



3 Minute Paper Series



Format

- Presenter shares a piece of existing research as a short presentation
- Inspired by the “3 Minute Thesis (3MT) format

Today

- Short summary of Roesler (2021)
- Longer discussion of related issues

Roesler, D. (2021). When a bug is not a bug: An introduction to the computer science academic vocabulary list. *Journal of English for Academic Purposes*, 54, 101044



3 Minute Paper

Patrick Fitzgerald | 4 February 2026 | BALEAP STEM SIG

Vocabulary Lists – some context

General word lists – GSL (General Service List - 1953) ... NGSL and *NewGSL* (New General Service List - both 2013)

Academic word lists – UWL (University Word List – 1984) ... AWL (Academic Word List – 2000) ... AVL (Academic Vocabulary List – 2013)

Specialist word lists (computing) – CSWL (Computer Science Word List - 2013) ... CSAVL (Computer Science Academic Vocabulary List - 2021)

CSWL (2013) → CSAVL (2021) improvements

- (1) Aim to include high frequency dual meaning words (CSWL removed all words from the General Service List and the Academic Word List eg 'string')
- (2) Use lemmas rather than word families (to reduce unnecessary learning burden on students)
- (3) Validate disciplinary specific meaning
- (4) Include frequency rank information

Corpus and process

Created main corpus (3.5m tokens) and test corpus (.7m tokens) from computer science text books and journal articles – lemma extraction using Lancsbox

Selection criteria:

1. frequency items >100 in main corpus
2. meaning in computer science OR greater frequency than a general English corpus
3. distribution across a range of computer science topic areas (sub-corpora)

20 Topic based sub-corpora (CSAC1/CSAC2)



From text books (12/10)

Computer systems organization

Data structures*

Algorithms*

Human-centered computing

Probability and statistics*

Mathematics of computing

Programming*

Operating systems*

Software and its engineering

Theory of computation

From Journal articles (142/42)

Computer systems organization

Computing methodologies

Hardware

Human-centered computing

Information systems

Mathematics of computing

Networks

Security and privacy

Software and its engineering

Theory of computation

Vocabulary coverage

using .7m token test corpus

Wordlist(s)	Wordlist Size (lemmas)	% coverage of test corpus
Computer Science Academic Vocabulary List (<i>CSAVL</i>)	904	16%
Computer Science Academic Vocabulary List + supplementary list (<i>CSAVL</i> + <i>CSAVL-S</i>)	1606	19.9%
New-General Service List (<i>NewGSL</i> + <i>CASVL</i> + <i>CASVL-S</i>)	3918	94.77%
General Service List (<i>GSL</i>), Academic Word List (<i>AWL</i>) + Computer Science Word List (<i>CSWL</i>)	8489	95.49

Limitations of the CSAVL

1. Size – 3.5m tokens (eg limited size of source corpus)
2. Possible mismatch with the reference general English corpus (BNC)
3. No discrimination between intended meanings (eg memory_n could refer to computer memory or human memory)
4. No hyphenated words (eg real-time) – no collocations

List availability and tools

CSAVL main list and supplementary list -

<https://www.eapfoundation.com/vocab/academic/other/csavl/>

Highlighting and gapfill tools for use with text selections -

<https://www.eapfoundation.com/vocab/academic/other/csavl/highlighter/>

Overview of specialist word lists –

<https://www.eapfoundation.com/vocab/wordlists/overview/>

Field-
specific/
technical

- SWL (Science Word List) 2007
 - BWL#1 (Business Word List #1) 2007
 - PWL (Pharmacology Word List) 2007
 - MAWL (Medical Academic Word List) 2008
 - AgroCorpus List 2009
 - BEL (Basic Engineering List) 2009
 - BWL#2 (Business Word List #2) 2011
 - CSWL (Computer Science Academic Word List) 2013
 - CAWL (Chemistry Academic Word List) 2013
 - MAVL (Medical Academic Vocabulary List) 2015
 - NAWL (Nursing Academic Word List) 2015
 - EAWL (Environmental Academic Word List) 2015
 - EAWL (Economics Academic Word List) 2019
 - CSAVL (Computer Science Academic Vocabulary List) 2021

 - MSVL (Middle School Vocabulary Lists) 2015
 - SVL (Secondary School Vocabulary Lists) 2018

 - HSWL (*Hard Science Spoken Word List*) 2018
 - SSWL (*Soft Science Spoken Word List*) 2018
- CSMWL (Computer Science Multi-Word List) 2013
 - SPL (Secondary Phrase Lists) 2018
 - NCL (Nursing Collocation List) 2020

References

Brezina, V., & Gablasova, D. (2015). Is there a core general vocabulary? Introducing the new general service list. *Applied Linguistics*, 36(1), 1–22. <https://doi.org/10.1093/applin/amt018>

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213–238

Lancsbox - <http://corpora.lancs.ac.uk/lancsbox/index.php>

Minshall, D. E. (2013). A computer science word list. Master's dissertation, University of Swansea. Available from: https://qmlanguagecentre.on-rev.com/baleap/archive/media/uploads/dissertation-awards/2014/Daniel_Minshall_Dissertation_Final_Draft.pdf

Roesler, D. (2021). When a bug is not a bug: An introduction to the computer science academic vocabulary list. *Journal of English for Academic Purposes*, 54, 1010443(2), 215-229

Xue, G., & Nation, I. S. (1984). A university word list. *Language learning and communication*

Questions ?

Any questions arising from the presentation?



Discussion

1. Are academic vocabulary lists generally a useful tool for EAP teachers?
2. To what extent should EAP teachers and materials developers concern themselves with subject specific vocabulary in STEM related areas?
3. In what ways can subject specific vocabulary lists be used by teachers of STEM EAP and materials developers?

Stay Connected

Website: baleapstemsig.org

**Next session:  Spotlight Session
– March 2026**

**Join our mailing list & share your
ideas #baleapstemsig**

