early years of analytic philosophy, the terminology is not to be found



it doesn't seem to be true

Don't think, but look! (Pl 66)

