





As the United States moved ever further in the gesellschaft direction, gesellschaft-adapted cultural features [...] showed a quantitative increase, whereas gemeinschaft-adapted cultural features [...] showed a quantitative decrease.

Greenfield (2013)





These findings signify that books as cultural products reflect human ecology. They also signify that cultural features can be indexed by word-use frequencies, which, in turn, reflect what is prioritized by a population.

Greenfield (2013)



I'm not arguing that her theory is wrong, or that the Google ngram datasets don't contain supporting evidence. But it's going to take a much more careful and systematic analysis of the lexico-historical data to convince me.



Liberman (2013)

How can diachronic corpus data give us reliable insights about social change?

Overview

Problems

- Fillmore's screwdriver
- Flach's grammar lesson
- the rising sea levels of Koplenig and Müller-Spitzer

Overview

- Problems
 - Fillmore's screwdriver
 - Flach's grammar lesson
 - the rising sea levels of Koplenig and Müller-Spitzer
- Approaching causation, culture, and constructional change
- The English make-causative
- Fading authority and its consequences
- What a case from Dutch predicts for English
- What data from COCA and COHA suggests
- Conclusions



















Problems summarized

- Fillmore's screwdriver
- Cultural features cannot be indexed by word-use frequencies.
- Flach's grammar lesson

 Word-use frequencies have to be controlled for grammatical contexts.
- the rising sea levels of Koplenig and Müller-Spitzer
 Correlations in time-series data can be spurious if they are not
- controlled by correlating changes, rather than absolute values.

The English make-causative





You make a grown man cry!

You make the rockin' world go round.





It made my eyes be gettin' blurry.















• Hypothesis:

- Fading authority has left its mark on the way the English *make*-causative has been used over the past 150 years.
- Predictions:
 - P1: Uses of the *make*-causative with the meaning of compulsion should decrease.
 - P2: Uses of the *make*-causative with animate causers should decrease.
 - P3: Uses of the make-causative in which both causer and causee are animate should decrease.

Data

- Near-exhaustive retrieval of examples from the COHA (Davies 2010), 1860-2000
 - [make] [NP with up to three words] [V_INF]
- n = 38'597
- Annotation
 - verb (1681 types)
 - causer animacy (animate, inanimate)
 - causee animacy (animate, inanimate)













- · How does this automatic positioning in semantic space work?
- What work steps are involved?

I thought he was going to

he did nothing but

they always, to some extent, have to with you? What do you

| watch Julie act | always want to

sheds more tears for England. I shoulder for his wife to

cry

cry

cry

cry

cry

cry cry

- What work steps are involved?
- · We take a set of verbs and retrieve frequencies of context items from a corpus.



. There's an interesting point in there.

Tor / MRS JERVIS enters. all the time – right out of character he told me. She is very funny each time I see the video on. But one thing leads

stop words

wolf. They have to advise for? MRS JERVIS enters.



cry

cry

cry

shoulder for his wife to

for? MRS JERVIS enters. all the time – right out of character he told me. She is very funny each time I see the video on. But one thing leads

interesting point thought going cry wolf advise MRS JERVIS enters. always extent cry cry all time right character told very funny nothing cry watch Julie act always want cry sheds more tears England shoulder wife cry each time see video one thing leads cry stop words removed





- X X X X cry context laugh context smile context throw context context ite context iter requent frequency 58 52 46 29 21 21 20 20 20 20 19 17 17 17 17 16 16 16 16 14 context ite back tried made people make even smile help makes each trying wave tries reason more little began requency 26 17 13 12 12 12 11 11 10 10 9 7 requency 238 135 130 51 48 47 47 44 39 38 36 33 33 33 26 25 25 22 20 20 20 out began started want starts wanted laugh heard make made now makes begins people one never cry make people made cry one laugh out makes loud even started know both now joke began again out up away going ball back around want down people over one more money hands things know 69 57 49 24 22 21 20 19 15 15 13 13 13 13 12 12 12 12
- What work steps are involved?
- We take a set of verbs and and retrieve frequencies of context items from a corpus.



• We construct a table of frequencies with the verbs in the columns and the context items in the rows.





		cry	laugh	smile	throw	
	out					
X	began					
ALL CONTEXT ITEMS	started					
	want					
	starts					
	wanted					

		cry	laugh	smile	throw	
ALL CONTEXT ITEMS	out	69 🗲	20	2	238	
	began	57 🗲	▶ 16	7	0	
	started	49 🗲	• 17	0	0	
	want	24 🗲	6 🔶	5	39	
	starts	22 🗲	> 10	1	0	
	wanted	21 🗲	▶ 4	0	11	
summed up differences = semantic distance						

- What work steps are involved?
- We take a set of verbs and retrieve frequencies of context items from a corpus.



- We construct a table of frequencies with the verbs in the columns and the context items in the rows.
- We transform the raw frequencies with a collocation measure (PMI).

				- K	
		cry	laugh	smile	throw
1000	out	69	20	2	238
X	began	57	16	7	0
ALL	started	(49)	17	0 🔶	> 0
ITEMS	want	24	6	5	39
	starts	22	10	1	0
	wanted	21	4	0	11
	problem: comparisons of low-frequency words – very small differences				
	solution: transformation with Pointwise Mutual Information				



- What work steps are involved?
- We take a set of verbs and retrieve frequencies of context items from a corpus.



- We construct a table of frequencies with the verbs in the columns and the context items in the rows.
- We transform the raw frequencies with PMI.
- Visualization.





















• Hypothesis:

- Fading authority has left its mark on the way the English *make*-causative has been used over the past 150 years.
- Predictions:
 - P1: Uses of the *make*-causative with the meaning of compulsion should decrease.
 - P2: Uses of the *make*-causative with animate causers should decrease.
 - P3: Uses of the *make*-causative in which both causer and causee are animate should decrease.





How can diachronic corpus data give us reliable insights about social change?

- They can do so if:
 - we base our hypotheses on theories for which there is independent, non-linguistic evidence,
 - we focus on linguistic phenomena that are obligatory or that involve tightly constrained choices,
 - we account for differences in semasiological frequency and onomasiological frequency,
 - we apply statistical methods that safeguard against spurious significance.



