

Causation, Culture, and Constructional Change

Martin Hilpert
Université de Neuchâtel

Large diachronic corpora such as the COHA or Google Books offer data that is tempting in its promise: In the words of Michel et al. (2010), it can “provide insights about fields as diverse as lexicography, the evolution of grammar, collective memory, the adoption of technology, the pursuit of fame, censorship, and historical epidemiology”. Studies that have pursued this line of thought have received widespread attention in the media, but have been met with reservation in the linguistic community. For example, Greenfield (2013) takes frequency trends of words such as *choose* and *get* as evidence for increasing individualism and materialism in American culture. Liberman (2013) comments on the paper as follows: “I’m not arguing that her theory is wrong, or that the Google ngrams 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’s comment raises the central question for this talk: How exactly can diachronic corpus data be analyzed in order to yield reliable insights about social change?

This paper addresses the issue of how culture can be analyzed on the basis of diachronic textual data in two parts. The first, theoretical part surveys a series of problems that need to be controlled for in analyses of diachronic textual data. These problems include the increasing risk of spurious correlations in large datasets (Koplenig and Müller-Spitzer 2016), the problem of the non-equivalence of onomasiological frequency and semasiological frequency, and the difficulty of disentangling cultural change and grammatical change (Szmrecsanyi 2016).

The second part implements these ideas in a study of change in the English make-causative (Kemmer 2001). As a grammaticalized expression of authority, the construction lends itself to a study of social change: Examples such as *She made the boys clean up their room* verbalize that a causer prompted a causee to perform a coerced action. If American culture becomes less authoritarian, as has been argued by Greenfield (2013) and others, examples such as the one above should recede in favor of uses such as *That music made me smile*, which involve inanimate causers and non-coerced actions. Data from the COHA is retrieved to track the history of the make-causative in terms of several semantic parameters, including animacy of causer and causee and the semantics of the verb that expresses the caused action.

References

- Greenfield, P. M. (2013). The changing psychology of culture from 1800 through 2000. *Psychological Science* 24, 1722-1731.
- Kemmer, Suzanne. (2001). Causative Constructions and Cognitive Models: The English Make Causative. *First Seoul International Conference on Discourse and Cognitive Linguistics: Perspectives for the 21st Century*, 803-846. Seoul: Discourse and Cognitive Linguistics Society of Korea.
- Koplenig, Alexander/Müller-Spitzer, Carolin (2016): Population size predicts lexical diversity, but so does the mean sea level – one problem in the analysis of temporal data. *PLOS ONE* 11(3).
- Liberman, Mark. (2013). The culturomic psychology of urbanization. Post on LanguageLog.
- Michel, Jean-Baptiste et al. 2010. Quantitative Analysis of Culture Using Millions of Digitized Books, *Science* 331: 176-182.
- Szmrecsanyi, Benedikt (2016). About text frequencies in historical linguistics: disentangling environmental and grammatical change. *Corpus Linguistics and Linguistic Theory* 12(1): 153-171