

Beyond Bundles: a new approach to vocabulary selection for English for Academic Purposes

Geraint Paul Rees
InfoLex, Universitat Pompeu Fabra
geraintpaul.rees@upf.edu @ReesGP

Introduction

This study deals with the selection of vocabulary for English for Academic Purposes (EAP) courses. EAP courses are important in both **professional** and **economic** terms:

- Govern students' access to education
- Important source of revenue for the publishing industry

There are currently two principal approaches to the selection of vocabulary for EAP, but both are problematic.

1. Single-word lists of general academic vocabulary

Examples:

- Coxhead's (2000) *New Academic Word List*
- Gardner & Davies's (2013) *New Academic Vocabulary List*.

Problematic because:

- No single wordlist could meet the needs of all EAP students
- Words have distinct meanings in different academic disciplines
- Word meaning is conditioned by collocates

2. The general academic 'Lexical Bundle' approach

Strings of words are selected according to their frequency of occurrence in a corpus (e.g. Biber & Conrad, 1999; Simpson-Vlach & Ellis, 2010).

Accounts for collocates.

Problematic because there is no consideration of the semantic properties of collocations when extracting bundles from text

Hypothesis: The lexical bundle approach obfuscates differences in meaning and use of vocabulary between academic disciplines.

Resources

- **Academic Journal Article Corpus 2 (AJACx2)** – represents language encountered by students.
- **ATLAS.ti** – annotation software
- **The Pattern Dictionary of English Verbs (PDEV)** (Hanks 2001) – represents patterns of general English (GE); annotation aid.
- **The Sketch Engine** (Kilgarriff, Rychlý, Smrč, & Tugwell,

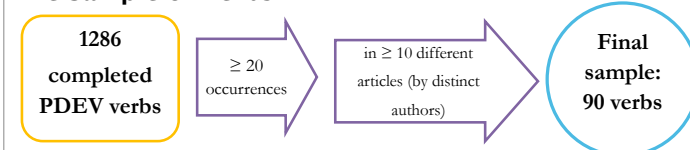
The Corpus

8,088,429 words in three discipline specific sub-corpora

- **History** 2,840,024 words
- **Microbiology** 2,606,594 words
- **Management Studies** 2,641,811 words

Balanced corpus, 11 years of articles from journals which student of the above disciplines are likely to consult.

The Sample of Verbs



Method: Corpus Pattern Analysis (CPA) (Hanks, 2004)

CPA = mapping meaning onto text.

CPA is a practical application of the Theory of Norms and Exploitations (Hanks, 2013). According to Hanks (2013: 167),

Language consists of a constantly moving and developing double helix of rules governing linguistic behaviour: normal uses and exploitations of normal use.

Collocates are arranged in lexical sets according to their collocational preference. Then mapped to syntactic structures as colligations:

[[Person]] grasp [[PhysObj]] (14%)

Implicature: **[[Person=Animate]] seizes [[PhysObj]] and holds it firmly.**

Lexical Alternation: **[[Person]] ↔ {hand, finger}**

Other Clues: **{in [POSDET] hand}, {by [DET] arm}**

Example: He grasped the handle of the door with one hand, and that of the spoon with the other.

Norms and their exploitations are discipline-specific and can be calculated from corpus lines by examining contextually determined default interpretations.

Some Typical Findings: a CPA Analysis of *accept*

Discipline Specific Syntactic Alternations

In Microbiology

[[Human | Institution]] accepts (37.5% MB vs. 21.6% GE)

Implicature: **[[Human | Institution]] agrees that [[that clause]] is true or correct**

Syntactic Alternation: **{it} is ({generally}{well}{widely})**

accepted {that} (100% in MB)

Example: It is generally **accepted** that compounds causing mutations in one type of cell should also be considered mutagenic for other cells.

Discipline Specific Semantic Type Alternations

In History

[[Human | Institution]] accepts

[[Proposition | Concept | Eventuality]] (67.3% HIST vs. 51.2% GE)

Implicature: **[[Human | Institution]] agrees that**

[[Proposition | Concept | Eventuality]] is correct and does not need to be contested

Semantic Type Alternation: **[[Human | Institution]] ↔**

[[Location]]

Example: Towns and territories **accepted** the reformation.

Discipline Specific Semantic Prosody

In Management

[[Human | Institution]] accepts

[[Proposition | Concept | Eventuality]] (59.3% MAN vs. 51.2% GE)

Implicature: **[[Human | Institution]] agrees that**

[[Proposition | Concept | Eventuality]] is correct and does not need to be contested

Predominant semantic prosody: **[[Eventuality = Negative]]**

Example: Employees are more **accepting** of an unfavourable outcome

Conclusion

A close examination of the semantic properties of verbs and their collocates makes it possible to elucidate differences in meaning that would not be apparent if a lexical bundle approach were employed. This study is a first step towards a better understanding of discipline-specific behaviour of verbs, and guidelines for their treatment in teaching.

References

- Biber, D., & Conrad, S. (1999). Lexical bundles in conversation and academic prose. In H. Hasselgård, S. Johansson, & S. Oksefjell (Eds.), *Out of Corpora: Studies in Honour of Stig Johansson* (pp. 181–190). Amsterdam: Rodopi.
- Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, 34, 213–238.
- Gardner, D., & Davies, M. (2013). A New Academic Vocabulary List. *Applied Linguistics*, 35(3), 305–327.
- Hanks, P. (2001). The Pattern Dictionary of English Verbs. WEB, Masaryk, Czech Republic: Masaryk University, Czech Republic. Retrieved from <http://www.pdev.org.uk>
- Hanks, P. (2004). Corpus Pattern Analysis. (G. Williams & S. Vessier, Eds.) *Proceedings of the Eleventh EURALEX International Congress* (pp. 87–97). Université de Bretagne-Sud, Lorient: EURALEX.
- Hanks, P. (2013). *Lexical Analysis: Norms and Exploitations*. Cambridge, Mass: MIT Press.
- Kilgarriff, A., Rychlý, P., Smrč, P., & Tugwell, D. (2004). The Sketch Engine. In G. Williams & S. Vessier (Eds.), *Proceedings of the Eleventh EURALEX International Congress* (pp. 105–116). Université de Bretagne-Sud, Lorient: EURALEX.
- Simpson-Vlach, R., & Ellis, N. (2010). An Academic Formulas List: New Methods in Phraseology Research. *Applied Linguistics*, 31, 487–512.