
Martin Hilpert
Freiburg Institute for Advanced Studies
martin.hilpert@frias.uni-freiburg.de

Watching language change: A new visualization technique and its theoretical implications

Introduction

In recent years, many historical corpora have become available that allow the study of linguistic change in real time (de Smet 2005, Davies 2008, 2010, inter alia). Although these resources have caused a surge of interest in diachronic corpus linguistics, many studies still limit themselves to analyses of rising or declining text frequencies. Meanwhile, diachronic corpora have much more information to offer. The present paper illustrates this with a new corpus-linguistic application – the flipbook technique.

The flipbook technique creates dynamic visualizations of language change. Like its real-world model, a flipbook of language change consists of a series of pictures in which each picture slightly differs from the previous one. Based on data from COHA (Davies 2010), two case studies illustrate recent changes in American English. The first study visualizes structural change in the behaviour of complement-taking predicates; the second one shows semantic change in the paradigm of modal auxiliaries.

Methodologically, flipbooks are created on the basis of data representing the same linguistic phenomenon across fifteen sequential decades. The data for each decade is submitted to a multi-dimensional scaling analysis (Kruskal & Wish 1978) that produces a two-dimensional plot. The overall result is a series of fifteen plots that gradually change from one to the next, thus revealing how a linguistic phenomenon has developed diachronically. It is argued that flipbooks of MDS plots offer an effective representation of complex linguistic changes that can not only guide exploratory analyses, but also speak to theoretical issues.

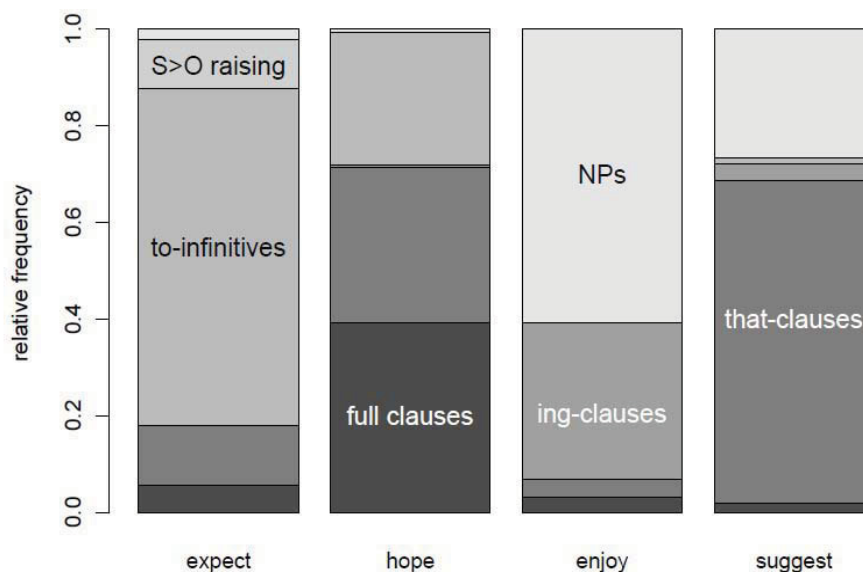
An illustration: complement-taking predicates

Verbs such as *expect*, *like*, or *imagine* project a syntactic complement structure that may take different shapes. The phrase *I expect* can be followed by phrases such as *a visitor*, *to hear from John*, or *that John will win*, amongst others. The present analysis distinguishes between the six different subcategorization frames shown in (1).

- | | | |
|-----|--|---|
| (1) | full clauses
<i>that</i> -clauses
<i>ing</i> -clauses
<i>to</i> -infinitives
subject-to-object raising
noun phrases | I suggest <u>we do nothing</u> .
I think <u>that John will win</u> .
I enjoy <u>knitting sweaters</u> .
I expect <u>to hear from John</u> .
I want <u>John to be our next president</u> .
I hate <u>broccoli</u> . |
|-----|--|---|

For 46 different verbs, the relative frequencies of these subcategorization frames have been retrieved from COHA. Figure 1 shows that complement-taking predicates exhibit a particular profile with regard to their preferred subcategorization frames: *expect* chiefly occurs with *to*-infinitives, whereas most examples of *enjoy* contain either an NP or an *ing*-clause.

FIGURE 1. RELATIVE FREQUENCY PROFILES FOR FOUR COMPLEMENT-TAKING PREDICATES.



With such frequency profiles, MDS can determine the relative similarities between larger sets of complement-taking predicates. This is done separately for the data from each COHA decade, such that the overall result is the series of 15 MDS maps shown in Figure 2. The graphs clearly distinguish between verbs with a preference for *that*-clauses (in the upper left), verbs with a preference for *to*-infinitives (upper right), and verbs that primarily take NP complements (lower middle). Amongst other things, the flipbook shows that the verb *confirm* abandons a preference for NP complements in favour of *that*-clauses from the 1930s onwards.

References

- Davies, M. 2008. The Corpus of Contemporary American English (COCA): 400+ million words, 1990-present. <http://www.americancorpus.org>.
- Davies, M. 2010. The Corpus of Historical American English (COHA): 400+ million words, 1810-2009. <http://corpus.byu.edu/coha>.
- De Smet, H. 2005. A corpus of Late Modern English texts. *ICAME Journal* 29, 69-82.
- Kruskal, J. and M. Wish. 1978. *Multidimensional Scaling*. London: Sage.
-