

TRACING SIGNIFICANT CHANGE IN 17TH-CENTURY ENGLISH LEXIS: THE CIVIL-WAR EFFECT

Table 1. Number of words whose frequencies are significantly different in the two periods of the CEEC (1600–1639 and 1640–1681), at various significance thresholds α (n = 46,440).

α	Log-likelihood ratio test	Bootstrap test	Both
0.01	2,685 (6 %)	2,365 (5 %)	2,199 (5 %)
0.001	1,400 (3 %)	1,209 (3 %)	1,108 (2 %)
0.0001	937 (2 %)	759 (2 %)	722 (2 %)

For more information, see Lijffijt et al. (forthcoming 2012).

REFERENCES

- CEEC = *Corpus of Early English Correspondence*. 1998. Compiled by Terttu Nevalainen, Helena Raumolin-Brunberg, Jukka Keränen, Minna Nevala, Arja Nurmi & Minna Palander-Collin at the Department of English, University of Helsinki.
<http://www.helsinki.fi/varieng/CoRD/corpora/CEEC/>
- Dunning, Ted. 1993. “Accurate methods for the statistics of surprise and coincidence”. *Computational Linguistics* 19(1): 61–74.
- Google Books Ngram Viewer. <http://books.google.com/ngrams/>
- HT = *Historical Thesaurus of the Oxford English Dictionary*. 2009. Edited by Christian Kay, Jane Roberts, Michael Samuels & Irené Wotherspoon. OED Online.
<http://www.oed.com/thesaurus>
- Kilgarriff, Adam. 2001. “Comparing corpora”. *International Journal of Corpus Linguistics* 6(1): 97–133.
- Lijffijt, Jeffrey, Tanja Säily & Terttu Nevalainen. Forthcoming 2012. “CEECing the baseline: Lexical stability and significant change in a historical corpus”. To appear in the proceedings of the Helsinki Corpus Festival.
<http://www.helsinki.fi/varieng/journal/>
- Lijffijt, Jeffrey, Terttu Nevalainen, Tanja Säily, Panagiotis Papapetrou, Kai Puolamäki & Heikki Mannila. Forthcoming. “Significance testing of word frequencies in corpora”. Submitted to *International Journal of Corpus Linguistics*.
- Michel, Jean-Baptiste, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K. Gray, The Google Books Team, Joseph P. Pickett, Dale Hoiberg, Dan Clancy, Peter Norvig, Jon Orwant, Steven Pinker, Martin A. Nowak & Erez Lieberman Aiden. 2010. “Quantitative analysis of culture using millions of digitized books”. *Science* 331(6014): 176–182.
- Paquot, Magali & Yves Bestgen. 2009. “Distinctive words in academic writing: A comparison of three statistical tests for keyword extraction”. *Corpora: Pragmatics and Discourse*, ed. by Andreas H. Jucker, Daniel Schreier & Marianne Hundt, 247–269. Amsterdam & New York: Rodopi.
- Raumolin-Brunberg, Helena. 1998. “Social factors and pronominal change in the seventeenth century: The Civil-War effect?” *Advances in English Historical Linguistics (1996)* (Trends in Linguistics: Studies and Monographs 112), ed. by Jacek Fisiak & Marcin Krygier, 361–388. Berlin: Mouton de Gruyter.

