

tPM: Selected Bibliography

The Prime Machine contains a range of different features which draw on previous research and software tool development from the field of corpus linguistics.

This selected list of wider reading is in no way meant to represent all the major influences which have had an impact on the development of tPM, but these should give a good overview of some of the background to the different ways of looking at data which are available in this tool. For a more detailed bibliography for V1 see Jeaco (2015) listed below.

The Theory of Lexical Priming

Hoey, M. (2005). *Lexical Priming: A New Theory of Words and Language*. London: Routledge.

Hoey, M., & O'Donnell, M. B. (2008). Lexicography, grammar, and textual position. *International Journal of Lexicography*, 21(3), 293-293.

Pace-Sigge, M. (2013). *Lexical Priming in Spoken English Usage*. London: Palgrave Macmillan.

Background to *The Prime Machine*

Jeaco, S. (2015). *The Prime Machine: a user-friendly corpus tool for English language teaching and self-tutoring based on the Lexical Priming theory of language*. Unpublished Ph.D. dissertation. University of Liverpool. Retrieved from <https://livrepository.liverpool.ac.uk/2014579/>

Jeaco, S. (2017a). Concordancing Lexical Primings. In M. Pace-Sigge & K. J. Patterson (Eds.), *Lexical Priming: Applications and Advances* (pp. 273-296). Amsterdam: John Benjamins.

Jeaco, S. (2017b). Helping Language Learners Put Concordance Data in Context: Concordance Cards in The Prime Machine. *International Journal of Computer-Assisted Language Learning and Teaching*, 7(2), 22-39.

Key inspirations from other software for many of the main software features of tPM¹

Anthony, L. (2004). *AntConc: A learner and classroom friendly, multi-platform corpus analysis toolkit*. Paper presented at the Interactive Workshop on Language e-Learning, Waseda University, Tokyo.

Cobb, T. (2000). The Compleat Lexical Tutor, from <http://www.lextutor.ca>

Davies, M. (2008-). The Corpus of Contemporary American English (COCA): 520 million words, 1990-present. Retrieved 3 April, 2017, from <http://corpus.byu.edu/coca/>

¹ Of course I have been influenced by so many other great tools – my PhD thesis explored many of these, and each corpus linguistics conference I attend continues to inspire me through web tools, software tools, applications and research questions.

Rayson, P. (2008). From key words to key semantic domains. *International Journal of Corpus Linguistics*, 13(4), 519-549.

Rayson, P. UCREL Log-likelihood and effect size calculator Retrieved 04 June, 2018, from <http://ucrel.lancs.ac.uk/llwizard.html>

Scott, M. (2010). *WordSmith Tools (Version 5.0)*. Oxford: Oxford University Press.

Key inspirations from reading for many of the main software features of tPM

Garretson, G. (2010). *Corpus-Derived Profiles: A Framework for Studying Word Meaning in Text*. Unpublished Ph.D. dissertation. Boston University.

Scott, M., & Tribble, C. (2006). *Textual Patterns: Key Words and Corpus Analysis in Language Education*. Amsterdam: John Benjamins.

Key inspirations for Data-Driven Learning

Johns, T. (1991). Should you be persuaded: Two samples of data-driven learning materials. In T. Johns & P. King (Eds.), *Classroom Concordancing* (Vol. 4, pp. 1-13). Birmingham: Centre for English Language Studies, University of Birmingham.

Tsui, A. B. M. (2004). What teachers have always wanted to know - and how corpora can help. In J. M. Sinclair (Ed.), *How to Use Corpora in Language Teaching* (pp. 39-61). Amsterdam: John Benjamins.

Key inspirations for handling of literary texts and some DIY tools

Mahlberg, M. (2013). *Corpus stylistics and Dickens's fiction*: New York ; Routledge, 2013.

Mahlberg, M., Stockwell, P., Joode, J. d., Smith, C., & O'Donnell, M. B. (2016). CLiC Dickens: novel uses of concordances for the integration of corpus stylistics and cognitive poetics. [Article]. *Corpora*, 11(3), 433-463. doi: 10.3366/cor.2016.0102

Some background to some of the tagging systems used for the pre-processed online corpora in tPM²

Garside, R., & Smith, N. (1997). A hybrid grammatical tagger: CLAWS4. In R. Garside, G. Leech & A. McEnery (Eds.), *Corpus Annotation: Linguistic Information from Computer Text Corpora* (pp. 102-121). London: Longman.

² You can't tag DIY Corpus texts with these taggers in tPM; but the ready-made corpora have been tagged using these taggers.

Nini, A. 2014. Multidimensional Analysis Tagger 1.1 - Manual. Retrieved from:

<http://sites.google.com/site/multidimensionaltagger>

Rayson, P., Archer, D., Piao, S., & McEnery, T. (2004). *The UCREL semantic analysis system*. Paper presented at the Beyond Named Entity Recognition Semantic Labeling for NLP Tasks Workshop, Lisbon, Portugal.

References and links for the resources used to make links in the corpora lexicons for search suggestions

CC- CEDICT (<https://www.mdbg.net/chinese/dictionary?page=cedict>)

FrameNet (<http://framenet.icsi.berkeley.edu/>)

Fillmore and Baker (2010) "A Frame Semantic Approach to Linguistic Analysis" in Heine and Narrog (eds.) Oxford Handbook of Linguistic Analysis.

WordNet (<https://wordnet.princeton.edu/>). WordNet 2.0 Copyright 2003 by Princeton University. All rights reserved.

Note

This list is by no means complete, and this list is **not** intended to give an indication of what is considered **important** or the **most up-to-date**, but I believe these texts and resources in particular have influenced the development of tPM.

Support

The Prime Machine is still undergoing development.

For further information see <http://help.theprimemachine.com>

Last Updated: 12 September 2018