

Productivity and schematicity of the *way*- construction in Late Modern English

Florent Perek
University of Birmingham

Diachronic construction grammar

- Language change consists either of constructionalization or constructional change (Traugott & Trousdale 2013)
- Three aspects of constructional change:
 - Compositionality: how semantically transparent a construction is
 - Productivity: the range of lexical items that may occur in it
 - Schematicity: the level of detail (esp. semantic) with which the construction is stored; defines restrictions on use
- This talk focuses on how the last two can be characterized from corpus data

Productivity and schematicity

- Productivity can be observed in corpus data; what about schematicity?
- The two are commonly thought to be interrelated
 - More schematic constructions have more schematic slots: fewer constraints on the lexical items that can be used
 - Conversely, the occurrence of more diverse items makes a slot more schematic

Productivity and schematicity

- By no means a one-to-one relation
 - Compatibility of an item with a schema does not mean that it will necessarily be attested
 - Conversely, new coinages can happen outside of a schema, e.g., by analogical extension
- Schematicity vs. productivity ~ licensing vs. coining
- How to characterize schematicity when only attested types are observable?

Case study: the *way*-construction

- Verb *one's way* PP, e.g., *He pushed his way through the crowd*
- Describes motion of the subject referent
- Three senses of the construction:
 - Path-creation: the verb describes what enables motion
They hacked their way through the jungle.
 - Manner: the verb describes the manner of motion
They trudged their way through the snow
 - Incidental-action: the verb refers to some co-occurring action unrelated to motion
He whistled his way across the room

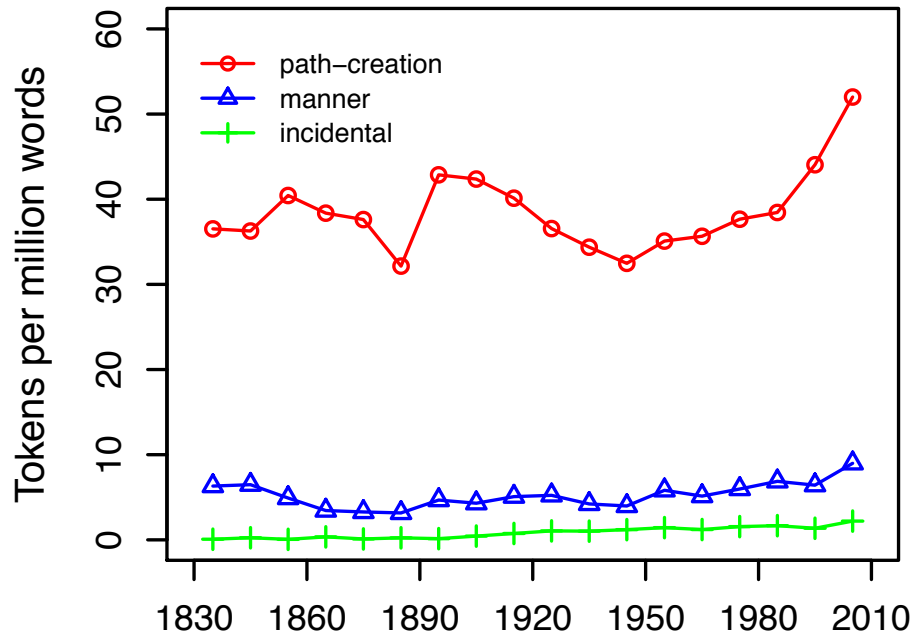
The *way*-construction in diachrony

- Previous research mostly focused on the origins of the construction (Israel 1996, Traugott & Trousdale 2003)
- Little discussion of the recent history of the construction (19th–20th)
 - Grammatically stable since the 19th century
 - Good case for the study of syntactic productivity
- Exception: Mondorf (2011)
 - But her focus is on the comparison with the *self*-resultative construction (e.g., *He worked himself to exhaustion*)
 - Only ten verbs, few datapoints

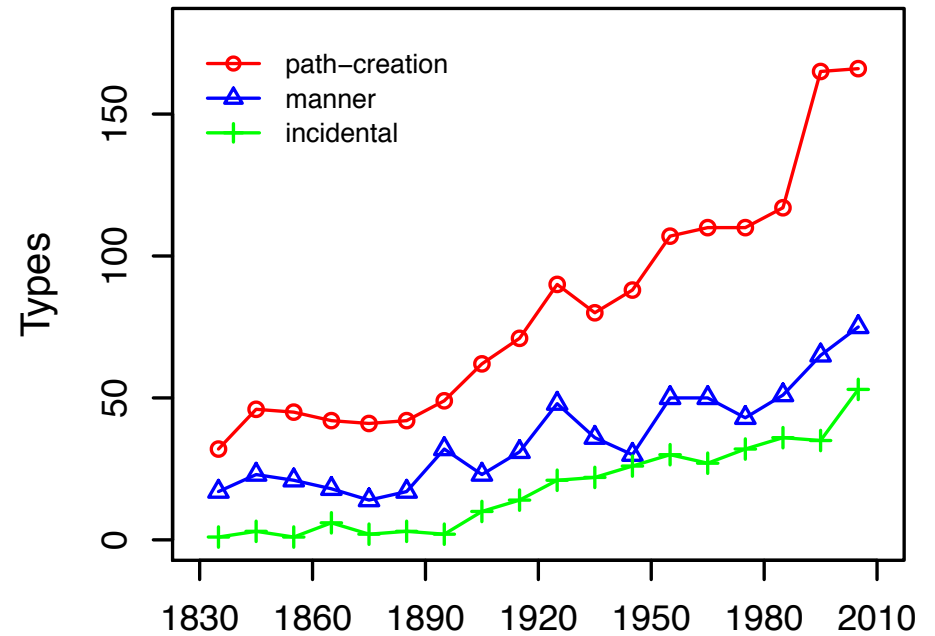
Data

- All tokens of “V Poss *way* Prep” between 1830 and 2009 extracted from the Corpus of Historical American English (COHA, Davies 2010)
- Manually filtered, annotated for constructional meaning: path-creation, manner, incidental-action
- Quantitative measures of productivity
 - Token frequency: how often the construction is used?
 - Type frequency: with how many different verbs?

Token frequency (per million words)



Verb type frequency



- Token frequency is relatively stable
- Steady increase in type frequency: the construction is used with more and more different verbs

Type frequency

- Type frequency reflects the lexical range of a construction
- But it is a purely quantitative measure
 - Only indirectly related to semantic diversity
 - No account of how *different* items are
- Questions:
 - What kinds of verbs joined the distribution?
 - Did it become more semantically diverse?
 - Are there particular semantic domains favored by the construction?

Distributional semantics

- Most studies in DiaCxG draw on semantic intuitions
- This paper takes a different approach: using distributional semantics to measure semantic similarity (Perek 2014, 2016)
- Words that occur in similar contexts tend to have related meanings (Miller & Charles 1991)
 - “You shall know a word by the company it keeps.” (Firth 1957: 11)
- Therefore, a way to characterize the meaning of words is through their distribution in large corpora
- Widely used in computational linguistics
- Benefits:
 - Fully automatic
 - Data-driven and objective

Distributional semantic model

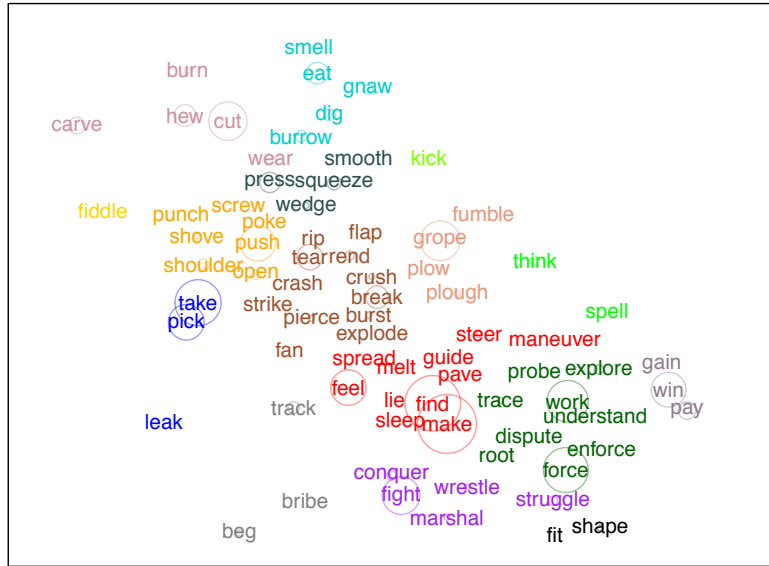
- The more frequent collocates are shared by two words, the more similar they will be considered
- “Bag of words” approach
 - Extraction of lexical collocates of each verb in a 5-word window from a large corpus
 - Each verb is assigned an array of values (= a vector) derived from co-occurrence frequencies
 - High correlation between vectors is an indication of semantic relatedness

Distributional semantic plots

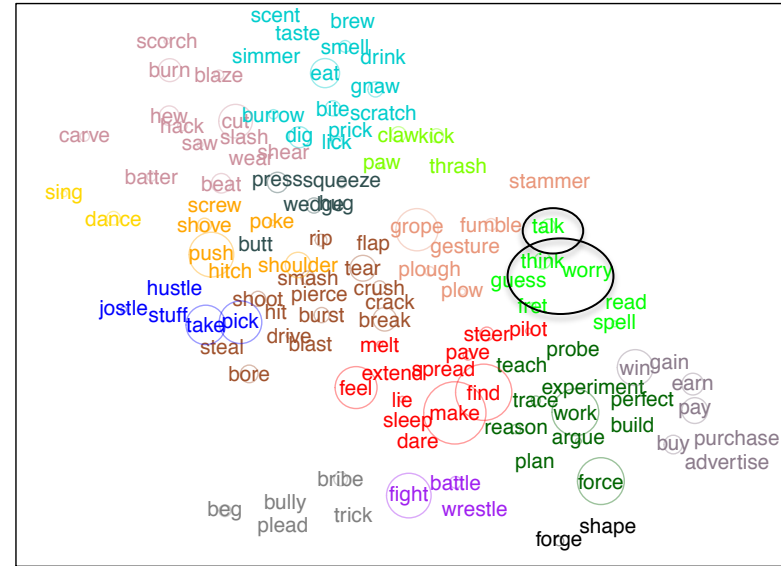
- Output: pairwise distances between verbs
- Define a semantic space that can be plotted for visualization
 - By means of *t*-Distributed Stochastic Neighbor Embedding algorithm (*t*-SNE) (Van der Maaten & Hinton 2008)
 - Places objects in a 2-dimensional space such that the between-object distances are preserved as well as possible
 - Superior to multidimensional scaling (MDS) for dense spaces with many dimensions
 - Distance matrix converted to coordinates for each verb
- Semantic domain of the construction plotted for four time periods: 1830-1879, 1880-1929, 1930-1969, 1970-2000.
- Three senses of the construction examined separately

The path-creation sense

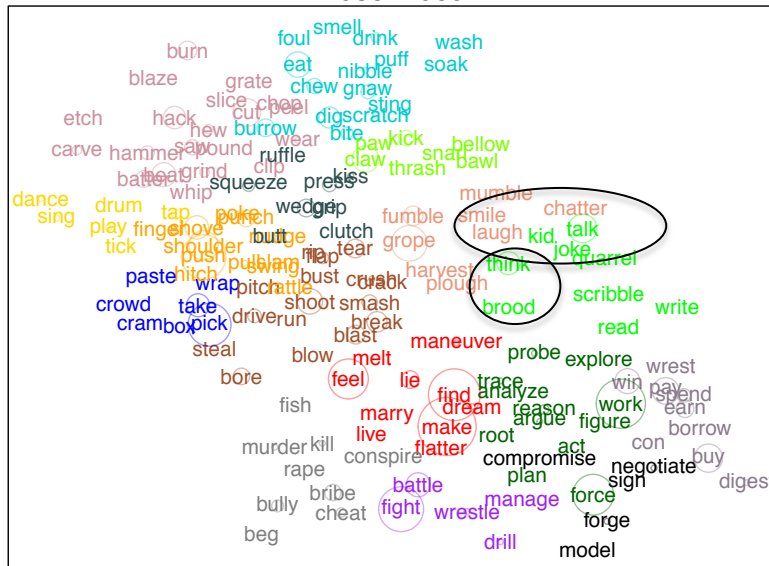
1830-1879



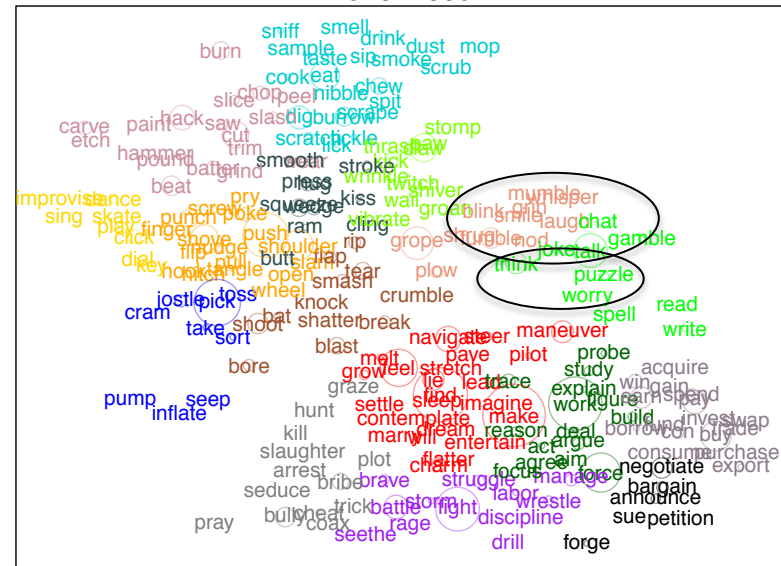
1880-1929



1930-1969



1970-2009



From period 3 onwards: social interaction (*chat, chatter, joke, kid, nod, quarrel, talk*), emotion (*grin, laugh, smile, shrug, laugh*), cognition (*brood, fret, puzzle, think, worry*)

The path-creation sense

- Initially centered on forceful actions
 - Pushing, hitting, cutting, etc.
 - Other regions are more sparsely populated
 - In line with the diachronic origin of the construction: literal creation of a physical path
- Over time, the distribution becomes more even
- More abstract verbs: interaction, commerce, cognition, etc.

The path-creation sense

- The new verbs tend to refer to increasingly unusual ways to cause motion
- They are especially prone to cause abstract, metaphorical motion, e.g.:

[T]hey talk about Uncle Paul having **bought his way into the Senate!**

By the time he was four he could **spell his way through his book** with only occasional pauses for breath.

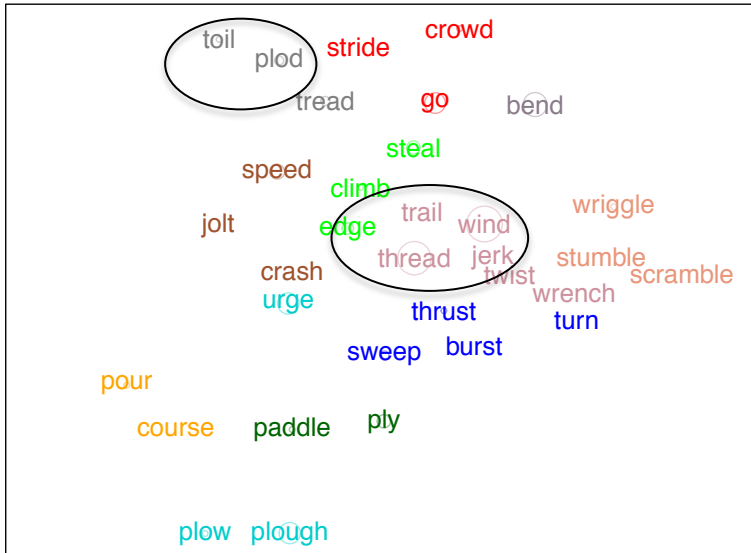
I sit and watch [...], **grazing my way through a muffuletta.**

I saw Wallace Shawn [...] **lisping his way through a mournful monologue.**

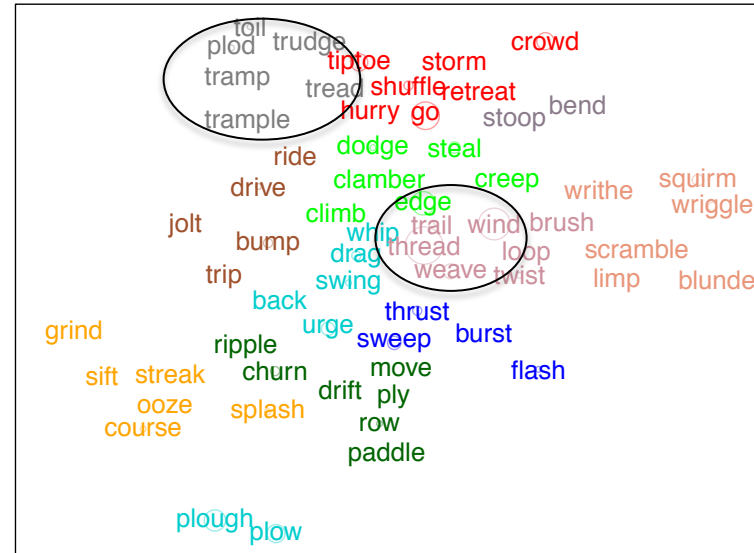
- Can be seen as reflecting an increase in the schematicity of the motion component of the construction

The manner sense

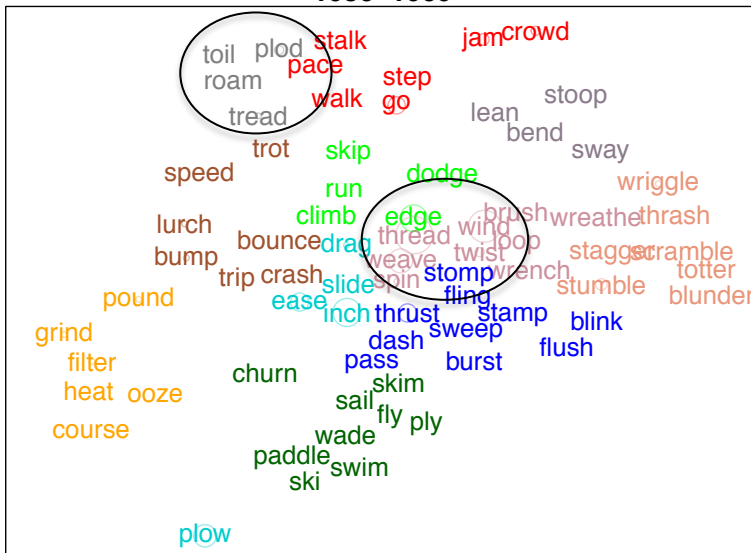
1830-1879



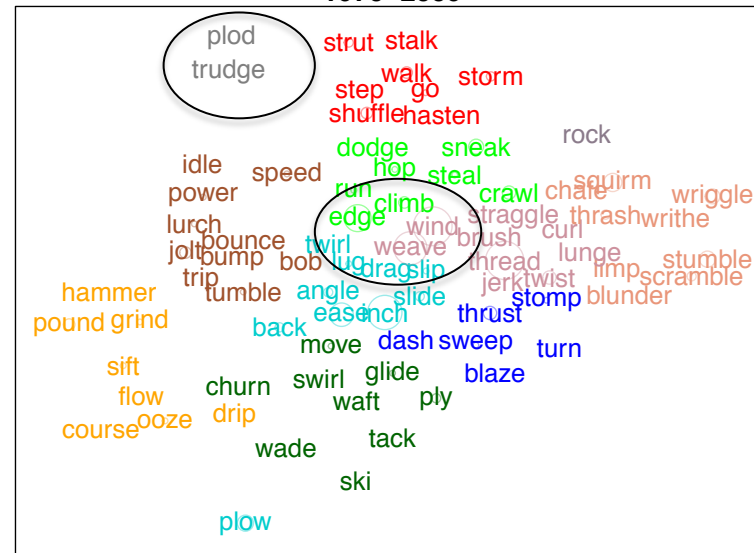
1880-1929



1930-1969

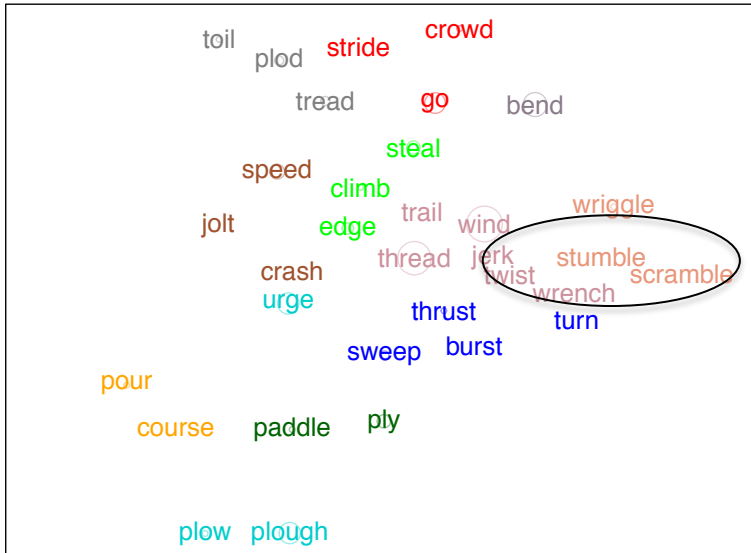


1970-2009

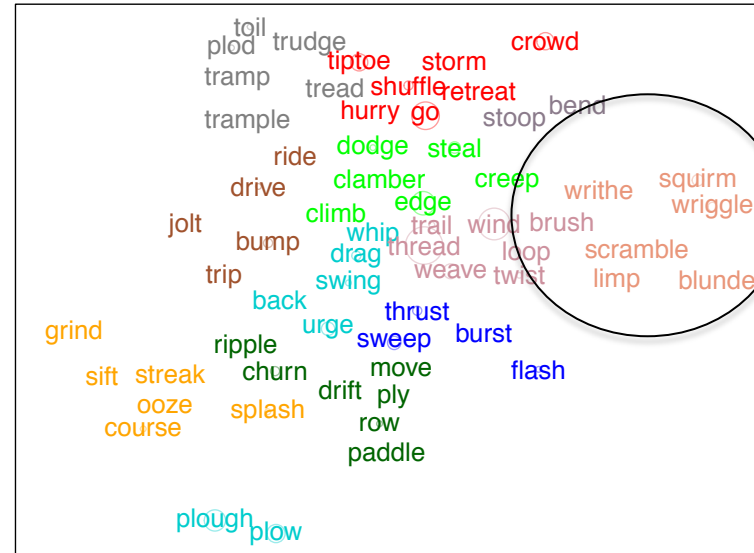


Verbs describing slow, indirect, or difficult motion: *thread*, *trial*, *weave*, *wind*, *plod*, *toil*, *tramp*, *trudge*.

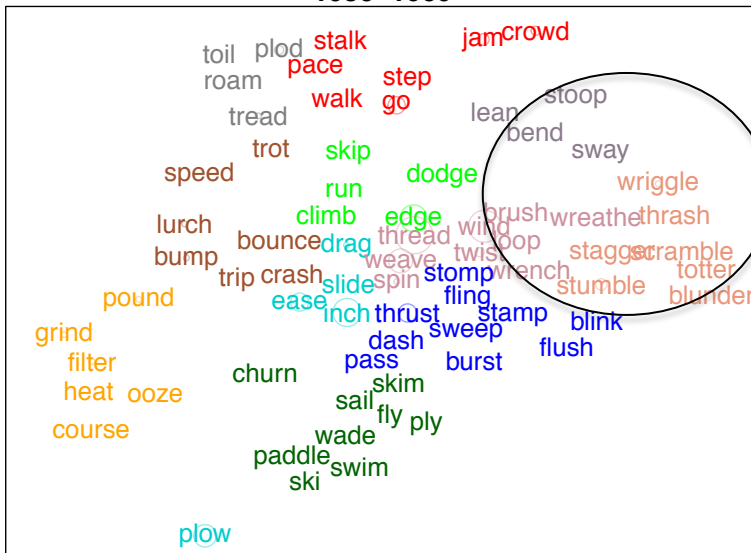
1830–1879



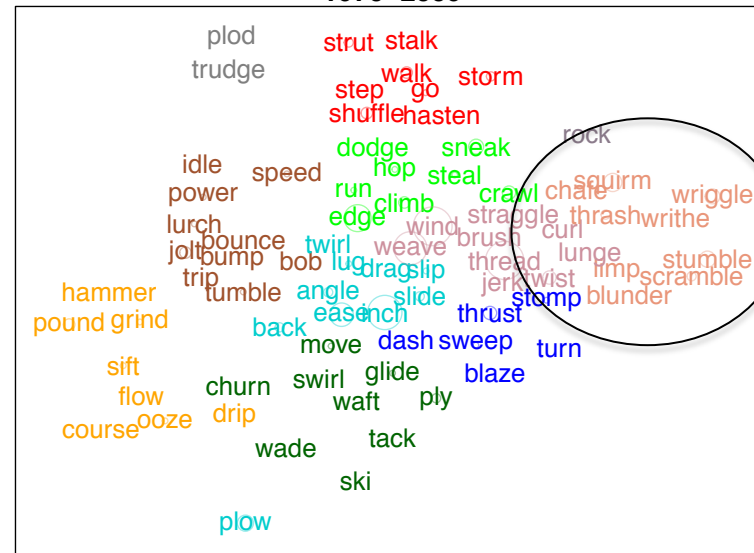
1880–1929



1930–1969



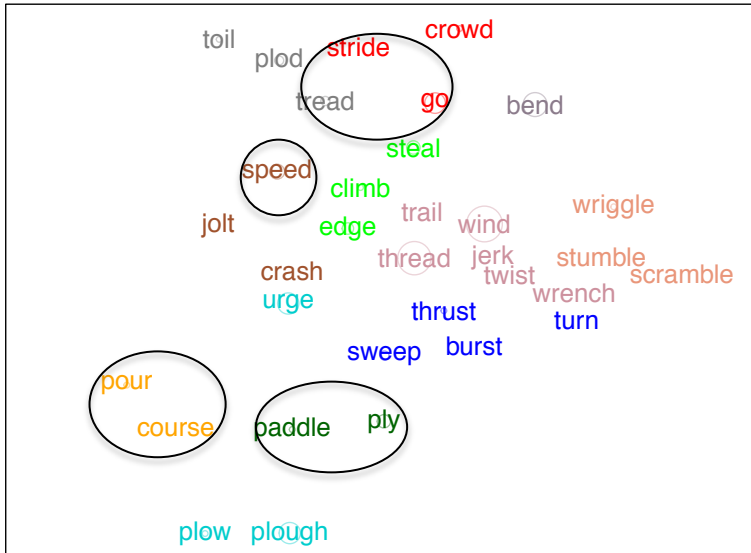
1970–2009



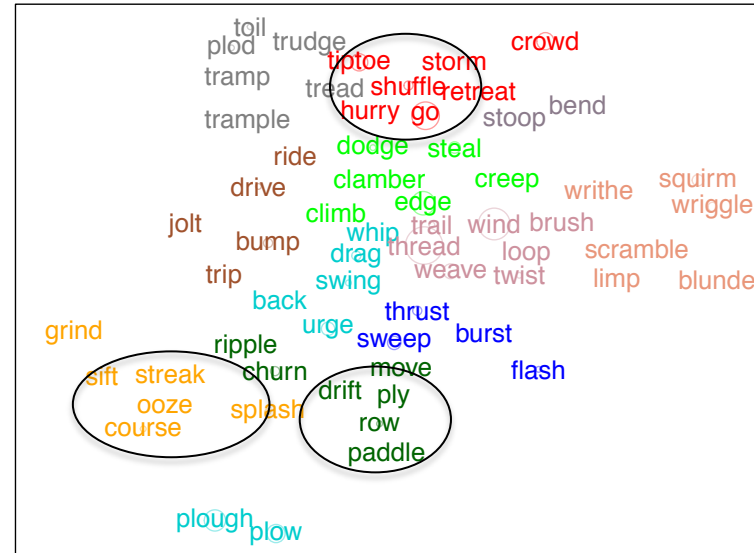
Clumsy or unsteady motion: *blunder, limp, scramble, stagger, stumble, totter*

Surrounded by verbs that encode body movements to facilitate motion: *bend, jerk, lean, lunge, stoop, thrash, twist, wrench, wriggle, writhe*

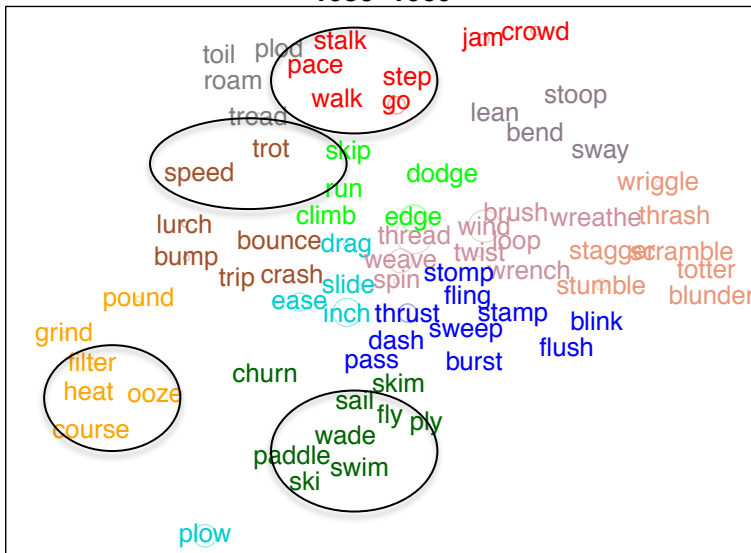
1830–1879



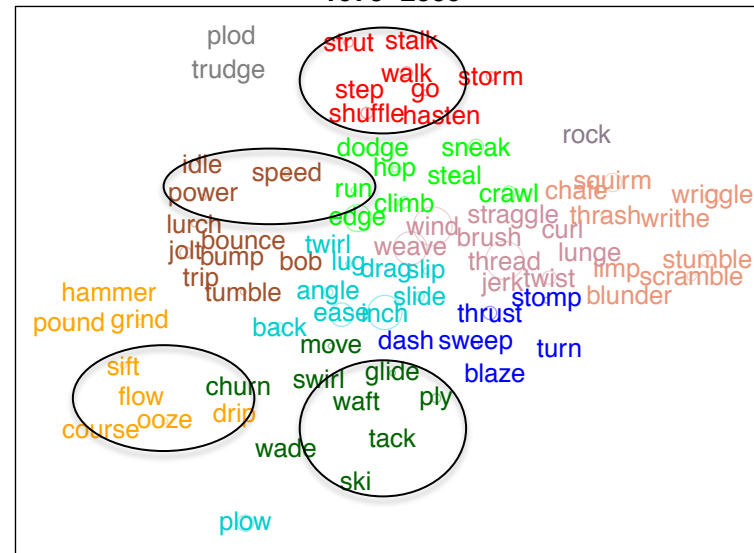
1880–1929



1930–1969



1970–2009

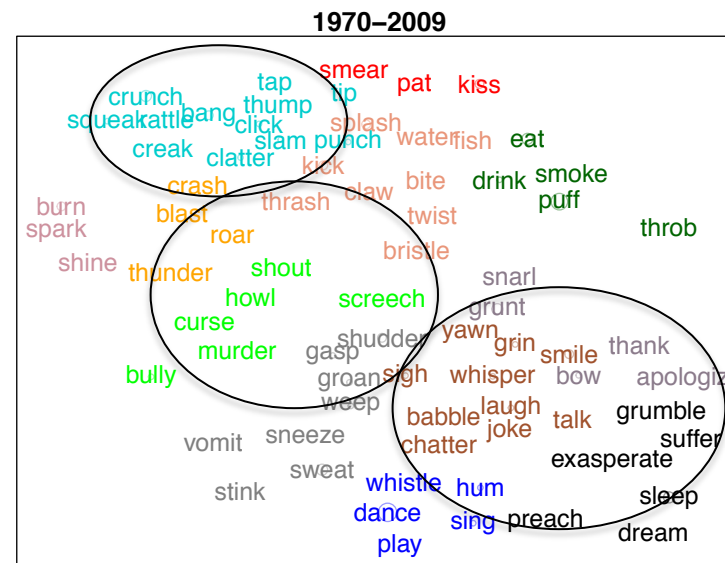
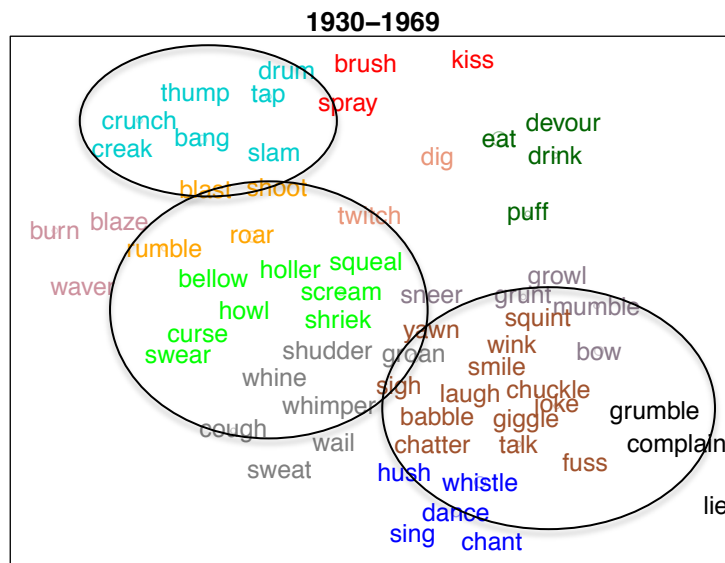
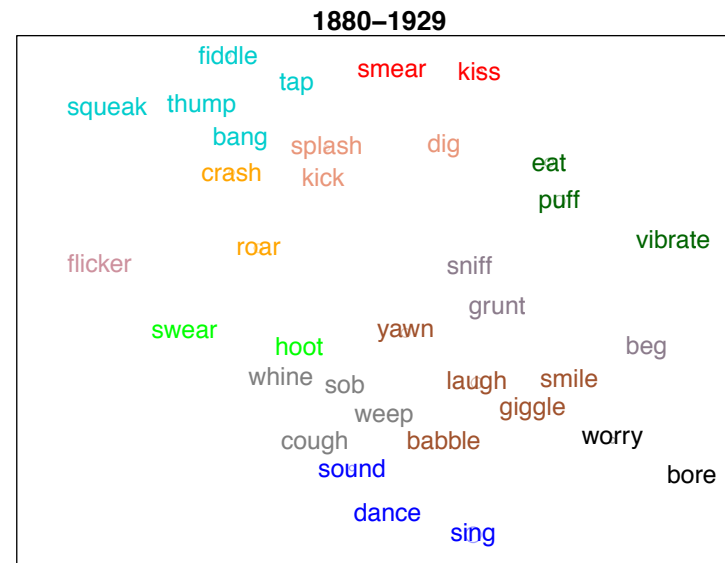
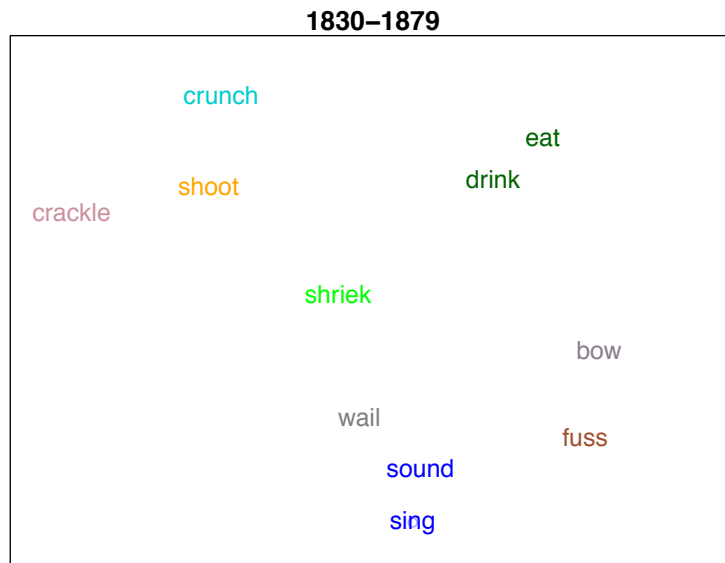


More 'neutral' manners of motion: walking (*stride, strut, tiptoe, walk, ..*), rapid motion (*power, run, speed, ..*), liquid motion (*course, drip, sift, ooze, ..*), vehicle/theme (*fly, paddle, ply, sail, ski, ..*)

The manner sense

- Difficult motion = semantic 'core' of the construction (Goldberg 1995)
 - Stable throughout the last 180 years
 - Several high-frequency members
 - Source of productivity: unsteady motion, body movements
- Non-difficult motion more prominent in later periods
- Likely interpretation: increase in schematicity of the verb slot, from difficult motion to general manner of motion

The incidental-activity sense



High semantic diversity from the start

Sound emission dominates in later periods: noise (*bang, creak, crunch, thump*), speech (*babble, chatter, joke, swear, talk*), cry (*bellow, howl, roar, scream, shriek*), other human sounds (*grunt, sigh, sob, whine, yawn*)

The incidental-activity sense

- Likely to be highly schematic from the start
 - No clear semantic core in the distribution
 - High semantic diversity from the start
 - Constructional meaning is rather open
- Increase in productivity, no increase in schematicity
 - Prominence of sound emission in mid and late 20th century
 - Probably explained by higher compatibility with schema
 - But not from the start (*contra* Israel 1996)

Conclusions

- Productivity and schematicity of the *way*-construction
 - Substantial distributional changes in the 19th and 20th centuries for all three senses of the construction
 - Likely to correspond to higher schematicity for two senses
- The usefulness of distributional semantic plots
 - Close examination of the semantic spread of a construction over time
 - Makes it possible to inform hypotheses about schematicity

Thank you!

florent.perek@gmail.com

<http://www.fperek.net>

- Davies, M. (2010). *The Corpus of Historical American English: 400 million words, 1810-2009*. Available online at <http://corpus.byu.edu/coha/>
- Firth, J.R. (1957). A synopsis of linguistic theory 1930-1955. In *Studies in Linguistic Analysis*, pp. 1-32. Oxford: Philological Society.
- Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Israel, M. (1996). The way constructions grow. In A. Goldberg (ed.), *Conceptual structure, discourse and language*. Stanford, CA: CSLI Publications, 217-230.
- Miller, G. & W. Charles (1991). Contextual correlates of semantic similarity. *Language and Cognitive Processes*, 6(1), 1-28.
- Mondorf, B. (2011). Variation and change in English resultative constructions. *Language Variation and Change*, 22, 397–421.
- Perek, F. (2014). Vector spaces for historical linguistics: Using distributional semantics to study syntactic productivity in diachrony. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, Baltimore, Maryland USA, June 23-25 2014*.
- Perek, F. (2016). Using distributional semantics to study syntactic productivity in diachrony: A case study. *Linguistics*, 54(1), 149–188.
- Traugott, E. & G. Trousdale (2013). *Constructionalization and Constructional Changes*. Oxford: Oxford University Press.
- Van der Maaten, L. & Hinton, G. (2008). Visualizing Data using t-SNE. *Journal of Machine Learning Research*, 9, 2579-2605.