A Phraseological Multi-discipline Approach to Vocabulary Selection for English for Academic Purposes

Geraint Paul Rees

TESI DOCTORAL UPF / 2018

DIRECTORA DE LA TESI

Dra. Janet-Ann DeCesaris Ward

DEPARTAMENT DEPARTAMENT DE TRADUCCIÓ I CIÈNCIES DEL LLENGUATGE

Universitat Pompeu Fabra Barcelona

To my mother and father

Acknowledgements

I would like to extend my gratitude to my supervisor Doctor Janet DeCesaris who persuaded me to pursue to doctoral studies at UPF and whose advice and infectious enthusiasm has inspired me to reach this point. I would also like to thank the members of the IULA, who generously awarded the funding which allowed me to undertake this project.

A special thank you to my parents, Gareth and Jane, for their gentle encouragement and unwavering support.

I am indebted to my wife Dina whose love, kindness, and continued patience and understanding I hope to somehow repay.

I would like to thank all the members of InfoLex, especially Doctors Elisenda Bernal and Sergi Torner, not only for the advice and encouragement they have provided, but also for making me feel part of the group.

It is not possible to list all the friends who have made these past three years so enjoyable, however, I would like to give special thanks Alba and Blanca for making me feel so welcome. I could not have done this without you, and I am certain it would not have been nearly as much fun.

A final thank you to all those who have been so generous with their time and knowledge throughout the process. This thesis would not have been possible without you.

Abstract

This study is motivated by concerns about the adequacy of current corpus-based methods for the selection of vocabulary in studies of English for Academic Purposes (EAP) lexis. It is hypothesised that lists of general-academic vocabulary cannot reflect differences in word meaning between disciplines, and furthermore, that discrete-item wordlists cannot account for the role of context in conditioning meaning. The relatively recent turn towards discipline-specific lists of phrases represents a positive development in this regard. However, its impact is limited by the methods of phrase extraction typically employed.

These beliefs are tested via an innovative corpus-based experiment which compares the syntagmatic patterns of frequently occurring verbs in a corpus of research articles from the disciplines of history, microbiology, and management studies. The results demonstrate that, in many cases, the prototypical meaning of a given verb varies according to the discipline it is found in. Moreover, in order to fully appreciate these differences, a means of phrase extraction which accounts for both syntactic and semantic concerns is necessary.

In addition to the methodological contribution represented by the experimental procedure, the study demonstrates how approaches to language which might be termed phraseological provide plausible explanations for many of the differences in verb behaviour observed in the corpus. From a practical perspective, the combination of the findings with best practice in EAP pedagogy and lexicography allows the creation of guidelines for, and an illustrative example of, useful lexicographical resources for the EAP community.

Resum

Aquesta tesi tracta de la selecció de vocabulari per a cursos i materials d'anglès per a finalitats acadèmiques. Neix dels dubtes sobre l'adequació dels mètodes basats en corpus que s'usen actualment per seleccionar vocabulari en estudis de lèxic d'anglès acadèmic. La hipòtesi que planteja és que les llistes de vocabulari acadèmic generals no poden reflectir les diferències de significat que hi ha en les paraules segons la disciplina en què es troben, i, a més, que les llistes de paraules descontextualitzades no permeten valorar el grau en què el context condiciona i determina el significat. En els últims anys, s'ha privilegiat oferir llistes de sintagmes de disciplines específiques com un desenvolupament positiu, tot i que l'impacte real està limitat pels mètodes d'extracció que s'utilitzen habitualment.

Les hipòtesis es posen a prova mitjançant un experiment que compara els patrons sintagmàtics de verbs d'ús freqüent en un corpus d'articles d'investigació de les disciplines d'història, microbiologia i administració i gestió d'empreses. Els resultats demostren que, en molts casos, el significat prototípic d'un verb varia en funció de la disciplina en què es troba. A més, per valorar adequadament aquestes diferències, cal que l'eina d'extracció de sintagmes inclogui tant qüestions sintàctiques com semàntiques.

A més de la contribució metodològica que representa la metodologia basada en corpus, la tesi demostra que els enfocaments de l'estudi del llenguatge que es podrien denominar fraseològics proporcionen explicacions plausibles per a moltes de les diferències de comportament verbal que s'observen al corpus. Des d'una perspectiva aplicada, la combinació de resultats pot incidir en millors pràctiques didàctiques i lexicogràfiques de l'anglès acadèmic i crear pautes i recursos més útils per als aprenents.

PREFACE

Since the early 2000s there has been an increasing interest in the processes involved in selecting vocabulary for English for Academic Purposes (EAP) materials and courses. This reflects the current status of English as a lingua franca in academic contexts (Jenkins, Cogo, & Dewey, 2011) and the growing importance of EAP in the academic, economic, and professional lives of countless students around the globe. Success in standardised proficiency tests of academic English such as the TOEFL or IELTS[1] is a prerequisite of graduation from higher-education institutions around the world and, for non-native English speakers, a prerequisite of entry to tertiary education in most English speaking countries. Previous research in this area has criticised the notion of a general-academic lexis present across all disciplines (Q. Chen & Ge, 2007; Hyland & Tse, 2009; Y. Li & Qian, 2010; Rees, 2013, 2016). It has also highlighted the need to consider word context in the study of academic lexis, adopting a phraseological approach which not only accounts for both syntactic and semantic properties, but is also appropriate for the analysis of large amounts of usage-based data. This thesis aims to fulfil this need. Although it deals with English, specifically for academic purposes, its methods and conclusions are applicable to specific purpose language courses and materials for other languages and in other domains.

This introduction gives an overview of the historical development of EAP and outlines some of the problems presented by current approaches to vocabulary selection. These problems motivate the objectives of the present study which are set out below. The remainder of this thesis is split into three principal sections:

The first section presents concepts which are key to the present study: Chapter 1 begins by outlining a number of approaches to linguistic analysis which may be considered phraseological. It concludes by putting forward a practicable theoretical framework for a corpus-based study of collocational behaviour of verbs across three academic disciplines. The second chapter outlines some important concepts in the teaching and learning of EAP vocabulary. It argues that lexicographical resources have a key role to play in these processes. The third chapter deals with the treatment of EAP vocabulary and phraseology in lexicographical resources.

The second section concerns the corpus-based experiment at the heart of this thesis: Chapter 4 puts forward general hypothesises tested in this experiment and the assumptions which underlie them. In particular those concerning the nature of academic language and the decision to concentrate on verb collocations rather than other parts of speech. The fifth chapter outlines the corpus-based methodology employed in the experiment.

The final section contains a discussion of the results obtained: The presentation and analysis of quantitative results occurs in Chapter 6. Chapter 7 comprises an in-depth qualitative analysis of a selection of the verbal collocations examined in the experiment. It also discusses how its findings might best be represented in lexicographical resources for EAP users, taking into account research on the teaching and learning of English vocabulary and learner dictionaries. The final chapter presents a summary of key findings and opportunities for future research and lexicographical advances.

[1] Test of English as a Foreign Language (ETS, 2017); International English Language Testing System (IELTS, 2017).

The interest in the selection of vocabulary for EAP is due in part to the circumstances in which the field of teaching of English as a foreign or second language developed. Technological advances, which from the 1960s onwards made the study of large corpora feasible, have also played a role.

Historically, the teaching of English as a foreign or second language as a profession has been a predominantly British and North-American phenomenon. In general, until the late 1800s the teaching of foreign languages in Europe usually took place in primary and secondary school contexts where translation, and the rote learning of grammatical rules dominated. The reform movement of the late 1800s, in particular the monolingual-target-language approach of the direct method, represents the establishment of English and other foreign language teaching, outside secondary school contexts, as a discipline in its own right. Howatt (1984) cites three basic founding principles of the movement "the primacy of speech, the centrality of the connected text […], and the absolute priority of an oral methodology in the classroom." (Howatt, 1984, p. 172). To the extent that it represented a shift away from the analysis of grammar rules towards the use of language for communication in the classroom (N Schmitt, 2000), the direct method can be viewed as a recognition the importance of vocabulary in language learning.

Further reform came later in the 1920s and 1930s, against a background of increasing demand for English teaching in non-school contexts. In the UK this demand was driven by refugees fleeing political upheaval in central Europe. Outside the UK demand was driven by the need to teach English as a foreign language in the Empire (later the Commonwealth), and elsewhere as a means of facilitating foreign trade. This increased demand was a catalyst for the creation of new materials for the new methodology. Harold Palmer's pedagogical grammars, *A Grammar of Spoken English* (GSE; H. Palmer, 1924) and *A Grammar of English Words* (GEW; H. Palmer, 1938) can be seen as a response to this demand. From a lexicographical perspective, Hornby, Gatenby, and Wakefield's *Idiomatic and Syntactic English Dictionary* (ISED; Hornby, Gatenby, & Wakefield, 1942), started in 1937, highlights the importance of vocabulary to the direct method.

Somewhat paradoxically, the direct method placed great emphasis on instruction and interaction in the target language yet lacked specific guidelines on which lexical items should be studied. The Vocabulary Control Movement attempted to fill thus gap. The Carnegie Report (Faucet, West, Palmer, & Thorndike, 1936) represents an attempt to systematically select vocabulary for beginner-level reading material for English language learning. It is a pre-cursor to West's (1953) *General Service List* (GSL). The GSL is a list of 2000 words needed for survival in a general English environment. Its words are selected according to frequency of occurrence and other criteria. It is a tangible result of the Vocabulary Control Movement's work. Further studies have found that 2000 words would be a reasonable general survival list (Shonell, Meddleton, & Shaw, 1956). However, others have found as little 120 to be sufficient (Nation & Crabbe, 1991). For more advanced tasks such as reading authentic texts 8000-9000 word families (Nation, 2006), groups of words organised on the basis of their inflections and their derivations, are said to be sufficient. For academic purposes, such as reading university textbooks, Hazenburg and Hulstijn (1996) suggest that 10,000 word families are necessary.

The 1960s were characterised an expansion of higher education institutions in Anglophone countries and an influx students who were not native speakers of English. This lead to the establishment of English Language Teaching (ELT) and, later, in the 1970s, EAP as specialisms (Jordan, 2002). Following in the tradition of the Vocabulary Control

Movement, a several studies on vocabulary selection were conducted. Campion and Elley (1971) is the first of several attempts to create a list of academic vocabulary with reference to real academic language, and like Praninskas (1972) is based on a manually compiled corpus of academic language. Lynn (1973) and Ghadessy (1979) adopt an alternative approach which involves noting those words which students have annotated in their textbooks. Xue's and Nation's (1984) *University Wordlist* is a fusion of the four previous lists.

Technological advances from the early 1960s onwards allowed the large scale examination of corpora (Leech, 1991). However, corpus-based studies of EAP lexis were rare until the 2000s. Coxhead's (2000) *New Academic Wordlist* (AWL) is the first wordlist to be compiled a using a machine-readable corpus. It is the attempt at developing an academic wordlist which has received the most attention over the last decade. This attention is presumably due in part to the AWL's pioneer status, and the belief in the viability of learning specialised vocabulary as a "shortcut" (Nation, 2013, p. 19) to understanding language in subject domains. The AWL is a list of 570 headwords compiled from a 3.5 million-word corpus of written academic text by examining the range and frequency of words outside the first 2,000 most frequently occurring words of English, as described by West's (1953) GSL. Coxhead (2000) makes a distinction between, general-service vocabulary, technical vocabulary, and academic vocabulary. The GSL, comprised of "the most widely useful 2000 word families" (Coxhead, 2000, p. 213), contains general-service vocabulary.

The AWL has been generally well received, however, several studies have highlighted shortcomings in Coxhead's (2000) approach. Firstly, the AWL's relationship with the GSL has been scrutinised. According to Schmitt (2010) some GSL items were collected as early as 1934 and thus its contents are in need of refreshing. Despite its age, some GSL items are still relevant to students of general English. This is particularly true of the first 1000 higher frequency items. After this point the usefulness of the items has been questioned (Nation & Kyongho, 1995). Perhaps, due to its age and the absence of computerised profiling tools in 1953, the GSL also suffers from other inconsistencies such as the inclusion of British but not American spellings, the absence of related word forms (i.e. the inclusion of *tour* and *tourist* but not *tourism*), pluralisation inconstancies, the inclusion of archaic terms, and the absence of newer technologies (e.g. *radio* is included but not *television*) among others. Some of these idiosyncrasies are likely to be reflected in the AWL. For Hancioğlu, Neufeld, and Eldridge (2008), the AWL and GSL combined "offer a package more or less suggesting that the basis for survival in an academic environment is knowledge of the 2000 word families of the GSL plus the 570 word families of the AWL" (p. 461). Coxhead's (2000) characterisation of vocabulary as either general-service, technical, or academic supports this suggestion. However, the length of the GSL coupled with the inconsistencies outlined above lead to the inclusion, in the AWL, of vocabulary which some would not consider academic in any respect. Hancioğlu and colleagues (2008) point to items such as *study* which occurs in the GSL but not in the AWL, and *drama* which occurs in the AWL but not in the GSL. From this they caution that the classification of vocabulary into mutually exclusive lists is motivated by convenience rather than pedagogical concerns.

Similar criticisms might be made of the practice of EAP in general. Firstly, many of the subject areas it covers are not in the strict sense academic but rather vocational. The Teaching of English in Academic Contexts (TEAC) might be a more appropriate label since it takes into account those subjects tackled at university which have not traditionally been considered academic. Secondly, it might be claimed that EAP cannot feasibly represent the language used in all possible subject areas. The latter concern has led to the creation of English for Specific Academic Purposes (ESAP) (Hyland, 2016b). In spite of

these concerns about its conceptual adequacy, in practice general EAP certainly exists. At this very moment, there are hundreds if not thousands of EAP courses taking place and an abundance of teaching and lexicographical materials aimed at such courses.

Like the notion of general EAP, the notion of a general academic-lexis present across all disciplines has been criticised. Intuitively, it seems unlikely that a history student would require the same vocabulary as a management or business studies student. Data from studies examining the percentage coverage of the AWL in corpora of texts dealing with different subjects would seem to support this view. For example, Cobb and Horst (2004) find coverage of around 14 percent for the history sub-corpus of the learned section of the Brown Corpus (Kučera & Francis, 1978), but only around seven per cent for the zoology and anatomy sub-corpora. However, since word distribution in natural language texts is non-random it is possible that comparing any texts would produce different coverage statistics. Adopting a slightly different approach employing range, frequency, and collocation measures, Hyland and Tse (2009) dispute the notion of a general academic vocabulary. They demonstrate that the great majority of the 570 AWL families have irregular distributions across the fields of engineering, sciences, and social sciences.

Through its inclusion of word families the AWL accounts for the relations between different word-forms in the same families. Nevertheless, Schmitt and Zimmerman (2002) caution that advanced learners of English are typically only able to make connections between some members of word families taken from the AWL. Concerns about the adequacy of word families as a basis for EAP vocabulary selection underlie the creation of more recent academic wordlists. The *New Academic Vocabulary List* (AVL) (D. Gardner & Davies, 2014) was determined using lemmas, rather than word families.

If finding a means of representing relationships between related word forms represents a challenge to discrete-item wordlists, the question of how to account for the relations of words and their syntagmatic context is more problematic still. This question is representative of a general trend in lexical studies from the study of words in isolation towards the examination of what could be termed multi-word-units (MWUs) or phrases. Wordlists are not immune from this trend. In her work with English learner corpora Paquot (2007) argues that although the AWL is a useful tool as regards learners' receptive skills, students' productive needs would be better served if a corpus-based approach were used to select which words and word sequences to teach.

MWUs realise many different purposes in language use. It is hardly surprising, then, that a variety of terminology abounds to describe these units (O'Keeffe, McCarthy, & Carter, 2007; B. J. Richards, Malvern, & Meara, 2009; N Schmitt, 2010; Wray, 2002). They might be more usefully classified by examining the ways they are identified or extracted in corpus-based studies. Schmitt (2010) highlights four approaches to identification of MWUs. The first concerns their identification in L1 acquisition studies and involves the measuring of repetitions and pauses in children's speech. Another approach is frequently described as form- or distribution-based. It involves identifying strings of texts of a given length which occur above a given frequency threshold. Biber, Johansson, Leech, Conrad, & Finegan (1999) term these sequences lexical bundles. The lexical bundle approach is, perhaps, the most widely employed means of phrase extraction in EAP studies. It has been productively employed in the creation of academic phrase lists (Hsu, 2014; Simpson-Vlach & Ellis, 2010). An obvious drawback of any approach to phrase extraction based on the form or frequency of an MWU is that meaning is essentially a secondary consideration. A third approach uses statistical strength of association measures to determine how strongly pairs

of words are associated. This is the basis on which the collocational networks in Geoffrey Williams's (1998) study of lexis in plant biology research articles are calculated. Finally, there is the phraseological approach which examines MWUs in terms of transparency and substitutability, that is to say, how they are constrained by semantic and pragmatic collocation restrictions. To date, there have been few studies of academic vocabulary which adopt this approach. Those that do usually only involve the analysis of a small number of corpus-based examples. For example, Hyland and Tse (2009) examine the semantic behaviour of *analyse* across academic disciplines. This paucity of studies may stem from the difficulty of operationalising the criteria of transparency and substitutability (Nesselhauf, 2005) and the reliance on human analysts which makes the process subjective and extremely labour intensive (Schmitt, 2010). It is worth noting that the distinction between these four approaches is not universally accepted; Granger and Paquot (2008) consider all these approaches phraseological. This view is shared in the present study where the terms 'phrase' and 'MWU' are preferred and employed fairly interchangeably. It is hoped that the precise type of unit being referred to is apparent from the context.

The lack of a consistent approach to MWUs is also reflected in dictionary design. This is particularly worrying given the important role phraseology has in native-like competence (Howarth, 1998; Nattinger & DeCarrico, 1992). Siepmann (2008) claims that when dealing with MWUs most monolingual learners' dictionaries focus on non-compositional expressions such as *carry the can,* at the expense of semantically orientated collocations. He hypothesises that there is a link between the inconsistent treatment of MWUs in semasiological learners' dictionaries and language teachers' and students' low awareness of them. He also points out that learners use fewer phrases more often, overuse certain phrases for key pragmatic functions, show preference for one-word markers, and tend to over-generalise with certain phrases. He attributes these errors to the discrete-word based methodology employed in the vocabulary sections of EFL/ESL textbooks which encourages the learning of "separate items which may become paired in rather haphazard fashion" (Siepmann 2008, p. 195). The treatment of MWUs in bilingual lexicography is similar. Pecman (2008) claims that bilingual phraseology has traditionally come second to its monolingual counterpart. While outlining an electronic French - English dictionary of routine formulae aimed at French academics, she calls for the creation of bilingual domain-specific phraseological dictionaries citing the domain-specific combinatorial properties of words, domain-specific style conventions, and the existence of related fields of knowledge that share the same phraseological features.

In light of this brief review of the literature the principal objective of the present study is:

- To examine differences in meaning and use of vocabulary between academic disciplines which might be obfuscated by the distributional approaches to vocabulary selection habitually employed in English vocabulary studies.

The following sub-objectives logically follow:

- To demonstrate a feasible means of phrase extraction which accounts for both semantic and syntactic concerns.
- To examine how the behaviour of academic lexis relates to wider phraseological approaches.
- To create guidelines for producing, and an illustrative example of, a useful lexicographical resource for the EAP community.

The key to achieving these objectives is a corpus-based experiment which adapts aspects of Corpus Pattern Analysis (CPA) (Hanks, 2004); a lexicographical technique for mapping meaning to text. The interpretation of the results of the experiment with reference to research on current approaches to vocabulary selection, phraseological theories of language, and research on EAP teaching and vocabulary resources will result in methodological, theoretical, and practical contributions to the field of EAP lexis studies. Figure 1.0 gives a schematic representation of the intended contribution of the present study.

Corpus-based experiment	+	Research on current approaches to vocabulary selection	=	Contribution to methodology
	+	Research on phraseological approaches to language	=	Contribution to theory
	+	Research on EAP teaching and vocabulary resources	=	Contribution to EAP pedagogy/lexicography

Figure 1.0. The intended contributions of the present study

This introduction has argued that the historical development of EAP is intertwined with a concern for the selection of lexis. Through a brief review of the literature, to be further supplemented below, it has argued in favour of an approach to the selection of EAP vocabulary which not only takes into account differences in word meaning across disciplines, but is truly phraseological taking into account both semantic and syntactic concerns. With this in mind the first chapter outlines approaches to linguistic theory which might be considered phraseological.

Table of Contents

List of Abbreviations

AILA - Association Internationale de Linguistique Appliquée
AJAC - Academic Journal Article Corpus
AJACX2 - Academic Journal Article Corpus Second Version
AVL - *The New Academic Vocabulary List* (D. Gardner & Davies, 2014)
AWL - A *New Academic Wordlist* (Coxhead, 2000)
BAWE - The British Academic Written English Corpus (see Nesi & Gardner, 2012)
BCG - Berkley Construction Grammar (see Fillmore, 1988)
BIC - The Bayesian Information Criterion
BNC - British National Corpus (Oxford Text Archive, 2015)
CAQDAS - Computer Assisted Qualitative Data Analysis
CG - Cognitive Grammar
CPA - Corpus Pattern Analysis (Hanks, 2004a, 2013)
CPASO - Corpus Pattern Analysis Shallow Ontology
CxG - Construction Grammar (Goldberg, 2006)
EAP - English for Academic Purposes
EAPWI - English for Academic Purposes Writing Instruction
ELFA - English as an Academic Lingua Franca / English as a Lingua Franca in Academic Contexts
ELT - English Language Teaching
ERPP - English for Research and Publication Purposes
ESAP - English For Specific Academic Purposes
ESP - English for Specific Purposes
EURALEX - European Association for Lexicography
FG - Functional Grammar (See Thompson, 1996)
GSL - M. P. West *A general service list of English words* (West, 1953)
IELTS - International English Language Testing System
JCR - Journal Citation Reports (Thomson Reuters, 2017)
LL - log-likelihood
LUG - Linear Unit Grammar (Sinclair & Mauranen, 2006)
MLD - Monlingual Learners' Dictionary
MWU - Multi-word-unit
PAW - Published Acadedmic Writing
PG - Pattern Grammar (Hunston & Francis, 2000)
PED – Pocket Electronic Dictionary
RA - Research Article
RCG - Radical Contruction Grammar (Croft, 2001, 2005)
SBMLD - Specialist Business Monolinugal Learners' Dictionary
SE - The Sketch Engine (Kilgarriff, Rychlý, Smrž, & Tugwell, 2004)
SJR - The Scimago Journal Rank (SCImago, 2007)
TEAC - The Teaching of English in Academic Contexts
TOEFL - Test of English as a Foreign Language
WSDiff - Word Sketch Difference

List of Dictionaries and Pedagogical Grammars

Al-Mawrid (Al-Baalbaki, 1990)
BBI Combinatory Dictionary (Benson, Benson, & Ilson, 1986)
CULD - *Chamber's Universal Learner's Dictionary* (Kirkpatrick, 1980)
Cambridge Advanced Learner's Dictionary (Walter, 2005)
CIDE - *Cambridge International Dictionary of English* (Proctor, 1995)
COBUILD1 - *Collins COBUILD English Language Dictionary* (Sinclair, 1987)
COBUILD2 - Collins COBUILD English dictionary (Sinclair, 1995)
Collins COBUILD Grammar Patterns 1 (G. Francis, Hunston, & Manning, 1996)
Collins COBUILD Grammar Patterns 2: Nouns and Adjectives (G. Francis, Hunston, & Manning, 1998)
Collins English Learner's Dictionary (Carver, Wallace & Cameron, 1974)
Concise Oxford Dictionary (Sykes, 1976)
Diccionario esencial inglés-español español-inglés Diáfora (Dutton, Harvey, & Walker, 1981)
Dictionary of Modern Written Arabic (Wehr & Cowan, 1976)
Dictionnaire du français langue étrangère (Dubois, 1983)
E-Advanced Learner's Dictionary of Verbs in Science (see G. Williams & Millon (2014) for a summary of construction process)
GEW - *A Grammar of English Words* (H. Palmer, 1938)
GSE - *A Grammar of Spoken English* (H. Palmer, 1924)
Harper Collins Spanish-English English-Spanish Dictionary (Butterfield, 1990)
ISED - *The Idiomatic and Syntactic English Dictionary* (Hornby, Gatenby, & Wakefield, 1942)
LASD1 - *The Longman Active Study Dictionary of English* (Summers, 1983)
LASD3 - *Longman Active Study Dictionary* (Summers, 1988)
Longman Business English Dictionary (Dignen, 2000)
Longman Dictionary of American English (Gray, 1983)
LDOCE1 - *Longman Dictionary of Contemporary English* (Proctor, 1978)
LDOCE2 (Summers, 1987)
LDOCE3 (Summers, 1995)
LDOCE5 (Mayor, 2009)
Longman Dictionary of English Idioms (T. Long, 1979)
MEDAL1 - *Macmillan English Dictionary for Advanced Learners* (Rundell, 2002)
MEDAL2 (Rundell, 2007)
NMED - *New Method English Dictionary* (West & Endicott, 1935)
MWAD - *Merriam-Webster's Advanced Learner's English Dictionary* (Perrault, 2008)
Megido Modern (Levenston & Sivan, 1968)
(O)ALD1 - *Advanced Learner's Dictionary of Current English* (Hornby, Gatenby, & Wakefield, 1948) (formally ISED)
(O)ALD2 - *Advanced Learner's Dictionary of Current English* (Hornby, Gatenby, & Wakefield, 1963)
OALD3 - *Oxford Advanced Learner's Dictionary of Current English* (Hornby, Gatenby, & Cowie, 1974)
OALD4 (Cowie, 1989)
OALD5 (Cowther, 1995)
OALD8 (Turnbull, 2010)
OALD9 (Deuter, 2015)
OALDCD - *Oxford Advanced Learner's English-Chinese Dictionary* (Wehmeier, 2004)
Oxford Business English Dictionary for Learners of English (Parkinson & Noble, 2005)
Oxford Collocations Dictionary for Students of English (Runcie, 2002)
Oxford Dictionary of Current Idiomatic English (Cowie, Mackin, & McCaig, 1983)

Oxford Dictionary of English Idioms (Cowie, Mackay, & McCaig, 1993)
OED - *Oxford English Dictionary (online)* ("OED online," 2017)
OLDAE - *Oxford Learner's Dictionary of Academic English* (D. Lea, Duncan, & Bull, 2014)
New Oxford Dictionary of English (Pearsall, 1998)
LEAD - *Louvain EAP Dictionary* (see Granger & Paquot, 2010)
PDEV- *Pattern Dictionary of English Verbs* (Hanks, 2001)
Selected English Collocations (Kozlowska & Dzierzanowska, 1987)
Le Petit Robert (Robert, Rey, & Rey-Debove, 1986)
WNWD - *Webster's New World Dictionary* (Gurulnik, 1976)

List of Figures

List of Tables

1. A PHRASEOLOGICAL APPROACH

In linguistics and other disciplines theorists have long recognised the importance of sequences of words in language. Wray (2002) describes how, in the mid-nineteenth century, neurologist John Hughlings Jackson noted that aphasic patients were able to utter, prayers, rhymes, and greetings, yet were unable to produce new sentences. Writing at the beginning of the twentieth century, Saussure (2013) describes how "a very familiar sequence of significant units [...] becomes a single unit" (p. 209) in the mind. Jespersen (1924) puts forward the existence of speech instinct to explain why speakers often use multi-word formulas which they could not feasibly remember if their words are treated as separate items. Bloomfield (1933) highlights the difficulty in delimiting the border between bound forms, words, and phrases. For Firth (1937) highlighting unusual collocations is essential in order to characterise communication in a speech community. Hornby (1954) employs word patterns in an attempt to provide learners of English with practical usage information. Similarly, Hymes (1962) describes recurrent patterns and routines becoming conventionalised for groups and cultures. Bolinger (1976) conceives of language as providing the speaker with an "incredibly large number of prefabs" (p. 1). Charles Fillmore (1979) argues that mastery of formulaic language is crucial to communicative competence.
Given this longstanding understanding the importance of sequences of words it is surprising that ESP and EAP material creators have persisted with discrete-item wordlists. A feasible explanation, in practical terms, is that prior to the widespread availability of machine readable corpora creating lists of phrases involved a prohibitive amount of time and effort. This could certainly have been true in 1920s and 1930s when a great deal of the studies which lay the groundwork for West's (1953) General Service List were undertaken (Gilner, 2011). However, from the 1990s onwards machine readable corpora have been widely accessible thus another explanation is needed. The predominance of generative approaches in linguistics since the 1960s, at least in the Anglo-American world and Western Europe, offers a more convincing explanation.

Chomsky's (1965) characterisation of the lexicon as a repository of irregularities of the language is typical of the subordinate role attributed to lexis in such frameworks. In very general terms, on generative approaches single lexical items are combined by syntax. In other words, lexical items are simply the raw material with which syntax builds longer units. This might explain the thinking behind providing students with a list of single words with which they can build more complex units. To borrow Hunston and Francis's (2000) expression, on this approach "phrases are normally seen as outside the normal organising principles of language" (p. 21).

In *Remarks on Nominalization* (1970) Chomsky recognises some limitations of generative approaches as far as morphological transformation is concerned. He suggests idiosyncratic nominalisation (destroy → *destroy-ation) might be better accounted for by "fairly idiosyncratic morphological rules" (Chomsky, 1970, p. 271) rather than syntactic or phonological operations. Later developments within the generative framework can be regarded as evidence of a general dissatisfaction with the marginalised position of lexis. These include Kaplan and Bresnan's (1982) Lexical-functional Grammar which was developed from Bresnan's earlier (1978) work in the transformational framework. Hanks (2013) argues that Chomsky's (1981) Projection Principle, which in basic terms contends that the syntactic structure of a sentence is projected bottom-up from the lexical properties of its predicator rather than top-down from an abstract sentence, is an attempt to repair a fundamental theoretical inadequacy. More recently, Jackendoff (2002) has criticised the syntactocentricism of generative approaches. He proposes a theory of thematic relations,

which is still a fundamentally generative approach, as semantic and phonological levels are linked by an independent syntactic level.

The hegemony of generative theories began to wane in the late 1970s. The following decades saw the emergence of various approaches to the description of language which could be loosely termed phraseological. This chapter provides a by-no-means exhaustive overview of some of these approaches and the occasional reflection on how they might address EAP lexis. It will also focus of the role of metaphor in some of these approaches. Understood here, in the same way as M. Black (1962) and Lakoff and Johnson (1980), as a means of interpreting one concept in terms of another, metaphor is of great significance to a phraseological approach to EAP vocabulary selection. Lakoff and Johnson's thesis that the function of metaphor is to allow the interpretation of abstract concepts in terms of everyday experiences befits the purpose of academic language: the diffusion of knowledge. The following caveats should be borne in mind. Firstly, the goal of this chapter is not to provide a principled taxonomy of phraseological approaches to language, but rather to outline some theoretical concepts which could be useful in an examination of EAP lexis. Secondly, theories are not formed in a vacuum; there is considerable crossover between approaches. For these reasons, the delimitation of the following approaches should be regarded as somewhat impressionistic.

1.1 Lexicographical Approaches

A great deal of research on phrases has been carried out by lexicographers. This is due to the need to resolve the practical problem of how to include phrases in the alphabetic macrostructure of a traditional dictionary. Traditionally, lexicographers have addressed this problem in terms of idioms which are non-compositional and cannot undergo any transformations such as passivisation or pluralisation. However, many lexicographers have pointed out the limitations of such an approach. Mel'čuk (1995) and Moon (1998) have demonstrated that the set of phrases goes far beyond idioms. Moon (1998) outlines three types of fixed expressions (including idioms): anomalous collocates such as *by and large* which violate normal rules of English, formulae which include, sayings, proverbs, and similes; and metaphors, be they transparent, semi-transparent, or opaque.

Mel'čuk, (1995, 1988, 2006, 2013) distinguishes between two different types of lexical units: single-word lexemes, and multi-word phrasemes. Phrasemes are non-free phrases fixed in particular ways and to particular degrees. In true idioms, meaning cannot be derived from the constituent words, in collocations or semiphrasemes the meaning of one word can be derived from the general lexicon, however, the meaning of the other word is collocation dependent. Taking the example *crack a joke,* from Hunston and Francis (2000), the meaning of *joke* is derivable from the lexicon but the meaning of *crack* depends on the particular collocation. In 'quasi-phrasemes' the meaning of both collocates is transparent however the phrase has an extended meaning. For example *start a family* could mean have one's first child (Hunston & Francis, 2000). Finally there are pragmatemes which are conventionalised phrases chosen, as the name suggests, for pragmatic reasons over other possible phrases in given circumstances. Mel'čuk (1995 p. 176) gives the example of food packaging where *best before...* is habitually chosen instead of *best consumed before.*

As a result of their practical objective of including phrases in the alphabetical index of a dictionary, many lexicographers make the practically motivated distinction between free and non-free phrases. This has the consequence of underplaying the central role phrases play in language. Apresjan's (2000) (translation of 1979 Russian original) study of English

and Russian synonyms provides an alternative account of collocations based on the concept of "a co-occurrence constraint, which, though not binding, is observed in pedantic and literary discourse" (Apresjan, 2000, p. 5). This approach is somewhat problematic in practice since delimiting pedantic and literary discourse is by no means a straightforward task. Apresjan does however, like many other approaches outside the lexicographical sphere, at least in theory, give phrases a more central role in language than many other lexicographical approaches.

1.2 Cognitive Approaches

Many cognitive approaches envisage a central role for the phrase in language. Unlike formalist approaches which conceive of language as a self-contained system independent "of the speakers who use it and the purposes for which they use it" (Taylor, 2002, p. 6), cognitive linguistics regards "language as embodied in the overall cognitive capacities of man" (Geeraerts & Cuyckens, 2010, p.4). In common with other human behaviour, on a cognitive approach language plausibly involves the cognitive capacities of categorisation, figure-ground organisation, the construal of mental images, the creation of metaphor, the recourse to conceptual archetypes, automatization, the storage and computation of linguistic expressions, social behaviour and, above all, symbolic behaviour.

Theories of grammar which could be deemed cognitive include Cognitive Grammar (CG) (Langacker, 1987, 1990, 1991, 1999, 2008); the approaches presently known as Berkley Construction Grammar (BCG) (Fillmore, 1985, 1988; Fillmore, Kay, & O'Connor, 1988; Michaelis, 1994; Michaelis & Lambrecht, 1996); Construction Grammar (CxG) as conceived by Lakoff (1987) and later developed by Goldberg (1995, 2006), and Radical Construction Grammar (RCG) (Croft, 2001, 2005). All these theories contain important conceptual differences, but in many respects they are broadly compatible. There are also differences in the terms they employ. The discussion below employs a generalised terminology.

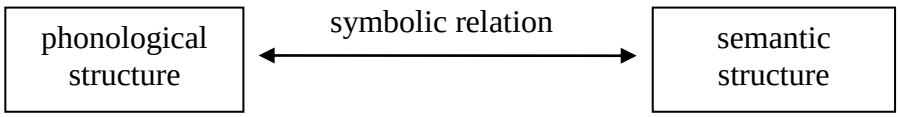

Figure 1.1. The organisation of language under the symbolic thesis

All these theories are, to varying extents, organised on the principles of the symbolic thesis. That is to say that they take that the fundamental role of language is to relate form and meaning. Under this thesis, any linguistic expression comprises three elements: First, as represented in Figure 1.1, an element pertaining to form which deals with syntactic properties, morphological properties, and phonological properties; second, an element which pertains to meaning which deals with semantic, pragmatic, and discourse-functional properties; and third a symbolic link which deals with the relation between the phonological and semantic units. Importantly, in contrast to generative grammar the relation between phonology and semantics is direct; there is no distinct level or syntactic module which mediates this relation. In general terms, under the symbolic thesis, there is no assumption of a one-to-one relationship between a single orthographic word and a unit of meaning.

This shared architecture conceals a number of important nuances. Chief amongst these is the status of syntactic elements. In BCG constructions are complex units derived from primitive atomic units with syntactic feature types such as [gf¬subj], which represents a unit which has a grammatical feature which is not a subject, or [cat v], which represents the syntactic category verb. In CxG, Goldberg (1995) takes a similar approach to the analysis of syntactic roles and relations in argument structure, employing, primitive grammatical relations such as Subject and Object and syntactic categories such as Verb. However, in contrast to BCG, for Goldberg (1995) an event's argument structure is derived from the event itself. Thus *rob/steal* has the participant roles *robber* and *victim*. The approach to syntactic relations posited by BCG and CxG has consequences for the types of units which may be considered constructions. In short, in order to be considered a construction in Goldberg (1995) an association of form and meaning must not be fully predictable from other form-meaning associations:

> C is a CONSTRUCTION iff $_{def}$ C is is a form-meaning pair $\langle F_i, S_i \rangle$ such that some aspect of F_i or some aspect of S_i is not strictly predictable from C's component parts or from other previously established constructions. (Goldberg 1995, p. 4)

In later work, Goldberg (2006) discards this requirement for non-compositionality on condition that a construction occurs with sufficient frequency to become entrenched. This modification brings CxG closer to the CG paradigm. Taylor (2002) and Langacker (2005) advocate a less restrictive conception of construction whereby any linguistic structure analysable into component parts is considered a construction. For Langacker (2005) the recourse, in Construction Grammar (CxG), to an autonomous syntax is one of the principle differences between the two approaches. Goldberg (2006) argues that this criticism is based on a misconception. She argues that labels such as "Sbj", "Obj", "N" and "V" in Goldberg (1995) do not refer to "irreducible grammatical primitives without corresponding meanings or function" (Goldberg, 2006, p. 221), but rather facilitate the relevant level of description to describe the form of particular construction in recognisable terms. However, she does take issue with the explicitly reductionist nature of CG. Taking up Langacker's metaphor "Grammar exists and needs to be described as such, like water (a particular configuration of oxygen and hydrogen atoms), it is however reducible to something more fundamental (configurations of semantic structures, phonological structures, and symbolic links)" (Langacker, 2005, p. 105), she questions the need for such a reduction: "water is clearly reducible to hydrogen and oxygen; however, no reductionist account of water is going to explain why water is wet, nor why it is used the way it is: to bath in; to drink etc." (Goldberg, 2006, p. 220). Similar debates about the degree of delicacy necessary to analyse phrases abound in most approaches which take a phraseological view of language.

In CG syntactic categories are reductionist in the sense that they are related to the construal of mental images. Thus, the category "Noun" represents the construal of a non-relational, a-temporal thing. A noun is this sense is basic in that it can be conceptualised without referring to another entity. A verb such as *runs* construes an entity as an event which presumes the existence of participants for that event. A different approach is taken in RCG the complex unit is regarded as a basic unit in which parts are defined in terms of their roles by categorisation in terms of formal features such as word order, patterns of contiguity, and specific morphemes or classes of morphemes in particular roles.

In addition to the status of syntactic elements, in the different cognitive theories of grammar there are differences in the relations between these elements. In CG such relations are conceived of in terms of predicate-argument relations which, in keeping with the symbolic thesis, are both syntactic and semantic. In CxG, Lakoff (1987) represents constructions which follow patterns of syntactic elements, lexical elements, syntactic conditions and phonological conditions. This allows for relations between individual elements and relations between whole constructions. Like Lakoff and Goldberg, Langacker (1987) regards valence as symbolic, however, a key difference between CxG and CG is that for the latter valence is also gradient.

In CG when an argument fills a role of a predicate it is said to elaborate the relevant substructure. As outlined above, *runs* presumes the existence of a runner. Thus the semantic structure for *runs* includes a schematic runner as a substructure. Valency is gradient as "One structure, D, is dependent on another, A to the extent to which A constitutes an elaboration of a salient substructure within D" (Langacker, 1987, p. 300). With simple expressions such as *John runs* this is straightforward. However, adapting an example from Croft (2007), more complex sentences such as *John runs on the track* present a greater challenge. Here, *on the track* is dependent on *run* because *run* elaborates the highly salient figure role of the locative relation *on the track*. On the other hand, *run* is autonomous relative to *on the track* because *on the track* elaborates only the weakly salient substructure of the running event. In CG any pair of conceptual structures is subject to this autonomy/dependence relation. With this in mind it is reasonable to conclude that on a CG approach all language is to some extent phraseological.

As its name suggests RCG posits a radically different approach to syntactic relations in constructions. It defines relations between the elements of a construction in semantic terms. Citing cross-linguistic evidence of mismatches between syntactic relations and semantic content, Croft (2001) rejects syntactic relations altogether. He instead posits the idea that morphosyntactic characteristics, which express syntactic relations in other theories, link components in the phonological pole to components in the semantic pole.

Cognitive theories generally organise their constructions in hierarchical taxonomic networks. There are, however, some key differences between in the manner in which they achieve this. BCG and CxG allow for meronomic as well as taxonomic relations between constructions. That is to say, that a part of a construction may itself be considered another construction. In addition, Goldberg (1995) posits a third type of link which deals with polysemy. For, her constructions that are identical syntactically but different in semantics are polysemic. One construction is central and another is an extension. This analysis also applies to metaphor as illustrated in section 1.2.1 below. Likewise, in CG there are three kinds of relation between units which are fundamental to an understanding of language: First: the 'vertical' relation between a schema, a less thoroughly specified unit, and an instance, a more thoroughly specified unit; Second, the 'horizontal' relation between parts and a whole; and third, a relation of similarity. The RCG approach is slightly different. There are not meronomic links between constructions, instead taxonomic links between different parts of different constructions.

In addition to their organisation in taxonomic hierarchical networks, another common characteristic which many cognitive theories of language share is their characterisation of constructions and the lexicon as opposite ends of same continuum. At one end of the continuum, there are complex constructions with many elements. At the other end, there are atomic, single element constructions. The patterns for the combination of such

elements also have the status of symbolic units and these occur on a continuum of schematicity. This absence of a discrete lexicon raises the question of how information enters and is stored in their hierarchical taxonomic network. Goldberg (1995) characterises two principal models for the representation of information; complete inheritance models and full-entry models. In BCG a complete inheritance model is employed information is represented non-redundantly at the highest and most schematic level its features are then inherited by its daughters. CxG, in contrast, Goldberg (1995) allows for a full-entry model where information is specified at all levels of the taxonomic hierarchy, however, she also allows for an important distinction positing the normal-inheritance model in which inheritance can be blocked if it conflicts with information in a specific instance. As Croft (2007) explains "Normal inheritance is a method for accommodating the fact that much of what we know about a category is not true of every instance of a category." (p. 488). Both Lakoff (1987) and Goldberg (1995) have recourse to normal inheritance.

A central tenet of frame semantics (BCG) is that understanding the meaning of a word involves accessing all related essential knowledge. Fillmore's oft cited (1976) example involves the word *sell*. In order to understand *sell* the frame for commercial transaction is needed: One person (the seller) has goods and wants to exchange them for money. Person two (the buyer) has the money and wants to exchange it for the goods. There is a relation of mutual exchange of goods for money (price). In a corpus-based examination of the collocational behaviour of *risk,* Fillmore and Atkins (1992) examine the viability of a frame-based dictionary which represents "word senses, relationships between the senses of polysemous words, and relationships between semantically related words" (p. 75) linked by frames. FrameNet (Fillmore, Johnson, & Petruck, 2003) is a practical lexicographic implementation of frame semantics its aim is "to document the range of semantic and syntactic combinatory possibilities (valencies) of each word in each of its senses" (Ruppenhofer et al., 2016, p. 7).

The situation in CG and RCG is quite different. Both approaches employ the usage-based model. This model, originally developed in morphology (see Bybee, 2001), maintains that language use, particularly frequency of use and the similarity of form and means determine the representation of grammatical knowledge in the mind of the speaker. Usage is of fundamental importance in CG; structures, regardless of their internal complexity, become entrenched through frequent successful use and gain unit status. These units are effectively dealt with by the speaker as integrated wholes. As for the role of similarity of form, novel expressions which lack unit status are meaningful if they are sufficiently similar to existing expressions with unit status. They can either be instances of more schematic units or undergo a process of assimilation through a relation of similarity to an established unit.

On the usage-based approach, individual linguistic experience is crucial in determining which multi-word-expressions count as units. Taylor (2002) describes how a person encountering the expression *tree hugger*, might initially attempt to interpret it along the lines of the general schema for similar expressions, for example, *dog lover* meaning someone who loves dogs, and *meat eater* would someone who eats meat. At first, the incongruence of such an interpretation would prove problematic. However, with repeated exposure to *tree hugger* the speaker might realise that the expression has a conventionalized semantic value. Data from a pilot-study carried out in preparation for the present study, provides an example of this phenomena in EAP. In a corpus of general English the expression *the document specifies that* predominates. It is unlikely that a person would have difficulty interpreting it as meaning something akin to *the document states*. However in a corpus of academic journal articles dealing with microbiology sentences such as *the mitochondria specify DNA*

predominate. Such sentences might prove problematic to students from other disciplines or those as yet unfamiliar with the technical term in microbiology. However, with repeated exposure they would come to realise that it has a meaning akin to *the mitochondria gives instructions of how the DNA should be assembled*. An important caveat here is that CG is primarily concerned with the way in which speakers conceptualise the word, language internal analysis such as this "must be regarded as symptomatic of meaning, not as meaning itself." (Taylor, 2002, p. 192). Taylor, nonetheless concedes, that, for natural language processing applications, the computational-statistical approach turns out to be compatible in principle with the cognitive grammar approach to issues of word meaning and disambiguation.

A taxonomic, hierarchical, usage-based account of language also allows a plausible explanation of the often polysemic relation between specialist and non-specialist vocabulary, a central concern of the present study. In linguistic categorical-hierarchies the basic level category is the one most people would choose when asked to designate an object. In general language contexts, such categories tend to be of high frequency of occurrence, while more schematic terms above the basic level are often more scientific (Taylor, 2002). Conversely, it seems reasonable to assume that less schematic categories below the basic level would be of lower frequency of occurrence and would be more technical. Crucially, "the level at which things are named varies according to a person's interests and purposes" (Taylor, 2002, p. 131).

In summary, the cognitive approaches outlined thus far offer a theory of language which is phraseological and usage based. The concept of entrenchment is particularly important as it provides a plausible account of how new phrases become increasingly fixed in the lexicon with increasing use. The idea that novel expressions are interpreted through analogy or similarity to established forms is also key. The idea of a taxonomic hierarchy of vocabulary offers an explanation for the often polysemic relationship between specialist and non-specialist vocabulary.

1.2.1 Cognitive Approaches to Metaphor

Metaphor has played a central role in the development of cognitive linguistics. There are a number of approaches to metaphor. The Lakovian theory of conceptual metaphors is perhaps the most well-known. On this approach, a more abstract domain is construed in concrete terms. For instance, the expression *we arrived at the conclusion*–which one might expect to find in academic writing–has the target domain of rational thought and the source domain of travelling. *Traveller* in the source domain maps to *thinker* in the target domain likewise *point of arrival* maps to *conclusion*. One of the conceptual metaphors which underlie this expression is A CONCLUSION IS A DESTINATION which in turn is an instance of the even more schematic underlying conceptual metaphors of IDEAS ARE LOCATIONS and THINKING IS MOVING. In a CG framework conceptual metaphors can be conceived of as schemas. The schema representing A CONCLUSION IS A DESTINATION is particularly entrenched since it is instantiated by a large number of different instances for example: *reach the conclusion, move towards the conclusion, come to the conclusion* etc.

The general direction of mapping from concrete source to abstract target is, for Lakoff and Johnson (1980, 1999), evidence that abstract domains must always be conceptualised through metaphor. There are a number of objections to this conclusion. Taylor (2002) cites the need for a prior conceptualization of the target domain if the hearer is to know which elements of the source domain map to which elements of the target domain. He also cites

the possibility that the use of an expression in one or more domains does not necessarily indicate metaphoricity but a phenomenon more akin to synonymy, where both expressions are instances of the same schematic meaning. The lexicalisation or conventionalization of metaphorical expressions over time is also problematic. It is difficult to conceive of widely used conventional expressions in terms of metaphorical mapping. In fact the expression *come to a conclusion* is so conventionalised that it is doubtful that an unprompted speaker would recognise the traveller metaphor (Taylor, 2002).

For Lakoff and Johnson conceptual metaphor, the construal of one domain in terms of another, is fundamental to language. The examples such as: *I went from the hotel to the airport, the inheritance went from George to Philip, the lights went from red to green,* are motivated by the conceptual metaphors STATES ARE LOCATIONS and a CHANGE OF STATE IS A CHANGE OF LOCATION. However, Langacker (1986) rejects the notion that change of state expressions are based on metaphor positing instead a schematic notion of change which can be instantiated in number of different domains.

There are a number of other approaches to these phenomena which are also compatible with cognitive linguistics. An alternative analysis of cross-domain similarities involves image schemas. An image schema is "a recurring, dynamic pattern of our perceptual interactions and motor programs that gives coherence to our experience" (Johnson, 1987, p. xiv). For example, Gibbs and Colston (1995) provide an account of how the physical act of maintaining balance results in an image schema which is extended to other domains such as psychological judgment (a balanced personality) or judgments (a balanced opinion) amongst others.

Talmy's (1988) system of force dynamics also allows a plausible account. Force dynamics concerns the interaction of entities in dynamic situations. The most straightforward of these situations involve two entities the Agonist, naturally disposed to rest, and the Antagonist who is able to exert a force on the Agonist which may overcome this disposition to rest.

1. The ball kept rolling because of the wind blowing on it.
(The wind overcomes the ball's natural disposition towards rest)
2. The ball's hitting it made the lamp topple from the table.
(The ball comes to impinge on the lamp overcoming its natural disposition towards rest)
3. They kept me waiting
4. I kept myself from responding in kind
5. That can't be true

The force-dynamic interactions are obvious in examples (1) and (2) which relate to concrete entities. However, the remaining examples illustrate how force dynamic notions might apply to other domains. For example, the social domain, in (3) the agonist counteracts the antagonist's natural inclination to wait; the psychological domain, in (4) where force-dynamic conflict between two aspects of a person's psyche; the epistemic domain, in (5) where the force of reason overcomes a fact's appearance of being true. On this reading, force-dynamic interactions can plausibly be regarded as unifying schema for patterns of interactions which can be instantiated in different domains.

Glucksberg and Keysar (1990) argue that *a is b* metaphors are not, as traditionally believed, similes, but rather "class-inclusion assertions" (p. 3). Following this line of argument, the

expression "my job is a jail" sets up a new super ordinate category containing jail and certain jobs schematic for "situations that are unpleasant confining, difficult to escape from, unrewarding and so on." (Glucksberg, Keysar, & McGlone, 1992, p. 578), jails and certain jobs are instances of this category.

Fauconnier and Turner's (1998) theory of conceptual blending follows a similar logic. Taking again "my job is a jail" as an example, a generic space involving any unpleasant confining, difficult to escape from, unrewarding situation is created. The input spaces are prison and job. The resulting blend contains element of both input domains and acquires a new logic that is not present in either of them. For example, following Taylor (2002) if someone is imprisoned in their job they have the option of escaping, but in the blend job and jail, the manner of escape is quite different from the manner of escape from a literal jail.

All of the above provide plausible accounts for a function which is particularly important in academic language: The expression of abstract concepts in concrete terms.

1.3 Functional Approaches

Like the approaches examined above, Functional Grammar (FG) has its genesis in frustrations with generative approaches. For proponents of FG these frustrations centre on the inability of generative grammar, with its emphasis on propositional meaning, to account for pragmatic uses of language and the influence of context beyond the immediate sentence. In order to address these concerns Haliday (1994), among other proponents of FG, proposes that analysis of language take place in terms of three metafunctions: Interactional, experiential, and textual. These metafunctions correspond to three components of the grammar each with its own systems of choices (Figure 1.2). The meaning of message is the result of the choices in all the relevant systems of the three components.

A functional analysis of language necessitates going beyond a view of words as discrete units of meaning. Halliday (1994) accounts for such a view with groups which are roughly analogous to phrases in generative grammar. These groups are the constituent parts of clauses. A clause can be viewed as having a series of functional slots in the interpersonal component for example: Subject, Predicator, Object, and Adjunct, which can be filled with groups with certain structural qualities. In addition to the three components, the Rank Scale (Halliday, 1994) establishes the relation between the different levels of analysis in FG. Clauses are made up of one or more groups, which in turn comprise one or more words, which in turn consist of one or more morphemes. While, at first sight, the relationship between clause and groups may seem little different from the recursive structure in generative approaches, it is the labelling of these groups according their functional rather than structural role which points to the existence of lexical units of meaning which go beyond the single word.

These structures are linked to data in terms of exponence relations which relate the abstract categories of the theory to the data. In many other approaches this is called realization. Words are exponents of grammatical level units, which are similarly semantic level categories. Theoretical classes of the same category are also linked by exponence relationships for example DETERMINER is an exponent of the more abstract WORD. Another Hallidean notion which hints at a central and organising role for phrases in

language is that of the scale delicacy, in other words the level of detail and particularity an exponent represents**.** This allows the analysis of embedded phrases.

Type of structure	Who	's	taken	her calculator?
Experiential →	Actor		Process	Goal
Interpersonal →	Subject	Finite	Predicator	Complement
Textual →	Theme		Rheme	

Figure. 1.2. The three kinds of structure in the clause reproduced from Thompson (1996, p. 32)

A description of language based on metafunctions also provides an account of language variation across register and genre. Halliday and Hasan (1989) define register as "variation according to use" (p. 41). They cite three main dimensions of variation which characterise a register. Each of these dimensions corresponds to a metafunction. The field describes what is being talked about and corresponds to the experiential metafunction. The tenor relates to how the language is used in interaction. The mode refers to whether language is being written or spoken and corresponds to the textual metafunction. Genre is the relation of register to purpose, it relates to what speakers and writers do with language and how they organise language to achieve that aim (Thompson, 1996).

Halliday (1994) has argued against regarding language only in terms of discrete constituents each with a clearly identifiable meaning. There is evidence to support this argument from each of the three metafunctions of language. An examination of modality demonstrates how interpersonal meaning does not involve a one-to-one link to a constituent but is instead extended over the whole clause (Thompson, 1996). Take for example the following sentences from (Thompson, 1996, p. 56):

> He has said nothing to me about that.
> He hasn't said anything to me about that.

In the first sentence it is the complement *nothing* which expresses negative polarity, while in the second sentence this is expressed by the finite *hadn't* with the mood adjunct *anything*. In an examination of hedging in academic research articles Hyland (1998, p. 135) provides examples of a similar phenomenon related to the textual metafunction.

An examination of evaluation provides further evidence of the unstable relationship between single words and meaning, as well as the primordial role of lexis in this function. Thompson (1996) claims "much of evaluation is expressed by lexical choices and there are few grammatical structures which can be seen as having evolved with a primarily evaluative function" (p. 65). This is particularly apparent in the case of swearwords and obscenities which according to Halliday (1979) "may occur at any or all points in the clause; it does not matter what segments they are attached to – many writers have noted that such elements readily occur even in the middle of word" (p. 66)

Analyses of the experiential metafunction involve labelling clause constituents as processes, participants, and circumstances. Processes can be categorised according to their transitivity Table 1.1 gives an overview of process type, core meaning, and corresponding participants.

Table. 1.1
Overview of process types reproduced from Thompson (1996, p. 102)

Process type	Core meaning	Participants
material	'doing', 'happening'	Actor, Goal
mental:	'sensing'	Senser, Phenomenon
perception	'perceiving'	
cognition	'thinking'	
affection	'feeling'	
relational:	'being'	
attributive	'attributing'	Carrier, Attribute
identifying	'identifying'	Identified, Identifier/Value, Token
verbal	'saying'	Sayer, Receiver, Verbiage, Target
behavioural	'behaving'	Behaver
existential	'existing'	Existent

The participant role Range, an "element that specifies the range or scope of a process" (Halliday, 1994, p. 146), provides slightly more speculative evidence of the unstable relationship between single words and meanings. In essence, range works together with a verb to express process when the verb itself 'lacks' something of the meaning. In many cases, such as in the first sentence below, this is a similar phenomenon to phrasal verbs. In this sentence *birth* is labelled as Range and works together with the general verb *give* to complete the material process give birth. In the second sentence *the lake* works together with the verb *cross* to complete the process of movement.

> She gave *birth* to a baby boy.
> The crossed *the lake* by boat.

As implied by its name, the textual metafunction of language deals with clauses in the wider context of a text as a whole. As such, it follows that analyses of this function do not assume that individual words are discrete units of meaning later arranged by syntax. The textual metafunction consists of two principal units the Theme and the Rheme. The Theme, for Halliday (1994) "the starting point of the message" or "the ground from which the clause is taking off" (Halliday, 1994 p. 38) and for P. Fries (1995) a framework for the interpretation of what follows. Rheme is a somewhat controversial concept (see U. Fries, 1984). Nonetheless, an analysis of Theme elucidates a number of ways in which a speaker can group together one element of the message as a single constituent. Halliday (1994) terms the first of these Thematic Equatives, traditionally they are also known as pseudo-clefts.

> [What happened] was that Benjamin Lee Whorf picked up Boa's example and used it.
> [Benjamin Lee Whorf] picked up Boa's example and used it.

The pair of sentences above taken from (Thompson, 1996, p. 126) could be said to have the same meaning in so much as they both contain all the components of same message. However, the second sentence does not use any words of the theme. Thompson claims that this demonstrates that almost any combination of meaning components can be grouped to form a theme.

While traditional generative approaches offer accounts for these operations and others like them, the differentiation made between deep structure and surface structure suggests that they may be regarded as afterthoughts. FG makes no such distinction and such operations are regarded as a central part of the language.

1.3.1 Metaphor in Functional Grammar

Metaphor is central to functional grammar. Halliday (1994) rejects the traditional distinction between metaphorical expressions and literal expressions. Instead, making a distinction between mental and external representations, he prefers the term 'congruent' which Thompson (1996) informally defines as "closer to the state of affairs in the external world" (p. 165). On an FG approach metaphor is extended to grammar because 'things' are generally congruently encoded by nouns while 'happenings' are generally congruently encoded by verbs. However, for metaphorical encodings this is not the case. The logical outcome of this is the concept of grammatical metaphor: "The expression of meaning through a lexico-grammatical form which originally evolved to describe a different kind of meaning" (Thompson, 1996, p.165). Importantly, and in a similar way to CG as conceived by Langacker (1987), metaphor is present in all language. It is a characteristic which gives language its flexibility; all language is congruent or metaphoric to some degree. In practice it is not possible to separate meaning and expression in absolute terms while a congruent wording of a metaphorical expression might refer to the same state of affairs the writer or speaker has chosen the metaphorical reading for a reason, to simply regard the metaphorical expression as an equivalent of its congruent representation is to obscure much of the meaning the speaker wanted to convey. Halliday and Martin (1993) demonstrate how an analysis of grammatical metaphor in scientific writing calls into question the epistemic basis of its supposed objective stance.

1.4 Corpus Approaches

Another group of approaches to language description which could be termed phraseological hail from corpus linguistics. It is worth pointing out, that there is some debate surrounding the epistemological status of corpus linguistics. Some consider it a theory, for instance Teubert (2005) and Tognini-Bonelli (2001), while for others, including McEnery, Xiao and Tono (2006), it is primarily a methodology.

Irrespective of this debate, corpus linguistics has some central tenets which have implications for the way in which we conceive of language. Foremost amongst these is that, unlike cognitive approaches, for many practitioners corpus linguistics is not chiefly concerned with the mental representation of words, traditionally it is more concerned with text; many corpus linguists reject the Saussurian distinction between *langue* and *parole* (Carter, 2004). Instead corpus linguistics often takes a distributional approach in which frequency plays a fundamental role. The outcome is a view of language which could be considered phraseological. For example, in practical application of a corpus methodology, Biber et al. (1999) extract frequently occurring 'lexical bundles', that is, frequently occurring strings of words.

Dealing with more theoretical concerns, Sinclair (1991) posits the existence of the Idiom Principle, a tendency for the co-selection of words which gives rise to collocations and idiomaticity. Its counterpart the Open-choice Principle describes the tendency of words to occur freely the only constraint on their occurrence being grammaticalness. In practice,

most usage lies between these two principles. Sinclair's (1996b, 1998) attempts to address the theoretical question of the nature of the word as a unit of meaning, are the logical extension of such principles. In essence, using corpus techniques Sinclair demonstrates that units of meaning do not map to single orthographic words but rather phrases which he terms lexical items. Sinclair (1998) outlines a model of lexical description which accounts for this phenomenon. The model involves five categories of co-selection for words in a lexical item. The first two categories; the core, and the semantic prosody are obligatory. The remaining three categories; collocation, colligation, and semantic preference are optional. Sinclair demonstrates the model with reference to the verb *budge* which represents the core. In an examination of corpus lines he demonstrates how the prototypical dictionary definition of *budge*: "to (cause to) move a little" (LDOCE in Sinclair, 1998 reproduced in Sinclair, 2004b, p. 142) is not adequate. In the majority of instances in the corpus *budge* exhibits negative semantic prosody demonstrating colligation with negative *not* or *n't* in other cases there is colligation with *to* which in turn collocates with forms of the lemma *refuse*. Similarly, where other collocates occur there is a strong semantic preference for refusal.

A similar approach could be adopted with EAP vocabulary. In the data from a study of corpus based collocational behaviour of academic words (Rees, 2013; 2016). The verb *stress* exhibits negative semantic prosody in texts from a corpus of management studies articles, while in a corpus of microbiology articles it is neutral.

Concerns about the nature of the word as a unit of meaning also raise questions about the epistemic status of language. Sinclair (1996) argues that single-word term-bank style lexicons support the idea of language as a carrier of message. In contrast, there is a vision of language where form and message are conceived of as inseparable interactions between words and other sections of texts which implies that words cannot be regarded as asynchronous stable units of meaning. To account for this view, Sinclair posits the existence of the empty lexicon. Here, the lexicon is empty to the extent that words have no a priori meaning, instead, meaning is constructed over time through usage. Usage data is obtained through the study of corpora in which "meanings may arise from the loose and varying co-occurrences of several words, not necessarily next to each other" (Sinclair, 1996, reproduced in Sinclair, 2004c, p. 160) as a consequence entries cannot be specified in advance because of the vast number of possible relations between words in their formation of lexical items.

Hoey (2005) also posits a usage-based account of the lexicon. Inspired by the psycholinguistic notion of priming he argues that collocation can only be accounted for by assuming that every word is mentally primed for collocational use. His theory of lexical priming is a usage-based theory in which as words are encountered in speech or text, they are acquired in conjunction with the context and co-text in which they occur. In this way a speaker's knowledge of a word includes its collocational behaviour in a given context.

G. Williams's (2008) notion of collocational resonance is related to lexical priming. Collocational resonance is based on the idea that users transfer semantic and etymological aspects of previously-encountered language, although they may not always be conscious that they are doing so. This priming effect can occur diachronically and synchronically. It can be traced by deriving collocational networks from corpora. This involves using statistical criteria to automatically extract the salient collocates of a given node word. The branches of this network are in turn linked to other node words. By comparing the semantic characteristics of the collocates extracted for a given node, G. Williams (2008)

demonstrates how phraseology from the King James Bible resonates with Shakespeare. Similarly, G. Williams and Millon (2014) demonstrate how, through an examination of the collocational networks of *probe* in general and specialised corpora, semantic transfers between general and specific domains can be revealed. Somewhat confusingly, Hanks (2006, 2013) uses the term resonance, in a similar yet distinct way, to refer to semantic transfer from the literal to metaphorical.

Sinclair's (2001) Lexical Grammar is essentially a call to reconsider traditional pre-corpus grammar categories and create a corpus based grammar which is based on corpus evidence instead of a priori assumptions about parts of speech. Sinclair and Mauranen's (2006) Linear Unit Grammar (LUG) is perhaps the most complete example of such a grammar.

Inspired by Brazil's (1995) work on phonology, LUG primarily deals with corpora of transcribed spoken language but it can also be applied to the written mode. Sinclair and Mauranen postulate the existence of 'natural units': so-called chunks of language which have psycholinguistic salience, in so much as, humans "acquire the ability to see a complex multi-layered sentence as a string of chunks, discernible in text" (Sinclair and Mauranen, 2006 p. 130). Their descriptive model involves separating text into chunks. These chunks are then classified as different types of elements of which there are two basic types. M elements deal with what is being talked about while O elements are concerned with organizing discourse. O elements can be further classified into OI elements which organize interactivity between speakers and OT elements which deal with textual coherence. M elements can be further classified into several types according to their role in the development of topic matter in a conversation. Straightforward M elements are complete units which require no additional element to be complete. Fragmented message elements such as hesitation markers, e.g. *umm*, and repetitions, *the the the*, are labelled MF. M- elements are those which, unlike MFs, clearly contribute to the shared knowledge while raising the impression that they will be completed by another interlocutor. They complement are +M units which supply the appropriate completion material. MR elements represent reformulations of previous message elements while MA elements are 'bumps' in a speakers turn unlike MF elements they are not abandoned completely but instead replaced with new starts. The final stage of the procedure involves the recombination of elements to create an output conventional enough for further grammatical analysis.

The recombination process involves eight operations. The first involves the removal the OI elements. This operation is repeated for MF elements. Then, those MA elements which refer to something fragmentary or are removed while incomplete MAs are then joined with +Ms. Similarly, M- and +M are also joined. Next MS units are added to the nearest previous M to leave long stretches of M elements interspersed with OT elements signifying the relation between them. The next task involves merging MR elements with the M elements of which they are reformulations. The next in the recombination process is a final adjustment process aimed at checking whether the removed M and O elements contained information important to understanding. In the final product spoken text is adjusted to written norms.

There are a number of limitations to the scope of LUG; unlike cognitive and construction grammars it makes no claims to comprehensively describe all grammatical knowledge. Instead its authors see it as a theoretically valid and principled ante-step towards the obtaining of conventionalised data from inaccessible modes of language, i.e. speech, which may then be examined using other approaches.

Hunston and Francis's (2000) Pattern Grammar (PG) is perhaps the most well-known corpus-driven approach to the lexical grammar of English. Unlike LUG, it relies on pre-corpus grammatical categories and has considerably more descriptive power. PG was developed during the creation of second edition of the *Collins Cobuild English Dictionary* (COBUILD2; Sinclair, 1995) and the associated *Grammar Pattern* Series (G. Francis, Hunston, & Manning, 1996, 1998) A pattern is defined by Hunston and Francis (2000) as:

> A phraseology frequently associated with (a sense of) a word, particularly in terms of the prepositions, groups, and clauses that follow a word. Patterns and lexis are mutually dependent, in that each pattern occurs with a restricted set of lexical items, and each lexical item occurs with a restricted set of patterns. (p. 3)

Unlike LUG, PG does employ pre-corpus categories; however Hunston and Francis (2000) reject structural labels such as object, complement and adjunct in favour of "superficial word class labels" (p. 45) such as v, n, adj, adv, that (defining that clauses), ing (ing forms), to –inf (infinitives), and wh (wh- and how questions). For the authors, the labels are superficial insomuch as word class itself is best identified on the basis of the behaviour of words in patterns. Their approach involves searching for a node word in a corpus then using these labels to annotate corpus lines found for each node word. The result of this process is a series of patterns associated with a given node word.

The corollary is that a pattern can be associated with a number of different words which often display similar meanings. Words which are associated with patterns can be arranged in 'notional groups' according to their meaning. In G. Francis et al. (1996) the category of logical relations, which in Halliday's (1994) terms "set up relations between two separate entities" (p. 199) be they concrete nouns or nouns which themselves realise complex processes and ideas, forms one such notional group. Halliday, explains this type of complex realisation in terms of grammatical metaphor, specifically "ideational metaphor" which, for him, is typical of academic writing. Hunston and Francis also envisage a role for notional groups in ELT.

Further information about meaning in particular notional groups can be gleaned by mapping participant roles, such as those employed by Halliday (1994), on to patterns. The configuration of participant roles is dependent on both pattern and lexical item. Similarly, Hunston and Francis (2000) argue that, to a certain extent their research supports the general assertions of valency grammarians such as Levin (1993) to the extent that it is possible to classify verbs in groups according to their syntactic patterns and general meaning. However, they argue that using corpus derived data rather than data which is the result of introspection brings to light the limitations of Levin's conclusions.

In contrast to Levin (1993), Hunston and Francis (2000) do not regard association of pattern and meaning as completely deterministic. Moreover, in contrast to Sinclair (1991), they do not conceive of a one-to-one correspondence between pattern and meaning. In an examination of the literature they identify three theoretical positions regarding this question. The 'weak' position that holds that the association between words exhibiting a certain pattern and meaning is not entirely random; Owen (1993) might be regarded as a proponent of this position. Hunston and Francis (2000) situate their view as an intermediate position which suggest that: "given a list of words occurring with a particular pattern, the majority will be divisible by most observers into reasonably coherent meaning groups" (p. 86). The strong view, consistent with that of Sinclair, posits a one-to-one relationship between meaning and pattern.

The idea of analogy as a mechanism for language change might account for this inconsistent relationship between meaning and pattern. In a similar manner to CxG and FG approaches analogy plays an important role in PG. Hunston and Francis (2000) speculate that speakers begin to use patterns prototypically associated with words which have particular meanings, with different words which have similar meanings to the prototypical words through a process of analogy. By means of example the authors hypothesise that although typically and most frequently used with the pattern *V n with n*, e.g. 'to provide someone with something', the verb *provide* sometimes, albeit infrequently, occurs in the pattern *V n to n* e.g. 'to provide something to someone' by analogy with *give*. While prototypical uses form the core of notional groups, analogous uses form part of an area of 'flux' beyond. Hunston and Francis (2000) demonstrate how many words in 'flux' can be subsumed by more general meaning categories or explained in terms of deliberate exploitations of semantic prosody.

Although it plays an important role in PG, frequency of occurrence alone does not provide sufficient evidence to indicate patterns. This contrasts with a view taken by many corpus linguists and calls into question the utility of approaches, such as the lexical-bundle approach, which rely solely on frequency of occurrence. Firstly, a naïve or injudicious interpretation of concordance lines selected on the basis on co-occurrence can conceal the word a pattern belongs to. The corpus lines below adapted from Hunston and Francis, (2000, p. 71) illustrate this problem. Although at first glance the two lines could be interpreted as pertaining to the pattern *N that*, on Hunston and Francis's analysis the second line pertains to the pattern *V that* in this case with the verbal slot being filled by *prove*.

I felt an incredible amount of satisfaction that the perpetrators
It's easy to <u>prove</u> to your oven satisfaction that you are worth many times

Another argument against relying solely on frequency to identify patterns involves the question of the dependency of the node verbs and it collocates. Hunston and Francis (2000) argue that *train as an architect* or *train in modern dance,* are instances of the pattern *verb as N* or *verb in N*, while *train at [location], train for [time period], train with [person] train in [location]* are not patterns. This is because while the choice of *as* or the first instance of *in* have a mutual dependency with the verb *train*, in the other instances there is no such dependency. In Sinclairian terms *trained as* and the first instance of *trained in* are selected in accordance with the Idiom Principle. In many cases the dependency relations in a corpus line are not so easily delimitated.

In addition to the role of frequency, PG also differs from other corpus-driven approaches to the description of language in the extent to which it seeks to generalise. Whilst, the comparison is not entirely valid since the grammars were conceived of with very different purposes in mind, PG offers a more delicate description of language than LUG. In contrast, PG is somewhat more general than Sinclair's approach to language which seeks to describe the behaviour of individual words and the patterns of specific lexical items. Hunston and Francis (2000) touch on the limitations of PG, highlighting the possibility of using the approach to describe certain patterns in more detail and the problems that such description involves. The pattern ADJ *in n*, in a sentence such as "She is adamant in her refusal to make any statement" (G. Francis et al., 1998, reproduced in Hunston and Francis, 2000, p. 78), is often used with adjectives which indicate strong beliefs. The nouns involved usually relate to a way of thinking i.e. *belief* or a way of talking i.e. *support*, *opposition*, or *refusal*. The noun group often starts with a possessive determiner, in this case her. With reference to Halliday's (1994) grammatical metaphor, Hunston and Francis (2000) argue

that the sentence might be "a metaphoric representation of the congruent *She adamantly refused to make any statement*" (p. 79). The words *success* and *failure* present a similar problem. In G. Francis et al., (1998) they are grouped under the pattern *N as n*, but this essentially only the case when the pattern occurs as part of a more specific pattern as highlighted in the sentences below from G. Francis et al., (1998).

> *I knew I could be a success as a fighter and a human being.* (v-link N as n)
>
> Bella is modest about her success as Young Designer of the Year. (poss N as n)
>
> We need more research reports on the failure of men as fathers and housekeepers and on the growing burdens imposed on working mothers. (N of n as n)

These, and other similar problems, are a question of a trade-off between descriptive specificity, and simplicity in representation. This issue depends a great deal on the intended purpose of the description. The following two chapters will examine in more detail the question of EAP learners' needs and how they might be met in lexicographic resources.

Since a pattern may be represented as belonging to any of its constituent lexical items, PG can be viewed as either a constituent hierarchical grammar, or a linear one. Figure 1.3 below, taken from Hunston and Francis (2000, p. 202) shows a hierarchical representation of a sentence.

<table>
<tr><td colspan="10"></td><td colspan="2">V</td><td colspan="2">adj</td></tr>
<tr><td colspan="10"></td><td colspan="2">v</td><td colspan="2">ADJ</td></tr>
<tr><td colspan="10">Byers</td><td colspan="2">pleaded</td><td colspan="2">guilty</td></tr>
<tr><td>V</td><td colspan="11">that</td></tr>
<tr><td colspan="5"></td><td colspan="2">V</td><td>n</td><td>from</td><td colspan="3">n</td></tr>
<tr><td colspan="8"></td><td>PREP</td><td colspan="3">n</td></tr>
<tr><td></td><td>poss</td><td>N</td><td colspan="3">v</td><td>PRON</td><td colspan="2"></td><td>n</td><td>N</td></tr>
<tr><td></td><td>DET</td><td>n</td><td>MODAL</td><td>inf</td><td colspan="3"></td><td>DET</td><td colspan="2">n</td></tr>
<tr><td>believing</td><td>his</td><td>candour</td><td>would</td><td>save</td><td>him</td><td>from</td><td>the</td><td>death</td><td>sentence.</td></tr>
</table>

Figure 1.3. A representation of a hierarchical analysis of the sentence "Byers pleaded guilty believing his candour would save him from the death sentence"

Following the work of Brazil (1995) a linear representation is also possible since each word that has a pattern might be seen to prospect the elements of that pattern. This linear representation of pattern avoids the problematic question of how to account for embedded clauses, for example, in Hallidean terms there is no need to invoke a rank shift. Another advantage of a linear interpretation is that patterns which begin with the dummy subject *there* are can be better accounted for.

Before moving on to examine the treatment of metaphor in corpus linguistics, it is worth mentioning the partial parsing or shallow parsing approaches adopted in NLP and information retrieval which might also be considered phraseological. It is debatable

whether these approaches fall within the remit of corpus linguistics. On one hand, as with many other corpus approaches their primary concern is the surface structure, on the other hand, unlike most corpus approaches they employ a top down approach from grammar to text (see for example Coniam, 1998; Hasegawa, Sekine, & Grishman, 2004; H. Yang, 1986).

1.4.1 Corpus Approaches to Metaphor

Unlike CG and CxG corpus linguistics has traditionally been more concerned with text rather than the mental representation of words. Most corpus treatments of metaphor employ corpus linguistics techniques as a methodological tool for the study of metaphor under other theoretical frameworks. Deignan (2005) provides a detailed review of the use of corpus techniques to study metaphor in cognitive linguistics, and discourse analysis.

1.5 Psycholinguistic Approaches

Proponents of all the approaches addressed thus far in this chapter have all, to some extent, emphasised psycholinguistic plausibility. There are, however, very few empirical studies which support these claims. Moreover, there is disagreement between those studies. Bolander (1989) claims phrases allow the memory to economise on the processing effort needed. Memory capacity has been linked to fluency, N. Ellis (1996) and Towell, Hawkins, and Bazergui (1996) claim that phrases promote fluency as they rely on declarative knowledge, which does not require the speaker's attention, rather than procedural knowledge which does. Pawley and Syder (1983) also make a link between phrases and levels of fluency which, they claim, would not be possible if speakers were solely reliant on individual words. They argue instead that a series of prefabricated strings and frames is used to reduce the workload of the memory. The few experiments carried out to test this hypothesis have presented conflicting results. Schmitt, Grandage and Adolphs (2004) use a psycholinguistic dictation task to establish how native and non-native speakers reproduce a variety of phrases. From this they conclude that there is not enough evidence to claim that phrases are stored in the memory as whole units. In contrast, Underwood, Schmitt and Galpin (2004) employ an eye-movement tracker to demonstrate that phrases offer an advantage in terms of processing effort. Schmitt and Underwood (2004) employ a self-paced reading task to measure recognition times for the individual words in phrases for native and non-native speakers. While no difference was found between the two groups overall, non-native speakers recognised words which formed part of phrases more quickly than those words which did not.

1.6 The Theory of Norms and Exploitations

The present study investigates the phraseological behaviour of verbs across academic disciplines and explores how they might be adequately represented in lexicographical resources for EAP users. With this aim in mind this chapter has outlined a number of approaches to the analysis of language which, in contrast to traditional generative approaches, could be considered phraseological. An examination of lexicographical approaches dealing with so-called fixed phrases brings to light the difficulties in delimiting such phrases from language in general. Cognitive approaches provide a plausible yet speculative account of how MWUs might be represented in the mind and countenance a key role for metaphor. Corpus approaches, provide a means of analysing real language thus avoiding the hazardous practice of inventing examples. Theoretical insights gained from

corpus studies are also important. This is particularly true of the Idiom Principle which situates the phrase at the heart of language as a central organising principle. Equally important is the emphasis that many corpus linguists place on paradigmatic and inter-textual matters. These matters are also of great importance in functional linguistics which demonstrates the possibility of examining language systemically at different levels of granularity. The conclusion of this brief overview is that an ideal theoretical framework for the study of the phraseological behaviour of verbs across academic disciplines would be systemic, evidence based, deal with meaning at a paradigmatic as well as sentential level, provide an adequate account of metaphorical language, and prove practicable with large amounts of data. Going full-circle back to a lexicographical approach Patrick Hanks's (2013) Theory of Norms and Exploitations (TNE) along with its practical lexicographically orientated counterpart Corpus Pattern Analysis (CPA) represents such a framework.

In common with Hoey (2005), lexical primes play a primordial role in TNE. Words do not have meanings in isolation; instead, in a similar fashion to Hunston and Francis's (2000) pattern grammar, they have a number of different meaning potentials which can be activated by different collocational patterns. In TNE collocates are arranged in lexical sets according to their collocational preference. In the Firthian mode, these lexical sets are mapped to syntactic structures as colligations. The resulting complex patterns, similar to form/meaning pairs of construction grammar also have a cognitive element as they represent the speaker's beliefs about the world. Hanks (2013, p. 215) conceives of language as a "double-helix" consisting of rules governing conventional collocation of words, intertwined with a set of rules specifying the ways in which those norms are exploited. These elements are not strictly delimited; they are poles at opposite ends of a continuum as illustrated in Figure 1.4.

Norms → Alternations → Exploitations

Figure 1.4. Hanks's continuum of rules governing collocation

In TNE, corpora play an essential role. Not only are they the means by which lexical sets are mapped to syntactic structures as colligations through a lexicographical technique known as Corpus Pattern Analysis (Hanks, 2004; Hanks & Pustejovsky, 2005), but they are also the source of evidence from which many of the insights on which TNE is based are derived. This will become apparent in the remainder of this section which deals with norms, exploitations, and alternations in greater detail.

For Hanks (2013), in a reasonably sized corpus, normal usage is immediately apparent from the large number of similar collocational and syntactic patterns. These patterns can be elucidated by examining a number of phenomena including the interplay of a collocate's extensional and intensional characteristics, the correlation between two or more lexical sets in different clause roles for a given keyword, and what collocates indicate about the key or base word's presuppositions and implications, contextually determined default interpretations, and semantic prosody.

There are two aspects of norms which are especially relevant to a study of the behaviour of vocabulary in texts from different academic disciplines. Firstly, norms can be domain-specific; in certain domains collocations containing a given word might display different characteristics, for example semantic types, or semantic prosody which, in Louw's (1993, p. 157) terms is "consistent aura of meaning with which a form is imbued by its collocates", than it would in general language. Norms also change over time, thus uses of language that were once novel become conventionalised through repeated use over time in a discourse community. G. Williams (2002) demonstrates how this evolution of meaning presents

problems for corpus studies of specialised language which is subject to more rapid change in and creation of meaning and use than general language.

In general humans are predictable in their language use. Normal usage accounts for the vast majority of language. However, we are capable of novelty in expression. In TNE this capability is accounted for by the mechanism of lexical exploitation by which old things can be said in new and interesting ways and new meaning can be created ad hoc (Hanks, 2013, p. 211). Exploitations are situated at the opposite end of the continuum to norms and can take a variety of forms. Hanks offers a typology of linguistic phenomena which can sometimes be regarded as exploitations. This includes, ellipsis, an exploitation which can be employed to change the focus of verb meaning; the use of anomalous collocates, in other words "non-canonical members of a lexical set" (Hanks, 2013, p. 169) as a rhetorical device; and semantic-type coercion, a concept from Generative Lexicon Theory (Pustejovsky, 1995), which explains how some words are coerced into having different semantic types by their context.

Metaphors and similes are particularly important exploitations. In common with many cognitive linguists, Hanks attributes a central role to metaphor in language. In Hanks (2010) he identifies the criterion of semantic resonance as central in the disambiguation of metaphorical and literal word senses. Conventionalised metaphors abound in everyday language and evidence the fluid nature of norms and exploitations. Over time and with repeated use in a discourse community metaphors which were originally novel exploitations of language use are conventionalised and become secondary norms. Exploitations also play a key role in rhetoric. However, that is not to say that all rhetorical tropes are exploitations. For example, like a great deal of metaphors, many examples of metonymy and synecdoche have become secondary norms.

Hanks (2013) states "some norms are more normal than others; some exploitations are more outrageous than others" (p. 316). Alternations represent a middle ground where one element can be substituted for another with no change in meaning. Be they lexical (e.g. the substitution of a lexical word), syntactic (e.g. active or passive), or semantic (e.g. changes in focus, you can talk about calming someone or alternatively, with a slightly different focus, about calming someone's anxiety (Hanks, 2013, p. 316)).

From a cognitive perspective lexical alternations are important. They allow users to say different things in a manner which is still intelligible to other users because it closely corresponds with other prototypical or schematic phrasal patterns. In effect, the hearer or reader is primed to understand the lexical alteration by the underlying phrasal pattern. Semantic-type alterations are regular and widespread. Take for example the sentence pair below which illustrates an alternation in subject position between [[Human]] and [[Institution]]:

> The mayor announced a new policy.
> The city council announced a new policy.

The active/passive alternation is the most frequent syntactic alternation in English it is widely used to change the focus of a sentence, for example the passive is often employed to emphasise an event rather than an agent. Causative/inchoative alternations are also very frequent with some verbs they are quite unexceptional have very little impact on sense while with others they, are less frequent and change the sense of a verb a great deal. Within the framework of TNE, the latter can be regarded as exploitations rather than alternations.

Other common syntactic alternations include indirect object alternation and reciprocal constructions. These are widespread but have very little influence on meaning. The same can be said for a number of ellipsis alternations where the meaning can be implied from the wider context of the text. When such ellipsis occurs frequently with a given lexical item it is regarded as a norm. Conversely, when it is relatively infrequent for a given word it can be regarded as an alternation.

The regularity and widespread nature of alternations raises the question of whether it is prudent to represent them in lexicographical resources. Hanks (1994) regards this question as an instance of the debate between what he characterises as lumpers and splitters. Lumpers are those who prefer to group senses under a more general superordinate while splitters prefer to make finer subdivisions. Ultimately, the decision depends on the needs of the user. In regards to EAP lexicography this question is discussed further in Chapter 3.

1.7 Relation of TNE to Other Approaches

This chapter section aims to examine the relationship between TNE and some of the phraseological approaches outlined above. However, before doing so it seeks to situate TNE in relation to two currents of linguistics which have been dominant for the last century, namely, European structuralism and generative linguistics.

Hanks (2013) regards Saussure's four basic distinctions between langue/parole, paradigmatic/syntagmatic relations, the relation between the sign/signified, and synchronic/diachronic relations as "four essential planks in the platform on which studies of meaning in language may be based" (p. 351). However, he sees some gaps in the Saussurean account which can be filled using TNE.

While in broad terms TNE can be regarded as compatible with Saussurean structural linguistics, the same cannot be said of Chomiskian generative linguistics. From the outset Chomsky (1957) makes it clear that the goal of the generative programme is to explain idealised linguistic competence rather than actual performance. This distinction lies at the root of the differences between TNE and generative approaches. In Chomsky (1965) lexical items are the terminal nodes in top-down parse trees representing the syntactic structure of a sentence. Lexical items are selected in accordance with subcategorization rules which restrict the item's subcategorization to words with certain properties, for example animate, inanimate, in the case of nouns. Hanks regards the projection principle (Chomsky, 1981) as an attempt to assign a more important role to the lexicon in the generative paradigm. Instead of a top-down approach from an abstract grammatical sentence 'S' to lexical items at terminal nodes syntactic well-formedness is determined by a set of rules associated with lexical items which are projected up the parse tree. Using corpus data Hanks (2013) demonstrates that actual non-idealised language from a corpus is better explained in terms of selectional preferences than selectional restrictions; in other words, a prototypical norm with alternations and exploitations. This criticism also holds for Chomsky's (1995) Minimalist Program. For Hanks TNE and the Minimalist Programme are fundamentally incompatible since TNE assumes that the goal of language is to say meaningful things rather than create syntactically well-formed sentences.

In accordance with the Projection Principle (Chomsky, 1981) the syntactic structure of a sentence is projected from the lexical properties of its predicator. In Lexical Functional Grammar (Kaplan & Bresnan, 1982), completeness and coherence conditions play much

the same role. For Hanks the principle congruency between TNE and Lexical Functional Grammar is the central place the latter assigns to functions which Hanks equates to clause roles in more traditional grammars; for him an essential part of any meaningful account of linguistic function. Another parallel is found in Bresnan's (2007) argument against the overreliance on intuitions in theoretical linguistics and calls for future research to take into account quantitative corpus research. This is one of the central tenets of TNE.

The work of Jackendoff while still broadly situated in the generative paradigm can be seen as more compatible with TNE than mainstream generative approaches. Jackendoff (1997, 2002) criticises the syntactocentrism of mainstream generative approaches. Jackendoff (1997) argues that it is psychologically implausible that all utterances are built up from individual lexical items instead he contends that "phrasal lexical items" (p. 158) and "constructional idioms" (p. 172) are stored in the lexicon as wholes. In a manner comparable to some cognitive approaches Jackendoff proposes a parallel architecture which consists of three principal interconnected modules: phonological, syntactic, and semantic. However the importance given to syntax in mediating the semantic and phonological levels limits its compatibility with many other cognitive theories and TNE.

In the context of the generative framework Pustejovsky's (1995) Generative Lexicon Theory, is perhaps theory which is most compatible with TNE according to Hanks (2013):

> It is compatible with a view of language as a probabilistic, preferential system. The lexicon is seen as consisting of a finite number of lexical items that can generate an infinite number of meanings. Words can be used meaningfully in an infinite number of novel contexts, activating different shades of meaning in different contexts. (p. 378)

A number of concepts from Generative Lexicon Theory are important in TNE: A hierarchical ontology of semantic types, for example [[Human]], [[Institution]], [[Document]], is fundamental to CPA; the notion of argument structure, compatible with Haliday (1994), involves identifying the participants; event structure is required to identify the type of event being described, for example state, process or transition; and qualia structure deals with the relevant properties of the participants. Similarly lexical inheritance structure identifies what sort of things or people are involved in the event. There are four qualia for entities: The formal: What sort of thing is it? The constitutive: What is the relation of its parts? The telic: What is it for? And the agentive which denotes factors involved in the origin or coming into being of an object. The cognitive profiles used to examine nouns are a good example of the use of qualia in TNE. Figure 1.5 is a cognitive profile for one sense of *shower*.

Shower 2: is an artifact for pouring a continuous flow of water in droplets, simulating rainfall, down over a person:

- Typically, a shower is *provided* by an architect or house designer and *installed* by a builder, either in a *cabinet* in the *bathroom* of a house, or above the *bath*, or in a separate *shower-room*.
- An *en-suite shower* is one that is installed in a room adjacent to a *bedroom*.
- When installed correctly, a shower *works*.
- Types of shower: *electric* shower*, power* shower*, gravity-fed* shower [and various trade names]
- People *switch* (or *turn*) a shower *on* in order to *use* it and *switch* (or *turn*) it *off* after use.

Figure 1.5. A partial corpus-driven driven cognitive profile for the noun *shower* adapted from Hanks (2013, p. 137)

Lexical inheritance structure allows that the telic need not be repeated for every item in a hierarchical lexical set. Adapting an example from Hanks (2013) if *car* is a vehicle we can infer that it is a physical object and that its purpose, like that of all vehicles, is to transport people.

The principle difference between TNE on the one hand, and many lexicographical approaches on the other, is the role attributed to phraseology as an organising principle of language. As discussed above the practical objective of many lexicographical approaches necessitates a strict delimitation of what constitutes a phrase and what does not. In essence this necessitates the somewhat artificial distinction between free and non-free phrases. TNE makes no such decision a priori. Following Sinclair (1991), Hanks (2013) defines a collocation as "quite simply, co-occurrences of words in a text" (2013, p. 1) the meaning of a particular element of a collocation can be derived a posteriori through an analysis of collocational patterns.

TNE and meaning-text theory (MTT) (Mel'čuk, 1995, 1988, 2006, 2013) differ in fundamental ways. MTT aims "to sate the necessary conditions for all possible correct uses of the lexical item being defined" (Hanks 2013 p. 363). In TNE word senses are variable and highly contextually dependent. This means that such thoroughgoing definition of all possible correct uses is not possible.

In theory Apresjan's (2000) account of collocations based on the concept of "a co-occurrence constraint, which, though not binding, is observed in pedantic and literary discourse" (Apresjan, 2000, p. 5) is more compatible with TNE than MTT. The idea of "non-binding" discourse-type-dependent constraints chimes with the idea of context-dependent norms and exploitations. However, there is one important difference between Asperjan's work and that of Hanks: the central role of corpus data. Hanks demonstrates how some of the introspective examples used by Asperjan do not match up with corpus evidence. This is unsurprising as large-scale computerised corpora were not widely accessible at the time Apresjan conducted this work. However, it is testament to the problems of basing theories on data which is the product of introspection.

The Pattern Dictionary of English Verbs[2] (PDEV; first reported in Hanks & Pustejovsky, 2005) is an example of the result of a lexicographic application of TNE. The PDEV is a corpus-identified list of the patterns of English verbs. It was created by tagging the arguments of verbs with semantic types from a custom ontology. Each PDEV entry contains the corpus-derived CPA patterns for a given verb along with their implicatures and examples of use.

In general terms, many currents of cognitive linguistics primarily focus on the mental representation of linguistic items. The symbolic thesis in which "lexicon, morphology and syntax form a continuum of symbolic units, divided only arbitrarily into separate components." (Langacker, 1991, p. 275) is broadly compatible with the mapping of lexical sets to syntactic structures as colligations in TNE. Similarly, there are also parallels with the notion of pattern from TNE and CPA, and construction in construction grammar.

The importance of frequency in many cognitive approaches in the entrenchment of new structures provides a plausible mental mechanism for the conversion of novel exploitations into secondary norms over time in TNE. Similarly, many cognitive approaches afford metaphor a central and fundamental role in language. The same is true in TNE, as evidenced by the discussion of metaphorical exploitations as a mechanism for language change above.

In spite of these communalities Hanks (2013) regards TNE as incompatible with many cognitive approaches because of their recourse to acceptability judgments. In this regard he is particularly critical of FrameNet which he claims is driven by concerns of theoretical elegance rather than real data. FrameNet is not corpus driven, examples are imported a posteriori to fill node in the net. He does concede however, that in the case of Burgman and Lakoff (1988) acceptability judgments match well with data from corpora (Hanks, 2013, p. 382). He further speculates that corpus-driven lexical analysis could be employed to discover what Langacker (1991) describes as "conventionally sanctioned" (p. 158) usage of lexical items in a given domain.

As the summary above demonstrates, in Hallidean Systemic Functional Grammar words are not conceived of as discrete units of meaning. For this reason, it is hardly surprising TNE is compatible with many aspects of Systemic functional Grammar. The Rank Scale provides a framework in which the relationship of the interdependent clause roles of embedded phrases can be analysed. As Hanks (2013) explains "A clause or phrase can have the function of a word in particular circumstances; thus, a phrase such as *flesh eating* has the function of adjective when embedded in another phrase, *flesh-eating mammals*. This phenomenon is known as rank shift." (p. 395). Another Hallidean notion which is compatible with TNE is that of exponence. The first step of a CPA involves classifying each token in the text according to its part-of-speech. The Hallidean notion of delicacy is central to TNE. Taking the word *hazard* as an example Hanks (2013) explains:

> At a coarse level, the verb *hazard* denotes a speech act. Slightly more delicately, the speech act is associated with a cognitive process (thinking). Even more delicate, though necessary for distinguishing the verb from other speech-act verbs, is the fact that hazarding implies lack of confidence in the truth of the proposition (the guess) that is hazarded (p. 396)

[2] http://www.pdev.org.uk [Accessed 29/09/17] – All future references to PDEV refer to this resource. First reported in Hanks and Pustejovsky (2005)

Hanks equates delicacy with "exploitativeness". Novel exploitations occur at a greater level of delicacy than those which are in the process of becoming conventionalized.

The emphasis which Halliday places on the text at a supra-sentential level exposes a limitation of the TNE approach. In most circumstances analysis takes place at the level of corpus lines, however, for many phenomena, for example pronominal anaphora, it is necessary to examine larger sections of text.

TNE has its origins in dissatisfaction with more traditional theories of language. These are inadequate on two fronts. Firstly, as argued at the beginning of this chapter, they assign a marginal role to lexis and secondly, and equally importantly, their conclusions are founded on introspection rather than actual data. As a corpus-driven theory with origins in CPA, a compatible corpus-based lexicographical technique, TNE addresses this second concern.

As a corpus-driven theory, TNE is indebted to the work of Sinclair. In practice, TNE provides a means of documenting the Idiom Principle, providing an alternative to "doomed" (Sinclair, 1998, p. 5) single-word lexicons by differentiating possible combinations from normal combinations. The Sinclairian concepts of semantic preference and semantic prosody (see also Louw 1993) are particularly useful in this regard. Another concept from corpus linguistics which is employed fruitfully in TNE is lexical priming (Hoey, 2005). The patterns resulting from primings can be observed in corpora. Importantly, given the objectives of the present study, Hoey (2005) advises studying priming in discipline-specific corpora. Unlike TNE, lexical priming does not make use of statistical measures. For Hanks, this is an important oversight. Hanks (2013) argues that "this would seem a prerequisite for the understanding not only for the 'reinforcement' component of priming but also the cognitive-salience of rare but memorable primings, such as idioms and other rare but striking phraseology. Frequency alone is not enough." (p. 402). G. Williams and Millon's (2010, 2014) use of collocational networks to create a dictionary of science verbs is a practical application of certain aspects of priming in discipline-specific corpora that does make use of statistical measures of word association.

In addition to the concepts of the Idiom Principle and priming, the notion of chunking as set out in Sinclair and Mauranen (2006) is critical to TNE. The same can be said for the notion of pattern as set out in Hunston and Francis (2000). The remainder of this chapter outlines a practicable means of analysing such corpus patterns within a TNE framework.

Central to TNE is the idea that the norm/exploitation continuum of rules is reflected in corpora. A first glance at statistical measures or the context of a word reveal patterns which initially seem simple to formalise, but as the depth of analysis increases more exceptions are revealed. Hanks (2013) demonstrates how corpus-driven techniques can be used to calculate norms of usage and reveal and their exploitations. The norms and their exploitations for verbs can be calculated in a number of ways. For example, by examining the presuppositions and implications of a verb in context, or by examining contextually determined default interpretations, in other words, through an examination of those contextual factors which give a word a different meaning to its synonyms. The semantic prosody of a verb and the levels of delicacy of a verb's implicatures can also be used to calculate norms and exploitations. In the case of nouns, Hanks (2013) explains how corpus evidence can be used to build cognitive prototypes for keywords by searching for collocates in a five-word window to its left or right. Crucially for the purpose of the proposed study, some norms are completely domain specific while others are more prevalent in certain domains than others.

CPA might be seen as the formalised expression of the corpus analysis technique from which the insights underlying TNE were derived. It is a technique for mapping meaning onto words in text (Hanks, 2004a). It is primarily concerned with describing the syntagmatic and syntactic relationships of verbs (Cinková & Hanks, 2010). Hanks and Pustejovsky (2005) argue that since existing attempts at such description are based to varying extents on intuition they often represent no more than unsubstantiated assertions. As evidence, Hanks and Pustejovsky (2005) cite the creation of artificial semantic classes to fill empty nodes in the WordNet (Fellbaum, 1998) hierarchy, the omission of words when populating frames in FrameNet (Fillmore, Johnson, & Petruck, 2003), and Levin's (1993) exclusion of sentential verbs from her verb classes. They argue that since CPA is strictly corpus based it is not subject to such omissions or fabrications.

As its name suggests, CPA derives patterns from corpora. These patterns represent a prototypical usage, judged by the relative frequency of occurrence in relation to other patterns, to which a 'primary implicature', in other words a general meaning can be attached. Verb patterns consist of argument structure based on the clause roles of systemic grammar (Figure 1.6) along with their semantic values. Where necessary other features such as the presence or absence of a direct object or of a determiner are also stated.

> S - Subject (Agent in dependency grammar) - the semantic subject of the clause (omitted or introduced by the preposition by in passive realizations)
> P - Predicator (the verb, together with its auxiliaries if any)
> O - Object (direct or indirect; in CPA, 'direct object' includes the subject of passive sentences)
> C - Complement (a phrase that is co-referential either with the subject of the sentence, as in: He is happy; he is the President or with the direct object, as in They elected him President; it made him happy)
> A - Adverbial (usually a prepositional phrase, a particle, or one of a small set of adverbs, as in She drove to London, she drove home, she drove off).

Figure. 1.6. Clause roles in systemic grammar as presented in (Cinková & Hanks, 2010, p. 2)

> 1. [[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}
> EXX: He grasped the handle of the door with one hand, and that of the spoon with the other.
> He reached out wildly, trying to grasp the creature, but it had moved away.

Figure. 1.7. An example pattern, sense, and implicature for *grasp* (Hanks and Pustejovsky, 2005, p.10)

In the example pattern in Figure 1.7 double square brackets indicate a semantic type, which represent the prototypical properties shared by the lexical items over which they generalise (for example, [[Human]], [[Physical Object]], [[Abstract]]). Semantic roles are indicated after an equals sign. These should not be confused with semantic types. They are employed to account for distinctions in meaning between certain contexts, for example: "[[Human = Film Director]] *shoot*" and "[[Human = Sports Player]] *shoot*" (El Maarouf, Baisa, Bradbury, & Hanks, 2014 p. 1002). Curly brackets represent specific lexical types and phraseological groupings such as those given in 'Other Clues'. Parenthesis, which are not present in Figure. 1.7 indicate optionality. Originally, CPA was designed for natural language

computing (Hanks, 2004a) and its pattern conventions are intended to be machine readable. However, there are many examples of other applications of CPA in which patterns are read by humans and contain greater or fewer details than those given in Figure. 1.7. See for example the PDEV.

1.8 Summary

This chapter has outlined a number of approaches to the description of language which could be considered phraseological. Ultimately, considering the practical lexicographical aim of the present study, TNE, with its associated lexicographic technique CPA, represents the ideal framework in which to carry out this study. Like all the other theories addressed, it affords a central and organising role to phrase in language, it also provides a feasible account of the role of metaphor – a phenomena which is more central to phraseological approaches than syntactocentric approaches and which, when conceived of as a means of conveying abstract concepts in everyday language, has obvious relevance to academic language.

TNE holds several advantages over the other approaches examined in this chapter. Like all the corpus approaches discussed, it is data-driven and does not primarily rely on introspection. In the context of the present study with its aim of providing guidelines for the treatment of vocabulary for EAP courses its principal advantage over the other corpus approaches addressed lies in its practicability. TNE allows a corpus-based approach to the study of phraseology through CPA which deals with semantic as well as syntactic concerns. It is more practicable in terms of the processing time necessary with substantial amounts of corpus data than frame-semantic approaches such as those expounded in Fillmore and Atkins (1992) and evident in FrameNet (Fillmore et al., 2003).

Detailed procedures for a corpus-based experiment using CPA to examine the behaviour of verbs and their collocates in different academic disciplines are set out in Chapter 5. Chapter 4 deals with the assumptions underlying this experiment and the hypotheses it aims to test. Having examined theories which might account for a phraseological view of EAP language in the present chapter, it will be beneficial to examine how phraseology is treated in dictionaries and other lexical resources used in EAP teaching and learning. This is dealt with in Chapter 3. To complete the panorama, it is necessary to examine the processes involved in teaching and learning vocabulary for English for academic purposes and more generally the teaching and learning of phraseology. This is the concern of the following chapter.

2. TEACHING AND LEARNING EAP VOCABULARY

This chapter concerns the processes involved in teaching and learning vocabulary for English for academic purposes. Before addressing concerns specific to EAP lexis, it gives an overview of the issues involved in the teaching and learning of vocabulary in general, and more specifically the teaching and learning of phraseology.

Valid comparison of the findings of research in vocabulary acquisition in English is problematic. In general, human learning is influenced by a many idiosyncratic factors. Moreover, in English vocabulary acquisition, there is an enormous wealth of research which covers a range of learning contexts, ages, and proficiency levels. Similarly, there are a wide range of methods employed in studies of vocabulary acquisition to the extent that many studies of similar phenomena employ vastly different methods. With this caveat in mind the following chapter attempts to outline the principal findings in the field.

A great deal of primary and secondary sources were consulted in the elaboration of this chapter. However, particularly frequent reference is made to the work of Paul Nation; in particular Nation (2013). This is not because his views on the nature of vocabulary teaching and learning accord entirely with those which motivate the present study. In fact, the primacy of frequency as a criterion for vocabulary selection (Nation, 2013) and a monosemic view of words (Nation, 2016) run counter to a phraseological approach to vocabulary selection. That said, Nation (2013), and its wordlist focused counterpart Nation (2016), do provide an useful overview of developments in research on the teaching and learning of vocabulary over the past 40 years, and these differences in approach regarding frequency and polysemy provide useful counterpoints from which to argue for a phraseological approach. Equally useful, for the purposes of the present study, are the model of vocabulary acquisition, focuses for deliberate vocabulary learning, and suggestions for teaching activities Nation (2013) draws from a review of existing research.

2.1 Learning Vocabulary in General

This section sets out the groundwork for an in-depth discussion of vocabulary learning in EAP. First, it outlines several common conceptions of the nature of words in vocabulary acquisition studies. This is followed by a discussion of the importance of frequency. Then the question of what word knowledge involves is addressed. Finally, there is a discussion of vocabulary teaching including some of the strategies, techniques, conditions, and activities conducive to this task.

In vocabulary acquisition in general, a prerequisite for any study, and a common point of conflict concerns the decision on what can justifiably be considered a word. The consensus is far from universal, many well-known studies deal with the problem of how to count inflected forms by concentrating instead on the lemma (W. N. Francis & Kučera, 1982; Leech, Rayson, & Wilson, 2001; Thorndike & Lorge, 1944). The rationale underlying this approach is that once learners know the base form, acquiring the inflected form is relatively straightforward (Nation, 2013). While this argument holds for learners whose L1 has a similar inflectional system to English, in situations where the L1 inflectional system is drastically different this is unlikely to be the case. The question of how to deal with irregular forms and spellings is also moot on this approach. Another, concern is what the headword of the lemma should be, given that in authentic texts certain inflected forms tend to occur far more frequently that the base form (Sinclair, 1991). Derived forms pose a

further problem on the lemma approach. Many such forms are often closely related to a base form but are considered separate lemmas.

The notion of word families, a headword with all its inflected and related derived forms, addresses some of the problems posed by the lemma approach. However, word families also have certain shortcomings. Bauer and Nation (1993) point out that since learners' morphological knowledge develops with experience, a sensible word family for an advanced learner might be beyond a beginner. To counter this danger, they suggest setting up a scale of word families graded by level in terms of transparency of their member words. Even with this countermeasure there are studies which call into question the validity of word families for L2 learners. Several researchers have pointed out that word families often contain words with very different meanings (Durrant, 2009; D. Gardner, 2008; Martinez & Murphy, 2011). O'Keeffe, McCarthy, and Carter (2007) claim that there is no reason why learners should make the connection between forms of the same lemma. Similarly, in a lexicographically orientated study, Bogaards (2001) claims there is no experimental evidence to confirm the assumption that learners who know a headword would be able to work out the meaning of a derived form. Schmitt and Zimmerman (2002) find associating even closely related words is a difficult task for learners. In their study of German past participle formation in L1 speakers and L2 learners Neubauer and Clahsen (2009) suggest that high-proficiency learners store words as unanalysed wholes while L1 speakers undertake morphological parsing. In their study of lower-level students from Thailand, J. Ward and Chuenjundaeng (2009) find no evidence that knowledge of the headword of word family implies knowledge of other family-member words.

Central to the notion of lemma and word families in studies of vocabulary acquisition is the notion of frequency. Several researchers have a dogmatic belief in the importance of frequency and prioritise it over other factors which might be more beneficial to learners' vocabulary acquisition. Nation (2006) suggests that between 6,000 and 9,000 word families are necessary to gain 98 per cent coverage of a text. A number of studies equate the 98 per cent coverage level with a manageable amount of unknown vocabulary in reading comprehension (H. Hu & Nation, 2000; N Schmitt, Jiang, & Grabe, 2011; Van Zeeland & Schmitt, 2013). This claim is somewhat problematic since comprehension is not a question of quantity of known or unknown words but rather of understanding those words key to the message of the text. The semantic set a word belongs to should also be taken into account; as McCarten and McCarthy (2010) and O'Keeffe et al. (2007) argue, it would make little sense to exclude the low frequency days of the week *Tuesday* and *Wednesday* yet include the remaining days of the week because they occur at higher frequency. The question of how to treat frequently occurring phrases is also problematic. O'Keeffe et al. (2007) demonstrate how the phrase *all the time* occurs in the top 1,000 items in a mixed spoken and written corpus of 10 million words. In general terms, if it were treated as a single word it would be amongst the highest occurring in English.

In order to facilitate acquisition many researchers advocate the division of vocabulary into high-, mid- and low-frequency groups. Schmitt and Schmitt (2012) suggest that the high-frequency group contains 3,000 word families. Nation (2013) suggests that the group contains 2,000 word families which cover around 90 per cent of most texts. In their spoken and written corpus, O'Keeffe et al. (2007) suggest that 2000 words gives over 80 per cent coverage. Nation (2006) argues that learners should learn high frequency words as soon as possible in order to make the written and spoken texts accessible, elsewhere it is argued that the inability to fluently access high-frequency vocabulary has a detrimental effect on comprehension and the ability to produce speech and writing. He also argues that it is

worthwhile distinguishing between high-frequency word families and mid-frequency word families. The latter group comprises 7,000 word families from the third-thousandth to the ninth-thousand word families. Along with proper nouns these provide 98 per cent coverage of texts from newspapers and novels, are reasonably well known by most native speakers, and are fundamental to the unassisted reading of non-simplified text (Nation, 2013). Low-frequency words are those outside the most common 9,000 words in English. They are the most numerous group.

The insistence on frequency and the division of vocabulary into high-, mid- and low-frequency groups is somewhat problematic. While it is reasonable to suppose that students need to understand and produce high-frequency vocabulary, this should not necessarily entail a great deal of teaching time. Learners will be naturally exposed to high-frequency vocabulary because of its highly frequent nature. Teaching time would be better spent on more salient vocabulary. The division between high- and mid-frequency words presents further problems. Frequency and range are dependent on the nature of the corpus under study. The frequency of occurrence of country names in O'Keeffe et al.'s (2007) corpus provides a good example of this. Countries such as *America*, *France*, *Italy*, *India*, and *Ireland* which could be said to have close cultural ties to Britain, the source of the corpus texts, occur among the most frequent 2,000 words, while other countries with more distinct cultures such as *Spain* and *China* fall outside the high-frequency group. The ambiguous relation of low-frequency words to technical vocabulary is also problematic. Many researchers advocate the incidental learning of low-frequency words, however, if such words are fundamental to the specific purpose of an English for specific purposes course then it stands to reason that they should be taught deliberately.

In sum, frequency is an important factor to be taken into account in vocabulary teaching. There is little to be gained by spending a great deal of teaching time on words which occur very infrequently and/or have a meaning which can be easily guessed from context. However, an overreliance on frequency underestimates the importance of salient words which have meanings which cannot be guessed from context to comprehension. Moreover, the designation of words as high-, medium- or low- frequency often changes according to the corpus studied.

The questions of what precisely constitutes a word and which vocabulary merits teaching time are complex, the same is true of the question of what is involved in knowing a word. Central to this question is the concept of learning burden (Swenson & West, 1934). The learning burden of a word corresponds to the degree of difficulty it poses to the learner. In practice this means that L2 vocabulary which exhibits spelling (de Groot, 2006; Hamada & Koda, 2008), grammatical, and semantic patterns similar to those of the learners' L1 vocabulary, or previous knowledge of L2 or L3 vocabulary is easier to learn than vocabulary which exhibits distinct patterns. It is essential that teachers quickly and adequately assess the learning burden of a word so that they can concentrate on underlying patterns that need attention (Nation, 2013). Lexicographical resources have an important role to play in this process.

In broad terms a distinction is made between knowing a word productively, using it for speaking and writing, and knowing a word receptively, using it for listening and reading. However, this distinction is somewhat controversial since when listening and reading meaning is also produced. The alternative nomenclature of passive, for listening and reading, and active for writing and speaking are also sometimes employed (Corson, 1995; Laufer, 1998; Meara, 1990). Nevertheless, these terms are not entirely adequate as reading

and listening cannot be considered passive processes in the canonical sense of the term. Melka-Teichroew (1982) conceives of the passive/active distinction as a cline. Meara (1990) posits an active/passive distinction based on different types of associations between words. Productive vocabulary has many connections, both incoming and outgoing, with other words while receptive vocabulary is not activated by links to other words. Nation (2013) criticises this view claiming that it is possible to name an object in an L2 without making links to other L2 words. In spite of these disagreements the receptive/productive dichotomy is almost universally employed in language teaching courses and material design. It follows that this thesis with its practical and applicable objective should maintain this distinction.

In general, receptive learning is considered easier than productive learning. From a psycholinguistic perspective, N. Ellis and Beaton (1993) put forward several explanations as to why this might be the case. The first explanation involves the amount of knowledge required, while learners need only recognize the written and spoken form of a word for receptive use, productive use entails more precise knowledge of the word form and its collocation patterns. There is also the matter of practice; in many formal contexts where language is taught and learnt, such as language classes, learners tend to have more practice of receptive skills than productive ones. A further explanation, supported by tip-of-the-tongue experiments (R. Brown & McNeill, 1966), is that while in the early stages of L2 learning an L2 word has only one direct link to its L1 translation in the receptive direction, in the productive direction there are many more possible links from the L1 word to L2 collocates, synonyms, antonyms etc. Finally, there is the possibility that learners may know a word receptively but choose not to use it productively for socio-cultural reasons. For example, some people rarely use profanity but nonetheless understand the meaning of such words. In Corson's (1995) terms such vocabulary is said to be unmotivated. Empirical testing of the relative difficulty of productive and receptive knowledge is not a straightforward task. Test types inevitably risk confounding productive and receptive knowledge (Waring, 2002); a test of productive knowledge necessarily involves reading or listening to a prompt, similarly, a test of receptive knowledge necessarily involves producing a response to a test item.

This study defines productive and receptive knowledge following Nation's (2013) model. For Nation, knowing a vocabulary item involves the aspects form, meaning and use. In order to completely know a word or MWU a learner must know the answer to the questions posited in Table 2.1 below. All of the aspects listed; spoken, form and meaning, grammatical functions, etc. have productive and receptive components. Constraints on use are also an important part of the model.

Table 2.1 *What is involved in knowing a word or MWU* (adapted from Nation (2013, p. 27))

Form	spoken	R	What does the word or MWU sound like?
		P	How is the word or MWU pronounced?
	written	R	What does the word or MWU look like?
		P	How is the word or MWU written and spelled?
	word parts	R	What parts are recognisable in the word's or MWU's words?
		P	What word parts are needed to express the word MWU's meaning?
Meaning	form and meaning	R	What meaning does this word or MWU form signal?
		P	What word or MWU form can be used to express this meaning?
	concepts and referents	R	What is included in the concept?
		P	What items can the concept refer to?
	associations	R	What other words or MWUs does this make us think of?
		P	What other words or MWUs could we use instead of this one?
Use	grammatical functions	R	In what patterns does word or the MWU occur?
		P	In what patterns must we use this word or MWU?
	collocations	R	What words, MWUs or types of words occur with this one?
		P	What words, MWUs or types of words must we use with this one?
	constraints on use	R	Where, when, and how often would we expect to meet this MWU?
	(register, frequency...)	P	Where, when, and how often can we use this MWU?

Note: R= receptive knowledge, P = productive knowledge

There is a widespread distinction in vocabulary teaching between the direct teaching and learning of vocabulary; that is teaching and learning through tasks and procedures which have the specific aim of increasing vocabulary knowledge, and incidental vocabulary teaching and learning whereby increases in vocabulary knowledge occur as a by-product of tasks and procedures with other learning goals.

The value of direct teaching and learning of vocabulary has been widely questioned. This is particularly true for L1 learners. Some of the arguments put forward involve the L1 lexicon which is simply too vast for direct vocabulary teaching and learning to have any meaningful impact (D'Anna, Zechmeister, & Hall, 1991; Nation, 1993). Similarly, there is too much to learn about each individual word, this is particularly true if the validity of word families is called into question (Nagy, 1997). Teaching time is obviously important in direct teaching or learning of vocabulary. In order to achieve a significant effect on productive language use a great deal of teaching time is needed (McKeown, Beck, Omanson, & Pople, 1985), although, less teaching time is needed to achieve improvements in receptive use than productive (McDaniel & Pressley, 1989). It has also been argued that there are more effective means of learning vocabulary outside the classroom through incidental learning. Lexicographical resources can play a key role in facilitating such learning (Nation, 2013). Finally, some researchers have questioned the effectiveness of vocabulary teaching. For example, advocates of the reading hypothesis (D. Gardner, 2013) consider direct instruction too slow to account for vocabulary growth (Nagy & Herman, 1987) or argue that direct teaching of vocabulary might not lead to it being learnt in sufficient depth

(Krashen, 2010). However, according to Nation (2013), there are three key differences between and L1 and L2 vocabulary acquisition which favour the direct teaching and learning of vocabulary for L2 learners. Firstly, it is practical and feasible to teach the first 2000 to 3000 most frequent words. Secondly, in most contexts L2 learners have fewer opportunities to acquire vocabulary incidentally from input and output than L1 learners, this is particularly true at lower-proficiency levels. Direct vocabulary teaching could provide the springboard needed to reach the proficiency level where learners can learn from unspecified input; finally, since second language learners typically have less time for learning than L1 and incidental learning requires a good deal of time, direct vocabulary teaching provides a means of bridging the gap.

With these points in mind there is clearly a place for direct vocabulary teaching. However, there are three principal caveats. Firstly, attention should be directed to towards highly frequent and/or salient vocabulary items. Secondly, direct vocabulary instruction should occupy a small proportion of course time. Thirdly, direct instruction cannot deal effectively with those aspects of language which deal with real word experience and implicit knowledge. Lexicographical resources have a role to play in the direct teaching of vocabulary and in overcoming many of the factors which limit its success.

Having established the validity of teaching vocabulary the pertinent question is how best to go about this. The distinction between productive and receptive skills, although questionable form a theoretical perspective, is widely adhered to in practice (Nation, 2016). That said, it is possible to outline general strategies which apply to both receptive and productive knowledge. According to Nation (2013) a vocabulary learning strategy must involve a choice between several strategies, involve a complex procedure with several steps to follow, entail knowledge and benefit from practice, and ultimately lead improvements in vocabulary use and learning efficiency. However, as Gu, (2003b). points out, the effectiveness of these strategies varies considerably depending on the task, learner, and context. In addition personal attributes such as learning styles also play a role (Gu, 2003a). For example, in a study of on the effect of gender in a French L2 classroom at university level Zoubir-Shaw and Oxford (1995) found that female students are more likely to use guessing strategies than male students.

A number of researchers have attempted to outline taxonomies of vocabulary learning strategies. Perhaps the most extensive taxonomy is Schmitt's (1997) attempt which relies on Oxford's (1990) notion of social, memory, cognitive, and metacognitive categories. Similarly, Gu and Johnson (1996) identify beliefs about vocabulary learning, metacognitive regulation, skilful use of dictionaries, memorisation and activation strategies. They conclude that strategy combinations, rather than individual strategies, may make the difference in students learning. Zhang and Li (2011) adopt a factor analysis approach to arrive at a classification of vocabulary strategies involving cognitive, metacognitive, and affective aspects. Nation's (2013) taxonomy attempts to separate vocabulary knowledge from the sources of that knowledge. The resulting taxonomy consists of four general classes of strategies: planning, sources, processes, and skills in use.

All vocabulary strategy categories: planning, sourcing, processing, and using, include a number of activities in which dictionaries and other lexical resources play an important role. Successful learners employ strategies for choosing which vocabulary to focus on (Gu & Johnson, 1996) and where this vocabulary can be found (Barker, 2007). Dictionaries with frequency information, adequate examples and well selected lists of specialist vocabulary are vital in this regard. Similarly, finding information about words is also imperative.

Dictionaries can play a crucial role in helping students analyse words parts. This is particularly relevant to EAP vocabulary with its Greco-Latin affixes (Corson, 1995) and discipline-specific phraseology. Both these phenomena will be examined in more detail below. Specialist pattern and collocation dictionaries would be particularly helpful as regards the latter. Both monolingual and bilingual dictionaries are also useful in the strategy of establishing parallels with other languages, for example, and in avoiding false cognates. Processing strategies involve noticing, retrieving, and creative use strategies. Noticing strategies are essentially recording strategies, but nonetheless can still be used in conjunction with dictionaries and other lexicographical resources. Retrieval strategies involve recalling previously met items given a cue; they are superior to noticing strategies. Creative use strategies cover a wide range of strategies including semantic mapping, use in context, with collocations, and in sentences. All of which will be examined in more detail in the section on teaching and learning phraseology below. Learners adept at strategies in using skills are able to acquire large amounts of input at the correct level through extensive reading, watching movies, listening to audio recordings and participating in interactive situations. They also take opportunities to produce language adopt strategies to improve fluency. Nation and Yamamoto (2012) argue that students can do this autonomously without help from teachers. Dictionaries and other lexical resources play an important role here.

The question of how well learners use vocabulary learning strategies is a vexed one. This is principally due to problems associated with gathering reliable information about learners' strategy use. One approach involves learners self-reporting on the vocabulary learning strategies they employ via questionnaires or interviews (Barcroft, 2009; Fan, 2003; Gu & Johnson, 1996; Kojic-Sabo & Lightbown, 1999; Sanaoui, 1995; N Schmitt, 1997). This is problematic for a number of reasons, firstly, these interviews or questionnaires take place after the fact, and secondly, learners often say they have used vocabulary learning strategies even if this is not the case (D. Qian, 2004). Another approach involves gathering information about what learners are able to do often via think-aloud tasks as in Ahmed (1989), and Lawson and Hogben (1996) or annotations is vocabulary notebooks (Nation & Moir, 2008). In addition to requiring a great deal of time and resources, such studies are subject to the observer effect. A final approach involves observing what learners actually do when they are unaware of being observed or else unaware of the goal of the observation. This final approach avoids the biases of the previous approaches but is limited insomuch as it can only observe external indications of the strategies being used and provides little insight into the motivations of the learners' choices. Since these studies deal with a wide range of different contexts and students, and employ different methodologies, there is some variation in their findings however it is possible to conclude that successful learners are independent and use resources such as dictionaries well.

Before going on to examine the activities and procedures for learning and teaching each of the receptive and productive skills. It is important to highlight general techniques, activities, and procedures applicable to all four skills.

Good vocabulary-teaching techniques are goal orientated and give learners access to the required knowledge to achieve the goal. This knowledge may come from written or spoken texts, context, reference sources such as teachers or dictionaries (Nation, 2013), and other learners (Newton, 2013). In an effort to find the most effective incidental vocabulary learning activities, Hulstijn and Laufer (2001) turn to the construct of task involvement which is affected by three features: need - the requirement for the use of the target vocabulary in the task, search – the requirement for the learner to search for the meaning

of an item, and evaluation–the requirement that the learner assess the appropriateness of a word choice. A number of studies support the involvement load hypothesis: The greater the involvement load the more effective the learning (Keating, 2008; Kim, 2008; Laufer & Hulstijn, 2001). However, Nation and Webb, (2011) criticize the involvement load hypothesis for its omission of many factors involved in vocabulary learning and instead propose a technique feature analysis checklist which includes awareness, negotiation, repetition, among other factors.

There are certain conditions which facilitate vocabulary learning across all four skills. Key amongst there is motivation, if learners are motivated they are likely to be more exposed to and pay more attention to vocabulary which is likely to increase learning (Laufer & Hulstijn, 2001). Another important condition is noticing (R. Ellis, 1991; McLaughlin, 1990; Schmidt, 1990). Factors which affect noticing are a word's saliency in a text, previous exposure to a word, and the realization on the part of learners that a word fills a gap in their knowledge (Schmidt & Frota, 1986). Noticing also takes place in dictionary look up, deliberate study, guessing from context, and teacher explanation. R. Ellis, Tanaka, and Yamazaki (1994) and Newton (2013) find that when learners negotiate the meaning of a word they are more likely to learn it. There is a great deal of evidence which highlights the positive impact of definition on vocabulary learning in L1 (Brett, Rothlein, & Hurley, 1996; Elley, 1989) and L2 vocabulary acquisition (Hulstijn, 1993; Knight, 1994). The following chapter on lexical resources in EAP contains in depth discussion of defining techniques and styles, including the use of hypertext which has also been shown to be a condition which enhances vocabulary learning (Barcroft, 2002). An interest in and awareness of different component parts, form, usage, and relations of lexis on the part of learners is also beneficial to vocabulary acquisition (J. Scott & Nagy, 2004). Another key condition for the acquisition of vocabulary is the occurrence of retrieval. From a psychological perspective the idea underlying retrieval is that more information is stored than can actually be reported (Baddeley, 1997) and that some cue is necessary in order that reporting occur. From the perspective of vocabulary acquisition, productive retrieval involves the need to communicate a meaning and retrieving the form, while receptive retrieval involves meeting the form and retrieving its meaning (Nation, 2013). A number of studies suggest that repeated retrieval of a word is beneficial to acquisition at least in the short term (Nakata, 2017). It has also been claimed that the creative use of words also promotes learning (Joe, 1995). This creative use involves changes on a word's phraseology. To put this in TNE (Hanks, 2013) terms the exploitation of linguistic norms promotes vocabulary learning. It must be pointed out it is equally possible that this is a confusion of cause and effect and that creative use is in fact a sign of successful acquisition rather than a contributing factor. The levels of processing hypothesis states that "the more deeply and thoughtfully something is processes, the more likely it is to be retained" (Nation, 2013 p.113.). In this regard techniques such as instantiation–the recall of a particular instance or example of the meaning of a word (R. Anderson, Stevens, Shifrin, & Osborn, 1978)–and imaging–the deliberate creation of a visual image to promote learning–promote deep and thoughtful processing of vocabulary and by extension retention, are particularly beneficial to vocabulary learning.

There are a number of teaching activities which promote these conditions. Most involve rich instruction (I. Beck, McKeown, & Omanson, 1987) or rich scripting (McWilliam, 1998), that is, going beyond the immediate demands of the learning task to focus specifically explicitly on the teaching of a vocabulary item. Rich scripting activities include: examining a range of contexts and uses perhaps concordances or collocation dictionaries, semantic mapping activities (Sokmen, 1992), examining word parts and etymology (Ilson,

1983), 'What is it? activities' where learners find L2 synonyms or L1 translations for a word from contextual definitions (Nation, 1990), classification activities where learners place a new word in a lexical set, reflection activities where learners reflect on what it means to know a word. Dictionaries and other lexical resources have an important part to play in all of these activities.

In addition to these formal vocabulary teaching activities vocabulary is often taught via on-the-fly defining in lectures and stories, as part of the deliberate teaching of content in lectures. There are a number of ways in which this can be done; using gesture, pictures, diagrams or realia; via translation; defining and providing examples in L2 this is particularly common in academic contexts and is discussed in more detail in section 2.3 below. The efficacy of various defining styles and techniques in lexicographical resources is discussed in the following chapter.

2.1.1 Learning Vocabulary and Listening

In common with general studies of vocabulary acquisition, studies of vocabulary learning and listening place great emphasis on frequency and coverage. Here again it is important to note that there are many variables which differ between studies, for example, the circumstances of listening, the speed of speech, and the nature of the text. Additionally studies also vary considerably in how they conceive of and operationalise comprehension, all of which this means that the validity of direct comparisons between studies can be called into question.

Van Zeeland and Schmitt (2013) claim that learners need to understand at least 95 per cent of the running words in informal narratives in order to comprehend them a reasonable degree and to have a reasonable chance of guessing the meaning of unknown words from context. Breaking the 90 per cent threshold is not straightforward; the Zipfian nature of word frequencies means the learner has to acquire ever greater quantities of words for ever diminishing gains in terms of percentage coverage of running words. The challenge non-literal meanings present to learners should also be borne in mind. O'Keeffe (2012) demonstrates how in use *dampen*, in the Collins Bank of English, extends beyond the literal sense "to make something slightly wet" to uses with more abstract nouns such as *confidence*, *hopes*, *spirits*, *inflationary pressure* etc. (p. 242). The use of metaphor to describe the abstract in terms of the concrete is important in academic communication, where there is often a need to describe novel abstract concepts in terms of lived physical experiences.

There is some evidence for differences in coverage needed for comprehension between text types. Stæhr (2009) found 98 per cent coverage was needed for academic listening. Cummins (1986) framework of language proficiency could account for this difference. Figure 2.1 shows the approximate position of spoken academic and general face-to-face communication in Cummins framework in relation to its two continua: Cognitive demand and context embeddedness.

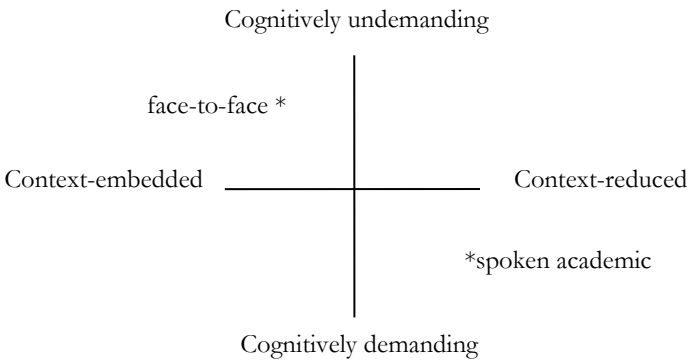

Figure 2.1. Cummins's framework of language proficiency

According to Goh (2000), listening presents four vocabulary related problems for learners. Firstly, learners do not recognise known words when they are spoken. Secondly, learners are unable to 'chunk', break down into individual words, a stream of spoken text. Thirdly, learners cannot form mental representations of words. Finally, learners sometimes understand individual words but are unable to grasp the message. There are a number of ways in which learners can be helped to overcome these problems. Principal amongst these is the provision of written input to accompany spoken text. This might take the form of receptive information transfer activities such as filling in the gaps on a diagram (D. Palmer, 1982) (incidentally, such tasks are a key component of the IELTS Academic listening task) using closed-captions with videos (Sydorenko, 2010); listening while reading (R. Brown, Waring, & Donkaewbua, 2008; Webb & Chang, 2012; Webb, Newton, & Chang, 2013); and holding quiz competitions (Manzo, 1970). Listening skills are also a key component of interactive tasks. There are a number of ways in which learners can be helped with this component. Newton (2001) advocates the use of pre-, during-, and post-task options for dealing with unknown vocabulary. Dictionaries and other lexical references can play a key role in this regard.

2.1.2 Learning Vocabulary and Speaking

Throughout this chapter the adequacy of frequency as a measure of vocabulary need or usefulness has been challenged. However, the majority of studies of vocabulary acquisition commonly recur to this criterion. This is the case with studies of the vocabulary used for speaking. Nation (2006) claims that a much smaller vocabulary is needed for speaking than writing. An obvious counter point to this argument is the influence of topic and degree of formality. A written note to a friend would surely contain less lexical variety than a lecture on history. Nation and Crabbe (1991) claim that 120 word survival list is sufficient for beginner learners who need L2 for tourism of short stays the L2 country. Nation (2013) advocates learning phrases in such lists as single items in order that the items are fluently produced, this is a tacit admission of the limitations of lists of words or word families in isolation.

Fluency is a significant problem in learning vocabulary for speaking. Nation (2013) outlines a procedure for improving fluency with early stage learners which involves a list of target vocabulary to which the teacher points to and the student repeats. The teacher and student then swap roles. Phraseological items play a fundamental role in the promotion of fluency. This will be further discussed in Section 2.2 below.

In addition to fluency, tip-of-the-tongue tests suggest that pronounceability has significant influence on learning difficulty. Learning difficulty depends on the similarity between individual sounds and suprasegmental elements in the L1 and L2, and the relationship between spelling and sound patterns. For Gathercole and Baddeley (1989) learners' ability to retain phonological patterns in short-term memory has a significant influence on vocabulary learning. Papagno, Valentine, and Baddeley, (1991) demonstrate that short term phonological memory as a limiting factor can be overcome by developing meaning based association techniques such as keyword techniques. Gupta and Tilsdale (2009) present a slightly more nuanced view where learning how to say a new word involves a mix of short-term, declarative and procedural memory.

At first glance, the overwhelming importance of pronunciation in the learning of words for spoken production appears to point to a limited role for dictionaries; however, dictionaries can play a key role in techniques for overcoming the limitations of short term phonological memory. Dictionaries can play a supporting role in activities such as semantic mapping, decision making, debates, information transfer, split information task, ranking, cooperation, retelling, and role play all of which can be employed to facilitate vocabulary acquisition through speaking.

2.1.3 Learning Vocabulary and Reading

As with listening, the question of the quantity of vocabulary required for reading is a vexed one. Particularly problematic is the question of what is understood by comprehension, moreover, if a learner does comprehend a text how can one be sure that this is because of what he or she has read and not because of pre-existing extra-linguistic background knowledge. One response to the former point has been to consider the existence of threshold of language knowledge delimiting the boundary for successful comprehension. For Nation (2013) there are at least two views of this threshold. The strong view regards the threshold as an absolute boundary between having enough language knowledge and successfully comprehending a text on one hand, and not having enough knowledge and not comprehending a text on the other. The weaker view of the threshold regards it as a probabilistic boundary; the chances of a learner having sufficient language knowledge for comprehension are proportional to the distance from the threshold. Adopting a probabilistic view Laufer (1989) investigates the relationship between knowledge of word tokens and comprehension. She found that an understanding of 95 per cent or more of the running tokens lead to significantly more 'reasonable' 55 per cent scores in a comprehension test. This was not the case with an understanding of 90 per cent of the tokens. Adopting a strong view, Hue and Nation (2000) find that a vocabulary coverage level of 80 per cent for acceptable comprehension of fiction text since no learner below this coverage level demonstrated adequate comprehension. Schmitt et al. (2011) determine a linear relation between vocabulary coverage and comprehension the greater the coverage the greater the reading comprehension.

Another alternative means of examining the difficulty of a text in reading comprehension is to examine the quantity of different words a learner might encounter within the text. Nation (2006) examines the vocabulary sizes needed to obtain coverage levels of 95 per cent and 98 per cent in various text types. He examined novels, newspapers, writing for children, children's movies, and spoken English. At a 95 per cent coverage level 4,000 word families suffice for all text types except spoken English where 3,000 word families is adequate. At a 98 per cent coverage level figures ranged from 6,000 word families in children's movies to 10,000 word families in writing for children. D. Gardner's (2004) study

of expository, and narrative texts also points to difference in between text types; the former containing more theme-related and so-called academic vocabulary than the latter.

Nation (2013) claims it is vocabulary size rather than text coverage which determines the critical vocabulary factor in comprehension. However, the present study argues that both approaches underplay the importance of context, be that at a phrase, sentence, paragraph, or intertextual level. That said, studies of the relation between vocabulary size, text coverage, and reading comprehension support the axiomatic claim that lack of vocabulary knowledge makes reading difficult.

Reading both necessitates vocabulary and is also a motor for vocabulary learning. Concerned with L1 reading, Perfetti and Hart (2002) posit the Lexical Quality Hypothesis where word knowledge is primordial to reading. They regard the process of reading as a causal circle whereby good vocabulary knowledge facilitates better comprehension which in turn lead to increased input and practice which results in better vocabulary knowledge (Perfetti & Hart, 2001). Figure 2.2 illustrates a more elaborate causal circle taken from Perfetti (2010).

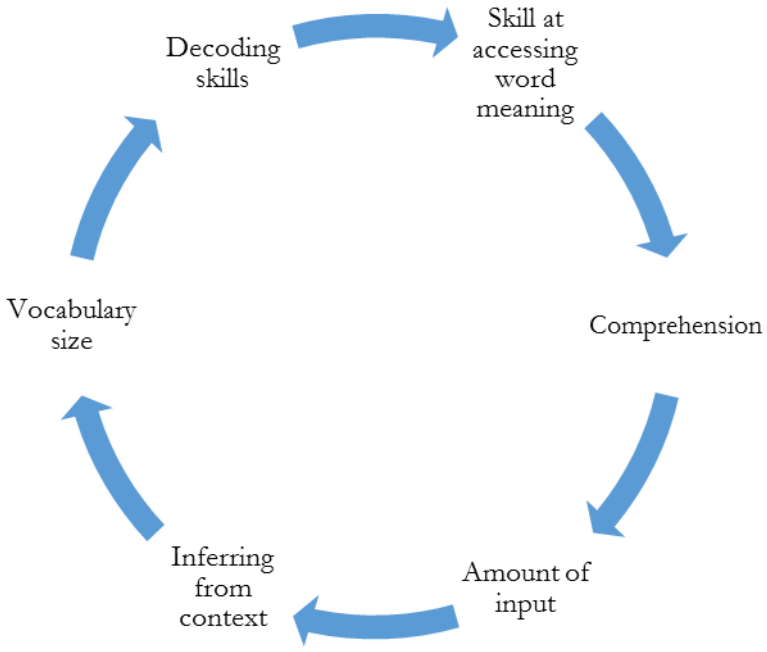

Figure 2.2. Perfetti's (2010) Causal Circle of Vocabulary Learning

In studies of both L1 and L2 language acquisition there is considerable debate about whether extensive reading is sufficient for vocabulary development. There is certainly evidence that learning vocabulary from context alone is possible. Studies of young L1 English speakers have found learning rates of 5 – 10 per cent of unknown words (Nagy, 1985; Nagy, Anderson, & Herman, 1987; Swanborn & de Glopper, 1999). The interval between reading and the administration of the test is a major factor in the variation. In an L2 context Horst, Cobb, and Meara, (1998) found learning rates of around 20 per cent of unknown words in long graded readers. Fraser (1999) found better retention rates for vocabulary inferred from context when reading when this was followed up by consulting a dictionary or an L1 translation or L2 synonym.

The words that learners acquire easily from context are those that represent familiar concepts. Nagy et al. (1987) finds learners have great difficulty learning new concepts. The

instantiation studies of Anderson and colleagues demonstrate that the concept of concrete core meaning is not sufficient for language, instead speakers need contextual help to instantiate. In reading comprehension much of the contextual help can be found in the phraseology of verbs. This is discussed in greater detail in Section 2.2.

It is clear that vocabulary learning can take place through extensive reading, however, acquisition rates are low and require that learners read a great deal of material for very little return. In many L2 learning contexts including many EAP contexts this is simply not feasible. Indeed, Nation (2013 p. 357) advises that "Second language learners should not rely solely on incidental vocabulary learning from context. There needs to be judicious attention to decontextualized learning to supplement and be supplemented by learning from context".

Meaning-focused output and language-focused learning activities are a good means of promoting such decontextualized learning. Just as more deliberate intentional focus on vocabulary results in greater learning than an incidental focus (Hulstijn, 1988) reading with deliberate learning activities is more effective in terms of vocabulary learning than reading by itself (Min, 2008; Paribakht & Wesche, 1996; Sonbul & Schmitt, 2009).

Irrespective of the type of reading, vocabulary acquisition from reading is most effective when texts are graded at the appropriate level. In many EAP and ESP contexts the use of simplified texts is not feasible. In such circumstances there are number of techniques teachers can adopt to support learners. Sutarsyah, Nation, and Kennedy (1994) demonstrate how narrow reading within a strictly defined topic area substantially reduces the vocabulary load, incidentally, it is an extremely practical activity for EAP courses. Elaboration of texts involves the addition of redundant elements to a text and especially clear signalling of the thematic structure using elements such as discourse markers. Some see elaboration as a more appropriate alternative to simplification (M. Long & Ross, 1993; Yano, Long, & Ross, 1994). Negotiation, that is group, or interactive reading has proved productive (Palincsar & Brown, 1986). As indicated in the section on general vocabulary teaching techniques intensive reading and direct vocabulary teaching can be justified for L2 learners. Some studies of pre-teaching have shown positive effects (McKeown et al., 1985). Use of vocabulary exercises with reading texts may also be fruitful. Paribakht and Wesche (1996) provides a framework for analysing vocabulary exercise in relation to the necessary conditions for learning they fulfil. Dictionaries and lexicographical references can play a supporting role in all of these techniques. However, the final technique, glossing, is a more intricately related with lexicographical concerns. The benefits and different styles of glossing will be discussed in the following chapter on phraseology and EAP in lexicography.

2.1.4 Learning Vocabulary and Writing

There is less research about the relationship between the writing skill and vocabulary knowledge than the other three skills. However, it is clear that vocabulary knowledge is an important factor in the perception of writing quality for both teachers and learners. From the teachers' perspective, vocabulary accounts for the greatest amount of variance on Jacobs, Zingraf, Wormuth, Hartfiel, and Hughey's (1981) ESL composition profile (Astika, 1993). University professors perceive lexical errors to be the most severe of all the errors committed by students (Santos, 1988). From the learner's perspective Leki and Carson (1994) claim that L2 learners see lack of vocabulary as having a major negative influence on the quality of their writing.

Read (2000) outlines a number of approaches for measuring the written vocabulary of language learners. One such approach involves discrete-point vocabulary tests. Laufer and Nation (1999) have developed such a test which uses five frequency-based group of words to distinguish between different proficiency levels. An alternative approach involves analysing texts in learner corpora using measures such as type/token ratio, lexical density among others (Laufer & Nation, 1995). Another means for measuring involves rating scales such as the vocabulary component of Jacobs et al.'s (1981) EFL composition profile and various institutional assessment rubrics. An alternative approach involves examining the lexical errors committed by learners in their writing. The latter two approaches are complicated by Ruegg, Fritz, & Holland's (2011) and Fritz & Ruegg's (2013) finding that many raters have difficulty distinguishing lexical errors from grammatical ones. This finding lends weight to a phraseological view of lexico-grammar; a central tenet of the present thesis.

Given that Leki and Carson's (1994) findings about students regarding lexical deficiencies as a major weakness in writing quality and Zhou's (2009) claim that learners are eager to increase their vocabulary size yet are unsure how to do so seem plausible, it is clear that this is an area ripe for further innovation. There is evidence to suggest that vocabulary acquisition across all four skills would benefit from phraseological vision of language. With this in mind, the following section examines important research on the acquisition of phrases.

2.2 Perspectives on the Acquisition of Phrases

In vocabulary acquisition research in general, valid comparison of findings is problematic because of the heterogeneous nature of the groups under study and the methods employed. In studies of the acquisition of phraseology, the situation is complicated further by the diverse and often contradictory notions of what can be considered phraseological. There are an abundance of terms for what are often unsatisfactorily referred to as multi-word units (MWUs). Wray (2000) counts over fifty such terms many of which are overlapping in scope. In English vocabulary research MWUs can be understood in one of four ways: Firstly, they are sometimes understood as strings of words which researchers intuitively judge to frequently occur together, for example, 'most of the time' (Yoiro, 1989). Secondly, they can be understood as collocations in the strict sense of the 'Russian School' (Cowie, 1998) a fundamentally lexicographical approach, like that described in chapter two, where words in collocations take on senses they would not otherwise exhibit or are non-compositional and cannot undergo any transformations such as passivisation or pluralisation. Thirdly, the term can refer to all occurrences of a particular word or part of speech and its accompanying words. This is the approach taken by Granger (1998) in her study of French L1 learners' of English use of adverbs. Finally, MWUs can be identified on the basis of frequency of occurrence of items within the corpus being studied using statistical association methods. Granger and Paquot (2008), Gries (2008), and Kovalyuk (2016) make a distinction between theory driven approaches on the one hand, which include the two former understandings of MWUs, which they attribute to an Eastern tradition, and data driven approaches on the other, which include the latter two, which they attribute to a Western tradition. Walker (2011) demonstrates how these different definitions of MWUs have consequences for the way they are represented in learners' dictionaries. This issue is examined in greater detail in the following chapter.

Taking the different understandings of MWUs into account Nation (2013) outlines five characteristics for MWUs as they are understood in English vocabulary acquisition studies:

MWUs are recognised as variable in nature, however they have a predominant canonical form (Sinclair, 2004a). The combination of the individual words in MWUs is not arbitrary it is consistent with the grammatical and semantic use of the individual words. There are MWUs which are greater than the sum of their parts insomuch as their meaning cannot be fully understood by analysing the individual parts. There is some evidence that MWUs are stored in the lexicon as single choices however, this does not entail that they are not analysed, or that they are fixed. Siyanova-Chanturia, Conklin, & Schmitt (2011) suggest that collocations present clear processing advantage to both native and non-native English speakers. Finally, like single words MWUs have a communicative purpose, knowing this purpose is an essential component of knowing a MWU.

Studies of MWUs in language acquisition suggest that these units play an important role in language. Chapter 1 outlined a number of approaches to language analysis which could be termed phraseological insomuch as they reject the traditional divide between grammar and lexis. It is argued that these approaches better account for the facts of language than traditional generative approaches. The potential of the rejection of the hegemonic position of syntax to language pedagogy has long been recognized. Lewis's (1993) *the Lexical Approach* is essentially a call to subordinate grammar to lexis in language pedagogy. In spite of these insights from lexicography, cognitive linguistics, functional linguistics, corpus linguistics, and developments in language pedagogy, a large tranche of research in English vocabulary acquisition continues to focus on single word items. There is however some research in the field of English vocabulary acquisition studies which suggest a phraseological view of language.

A vision of the lexicon as a continuum, containing neither exclusively single-word items, nor fixed phrases is supported by studies which suggest that MWUs are stored as wholes. Arnon and Snider (2010) study frequency effects for four-word phrases of the type *don't have to worry*. In a study of native-speakers receptive use, they find that users process more frequent phrases faster, in a manner that cannot be reduced to the frequency of the component words. From this finding they conclude that language users store frequency information about whole phrases. Conklin and Schmitt (2008) compare reading times for formulaic sequences with comparable non-formulaic sequences for both native and non-native speakers. They found that both literal and idiomatic formulaic sequences were read more quickly than creatively generated language. Jiang and Nekrasova (2007) use an online grammaticality test to examine L1 and L2 speakers of English's use of formulaic and non-formulaic phrases of the type *as soon as*. The phrases are matched for word length and frequency. They find that results showed that both groups responded to the formulaic sequences significantly faster and with fewer errors than they did to non-formulaic sequences. Vilkaitè (2016) employs an eye-tracking experiment comparing receptive processing of adjacent collocations (e.g. *provide information*), non-adjacent collocations (e.g. *provide some of the information*) and control phrases. She demonstrates that collocations of both types are processed faster than control phrases. Although, adjacent collocations present a greater processing advantage than their non-adjacent counter parts. However, the results of these studies do not imply that knowledge of individual words is not important; the compositionality of MWUs is not arbitrary and the meaning of individual component words often plays a role in the meaning of the MWU (D. Liu, 2010; Walker, 2011). Understanding the meaning of individual parts facilitates the learning of MWUs (Boers & Lindstromberg, 2009; Bogaards, 2001; Webb & Hang, 2016). González-Fernandez's (2016) study of 144 Spanish L1 learners of English makes this relationship between knowledge of MWU and single-words quite clear. She uses a battery of tests designed to evaluate productive and receptive of knowledge of four components of word knowledge: The form-

meaning link, derivatives, polysemy, and collocation. Using a structural equation modelling analysis she finds that collocational knowledge interrelates with the other components of word knowledge.

Since they point to a redundant processing or dual-entry model of the lexicon, to some extent, MWUs account for the puzzle of native-like selection and fluency. Pawley and Syder (1983) propose a dual entry model of the lexicon to explain how native speakers choose a felicitous utterance out of the multitude of grammatical possibilities and under most circumstances produce clauses without hesitations. On their account the lexicon contains both individual words and, lexicalized sentence stems (fixed-MWUs) and semi lexicalized-sentences (variable-MWUs). In a study of second language French learners Towell, Hawkins, & Bazergui (1996) claim that increase in fluency is the result of learners memorising sequences of words.

There is evidence to suggest a usage-based entry model (Bybee, 2006). Several studies have found a relation between the learning of many types of MWU and the frequency with which and sequence in which they are met by the learner. N. Ellis, Frey, & Jalkanen (2009) demonstrate that native speakers recognize frequent verb argument collocations more frequently than less frequent ones. L. Qian (2015) and Siyanova-Chanturia (2015) draw similar conclusions for non-native speakers. N. Ellis, Simpson-Vlatch, and Maynard (2008) examine the relationship between corpus measures of collocation frequency, and mutual information (MI) (coherence) and both learners and native speakers' recognition, articulation and priming of formulaic expressions. They deem both measures to be psycholinguistically valid, however, they find that MI is a better indicator for native speakers while ESL learners' results show a stronger correlation with frequency measures. In an experiment involving artificial language, N. Ellis and Schmitt (1997) demonstrate that repetition of phrases involving novel long-range grammatical dependencies results in acquisition.

Despite these phraseological advances, the single-word approach predominates in English Vocabulary studies. This is particularly true as far as the selection of vocabulary for EFL and ESL courses and material. This is problematic since knowledge of a word's collocation patterns are an essential aspect to knowing about its form, meaning and use (Chang, Chang, Chen, & Liou, 2008; Lewis, 1993; Sinclair, 2004a; Stubbs, 2009).

Nation (2013) sets out some general guiding principles for the teaching and learning of MWUs. These include learning the principal components of the MWUs and how they can be varied in the unit. Examining how the meaning of the parts contributes to the whole. He highlights the importance of fluent recognition of and access to MWUs and the need to examine the use of well-formed MWUs in context. Lexicographical materials, with adequately selected and exemplified MWUs are essential in this regard.

This raises the question of how to select the MWUs to be taught. In practice, the selection of MWUs for teaching purposes has been rather haphazard. For example, Koprowski (2005) highlights the unprincipled approach to the selection of MWUs in ELT course books. With the exception of everyday phrases such as greetings, O'Keeffe et al. (2007) highlight a tendency to treat MWUs as the domain of higher-level learners even though many MWUs are at least as frequent as commonly taught words at lower-levels. In a study of MWUs in 24 Spanish and English textbooks López-Jiménez (2010) found that only 8.3 per cent of textbooks make the vocabulary selection criteria explicit and that treatment of MWUs in English textbooks is particularly scarce at lower levels. In theory, it is possible

delimit three principal approaches to the selection of MWUs: Form-based, meaning-based, and storage-based.

Form-based approaches are usually corpus-based. For bigrams Mutual Information (MI) (Church & Hanks, 1990), a score which compares the probability of two words occurring together with the probability of them occurring individually, is often employed. However this is problematic as MI tends to overestimate the significance of rare words and as is sensitive to data sparseness, thus it is best suited to use with very large corpora (Kilgarriff, 2001; Manning & Schütze, 1999). Cobb's (n.d.) *Compleat Lexical Tutor* and a number of other concordancing programmes aimed at classroom employ such measures. The frequency based selection of n-grams is more common in English teaching. In this field it is better known as the 'lexical bundle' approach (see for example Biber, Conrad, & Cortes, 2004). This approach involves searching a corpus for string of n words in length which occur above a certain frequency level. It is problematic because no semantic criteria are employed in this search, thus phrases with little teaching value such as *it is a* are often selected. Studies which employ this n-gram approach often have a second phase in which a human analyst sorts the output. Non-contiguous MWUs, for example phrasal verbs such as *look up something / look something up* present further problems for the lexical bundle approach. Cheng, Greaves, and Warren (2006) address this problem with variations on the lexical bundle approach which they dub skip-grams and concgrams. In general English Biber, Johansson, Leech, Conrad, & Finegan's (1999) *Longman Grammar of Spoken and Written English* places a great deal of emphasis on lexical bundles. Many EAP/ESP vocabulary studies use the lexical bundle approach. The advantages and disadvantages of this approach are outlined more fully in Section 2.3 on EAP and vocabulary teaching.

Meaning-based approaches usually involve the criterion of compositionality. Grant and Bauer (2004) make a distinction between core idioms, where the meaning of the parts of the MWU is not clearly related to the whole, figurative MWUs, and literal MWUs where the meaning is largely obvious from the parts. There is evidence for the influence of L1 on accurate and inaccurate idioms in L2 (Nesselhauf, 2003). Thus for language teaching purposes, Nation (2013) proposes applying the criterion of the existence of word-for-word parallels in the learner L1. MWUs which have no L1 parallel cause a great deal more difficulty for learners (Wolter & Gyllstad, 2011; Yamashita & Jiang, 2010). Idiom dictionaries are good examples of the application of the meaning based approach to MWU selection.

Storage-based approaches are concerned with the psychological representation of MWUs. Wray's (2002, p. 9) definition of formulaic sequences, as sequences that are "stored and retrieved whole from the memory at the time of use", hints at such an approach. The problem here is one of reliability and validity of evidence. Since direct observation of psychological phenomenon is not possible other identification and classification procedure must be used. Read and Nation (2004) discuss procedures involving intuition, corpus analysis, structural analysis, phonological analysis, and pragmatic functional analysis. They highlight the inherent advantages and pitfalls of each. In ELT the Phrasal Expression List (Martinez & Schmitt, 2012) is an attempt to identify morpheme equivalent MWUs, their starting point is a corpus-based search for frequently occurring n-grams. They then judge these strings using a criteria of morpheme equivalence, semantic transparency, and deceptive transparency (Laufer, 1989a) – "words learners think they know but do not" (Laufer, 1989a p.11); for example learners might hypothesise that *for some time* means 'for a short amount of time'. These three criteria are not obligatory. They are essentially used to confirm intuitive judgments.

Once MWUs have been selected for English courses the obvious question is how they should be taught and learnt. The following paragraphs will outline some general concerns regarding the teaching and learning of MWUs before going on to examine how they might be best taught and learnt in relation to the four skills of reading, listening, speaking and writing.

There is some evidence to suggest that idiomatic MWUs are more easily learnt than single word items. In a reading and translation test of Dutch native speaker learners of French, Bogaards (2001) found that learners are better able to retain idiomatic MWUs made up of well-known component words than unknown single-word items. These results relate to idiomatic MWUs, for example *faire la tête* (literally "make the head", which could be translated as "to sulk"), the extent to which they are generalizable to other types of MWUs is questionable. He also found that new meaning-related senses of lexical items are better retained than totally new lexical items. On the other hand, there is no difference between the retention of meaning-related and non-meaning-related senses of familiar forms.

Ackermann and Chen (2013) argue that productive use of collocations is more difficult than receptive use. As with learning single words there is evidence that MWUs can be learnt in context, especially if their component parts are known. Many studies have shown positive effects of repeated encounters. Use of lexicographical techniques such as underlining, bolding, colouring or glossing to capture students attention have been demonstrated to be effective, as has decontexutalisation. Explicit teaching and deliberate learning are also effective with MWUs. Nation (2013 p. 499) highlights five principal focuses for the deliberate learning of vocabulary in general and describes how they might be applied to MWUs.

An important focus of deliberate teaching involves encouraging noticing. O'Keeffe et al. (2007) suggest carrying out repeated listening activities some aimed at comprehending content others aimed at noticing chunks. Dictation related exercises are also useful in this regard as they involve chunking and holding chunks in working memory. These are not limited to listening and writing skills, 'read-and-look-up' (M. West, 1960), and 'delayed copying' (Hill, 1969) activities can also involve reading and speaking. Brainstorming and matching exercises in groups (D. Brown, 1974) can also prove useful in encouraging noticing of MWUs. However, Boers and Lindstromberg (2009) criticise MWU matching exercises which provide learners little guidance on how complete the task. They stress the importance of making students aware of the context of the source domain as well as formal factors. Another focus of deliberate learning involves rote learning. Nation (2013) argues that the procedure for rote learning of MWUs should be no different to individual words. He argues that Elgort's (2011) claim that rote learning of individual words produces both explicit and implicit knowledge is equally applicable to MWUs. A focus on mnemonic devices in MWUs facilitates the learning of the form and meaning of MWUs. Those MWUs which exhibit word repetition, rhyme, slant rhyme, alliteration, assonance, or consonance are likely to be learnt if learners notice these features (Boers and Lindstromberg, 2009). A focus on patterns within multiword units can also facilitate learning. Dilin Liu (2010) argues that MWUs, specifically the components of collocations, do not occur together arbitrarily but instead occur in accordance with their semantic and syntactic patterning. She attributes the difficulty L2 learners have in accurately using collocations to the difference in patterning as regards collocations in their L1 and advocates a pedagogical approach to collocations based on corpus data. Staying with collocations, Walker (2011) examines the collocational behaviour of semantically related verbs and nouns in a corpus of business English. He finds that many collocations can be explained

by the semantics of the individual words, semantic prosody, and metaphor. He contends that learning of collocations is facilitated if students notice such patterns. The use of concordancers is widely recommended as a means of drawing learners' attention to collocation pattern (Chang et al., 2008).

The final focus which benefits the learning of MWUs is the development of interpretative strategies. Such strategies are particularly useful with figurative MWUs. Grant and Boer (2004) and Grant and Nation (2006) outline a procedure for learning such items. The first step involves understanding the figurative meaning, the next understanding the literal meaning, the final step involves learners working out how the two meanings are related. This strategy is facilitated if learners are made aware of the domain from which the MWU originates, the possibility of false cognates, and are familiar with common metaphoric themes (Boers & Demecheleer, 2001). Dictionaries are particularly useful in this regard.

Drawing learner's attention to metaphor is beneficial to learning MWUs. In a study of Chinese L1 learner use of English phrasal verbs Liao and Fukuya (2002) find that learners tend to avoid figurative phrasal verbs to a greater extent than literal ones. Walker (2011) finds that verbal and nominal collocations can be explained by examining the metaphorical processes which they have undergone. Boers (2001) demonstrates that hypothesising about the origin of figurative MWUs benefits learning. Explanatory feedback after this hypothesising benefits learning further (Boers, Demecheleer, & Eyckmans, 2004). In a CALL procedure the identification of the origin of the figurative MWU, followed by learning its figurative meaning, then using it in a gap fill task was the most effective sequence (Boers, Eyckmans, & Stengers, 2007; Boers & Lindstromberg, 2009). This necessarily involves lexical resources which display MWUs in diverse contexts.

2.2.1 Learning Phrases and Reading and Listening

As with single words MWUs are an essential component of reading ability. At the same time, reading is also one of the principal motors for MWU learning. Kremmel, Brunfaut, & Alderson, (2016) study the influence of MWUs on reading comprehension test scores of 418 EFL learners. They find that a measure of MWU knowledge is a better predictor of reading comprehension score variance than traditional syntactic measures. They hypothesise that phraseological knowledge plays a critical, yet traditionally underestimated role in foreign language reading ability. Nist and Olejnik (1995) demonstrate that learners are better able to select the correct example of a use of a word when they see a definition and an example of the word in context. Prince's (1996) findings suggest that learning from phraseological context is of greater benefit to advanced learners than beginners. Several cloze-test based studies (Chihara, Oller, Weaver, & Chavez-Oller, 1977; Rye, 1985) indicate that clues for guessing word meaning from context are obtained from the immediate context, that is to say phraseology. Richard Anderson and colleagues (R. Anderson & Ortony, 1975; R. Anderson & Shifrin, 1980; R. Anderson et al., 1978) highlight the role of linguistic context, which includes a word's phraseology, in the creation of instantiations in readers and listeners. For them a word has no set core meaning but rather a range of meanings which are activated by context. To a certain extent this chimes with Hoey's (2005) Lexical Priming and Hank's (2013) TNE. Amnes (1966) provides a classification of context clues these were ranked by Rankin and Overholser (1969). The majority of these clues depend on a word's phraseology.

Shin (2009) claims listening provides more favourable conditions for incidental learning of MWUs than reading because spoken language makes use of MWUs more frequently than

written language. This claim is only valid with a very narrow definition of MWU; sequences such as *as well, you know, as much, I think* which occur with high frequency in speech. It does not apply to the present study which focuses on verb + noun collocations.

2.2.2 Learning Phrases and Speaking and Writing

Many theories of fluency attribute a primordial role to MWUs (Schmidt, 1992). From a theoretical standpoint this involves a full-entry model of the lexicon whereby, as in Goldberg's (1995) Construction Grammar, information is specified at all levels of the taxonomic hierarchy. There is evidence for the value of substantial repetition of MWUs in improving fluency in speaking (Durrant & Schmitt, 2010) and writing (Durrant & Schmitt, 2009). At a practical level the notion of exploiting MWUs to develop quick initial competence and fluency is nothing new. Take for example the use of holiday phrase books (H. Palmer, 1989). In a study of young Turkish L1 learners of English, Gheitasi (2016) finds that formulaic sequences are used strategically to economize processing effort and buy processing time. Wray's (2004) examination of an L1 English speaker giving a cookery class in Welsh, demonstrates the usefulness of formulaic sequences for quickly learning and fluently producing the language necessary in strictly defined situations. In this situation the typical early stage learner errors which occurred suggest that to some extent adult learners analyse MWUs. This might be considered yet more support for a dual entry model.

The majority of research on the acquisition of MWUs in English tests written production to some extent. In many cases this is not the principal aim of the research but rather an artefact of the procedures used to measure other skills. For example, acquisition and reading and listening are often measured via written tests. As far as deliberate testing of written production of MWUs is concerned, Čolovic-Markovič (2016) investigates the influence of explicit instruction on ESL learners' ability to produce topic-induced word combinations. She found that those learners who had received instruction significantly outperformed those who did not, in the use of topic-induced word combinations in an essay writing task. Schmitt and Redwood (2011) found a moderate collocation between a cloze-test score for productive use of phrasal verbs and their frequency of occurrence in the British National Corpus (BNC) (Oxford Text Archive, 2015)[3]. They conclude that the learning of phrasal verbs is related to their frequency of occurrence. However, this relationship is not particularly strong.

Prodromou (2008) indicates that native-like speech is characterised by the presence of MWUs. These chunks often have discourse-marking functions, serve as hedges to protect face and demonstrate politeness, or as markers of vagueness used to avoid sounding blunt or pedantic (O'Keeffe et al., 2007). Examples of MWUs with all these functions can be found in Shin and Nation's (2008) list of collocations in spoken English. O'Keeffe et al. (2007) highlight the importance of drilling and noticing activities in the adquisition of these types of chunks.

There are a vast number of studies which employ learner corpora to analyse learners' use of formulaic language. The majority of these learner corpora contain university student writing and fall within the remit of EAP, this perhaps because of these ease with which researchers working in tertiary education institutions can access such data. Perhaps the best known learner corpus is *The International Corpus of Learner English* (Granger, Dagneaux, Meunier, & Paquot, 2009) which contains argumentative essays written by L2 English

[3] Henceforth all mentions of the BNC refer to this resource.

students at university. Research employing this and other learner corpora will be discussed in the next chapter section on the acquisition of vocabulary in EAP.

2.3 The View from EAP

In their quest to delimit academic vocabulary many researchers have turned to frequency measures. The idea being that there is a group of 'semi-technical' or 'academic' vocabulary which lies between the 2,000 to 3,000 most commonly occurring words in English at one pole, and specialized technical vocabulary at the other. Proponents of this approach, point out that academic vocabulary is common to a wide range of academic texts (Barber, 1962; Campion & Elley, 1971; Coxhead, 2000; Praninskas, 1972; Xue & Nation, 1984). They also point out that it accounts for a reasonable number of words in academic texts. For example, there have been many studies of the coverage percentage of Coxhead's (2000) AWL. Cobb and Horst (2004) find coverage of 11.6 per cent in the Learned section of the Brown corpus (W. N. Francis & Kučera, 1979); Hyland and Tse (2007) find coverage of 10.6 per cent in a corpus Science and Engineering texts by professional and student writers; Q. Chen and Ge (2007) find coverage of around 10 per cent in medicine research articles; Konstantakis (2007) finds coverage of 11.5 per cent in business; Coxhead and Hirsh (2007) find 8.9 per cent in Science; J. Ward (2009) 11.3 per cent in engineering; Martínez, Beck, and Panza (2009) 9 per cent in agricultural sciences research articles; Vongpumivitch, Huang, and Chang (2009) 11 per cent in applied linguistics research articles; Li and Qian (2010) 10.5 per cent in finance.

A few studies have indicated that learners encounter greater difficulty with sub-technical vocabulary than with technical vocabulary (J. Anderson, 1980; S. Li & Pemberton, 1994; Trimble, 1985). Nation (2013) indicates that in contrast to specialist vocabulary which requires specialist background knowledge, academic vocabulary is "the kind of specialized vocabulary that an English teacher can usefully help learners with" (Nation, 2013 p. 293).

It is true that much of so-called academic vocabulary has an important role as discourse markers, variously called, 'discourse organisers' (Biber, 2006) 'stylistic formulas' (Yoiro, 1989), 'connectors' (Crewe, 1990), 'linking expressions' (Sinclair, 2004a) etc. Although there is some evidence that they occur with different frequencies and serve different purposes in different academic disciplines (Hyland, 2008b), these expressions occur frequently across a variety of academic texts and can be taught without recourse to specialist knowledge. As such there is little doubt about their teaching worth. However, beyond this restricted group the claims of proponents of a frequency based distinction between three levels of vocabulary are questionable. Trimble (1985) argues that academic vocabulary takes on extended and distinct meaning in different technical contexts. For example, *fast* in medicine has the sense of 'resistant to' while in mining it is used to describe solid ground. However, M. Wang and Nation (2004) dispute the influence of homography and homonymy in academic vocabulary. In a study of AWL words they found that while 60 of the AWLs word families contained homographs only three did not meet the list's range and frequency criteria. In contrast, the precursor to the present study (Rees, 2013), a study of the collocational behaviour of 57 AWL families across three academic disciplines found that 49 headwords exhibited different meanings in different disciplines and that many of these meanings would be problematic for learners accustomed to the general use. There are a number of factors which explain the discrepancy between these results. Firstly, M. Wang and Nation (2004) make a sharp and arguably unjustified distinction between homography and polysemy. Secondly, the extent to which their corpus represents academic language has been called into question (Durrant, 2013; Hyland & Tse, 2007). Thirdly, their recourse to a

dictionary of general English, *The New Oxford Dictionary of English* (Pearsall, 1998) results in definitions that are necessarily general and thus cannot reflect discipline-specific academic use.

On a practical level, a consequence of the emphasis on frequency has been a proliferation of frequency-based academic wordlists which respect the general, sub-technical, technical distinction. Campion and Elley (1971) is the first attempt to create a list of academic vocabulary with reference to real academic language, and like Praninskas (1972) is based on a manually compiled corpus of academic language. Coxhead's (2000) *New Academic Wordlist* was the first to be compiled using a machine readable corpus. *The New Academic Vocabulary List* (D. Gardner & Davies, 2014) employs a significantly larger machine-readable corpus and is determined on the basis of lemmas not word families. In contrast, some researchers do not employ the general/sub-technical/technical distinction. Lynn (1973) and Ghadessy (1979) note those words which students have annotated in their textbooks. Xue's and Nation's (1984) *University Wordlist* is the amalgamation of these notation lists and Campion and Elley's (1971) and Praninskas's (1972) manual corpus work. Several authors have rejected frequency based distinction in the selection of academic vocabulary. Ward (1999) argues that the three way distinction is not necessary for learners who have clear specialised goals. Taking another point of view, Paquot (2010) includes high-frequency words which play key a role in structuring academic texts in her *Academic Keyword List*.

More recently there has been a great deal of interest in the creation of discipline-specific wordlists; in engineering (J. Ward, 2009), agriculture (Martínez et al., 2009), environmental studies (J. Liu & Han, 2015), nursing (M. Yang, 2015), and medicine (Lei & Liu, 2016).
Further weight is given to the need to take academic discipline into account when selecting vocabulary for EAP by Sutarsyah et al. (1994). In the economics textbook examined in the study, 34 words accounted for 10 per cent of the total number of words, and the function word *you* occurred with higher frequency than in a general corpus. There are also a wealth of surveys non-native speaker English university students which highlight the problems posed by discipline-specific vocabulary (Berman & Cheng, 2010; Evans & Green, 2007; Evans & Morrison, 2011; Wu & Hammond, 2011).

The trend towards more discipline-specific academic wordlists is a positive advance. However, single-word wordlists cannot take phraseological concerns into account, nor can they deal with a great deal of discourse organizing vocabulary which plays a key role in academic language. These concerns have led to the creation of various lists of academic MWUs. Some deal with general academic vocabulary (Ädel & Erman, 2012; Biber & Barbieri, 2007; Biber et al., 2004; Kashiha, 2015; Pérez-Llantada, 2014). Others provide more discipline-specific lists of MWUs. For example, Cortes's (2004) study of lexical bundles in history and biology, Gledhill's (2000) study of collocations in science writing, Grabowski's (2015) study of lexical bundles in pharmaceutical articles, Marco's (2000) study of collocational frameworks in medical research papers, Le and Harrington's (2015) study of lexical bundles in applied linguistics research articles, Ward's (2007; 2009) study of collocations in Engineering, Hsu's (2014) formulaic sequence list for prospective EFL business studies students, William's (1998) study of collocational networks in plant biology research articles, and H. Yang's (1986) study of multiword terms in science texts.

As is the case with wordlists in general, the hegemony of frequency in vocabulary studies limits the usefulness of many of these MWU lists. The majority of such lists are created using an n-gram approach which in English vocabulary studies is known as the lexical bundle approach. Although this approach is straightforward to implement and has resulted

in interesting findings, it is limited insomuch as it only returns strings of words of a given length occurring above a given frequency threshold. Semantic considerations are not taken into account. In short, the consequences are phrases such as 'to do with the' or 'I think it was' which are intuitively not relevant or in Simpson-Vlatch and Ellis's (2010) terms not "psycholinguistically salient". Simpson-Vlatch and Ellis (2010) mitigate the problem of nonsensical collocations and those not worth teaching time by applying a teaching-worth measure. This results in more useful output but cannot account for the semantic information lost in the initial selection process. Similarly the *Academic Collocation List* (Ackermann & Chen, 2013) employs a mixed-method approach of statistical collocation and expert judgment to derive a list of lexical collocations which they claim is immediately useful to EAP teachers.

The lexical-bundle approach is also employed in a large body of research which exploits learner corpora. For reasons including ease of access to data, several of these corpora comprise writing and speech produced by university students as part of their studies thus, in broad terms, can be considered as attempts to represent academic language. Chief amongst these corpora is the *International Corpus of Learner English* (ICLE) (Granger et al., 2009). This has been used in studies of L2 error analysis and avoidance, particularly in regards to phrasal verbs. Inspired by Menachem and Laufer's (1985) study of error and avoidance with phrasal verbs by Hebrew L1 students of English, M. Chen (2013) employs the ILCE to compare Chinese, British and American students use of phrasal verbs; Mazaherylaghab (2013) looks at Iranian students and; Waibel (2007) German and Italian students.

Learner corpora are sometimes used as a basis for the vocabulary selection and wordlist evaluation. Durrant (2016) employs the British Academic Written English corpus (BAWE) (Alsop & Nesi, 2009) to evaluate D. Gardner & Davies's (2014) *New Academic Vocabulary List*. The use of learner corpora for vocabulary selection and wordlist evaluation is problematic since student writers are by definition writers in training and do not have the full range of linguistic resources necessary for their task. In the present study the view is taken that corpora containing texts written for publication represent a more adequate basis for the study of EAP lexis. The arguments supporting this view are presented in greater detail in Chapter 4.

Both the use of general-academic wordlists and the lexical-bundle approach might be motivated by practical concerns such as ease-of-use and interpretation rather than pedagogical concerns. This speculative point is backed up by Nation's assertion "that academic vocabulary the kind of specialized vocabulary that an English teacher can usefully help learners with" (Nation, 2013 p. 293), while the teaching of what is traditionally known as technical vocabulary is considered as outside teachers remit (Barber, 1962; Cowan, 1974; Higgins, 1966) since they do not have the body of technical knowledge within which to situate the term in a meaningful way. This fact, coupled with the crucial role such vocabulary plays in reading highlights the need for lexicographical resources with genuine 'technical' usage examples.

Traditional views of academic vocabulary emphasize its frequency of occurrence. However, frequency is not the only factor which has been considered. Corson (1985, 1997) posits the existence of a lexical bar which inhibits access to the great deal of academic vocabulary with Greco-Latin roots for learners whose L1 does not contain such roots. Others have attempted to define academic vocabulary in terms of the functions it carries out. For example, Strevens (1973) highlights the functions of discrimination, classification, inter-

relation, and explanation, while Martin (1976) presents a classification of vocabulary by its use in the research process, analysis, and evaluation. Meyer (1990) suggests that in academic contexts words such as *concerning*, *fact*, and *process* are becoming delexicalised and behave like function words. Similarly, McCarthy (1993) and McCarthy and Carter (1994) demonstrate how certain discourse patterns are associated with certain 'discourse-organizing words' such as *question*, *issues*, *assumption*, *position*, *case* and *situation*.

The findings of studies on the acquisition of EAP vocabulary have implications for the view of EAP and parts of speech under study in the present dissertation. The view of academic vocabulary taken in the present corpus-based study is not limited by the general, sub-technical, technical distinction. Instead, academic vocabulary is considered to be any vocabulary that occurs reasonably frequently in academic texts. The decision to concentrate on verbal collocations rather than those involving other parts of speech is supported by the delexicalisation process highlighted by Meyer (1990) and the existence of academic discourse-organizing words (McCarthy, 1993; Paquot, 2010). It is assumed that verbal collocations are more problematic for learners than function words.

The question of how to teach academic vocabulary depends whether its intended purpose is productive or receptive. However, Nation (2013 p. 298) proposes a three-step approach to the sequencing of academic vocabulary. The first step involves introducing learners to the 180 wide-range high-frequency items in the AWL through careful selection or simplification of texts. The second step involves 15 unmodified texts which would cover another 180 words of the AWL. The final step involves learners independently extensively reading large amounts of academic texts from their subject areas and beyond. This could be supplemented by the decontextualized learning of AWL items. The keeping of some kind of vocabulary record is also recommended, however, Leeke and Shaw (2000) in their study of the vocabulary learning habits of post-graduate EAP students in England, found that only around one third of students kept such a record.

Nation regards direct vocabulary learning key to non-native-speaker students being able to compete with native-speaker at an English medium university. Across all four skills, there is an emphasis on learning with activities which reflect academic contexts (Nation, 2013). For example reading, or listening to academic texts. Flowerdew (1992) and Bramki and Williams (1984) highlight the important role of definitions academic lectures and textbooks. In a study of EAP students File and Adams, (2010), found that isolated vocabulary instruction before reading lead to greater retention than vocabulary instruction integrated with reading. Both methods of instruction were more effective than incidental instruction. As for speaking, activities should reflect academic, presentations or discussions. Nikhat Shameem and Alison Hamilton-Jenkins (Peacock & Flowerdew, 2001) developed the idea of issue logs with the aim of promoting fluency. Issue logs involve a learner choosing a topic on which they become experts of the course of a pre-sessional EAP course. They improve their fluency with specialized vocabulary by providing oral and written reports on their chosen topic throughout the course. As far as writing is concerned, practice writing formal texts, literature reviews, and written summaries all provide students the opportunity to use vocabulary in an academic context. Dictionaries and other lexicographical resources can play an important role in all of these activities. The following chapter will discuss the treatment of EAP and phraseology in these resources.

3. PHRASEOLOGY IN LEXICOGRAPHICAL RESOURCES IN EAP TEACHING AND LEARNING

3.1 Introduction

The review of the literature in the previous chapter suggests that dictionaries play an important role in the teaching and learning of EAP. Having established the worth of dictionaries in EAP, the principal concern of current chapter is the treatment of phraseology in dictionaries and other lexical resources used in EAP teaching and learning. Before dealing with this specific concern, a more general examination how dictionaries are used in second language acquisition is undertaken. This is followed by a discussion of the representation of phraseology in dictionaries, and finally a look at issues specific to EAP lexicography. The present chapter will focus primarily on monolingual learners' dictionaries (MLDs) since these are the monolingual dictionaries most frequently used in the EAP context in production (Gilquin, Granger, & Paquot, 2007; Paquot & De Cock, 2010). However, in order to provide a more complete panorama of dictionary use in EAP, occasional mention will be made of research into other lexicographical resources including, collocational, bilingual, and bilingualised dictionaries, as well as term banks, and glossaries.

As an ante-step, it will be useful to situate this discussion in the wider context of research in lexicography. Hartmann (1987, p. 12) identifies four categories of research in lexicography:

1. Dictionary typologies – research into the categories of linguistic information presented in dictionaries and its relevance to particular user groups.
2. User typologies – research into groups of dictionary users and their expectations of dictionaries.
3. Need typologies – research into the contexts and purposes of dictionary use.
4. Skills typologies – research into look-up strategies and programmes and aids designed to promote these strategies.

In the MLD context Cowie (2002) also identifies two further types of research:

5. Comparisons of the merits of bilingualised dictionaries compared with MLDs
6. Studies considering at which stages of the learning process different types of dictionaries should best be employed.

The present chapter focuses, to varying extents, on all except the second of these categories since some assumptions are made about the nature of EAP and its users.

Defining the nebulous construct of EAP is far from straightforward. With reference to Flowerdew and Peacock (2001), and Jordan (1997), Hyland (2006) provides a traditional definition of EAP as "as teaching English with the aim of assisting learners' study or research in that language" (p. 1). In this broad sense it involves teaching from pre-tertiary (here understood as pre-sessional preparatory courses) to postgraduate levels, studies of classroom interactions, research genres, student writing, and administrative practice. Recent years have seen calls to broaden this definition and recognise the importance of school settings in the development of academic literacies (see Humphrey, 2016). The present study adopts the traditional view of EAP, concentrating on teaching and research that takes place

at universities and other similar higher education institutions. It might be suggested that the Teaching of English in Academic Contexts (TEAC) would prove a more appropriate label since it takes into account those subjects dealt with at university which have not traditionally been considered academic. However, in the present study the term EAP is used to describe this context in the same way it is frequently used elsewhere.

Even in this reduced context, given the wide range of situations in which EAP is used it would be unwise to make overarching generalisations about its users and uses. However, in the limited context of teaching and research at university, and leaving aside differences between national contexts, it is possible to make generalisations about the typical proficiency level of its users at each stage of study.

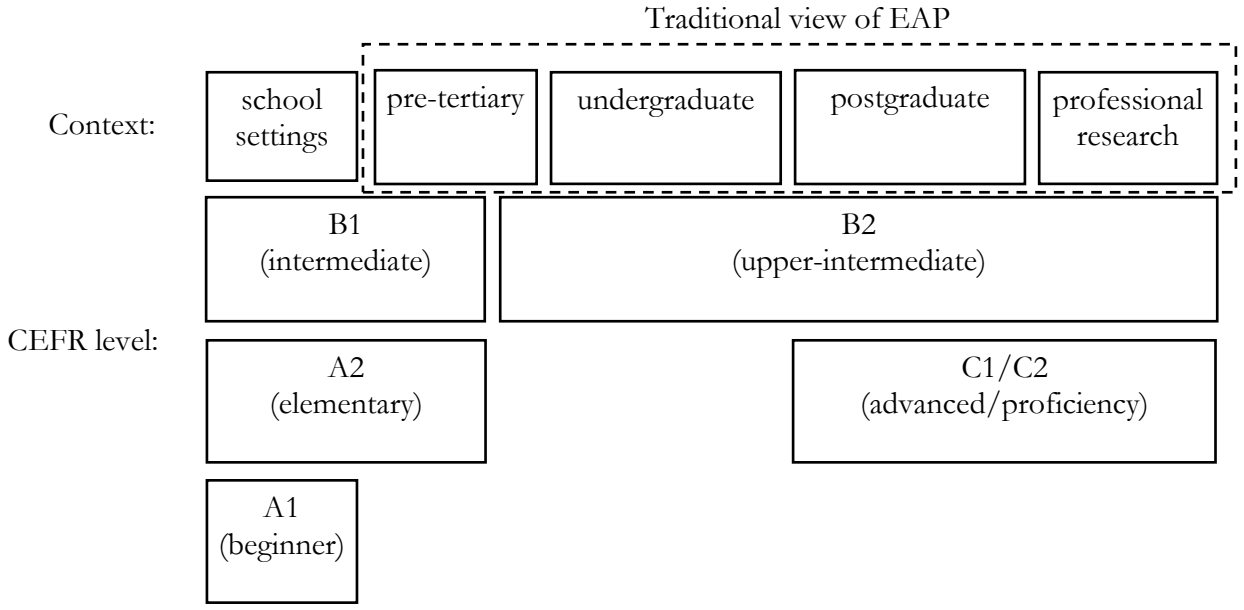

Figure. 3.1. EAP contexts and typical corresponding proficiency levels

Figure 3.1 above gives an overview of the contexts in which EAP is used and the typical proficiency levels of non-native-speaker-of-English users in these contexts. The proficiency levels (CEFR) shown for the traditional EAP contexts pertain to British universities. For pre-tertiary to postgraduate levels they were obtained by taking the typical standardised language assessment test scores (predominantly IELTS) required for entry to British universities (pre-tertiary IELTS 4.0 – 6.0, undergraduate IELTS 6+, postgraduate IELTS 6.5+) and converting these scores to CEFR reference levels using the scheme provided by Cambridge English Examinations (2017)[4]. This approach is far from ideal; firstly, the IELTS test predates the CEFR and therefore was not designed to map to the reference scale. Secondly, CEFR has some limitations when it comes to the assessment of academic language (de Jong & Benigno, 2016). These include a disproportionate weight attributed to speaking skills, an inability to make fine-grained distinctions at low (A1) and high (C1/C2) levels and most importantly; the fact that it was not designed to measure proficiency with academic language.

As far as non-native speaker professional researchers are concerned, publication in English has an increasing impact on the career advancement of academics from many countries.

[4] All estimations of proficiency levels given by the author henceforth were made with reference to this scheme

This is due in part to the emphasis put on publication in prestigious international journals by research evaluation systems (see Pérez-Llantada, Plo, & Ferguson, 2011 for discussion of the Spanish case). This has led to an increased interest in English for Research and Publication Processes (ERPP) at both a theoretical and pedagogical level (Flowerdew, 2013). Based on my personal experience, in this study it is assumed that most professional researchers who publish in English are able produce texts of at least a B2 level before proofreading. This does not mean to say that there are no users in this group with a higher proficiency-level. Assessing the proficiency level of professional academics whose only regular contact with English is in written form can be challenging since they sometimes display spikey profiles with productive skills often trailing behind receptive skills.

In the wider-EAP context, it is more difficult to identify the typical proficiency level of EAP users. For those countries with published and widely disseminated attainment targets in English for school leavers the minimum level is often specified at B1. However, in many cases the legitimacy and validity of such a target has been called into question, and in reality such targets are frequently not obtained (Díez-Bedmar, 2012).

In spite of these limitations, mapping typical proficiency levels of EAP users to typical EAP contexts allows the creation of a sampling frame to sort lexicographical studies which might reasonably be considered as pertaining to EAP from those which pertain to general English. It also allows the approximate pairing of lexicographical user-studies with proficiency level. This is necessary since many such studies take place in university contexts but do not provide clear information about the proficiency level of participants. It must be stressed, however, that the accuracy of this pairing is limited by differences between countries and the types of language courses studied. In the present study, it is assumed that all English language teaching which takes place at university is covered by the term EAP.
Having outlined how the present study relates to the wider context of EAP, it would be remiss not to situate it in the context of the historical evolution of MLDs. However, identifying historical stages of development of dictionaries is a matter of perspective. Cowie (2002) identifies three generations of learner dictionary whereas, more recently, Yamada (2013) posits five periods of development. The classification outlined below divides the development of MLDs and related research into their uses and users into five broad stages.

The first stage of MLD development is closely tied to the Vocabulary Control Movement. Dictionaries produced in this stage include the *New Method English Dictionary* (NMED; West & Endicott, 1935), *A Grammar of English Words* (GEW; H. Palmer, 1938), and the *Idiomatic and Syntactic English Dictionary* (ISED; Hornby, Gatenby, & Wakefield, 1942) which was later reissued under the title: *A Learners' Dictionary of Current English* (ALD1; Hornby, 1948). The GEW and ISED are primarily concerned with encoding, while the NMED is primarily a decoding dictionary; as such it is the exception in this group. The importance it places on reading with a controlled vocabulary reflects Michael West's teaching experience in Bangladesh under British rule where reading in English was the priority of his prospective civil-servant students. In the absence of previous dictionaries as a frame of reference, it is not surprising that the dictionaries produced in this stage were not created on the basis of studies of user behaviour or user-informed questionnaires. Instead, the first MLDs were research driven in the sense that they were based on the linguistic analysis conducted by their authors in their work on vocabulary control and pedagogical grammars. What little research exists based on these dictionaries takes the form of critical reflections about the value of aspects such as etymological and phonological information to learners (see for example Barnhart, 1962).

The second stage of MLD development is characterised by a greater concern for learners' receptive decoding needs than the previous generation (Cowie, 2002), as well as the genesis of research about users and uses. Notable MLDs from this stage include the second edition of the ALD1: *The Advanced Learners Dictionary of Current English* (ALD2; Hornby, 1963); The third edition of the ALD series: *The Oxford Advanced Learners' Dictionary* (OALD3; Hornby, 1974); and the first edition of *The Longman Dictionary of Current English* (LDOCE1; Procter, 1978). The latter marks the conclusion of Hornby's dominance in English monolingual lexicography and, with its reconstitution as an electronic database, the traces of the computer based lexicography that was to characterise the creation of a third generation of MLDs.

A third stage of MLD development is characterised by the emergence of corpus-based MLDs along with an increased interest in MLD research. The *Collins Cobuild English Language Dictionary* (COBUILD1; Sinclair, 1987) and a new edition of *The Longman Dictionary of Current English* (LDOCE2; Summers, 1987) represent the first corpus-based offerings. *The Oxford Advanced Learners' Dictionary* (OALD4; Cowie, 1989) on the other hand, represents the vestiges of the traditional introspective approach.

A period convergence towards user-friendly and corpus-based dictionaries (Rundell, 2006) characterises the fourth stage of MLD development. This was due in part to research carried out on the first generation of corpus based dictionaries. The year 1995 saw the publication new editions of existing series COBUILD2 (Sinclair, 1995), LDOCE3 (Summers, 1995), and the first corpus-based Oxford MLD the OALD5 (Crowther, 1995). It also saw the entry into the marketplace of a corpus-based MLD from Cambridge University Press: the *Cambridge International Dictionary of English* (CIDE; Proctor, 1995). The year 2008 marked the arrival of a North American dictionary to the MLD marketplace: *Merriam-Webster's Advanced Learner's English Dictionary* (MWAD; Perrault, 2008). For some it represents an outlier amongst MLDs not only in terms of its origins and but also in terms of its use of invented examples (Hanks, 2009).

The digital revolution characterised by widespread availability of personal computers and increasing access to the internet in the 1990s, is often regarded as a paradigm shift in MLD making. Yamada (2013) argues that this is not the case. Instead he argues that electronic editions of existing MLDs are essentially digitised versions of their paper counterparts be they in CD-ROM, or handheld pocket electronic dictionary (henceforth PED) form. The entry of Macmillan into the MLD marketplace with the *Macmillan English Dictionary for Advanced Learners* (MEDAL1; Rundell, 2002) is evidence of the co-existence of paper and electronic dictionaries. It marks the completion of the big-five British MLDs (published by Collins, Cambridge, Longman, Macmillan, and Oxford) and goes slightly further than its predecessors by adopting a dual-track approach combining the restraints of traditional paper-based dictionaries with the opportunities offered by electronic mediums. While the development of electronic dictionaries has not marked the end of paper dictionaries, there is evidence that it has led to greater heterogeneity in lexical resources available to learners. These include online-only specialist MLDs and glossaries, and a proliferation of wordlists and phrase lists. In recent years, digitisation has also provided a catalyst for a great deal of research on the use of dictionaries and language learning.

The following section of this chapter deals with general aspects of the use of dictionaries in the teaching and learning of second languages. The next major section addresses studies related to specific aspects of dictionary design; for example, defining style, examples, and other aspects of microstructure. An examination of the treatment phraseology in learner

dictionaries then follows. The chapter concludes with a consideration of dictionaries and EAP with special emphasis on discipline specificity.

3.2 Research into General Aspects of Dictionary Use in SLA

This section attempts to summarise research into users' dictionary preferences. Specifically, what type of dictionaries they employ for specific tasks. First it deals with research based on user questionnaires. Next follows research based on tests including those designed to examine the relationship between certain dictionaries and dictionary types; and reading comprehension and vocabulary acquisition. Finally, the section concludes with a summary of observation-based research into dictionary use.

3.2.1 Questionnaire-based Research

For many years research into second language dictionary use and users has involved questionnaires. Tomaszczyk (1979) administered questionnaires to a group of 449 Polish participants which included learners of English at university level, language instructors, and translators. Among this diverse group, dictionaries were most frequently used for translation, followed by writing, then reading. Use for speaking and listening was far less frequent. The great majority of subjects predominantly used bilingual dictionaries when looking for meaning, although somewhat paradoxically, they held monolingual dictionaries[5] in higher prestige. The preference for bilingual dictionaries is supported by another questionnaire-based survey (Baxter, 1980). Baxter administered a questionnaire about dictionary ownership to 342 students of English, at three proficiency levels, at three universities in Japan. He found that most preferred bilingual dictionaries. It is worth noting, however, that Baxter's questionnaire dealt with ownership not use and that names of dictionaries were not reported.

Not all questionnaire-based research investigates user preference between monolingual and bilingual dictionaries. Béjoint (1981) employed a questionnaire to examine the use of monolingual English dictionaries by 122 French students of English at the University of Lyon. 96 per cent of students surveyed owned a monolingual dictionary – principally the OALD3 (45 per cent), LDOCE1 (27 per cent), and *The Concise Oxford Dictionary* (Sykes, 1976) (14 per cent). It is worth noting that many participants had bought their dictionary following the recommendation of a teacher. 87 per cent of respondents regarded looking up meaning to be amongst the three most common reasons for using a dictionary, 25 per cent mentioned spelling and pronunciation, while etymology barely featured at all. The respondents primarily used the dictionary for decoding rather than encoding and showed a strong preference for use with the written rather than spoken mode.

Task type also plays a role in users' dictionary preference. In a study of 58 Italian university students of foreign languages, Marello (1987, 1989) inquired as to which information students looked for in their bilingual dictionaries. She found that 98 per cent chose meaning, and 70 per cent spelling. In contrast, when she asked what information they looked up in MLDs, meaning achieved the highest score with 51 per cent, followed by synonyms with 49 per cent, and grammar 36 per cent. The need to look up synonyms is more likely to arise in an encoding situation than a decoding one.

[5] In reporting this finding Tomaszczyk (1979) makes no distinction between a general monolingual dictionary and an MLD; in this case the OALD3 (see Cowie, 2002, p. 182).

The university year, and presumably proficiency level, of students also influences dictionary user preferences, Bareggi's (1989) study of MLD use by 50 first-year students and 20 third-year students of English at the University of Turin found that first year students turned to MLDs for help with pronunciation, while morphology and syntax guidance was rarely sought. When they needed information on meaning they indicated a preference for bilingual dictionaries. However the third-year students cited finding meaning as chief among the reasons for using MLDs. In a similar study, Battenburg (1989) administered a questionnaire to 60 non-native-speaker English students studying at the University of Ohio. Participant numbers were evenly distributed across three proficiency levels; elementary, intermediate, and advanced. The MLDs participants reported owing were the OALD3, LDOCE1 and the *Longman Dictionary of American English* (Gray, 1983). Across all the groups subjects were more inclined to use dictionaries to look up word meaning, while they all demonstrated a lack of interest in etymological information. No group reported much interest in dictionary use for listening and speaking, elementary and intermediate students favoured dictionary use for reading, and advanced students had a preference for dictionary use while writing. In broad terms the majority of the students came from Arabic or Chinese language backgrounds. Battenberg found no evidence to suggest that language background influenced behaviour in dictionary use.

In a more statistically sophisticated study, Lew (2004) investigated the relationship between the language proficiency level of 712 Polish L1 students of English at five proficiency levels and their evaluation of dictionaries. He hypothesised that as learners' proficiency level increased their general rating of dictionaries would decrease since they would be more aware of the faults in the dictionaries. This hypothesis held for the first three levels where satisfaction with dictionaries decreased as proficiency level advanced, however, students at the latter two levels gave dictionaries high ratings. Lew suggests that this might be a consequence of students switching from bilingual dictionaries to MLDs as they advance in their studies. This suggestion is supported by the finding that advanced users value MLDs more highly than bilingual dictionaries. According to Lew, international MLD makers tend to be better resourced than Polish-English bilingual dictionary makers thus the higher ratings for MLDs might be a reflection of better quality. This argument supports practice of concentrating resources on a small number of MLDs in an EAP context, rather than sharing resources among various L2–English bilingual EAP projects. Another possible explanation is the possibility that learners might lack the language proficiency to find fault with MLDs because MLDs tend to place greater linguistic demands on users than bilingual dictionaries.

Atkins and Knowles (1990) and Atkins and Varantola (1998) report on the EURALEX/AILA project. This large-scale European project which aimed to examine how learners of English used dictionaries: The effectiveness of different types of dictionaries in encoding, decoding, and translation. It also examined students' attitudes towards dictionaries, instruction in dictionary use, and opportunities for improvement in dictionary design. The study relied on three instruments: Firstly, a questionnaire was applied to in order profile dictionary use by French, Italian, German and Spanish speaking learners of English at various universities in Europe. The aim was to gather information about the type of dictionary they habitually used and the English instruction they had received. Then a language placement test was administered to assess participants' language proficiency. Finally, the Dictionary Research Test, a 44 item test including, gap-fill, matching and translations items, test aimed at assessing students' dictionary skills was applied. In total, taking all three tests into account, 723 complete responses were obtained. The results of questionnaire phase revealed that although 69 per cent of students received more than half

their instruction in English, 60 per cent had never been taught how to use a dictionary and only 14 per cent considered that they had received "precise and systematic instruction" (Atkins & Knowles, 1990, p. 384). There was considerable variation between speakers of different L1s. 80 per cent of French speakers claimed to have received no instruction, 68 per cent of Spanish speakers, 50 per cent of German speakers, and 44 per cent of Italian speakers. However, on the Dictionary Research Test, respondents who claimed to have received training in dictionary use did not use dictionaries significantly more than those who had not. Overall the respondents reported a preference for bilingual dictionaries. However a finer grained analysis reveals that language proficiency level and purpose of dictionary use are also factors. For example, respondents in the highest language proficiency band preferred a monolingual English dictionary to look up words in English text. Similarly, more proficient learners prefer a MLD when checking a word that is already known. In general, dictionary use decreased as proficiency increased, however, the opposite was true for monolingual dictionaries.

In addition to the proficiency level of the user, and purpose of use, the user's field of study or expertise is another variable which could feasibly influence dictionary preference. This is an important consideration given the aims of the present study. In a study of the dictionary use preferences of 320 Arabic L1 students at the University of Kuwait, Al Ajmi (1992) made a distinction between students of English (N = 240) and students studying sciences through the medium of English (N = 80). He also grouped students into two progress levels. He found that there was little difference between the dictionary use preferences of the two subject groups; both principally used bilingual dictionaries, *Al-Mawrid* (Dar el-Ilm Lil-Malayen, 1990) and *A Dictionary of Modern Written Arabic* (Wehr & Cowan, 1976), to search for meaning, followed by spelling. This reflects the results of Quirk´s (1974) study of general dictionary use by 220 students at a British university which discriminated between science and humanities students. Al Ajimi's findings also reflect Bareggi's (1989) conclusion that more proficient language users turn to MLDs to look for meaning to a greater extent than less proficient users.

Questionnaire-based studies of digital dictionary use are relatively rare. Most studies of user perception of digital lexicographical resources rely on analysis of log-files (Müller-Spitzer, Koplenig, & Töpel, 2011). Müller-Spitzer, Koplenig, and Töpel's online survey into online general dictionary use conducted in English and German in 2010 represents a rare example. Both Müller-Spitzer, Koplenig, and Töpel (2011) and Müller-Spitzer, Koplenig, and Töpel (2012) report on the findings of an analysis of 684 responses from linguists and non-linguists. They found that participants valued up-to-date content, clarity of presentation above the more innovative features, such as adaptability and multimedia elements, which the digital medium permits. Lew and de Schryver (2014) posit four possible explanations for these findings. First, that the respondents had a traditional conservative view of lexicography; Secondly, that they felt obliged to present a critical view of the novel features of digital dictionaries; and thirdly, that quality of content would logically always take precedence over quality of the interface since users will deal with a poor interface if high-quality content is available but not vice versa. Alternatively, they suggest that users may simply not be familiar with innovative features of digital dictionaries.

In summary, questionnaire-based research suggests various trends in the use of dictionaries by learners. Taken as a homogeneous group, learners seem to prefer bilingual dictionaries to MLDs. However, there is some evidence to suggest that this trend is reversed at higher proficiency levels. In the mapping of proficiency levels to EAP contexts set out above, EAP users are predominantly higher-proficiency level users. This suggests that the

guidelines for producing useful EAP vocabulary resources set out as an aim for the present study should concentrate on monolingual resources. In general terms, bilingual dictionaries are preferred to MLDs for decoding whereas MLDs are preferred for encoding. Although, here again, level has an effect since higher-proficiency learners profess a preference for MLDs for both task types. The only study to investigate the effect of learners' field of expertise or study (Al Ajimi, 1992) does not show significant influence for this factor on dictionary type preference. The few questionnaire-based studies of digital dictionaries suggest that their attraction lies in the ease of access they offer to information, rather than multimedia innovations. Somewhat puzzling, is the apparent contradiction between users' overall preference for bilingual dictionaries in use, and their perception of monolingual dictionaries as having higher prestige. However, Lew's (2004) suggestion that this perception is related to the language-proficiency of the users or the resources available to lexicographers represents a plausible explanation. In the first instance, it is possible that while users have the necessary language proficiency to perceive short-comings in bilingual dictionaries, the greater degree of linguistic complexity in monolingual dictionaries obscures similar shortcomings. In terms of resources, while the big-five MLD makers benefit from a high concentration of sales and resources per MLD, in the case of bilingual dictionaries these resources and sales are distributed among many L2s. This is particularly problematic for those lexicographers who work with less widely-spoken languages.

A number of objections have been raised to the use of questionnaires to investigate dictionary use. Firstly, questionnaires are only as reliable as their respondents. It is unlikely that respondent recollections of the last time they used a dictionary are accurate (Crystal, 1986). It is equally possible that respondents have preconceived ideas about what constitutes good dictionary use which influence their responses to questionnaire items (Hatherall, 1984). Secondly, in an MLD research context, it is unlikely that respondent and researcher share the same terms of reference (Nesi, 2000a). Crystal (1986) claims that the use of terms such as, *form*, *class*, *etymology* etc. in questionnaires tends to lead respondents to a limited set of responses to the detriment of more imaginative alternatives. In other words, use of specialist linguistic terms and meta-language which is unfamiliar to the respondent makes misinterpretation of the question a real possibility. This is especially true in surveys of lower proficiency EFL learners.

3.2.2 Test-based Research

3.2.2.1 Studies of the Effects of Dictionary Use on Reading Comprehension

An important group of studies focuses on the effect of dictionary use on reading skills. Bensoussen, Sim and Weiss's (1984) report on a series of studies involving first-year students at the University of Haifa is among the earliest to explore the relationship between dictionary use and reading comprehension test score. In a pilot study, 900 students were assigned to one of two groups of 450 participants. One group was assigned a monolingual dictionary, the other no dictionary at all. There was no significant difference in test score on a multiple-choice reading comprehension test. To investigate further, the authors conducted a study involving 91 students who used either a monolingual dictionary, bilingual dictionary, or no dictionary at all. Once again they found that dictionary type had no significant influence on the scores obtained in a complete-sentence-answer reading comprehension test. A procedure in which students had to indicate the words they intended to look up and those which they did, in fact, look up was conducted at the same

time as the test. It established that bilingual dictionary users looked up more words than monolingual dictionary users, and that users' proficiency level had no significant influence on the number of words looked up. Two further studies, involving 670 and 740 students respectively, found no overall significant relation between type of dictionary used, test score, or the time taken to complete a reading comprehension test.

Surprised by Bensoussan, Sim and Weiss's (1984) findings, Nesi (2000b) sought to replicate their study by conducting a pilot-study and two further studies investigating the effect of dictionary use on performance in a multiple-choice reading comprehension test. The pilot involved 20 students beginning an EAP course at Aston University in the UK. Four students used the OALD4 or 5 to complete the test the remainder chose to forgo use of a dictionary. The first study involved 83 overseas students at Warwick University in the UK coming to the end of an EAP course. 43 of these students were allowed access to the OALD4 and 5, LDOCE2 and 3 or the *Longman Active Study Dictionary* (LASD3; Summers, 1998), or a bilingual dictionary, while 40 were denied access to a dictionary. Dictionary users were asked to indicate which words they looked up. In line with Bensoussan, Sim and Weiss's study Nesi found no significant difference between type of dictionary used and test score, although, in contrast to the former study, participants who employed dictionaries to a greater extent, took longer to complete the test. In an attempt to explain these findings she conducted a second study which involved 65 overseas students at Warwick University. They were divided into two groups of equal ability. One group containing 31 subjects was denied access to a dictionary. The other group was given access to the OALD and asked to underline the words in the text they looked up. The group who had dictionary access took significantly longer to finish the task even if they showed no indication of dictionary use. Nesi posits a number of reasons to explain the difference between this finding and that of Bensoussan, Sim and Weiss. These include differences in efficacy of dictionary use and environmental factors such as background noise and the presence of a tutor.

The stand-out finding of Bensoussan, Sim and Weiss (1984), the scant influence of dictionary use on test score, might seem unexpected. However it could be explained by deficiencies in the texts employed in the study, in dictionaries themselves, or in students' dictionary skills. The questionnaire administered to students in order to clarify the results of the study points to the latter two scenarios. Respondents to the questionnaires reported a belief that simply looking-up a word would lead to an understanding of the text. They also reported difficulties with defining styles, particularly long definitions, and the use of meta-linguistic terms. Nesi (2000b) also considers that deficiencies in dictionaries played a role in her results; by comparing the relevant monolingual dictionary entries and test items, she demonstrates that in many cases the dictionary entries lacked the information necessary to answer the questions. One omission is related to scientific or technical language; in reference to the OALD, Nesi (2000b) states that "examples were limited in number and not always applicable to the technological topics of the New Scientist texts" (p. 67). Other problems include the omission of literal senses in favour of figurative senses, the inclusion of examples for some senses but not others. Bensoussan, Sim and Weiss (1984) give little information about the texts employed in their study. However in her study, Nesi concludes that although the texts employed were adequate, the test items tested general reading strategies rather than the understanding of specific words.

The research that has been carried out into the relationship between dictionary use and reading comprehension test performance would seem to suggest that dictionary use has no significant influence. However, more thorough examination of the findings of research in this area suggests that this is perhaps a reflection of methodological deficiencies which are

understandable given the complex relationships between reading comprehension skills and extraneous factors such as general reading and dictionary strategies, intelligence and overall language proficiency. That said the studies examined, particularly Nesi (2000b), highlight some possible dictionary deficiencies which could feasibly explain the negligible influence of dictionaries on reading comprehension performance. In the context of the present study the failure to adequately deal with technical or discipline-specific vocabulary is particularly pertinent.

3.2.2.2 Studies of the Effects of Dictionary Use on Vocabulary Acquisition

Several studies point to a positive effect for dictionaries on vocabulary acquisition. Such studies typically rely on participants reading a passage and then completing a delayed vocabulary test. Beyond EAP, in a study involving 44 Dutch L1 students of French in their first year of French at university, Bogaards (1991) studied the influence of three types of dictionaries: bilingual (not named), MLD: *Dictionnaire du français langue étrangère* (Dubois, 1983), general monolingual: *Le Petit Robert* (Robert, Rey, & Rey-Debove, 1986), and no dictionary use on performance in a translation task and on vocabulary retention in general. Bilingual dictionary users looked up the most words and produced the best translations, while those students who had no access to dictionaries produced the worst. In a delayed repeat test conducted 15 days later, users of the MLD had most success closely followed by users of the bilingual dictionary.

Unlike participants in the Bogaards (1991) study, participants in the EURALEX study (Atkins & Knowles, 1990; Atkins & Varantola, 1998) were free to choose the type of dictionary they wished to use. In line with previous studies most showed a strong preference of bilingual dictionaries. As a consequence, it is not possible to make reliable comparisons between monolingual and bilingual dictionary use on the basis of this study. However, with this caveat in mind, the data from the Dictionary Research Test suggests that monolingual dictionary lookups, including MLDs, were more successful than bilingual lookups irrespective of proficiency level or task type.

In a study of 293 EFL students at a Japanese university, Luppescu and Day (1993) found that students who used a bilingual dictionary (of their choice) while reading gained higher scores on a vocabulary test which immediately followed than those who did not make use of dictionaries. There are a number of problems with their study. Firstly, it is not clear whether the words tested were previously unknown to the students. Secondly, the test relied not only on students knowing the target words but also the words contained in the multiple choice answers which often seemed much more difficult than the target words (Nesi, 2000a). Looking beyond EAP once again, in a similar study involving 105 second-year students of Spanish at Central Michigan University, Knight (1994) controls for these problems. She employed computerised dictionaries and texts to examine the relationship between bilingual dictionary use and vocabulary acquisition. The dictionaries used were *Harper Collins Spanish-English English-Spanish Dictionary* (Butterfield, 1990) and *Diccionario esencial inglés-español español-inglés Diáfora* (Dutton, Harvey & Liddel, 1981). Access to dictionaries resulted in significantly higher scores on immediate and delayed vocabulary tests and comprehension tests. These tests were also complimented by recall protocols in which students displayed positive attitudes towards computerised dictionaries.

The question of whether the positive relationship between vocabulary acquisition and dictionary use holds for digital dictionaries is taken up by Laufer (2000). In a study of 55

Hebrew L1, advanced-level EFL university students in Israel she compares incidental vocabulary acquisition from reading on paper with an in-margin gloss, and reading on screen with the support of an explanation in an electronic dictionary (the source of the explanation is not specified). Crucially, in terms of vocabulary acquisition, the score achieved in a delayed-recall test for the 24 students who read on screen was significantly higher than that of the 31 who read on paper. Laufer takes advantage of the versatility afforded by digital dictionaries to offer students a number of dictionary formats. Students who looked up a translation, L2 definition and L2 example in the electronic dictionary achieved higher scores in a delayed-recall test than those who looked up L1 translation only, or the translation with English definition.

Y. Chen (2010, 2012) conducted a series of studies with similar aims to Laufer (2000), however, instead of comparing paper and on-screen dictionaries he instead compared the effect of using a bilingualised paper dictionary, or a bilingualised PED on vocabulary acquisition. Y. Chen (2010) involved 85 junior English majors studying at Putian University, Fujian, China. Students conducted a vocabulary exercise which involved using their dictionaries to choose the correct meaning of 10 low frequency words in a multiple-choice and sentence composition test. The paper dictionary used was *the Oxford Advanced Learner's English-Chinese Dictionary* (OALDCD; Wehmeier, 2004) there were various PEDs. The vocabulary exercise was followed by a retention test and a delayed retention test. The findings indicate no significant differences between scores on the retention tests for users of each dictionary format, however students using PEDs took significantly less time to complete the tests. Y. Chen (2012) is a larger scale study which adopts a slightly different approach. It involves 176 students from three classes of Chinese L1 English seniors from Putian University and another four classes of juniors from Xiamen University and Fujian Agriculture and Forestry University. In contrast to Y. Chen (2010), students first read an English passage and completed a reading task. One group of students was permitted a paper bilingualised dictionary (OALDCD), while another had access to an electronic version of the fourth edition of the OALDCD; a third group had no dictionary access. The reading task was followed by an immediate retention test and a delayed retention test. Findings indicate no significant difference in reading task scores between the paper and the electronic dictionary groups; however the latter group demonstrated better vocabulary retention.

The findings of all these studies demonstrate a positive relation between dictionary use and the acquisition of vocabulary. This supports the validity of using dictionaries to support many of the teaching procedures and techniques outlined in the previous chapter. The relationship between dictionary media and vocabulary acquisition is not clear-cut; however, there is, at least, no indication that digital dictionaries have a detrimental impact on vocabulary acquisition.

3.2.3 Observation-based Research

In addition to questionnaire and test-based research, observation-based research provides another means of investigating dictionary use. It involves collecting data while or immediately after dictionary users complete or have completed an observable task. Proponents of this approach claim it avoids the problems of participant bias implicit in questionnaire-based research (Hatherall, 1984). Some observation-based studies also offer an advantage over test-based studies insomuch as they involve tasks which better reflect contexts in which dictionary use naturally takes place. However, it must be noted that the act of observation necessarily brings unnaturalness to proceedings.

Ard (1982) examines students' use of bilingual dictionaries. His report lacks key demographic information, however, it does state that participants were Japanese and Spanish L1 high-intermediate level ESL students at the University of Michigan. The study involved students giving oral feedback on their use of a bilingual dictionary while writing compositions in class. His principal finding is that Spanish students use bilingual dictionaries more successfully than their Japanese counterparts despite having a lower writing ability. He attributes this to the greater likelihood of a one-to-one translation existing between Spanish and English than Spanish and Japanese.

In addition to first language background, observation-based studies point to proficiency level as a factor which influences dictionary use. In common with Ard (1982), Hatherall (1984) gives little demographic information about the participants in his pilot study. It appears they are all English native-speaking students of German at college level who conducted a written translation task while recording their dictionary use in written form. Key findings were that more advanced students tend to use the dictionary more often than less advanced students and that students tend to translate word-for-word and that this is encouraged by a view of language "as words in sequence rather than a system or systems of structures" (Hatherall, 1984 p. 188), in other words the non-phraseological view of language, adopted in many dictionaries. In common with Bogaards's (1990) study L2 involving Dutch and French university-level students of French, Hatherall (1984) found that students look up nouns rather than verbs in noun-verb expressions and that this is a contributing factor in unsuccessful dictionary use. Both Ard (1982) and Hatherall (1984) provide scant demographic information about participants. Their proficiency level at the time of the procedure is also unclear, nor is it apparent how this was assessed. This limits the generalisability of the findings the studies report. It is also worth noting that many years have passed since many of the studies mentioned in this chapter were conducted. Since dictionaries have undergone significant change in this time, the relevance of many of these findings to current lexicographical resources is limited.

Neubach and Cohen (1988) examine the use of three dictionaries; LASD1, *Collins English Learner's Dictionary* (Carver, Wallace & Cameron, 1974), and a bilingual English-Hebrew dictionary the *Meggido Modern* (Levenston & Sivan, 1968) across language proficiency levels. Their small-scale study involves six students on an EAP reading course at the Hebrew University of Jerusalem. The students are evenly distributed across three proficiency levels, high, intermediate, and low. It is not apparent what criteria were employed in assigning these levels. The study involves two tasks each followed by a structured interview. In the first task participants were given ten sentences in which a polysemic word was underlined. Participants had to look up the underlined word in the dictionary while giving an oral report on the look-up process. The final step of the task involves students translating the underlined words to Hebrew and explaining their rationale in choosing the particular meaning selected. In the second task the students read a 150-word text in which ten uncommon words had been underlined. Students were permitted to use any of the three dictionaries and had to make an oral report of the look-up process. Finally, they were asked to summarise in Hebrew what they had read. Key findings were that in monolingual dictionaries, users tended to read only the first definition. They also encountered problems decoding the vocabulary and terminology used in the definitions. More generally they experienced problems with alphabetical order, the presentation of the entries, frustration during the look up process, and uncertainty even when they had in fact found the correct meaning. Neubach and Cohen (1988) also highlight students' tendency to interpret the meaning of a word based-on pre-conceived notions of word meaning. Müllich (1990)

notices a similar tendency amongst German learners of English at high-school level which he terms sham-use of the dictionary.

Neubach and Cohen (1988) sought to elucidate the processing strategies used by university students in their dictionary use. As regards the relationship between language proficiency and dictionary use, they conclude that only advanced students benefitted from dictionary use. In this regard, their ability to place a word in context was key: "the higher-proficiency students generally had an idea of the semantic field of a polysemic word before attempting to search for it." (Neubach and Cohen, 1988, p. 11). Although these findings are thought provoking and provide some insight into EAP users' dictionary use, Neubach and Cohen's study suffers from a number of methodological problems (Nesi, 2000a). Chief amongst these is the emphasis that think-aloud protocols put on problematic dictionary use while underestimating successful dictionary use.

Although it did not take place in an EAP context, K. Harvey and Yuill's (1997) study of the COBUILD1 was a far larger-scale study than Neubach and Cohen (1988). It serves as a useful test of the generalisability of the latter's findings. It involved 221 intermediate students of English at language schools in Britain. It aimed to establish how learners consulted the dictionary and how successful this consultation was from the students' point of view. Students engaged in a writing task and recorded their dictionary use using flow charts. Findings suggest that learners use the dictionary to search for correct spellings, to see if a word exists, to find synonyms, to check on meaning, and to check grammar. A key finding as regards the latter two uses is that users tend to resort to example sentences rather than the dictionaries' grammatical coding schemes. This fits with Béjoint's (1981) findings which suggest a preference for example sentences over grammatical coding schemes. Similarly, in her chapter of the role of dictionaries in ELT, Summers (1988) suggests that the length of an entry particularly the ease with which example sentence can be found is a major factor in successful dictionary use.

Nesi and Haill (2002) conducted an observational study which drew on data gathered from 89 students with various L1s studying an EAP course in preparation for undergraduate studies at Oxford Brookes University. The study involved the analysis of a study skills task in which students had to choose a text and look up five previously unknown words in a dictionary of their choice. Participants also had to answer a series of questions related to the dictionary consultation. MLDs were by far the most consulted dictionaries. Over half of the 77 students who completed the task were unsuccessful in at least one look-up. Overall, 16.4 per cent of the total number of look-ups were analysed as unsuccessful. By comparing the participants' answers to the dictionary consultation questions with the dictionary entries for the words looked up Nesi and Haill (2002) concluded that participants have difficulty choosing the appropriate entries and sub-entries in dictionaries. This is often due to misunderstanding of definition, example, or other coding, or the ordering of elements of the microstructure.

In his response to Hatherall's (1984) rejection of questionnaires in favour of direct observation in dictionary research, Lew (2002) points out that several of Hatherall's criticisms of questionnaires are also applicable to observation based studies. In short, Lew (2002) argues that observation-based studies do not mitigate the possibility of participant bias, on the contrary they give rise to the observer's paradox, whereby the presence of the observer or the fact the participant knows he or she is being observed provokes a change in the behaviour under study. Some studies of digital dictionaries offer the opportunity of less intrusive observation. Log-file studies offer a method for research into dictionary use

(Bergenholtz & Johnsen, 2005; De Schryver, 2013; De Schryver, Joffe, Joffe, & Hillewaert, 2006; De Schryver & Joffe, 2004; Lew, 2011). Studies employing this method tend to focus on the frequency and type of items searched for in monolingual dictionaries (e.g. Müller-Spitzer, Wolfer, & Koplenig, 2015).

Until recently, studies of online second language dictionaries have relied upon traditional observation techniques. In a study involving ten students studying an elective English course at a university in Seoul, Korea, Chon (2009) employs a think aloud task which students complete while writing and looking up words in an online dictionary[6]. On analysis the author concludes that the bilingual dictionary lookup was often unsuccessful because users lacked confidence to employ the L2 translation or considered the translation inappropriate in context. The freedom permitted to students in choosing the writing topic allowed them to peruse alternative strategies to dictionary use such as avoidance or looking for a synonym. They seemed to prefer such strategies.

Eye-tracking represents another observation-based approach which has been fruitfully applied to research on dictionary use (Kaneta, 2011; Lew, Grzelak, & Leszkowicz, 2013; Müller-Spitzer, Michaelis, & Koplenig, 2014; Simonsen, 2009, 2011; Tono, 2011). However, it is better suited to the study of specific elements of dictionary microstructure and will be dealt with in the following chapter section.

In common with studies using questionnaires and test-based methods a key finding from observational studies includes the tendency of more advanced student's to recur to MLDs while beginners recur to bilingual offerings. Assuming the prototypical EAP student has an intermediate to advanced proficiency level, this is partial justification for the decision to concentrate on monolingual lexicographical resources in the guidelines and illustrative examples later in this dissertation. A less debatable conclusion from this overview of observation-based dictionary research is that specific elements of dictionary design are particularly problematic for users. These will be addressed in the following chapter section.

3.3 Research into Specific Aspects of Dictionary Use in SLA

3.3.1 Defining and Example Style

In addition to the observational studies discussed in the previous section, another group of observational studies is primarily concerned with defining styles. Miller (1984) and Miller and Gildea (1987) call the inclusion of examples in dictionaries into question. Their study set out to discover how L1 English 6th grade (age 10 to 11) children use and misuse dictionaries. The study relied on a look up and compose sentence (LUCAS) task. Through an analysis of the responses received and the dictionary entries consulted they conclude that many unacceptable sentences are the result of a lookup strategy. This strategy, which Miller and Gildea term Kidrule, involves users finding a familiar word or phrase within a definition and simply substituting the target word for that phrase.

In the second language paradigm, there are two approaches to the examination of the effectiveness of dictionary definitions: studies which involve observing learners as they use dictionaries; and more speculative studies which involve looking at common learner errors

[6] The dictionaries listed are defunct versions of Yahoo, Naver, and the thesaurus in Microsoft Word (no version information is provided)

and speculating about how these errors might be attributable to dictionary definitions. The provenance of data analysed means that some of these speculative studies are more speculative than others. Some involve errors produced by students in texts written specifically after dictionary consultation whereas others use student errors taken from texts which learners may have written for other purposes.

In a study involving errors form general learner texts, Nesi (1987) analysed the errors made by prospective non-native English speaker postgraduate students in a placement test at Aston University. With reference to the OALD, LDOCE[7], and *Chamber's Universal Learners' Dictionary* (CULD; Kirkpatrick, 1980) she attributes the ten examples of errors discussed to problems with definitions. Inspired by Nesi (1987), Meara and English (1988) analysed a collection of 1364 lexical errors from First Certificate examination papers (Approximate target-CEFR level B2). Since the papers were provided by the examinations syndicate it is not known whether they were completed by EAP students or students in more general contexts. They conclude that the majority of errors would not have been avoided had the learners successfully looked up the relevant entry in the LDOCE and that their failure to do this was often due to deficiencies in defining style. In a study involving sentences written by students, Jain (1981) employed an exercise in which learners had to decide whether a series of keywords could be interchanged in the context of a sentence. No information is given about the participants, but it is stated that they were given access to three dictionaries; the OALD3, CULD, and LDOCE1. All three of these studies, Nesi (1987), Meara and English (1988), and Jain (1981), attribute students' errors to problems with definitions including failure to account for socio-cultural context, polysemy, register, selectional and collocational restrictions, the use of one- word definitions which promote false equivalences, and definitions which conflate different parts of speech. Maingay and Rundell (1987) present an alternative viewpoint. They analysed essays by L2 English students in India at the Indian Institute of Technology in New Delhi, in Japan at the Senior High School in Hiroshima, and samples of essays by students who took the Cambridge Proficiency Exam in 1982 (Approximate target CEFR level C1/C2) and developed a typology of errors. No more information is given about the number or proficiency level of the participants. Citing their own analysis and an unpublished internal report from Longman dictionaries (A. Black, 1986)[8] they conclude that definitions usually include all the information needed for users to avoid errors, but this is often ignored by users. To remedy the situation they recommend the inclusion of longer examples and definitions which repeat information several times.

Acceptability testing, an approach which involves observing subjects as they evaluate different types of dictionary entry, is another means of evaluating defining styles in learner dictionaries (Crystal, 1986). MacFarquhar and Richards (1983) adopt this approach in a study involving 180 intermediate to advanced students from the Asia-Pacific region enrolled at the English Language Institute at the University of Hawaii. They judged the intelligibility of definitions from three MLDs with different defining styles; the LDOCE1 with its restricted defining vocabulary; the OALD which they claim has clear and unambiguous definitions for non-native speakers; *Webster's New World Dictionary* (WNWD; Gurulnik, 1976) which has an unrestricted defining vocabulary. Their analysis revealed that 51.5 per cent of the time students regarded the LDOCE1 definition as the most intelligible, 28.5 per cent the OALD3, and 20 per cent the WNWD. However, the authors are careful

[7] No information is given about the editions examined, however, taking the age of the study into account it is assumed that they are OALD3 and LDOCE1 and 2.

[8] There is an inconsistency in Maingay and Rundell (1987): The reference given in a footnote in is Black (1985), however the bibliography lists Black (1986)

to stress that they are only measuring the users' perception of the entries. Testing readability would require users completing a task. They call for further studies to compare perceived intelligibility of entries and the learning that takes place.

Many of the studies dealing with dictionary definitions already mentioned were motivated by the use of a controlled defining vocabulary in MLDs. Jansen, Margeai and Vanadroye (1987) conduct an analysis of the LDOCE1's defining vocabulary and point out potential problems for users including, homography, polysemy and the use of phrasal verbs. Adopting a different approach, Herbst (1986) analysed words included in LODCE1 definitions which fall outside the defining vocabulary. He concludes that that scientific and technical vocabulary poses a challenge to the controlled vocabulary approach.

The study reported in Nesi (2000c) and Nesi and Meara (1994) is, in part, a response to MacFarquhar and Richards's (1983) call for a task-based study of the readability of different types of dictionary definition. They examine the OALD4, with its traditional defining style; the LODCE2 with its controlled vocabulary; and the COBUILD1 with its sentence definition (they exclude the COBUILD's third column which contains grammatical information). LODCE users produced the lowest percentage of semantic errors, however COBUILD users produced fewer usage errors, closely followed by LODCE users. The main aim of the study was to shed light on the discrepancy between Maingay and Rundell's (1987), and A. Black's (1986) contention that the best way to help users avoid errors is by repetition and emphasis of information in definitions, and Miller and Gildea's (1987) finding that the situation in complicated by unexpected user behaviour such as Kidrule. In a study involving 52 students on a university-level English language study skills course, Nesi and Meara (1994) investigate the prevalence of Kidrule among L2 English speakers using learner dictionaries. Their experiment involved pairs of target-words, one high-frequency which they assumed the students knew, and one low-frequency which they assumed students did not know. Students had to use the target pairs to create a sentence. The two-word approach was intended to discourage students from simply copying from the dictionary. Since the experiment was computer mediated the number of lookups and time taken examining the definition could be easily recorded. In an analysis, which the authors admit was somewhat subjective, just under a quarter of the errors encountered could be attributed to Kidrule and, more specifically, to students' tendency to focus on the wrong part of the definition or even the wrong sub-entry in the dictionary. They also found four errors which could not be explained by Kidrule. They hypothesise that two of these student errors: The failure to employ information about grammatical and collocational restraints might be the result of dictionary design. For example, the inclusion of such information in parenthesis might lead students to think that it is unimportant. Other sources of error are phonological or orthographical confusion, for example, a confusion of *crime* with *climb*. They also attribute errors to students' pre-conceived notions about the meaning of target words in phenomena similar to Müllich's (1990) sham use of dictionaries. These findings are broadly in line with Jain's (1981) speculative study of the sources of dictionary error.

An overview of user-based research into dictionary definitions indicates that they could feasibly be responsible for a good deal of unsuccessful dictionary look-ups. As far as defining vocabulary is concerned although, users regard controlled vocabulary definitions as more intelligible and acceptable, they are by no means a panacea for the problems of definition in MLDs. It is difficult to define scientific and technical wording using controlled defining vocabularies. Like detailed and repetitious definitions, they do little to discourage less-than-ideal look-up strategies such as Kidrule. A lack of information about selectional

or collocational restrictions is particularly problematic for productive purposes. Section 3.4 below examines the treatment of collocation in learner dictionaries in greater detail.

The question of how to best to use examples in MLDs is closely related to defining style. There is a widespread view among lexicographers that the inclusion of illustrative examples is beneficial to the users of learners' dictionaries (Cowie, 1989; Creamer, 1987; Drysdale, 1987; Landau, 2001). However, there are relatively few empirical studies which put this view to the test.

A. Black (1986) conducts two experiments comparing the comprehension and retention of 24 target words defined in three different styles: Abstract dictionary style definitions, definitions with examples, and example only. The study was conducted in the UK with students preparing to take the Cambridge Proficiency Exam (approximate target CEFR level C1/C2). It is unclear whether this took place in an EAP context. There were sixteen students involved in the first experiment and 24 in the second. In both experiments, participants had to read three texts containing the target words then look up the definitions on index cards which were distributed amongst students in such a way that the three defining styles could be compared. They then had to complete a multiple-choice comprehension test. In experiment one, the students had access to the cards with the definitions while completing the test. In experiment two, they did not. In both experiments there was a highly significant difference between scores for defined and undefined words. However, there was no significant difference between example styles.

In a study of young native-English speakers' use of dictionaries, Miller and Gildea (1987) also call the inclusion of examples in dictionaries into question. The authors set out to discover if native English speaking 6th grade (age 10 to 11) students learned words more easily from illustrative examples than dictionary definitions. They employed a method in which participants had to look up a word provided to them and then compose a sentence using that word. The type of definition which participants had access to was controlled. The results suggest that example sentences were more useful to students. However, in some of the sentences there was evidence that participants had employed Kidrule, simply replacing the part of the example sentence they understood with the target word. They suggest that in some cases this is because the example sentences did not provide sufficient usage information. A further experiment from this study also calls into question the value of multiple example sentences since for Gildea and Miller (1987) "the acceptability ratings of sentences written after seeing one model sentence were the same as the sentences written on the basis of three" (paragraph 16). Since their subjects were young children acquiring English as an L1 the applicability of Miller and Gildea's finding to an EAP context is doubtful.

In another study outside an EAP context, Summers (1988) reports on two tests which involve students preparing for the Cambridge Proficiency Exam (approximate CEFR level C1/C2) at a language school in Cambridge, UK. The first test employed a reading passage with multiple-choice definition-matching questions to test the influence of examples on dictionary use in decoding. The second test involved students producing the keywords in sentences. Both tests demonstrated the benefit of dictionary use, in this case LDOCE2, on test performance compared with the control condition of no dictionary use. However, there was no significant difference between the three experimental conditions: Definition only, example only, and both example and definition.

In a context more relevant to EAP, Laufer (1993) investigated whether new words were best acquired with examples and definitions or definitions alone. Her study involved 43 first-year students, the majority L1 Hebrew speakers and some L1 Arabic speakers, beginning an EAP course at the University of Haifa. The test employed included eighteen words presented out of context nine of which were presented with definitions and nine with illustrative examples. In the first part of the test participants had to translate the words into their L1 and compose a sentence in English. The second part of the test was a repetition of the first this time with the words which were defined in the first test illustrated and vice versa. The results demonstrate that combined definitions and examples are more effective in production, while definitions are more effective than examples in comprehension.

Cumming, Cropp, and Sussex (1994) address MacFarquhar and Richards's (1983) call for a study which compares the intelligibility of entries and the learning that actually takes place. They compare the use of sentence definitions, specifically those found in COBUILD1, with more traditional phrasal definitions. The participants were 85 adult ESL students of an intermediate to advance level (no information is given on how this assessment was reached). In the experiment participants were shown four types of definitions: phrasal definitions without usage examples, phrasal definitions with usage examples, sentence definitions without usage examples, and sentence definitions with usage examples. Participants rated their initial familiarity with the target-word, and the definitions in terms of initial helpfulness in a scale of 0 to 100. They also completed a production task which involved writing a sentence with the target word and a comprehension task which involved deciding whether sentences containing the target word were correct or incorrect. The initial familiarity rating showed little correlation with production or comprehension measures. There was no significant difference for helpfulness, production and comprehension between each of the defining styles. However, sentence definitions were judged higher on the final usefulness scale than phrasal definitions.

Nesi (1996) (also reported in Nesi, 2000c) assumes a positive relationship between the quality of a dictionary and the speed at which it can be read and put into use. She examined whether definitions containing examples took longer to read, and were of more use productively than those without. Her study involved 40 non-native speaker students studying English at an advanced level in the UK. The students were given pairs of target words and had to write sentences containing these. The students were split into two groups. Group A was given a list of target words half of which contained examples, likewise group B, only this time the other half of the list was accompanied by examples. No significant difference between reading times or performance was found across the two groups.

Working in first language acquisition, Nagy and Scott (1997) used a definition acceptability judgment task to investigate the influence of defining style and the presence of examples in dictionaries had on young native-English-Speaking students' ability to spot errors while reading sentences. They gave 55 fourth-grade (aged 9 to 10 years) and 45 sixth-grade (aged 11 to 12 years) students the definition of a target word using various defining styles: traditional definitions, traditional definitions with an example sentence, and definitions in an informal style. Students then had to judge whether sentences containing the target word, some correct some with varying grades of error, were appropriate or not. Although the results show a small yet statistically significant positive influence on the identification of errors when example sentence are provided, this was not greatly different from chance.

Al Ajmi (2008) conducted an experiment involving 54 L1 Arabic students of English at Kuwait University. The students were split into two groups and asked to provide the Arabic equivalents for ten English keywords. The first group was given entries from OALD5 while the second group was deprived of illustrative examples. His findings suggest that illustrative examples had a negative impact on students' performance.

Although empirical studies of the influence of examples in dictionaries on both native speaker and non-native-speaker users provide little evidence in support of the received wisdom that the inclusion of examples in dictionaries is beneficial to the user, there are a number of convincing reasons to call these findings into question. The first group of reasons relates to problems with the studies themselves. For example, Nesi (2000) argues that A. Black's (1986) findings can be called into question citing the possibility that students guessed the meaning of words in the multiple-choice task. The presence of keyword context in both the reading text and the multiple-choice comprehension items makes guessing all the more possible. Additionally, many of the studies discussed did not control for the possibility of participants simply copying directly from the examples or did not specify how such occurrences were scored in production tests. For example, Nesi claims that Miller and Gildea's (1987) finding that with access to examples children produced a greater number of acceptable sentence is questionable since many children simply reproduced the example when asked to write a sentence using the target word.

Another possible methodological problem relates to subjects' initial familiarity with keywords, it is possible that in many cases the participants had to look up words that they were already familiar with. Many of the studies discussed attempted to control for this possibility using frequency-based wordlists or vocabulary tests. It should be clear from the discussion of the frequency-based approach in previous chapters that these lists and tests are of limited use, especially where studies involving a handful of keywords are concerned. Cumming, Cropp, and Sussex (1994) took an alternative approach, they controlled for the possibility of previous familiarity with an initial familiarity measure and they also attempted to stop participants from copying from the example by instructing them not to do so. However, they do not report on the extent to which this instruction was followed. Indeed, in studies of defining style, Nesi and Meara (1994) and Nesi (2000c) report evidence of students copying directly from examples in spite of measures intended to discourage them from doing so.

Many of the studies discussed treat word knowledge as an all-or-nothing phenomenon, it should be apparent from the discussion in the previous chapter that this is not the case; there are many aspects to knowing a word. This conception of word knowledge has implications for the validity of the findings in some of the studies discussed. For example, Nesi (2000c) concedes that under her coding system participants "who displayed some understanding of word meaning might still produce sentences coded as inappropriate if their word knowledge was less than complete." (p. 115).

Frankenberg-Garcia (2015) attributes the negligible influence of examples demonstrated in most studies to a failure do distinguish between receptive and productive examples. In a series of experiments she demonstrates that when dictionary examples written or included with productive or receptive use in mind are paired with corresponding tasks they have a significant positive effect on performance.

Another possible reason for the incongruence between the generalised perception of the benefit of examples and the results of the empirical studies discussed relates to the

examples themselves. Although most of the examples employed in the studies were corpus based, the use of a corpus does not prevent a common pitfall for lexicographers: the inclusion of creative examples rather than more obvious examples (Drysdale, 1987). For example, Nesi (2000c) suggests that the second example provided for *perpetrate* in LDOCE2; *it was the managing director who perpetrated that frightful statue in the reception area,* caused participants to produce sentences such as "last year winter perpetrated many horrible storms" (p. 113).

The prevailing teaching methodology might also account for the results of several of the studies. Nesi (2000c) attributes students going beyond structures permitted in examples and using target words creatively to "communicative language teaching, which may, in some of its manifestations sacrifice accuracy in an attempt to foster fluency and self-expression" (p. 115). Similarly, from her informal conversations with students she cites a lack of awareness of grammatical concepts such as transitivity and grammatical collocation. She also attributes this to the influence of communicative language teaching.

Thus far, although dealing with subjects at different proficiency levels, the studies of exemplification examined have been fairly generalizable in terms of user L1. However, there are a handful of studies which deal with aspects of exemplification with a specific L1 user group in mind. Lew and Dziemianko (2012) examines a possible mismatch between folk defining in Polish and the folk defining techniques which employ single *when*-clauses. The latter underlie definitions in several English MLDs. For instance, the example for *ascent* in *Cambridge Advanced Learners' Dictionary* (Walter, 2005) reads: *when someone starts to become successful*. Lew and Dziemianko (2012) report on a series of experiments to investigate whether native speakers of Polish could recognise the part of speech of English abstract nouns when presented with single-clause *when*-definitions found like those in MLDs. Their first study (Lew & Dziemianko, 2006a) involved 129 Polish L1 upper-intermediate or advanced learners of English at university level. The experiment consisted of two tasks involving 20 headwords. First the participants had to match the headwords to their Polish equivalents. Secondly, they had to write a sentence using the headword. In both tasks participants had considerably more success recognising the part of speech of the target items when they used traditional analytical definitions. A further study (Lew & Dziemianko, 2006b), this time involving 238 secondary school students of an intermediate proficiency-level who were asked to choose the correct translation of headword, suggests that the difference in successful use of *when*-definitions and traditional definitions could be mitigated by the inclusion of semantic class labels. Lew and Dziemianko (2012) compares the two defining styles using a definition matching task. The latter study involved 134 native speakers of Polish of an upper-intermediate to advanced proficiency level. No further contextual information is given. Results of the latter indicate traditional analytic definitions were marginally significantly more effective. Although the effect size observed was very low.

In his speculative study G. F. Huang (1985) compares common errors made by Chinese students in writing and entries from three MLDs: the LDOCE1, CULD, and OALD3. He argues that Chinese students' productive needs would be better served by examples which reflect verb agreements after expressions of quantity, and agreement between collective nouns and finite verb. He also makes a case in favour of space-consuming examples over abstract codes arguing that students appreciate their immediacy.

This brief summary of research on examples in learner dictionaries makes it clear that the issue is more complex than presence of examples or otherwise. The style of examples plays

a role. There is strong evidence that users seem to prefer COBUILD's sentence examples. It is also possible that a tendency to include creative rather than typical uses in examples might be detrimental the effectiveness of MLDs. Frankenberg-Garcia's (2015) research also indicates that the intended purpose, productive or receptive, of consultation also plays a role.

3.3.2 Visual Elements

A good deal of research has been conducted on the role of signposts, cues throughout the entry which are intended to help the user distinguish between senses, in dictionaries for language learners. In user studies involving traditional dictionaries, Lew (2010) employed a translation task involving 90 Polish A2 and B1 level high-school students, while Nesi and Tan (2011) employed a sense selection task involving 124 second and third year mixed-proficiency undergraduate students studying at a university in Malaysia. Both tasks point to the role signposts play in improving accuracy although neither study found evidence of signposts improving speed of retrieval. From the point of view of the lexicographer, DeCesaris (2012) highlights the challenges that the creation of signposts pose in the bilingual dictionary creation process.

In the digital realm eye-tracking has been employed to study signposts. In a study involving eight English L2 speakers from Tokyo Foreign Studies University, Tono (2011) used an eye-tracking experiment to examine how supporting devices such as signposts or menus, different types of grammar codes, and positions of target definitions influenced look-up behaviour in MLDs and bilingual dictionaries. In addition to finding that almost one third of lookups ended in failure, he found that signposts were more effective for higher-proficiency learners, while lower proficiency learners performed better with menus. Although scanning bilingual entries proved easier than scanning monolingual ones, there was no significant difference in performance on a meaning finding task between the two entry types. Bilingual, lookups resulted in a higher success rate than MLD lookups if information was presented at the beginning of an entry, however, in entries where information was not placed at the beginning or not obvious there was little difference. This calls the reasons underlying lower proficiency level users' preference for bilingual dictionaries into question.

Dziemianko (2016) employed an eye-tracking study to compare the efficiency of methods for presenting signposts in three online MLDs: LDOCE5, OALD8 and OALD9. The study involved 243 advanced-level learners of English (C1 in CEFR) at Adam Mickiewicz University in Poznań, Poland. The participants completed a sense selection and comprehension task. They were divided into three groups: 91 subjects did the test with signposts in crimson capitals ruled off from relevant senses by a crimson line, 84 students did the test in with signposts were printed in white capitals on a blue background. The remaining 68 participants did the test with signposts in lower case above a dark orange line. Dziemianko found the signpost highlighting in the LDOCE5 (white capitals on a blue background) to be the most efficient in terms of speeding up sense identification and retention. On the other hand, neither signpost style or sense distribution had a significant influence on participant's ability to identify the correct sense.

In addition to signposts, eye-tracking has provided pertinent insights into other aspects of digital dictionaries and language learning. In a study involving six foreign language majors at the University of Tokyo, Kaneta (2011) compares a traditional dictionary interface to a layered interface where illustrative examples including idioms and phrases are folded and

have to be opened via a menu. He found no significant difference between users performance for the two interfaces. However, users consulting the traditional interface referred to examples more frequently yet for shorter periods that those using the layered interface. In a study involving ten Polish learners of English at university, five at a B2 to C1 level and five at an A2 to B1 level. Lew et al (2013), examined how learners selected senses in polysemous bilingual dictionary entries. Overall, participants selected the correct sense 80 per cent of the time. High-proficiency users tended to examine every sense in the entry while low-proficiency participants tended to stop their search when they had found the correct sense. This raises an interesting parallel with Müllich's (1990) sham dictionary use. To their surprise the authors found no evidence that shorter entries lead to higher success rates. On the contrary, they noted the opposite tendency.

The digitisation of dictionaries has also offered the possibility of new forms of exemplification. Like several of the studies discussed in the previous chapter section, Lew and Doroszewska (2009) investigated the influence of different defining styles on vocabulary retention. What is novel about their study is that they also test the effectiveness of animated pictures as definitions in addition to L1 (English) definitions, L2 (Polish) definitions, and examples. Their participants were 56 native Polish-speaking school students (age 17 to 18) all of whom had a proficiency level in the A2 to B1 range. After a vocabulary pre-test, students read a short passage containing ten target words which the pre-test indicated were probably unknown. They then took a vocabulary retention test and finished by completing a short reading comprehension. The Polish L1 equivalent, alone or in combination with L2 (English) definition was the best predictor of retention as well as being the most consulted defining style. Although, few subjects chose to look up examples, a regression analysis suggested examples had a positive role. The authors express surprise that animations performed poorly with retention rates around half that of the other options.

Since they are unrestrained by printing costs, online dictionaries make the use of a range of colours a financially viable option. Dziemianko (2015) investigates the influence the colour of part-of-speech and syntactic labels have on the speed and effectiveness the search and retention of grammatical information in dictionaries. The study involved 219 B2 to C1 level participants who studied degrees in the Faculty of English at Adam Mickiewicz University in Poznan, Poland. It employed an online multiple-choice test to be completed after consulting one of two versions of a purpose-built e-dictionary. In one version the entries in black and white, in the other, labels were in colour. Results indicated that displaying the functional labels in colour significantly increased the speed and effectiveness of online dictionary search as well as improving retention scores.

The digitalisation of lexicography has blurred the lines between dictionaries and other lexicographical resources. The hypertextuality of digital lexicography permits its exploitation in many practical applications for language learning. This includes glossing, the automatic creation of exercises from dictionaries, and the creation of wordlists or term banks. Glossing takes place in many forms, Roby (1999) provides a detailed taxonomy of these. In general glosses have been shown to have a positive effect on vocabulary learning (Bowles, 2004; Hulstijn, Hollander, & Greidanus, 1996; G. Jacobs, Dufon, & Hong, 1994; Jung, 2016; Ko, 2012; Watanabe, 1997). This is mirrored in the digital realm where, in a meta-analysis of computer-mediated glosses, Abraham (2008) found a medium effect size for glosses on comprehension and a large effect size for vocabulary acquisition. De Riddler (2002) investigated hyperlink glosses in electronic texts and found that although highly visible highlighted and underlined hyperlinks lead to much more clicking, in terms of

improvements in vocabulary learning or comprehension scores, they offered no advantage over unmarked hyperlinks. The automatic creation of exercises using online lexicographical resources has been shown to be feasible. As its name suggests, Hayward's (n.d.) *AWL Highlighter and Gapmaker* employs the AWL to highlight or remove a text's so-called academic vocabulary. Cobb's (n.d.) *Compleat Lexical Tutor* draws on the AWL as well as WordNet (Fellbaum, 1998) to automatically create vocabulary practice exercises. Similarly, projects such as *Collocaid* aimed the integration of lexicographical data about academic collocations with word processing tools have presented a number of interesting perspectives on the visualisation of academic collocations (Frankenberg-Garcia, Lew, Roberts, Rees, & Pereda, 2017; Roberts, Frankenberg-Garcia, Lew, Rees, & Pereda, 2017).

3.4 Phraseology and MLDs in Context

English MLDs have always shown special concern for phraseology. This is evident in the dictionaries produced during the first historical stage of MLD production. For example, Cowie (2002) highlights the "exceptionally rich treatment of verb combinations" (p. 59) in Harold Palmer's GEW. Hornby's OALD1, with its classification of grammatical collocations in four basic structural types and use of brackets enclosing collocates to show collocational range of lexical collocations, showed the way for the specialist collocational dictionaries of the 1980s and 1990s. Along the road to these fully-fledged collocational dictionaries we find the *Longman Dictionary of English Idioms* (T. Long, 1979) the second volume of the *Oxford Dictionary of Current Idiomatic English* (Cowie, Mackin, & McCaig, 1983) which was later renamed the *Oxford Dictionary of English Idioms* (Cowie, Mackin, & McCaig, 1993). These works are principally concerned with the classification and collocatability of idioms. Collocational dictionaries, in contrast, are concerned with the collocatability of single words in general, since for reasons of space MLDs were not able to provide the wide coverage learners demanded (Herbst, 1996). According to Cowie (2002) *Selected English Collocations* (Kozlowska & Dzierzanowska, 1987) represents the first reliable dictionary of English collocations. In each entry collocates are grouped according to part-of-speech. The *BBI Combinatory Dictionary* (Benson, Benson & Ilson, 1986) is perhaps the most well-known and widely respected dictionary of English collocations. However, it is not immune from criticisms, Cowie (2002) criticises its inclusion of syntactic complementation information, which in his strict definition of collocation, has no place in a collocational dictionary. The present dissertation takes a much less conservative and more inclusive view of collocation. This is in part a pragmatic stance in which any information which might help the user in encoding while not overly burdensome in terms of space or complexity to a dictionary would be included.

In parallel to the emergence of specialist collocational dictionaries, the second and third stages of historical development of MLDs saw many advances in the treatment of phraseology. Developments in the OALD2 are characterised by a greater concern for users' decoding needs than those in its predecessor; as such the way in which verb patterns are represented is largely unaltered. However, Cowie (2002) finds evidence of more detailed analysis of noun and adjective complementation, some attempt to match constituent elements of patterns to functional categories, and the beginnings of a novel approach to phrasal verbs which would be further developed in the OALD3. The OALD3 saw significant changes to the treatment of verb patterns. Verbs were categorized according to major verb-types: copular, intransitive, mono-transitive, di-transitive and a number of sub-patterns which took into account the semantic type of the complement were introduced. In spite of these developments, the dictionaries reliance on abstract codes to refer to verb patterns impeded their usefulness (Cowie, 1984; Heath, 1982; Lemmens & Wekker, 1986).

Other notable phraseological developments include the insertion of prepositions as part of skeleton examples, more comprehensive complementation frames, and examples which include complex typical post-modification of the headword, and the marking in boldface of phrasal items that are invariable. Changes in the presentation of phrasal verbs continued from the OALD2. Phrasal verbs with fixed collocates in object position were listed immediately after the simple phrasal verb. Phrasal verbs were also listed in bold at the edge of each column. Like the OALDs the LDOCE codified verb complementation, but also introduced codes for subclasses of nouns, adjectives and their post-modifying compliments. However, the scheme has received criticism in some quarters for being excessively complex and for containing examples which do not indicate the range of collocational patterns. As far as the indexing of idioms is concerned the LDOCE adopts the same approach as the OALDs, it assumes users are able to analyse idioms in terms of transparency and compositionality, within entries idioms were listed in alphabetically numbered subsections. In contrast phrasal verbs were treated as separate entries. This resulted in the vexed question of how to distinguish between idiomatic and semantically transparent phrasal verbs.

The third stage of historical development of the MLD is primarily characterized by the increased use of computers. It also included many phraseological innovations. For example, complementation codes in the LDOCE2 are far less complex than its predecessor. COBUILD1 adopted an innovative extra-column approach for clause patterns and function labels and what could be termed a phraseological defining style in which the headword formed an essential part of the definition. These phrasal or sentence definitions often resemble folk definitions. The OALD4 represents a complete redesign of verb-pattern scheme of its previous editions. The dual-level of description it contains might cause problems even for advanced students. There were also phraseological innovations in its defining style with typical objects included in parenthesis in the definitions.

In spite of the many advances made in English learner lexicography, and the great deal of critical attention that the treatment of phraseology in MLDs and collocational dictionaries has received, there is relatively little research which deals with the uses and users of the phraseological aspects of lexicographical resources for non-native speakers.

3.4.1 Research on Phraseology in General MLDs

Paul Bogaards conducted a series of experiments dealing with look up strategies for phraseology in dictionaries. Bogaards (1992) set out to examine the effect of frequency on look-up strategies for multi-word expressions monolingual French dictionaries. His study involved 62 L1 French students at secondary level. They were given lists of fixed expressions chosen according to frequency rank and asked to underline the keyword they would use to look up the expression in a dictionary. This was followed by a post-test which was applied two weeks later and intended to ensure that frequency was indeed the explanatory factor. In the second test participants had to decide which word in a list of word pairs was the most frequent. Bogaards concludes, with some reservations, that French L1 students using monolingual dictionaries tend to look up the least frequent word when looking for multi-word expressions. Turning his attention to learners' dictionaries and assuming that MWUs are indexed by their base, Bogaards (1999) examined whether learners looked up a word using the base or the collocate. The experiment which involved 56 L1 Dutch learners of French in their first (N = 41) and third years (N = 15) of university. Although no statistical significance tests were employed, there was no obvious preference for base or collocate among first-year students while third-year students tended

to look up the base. Users showed a slight preference for nouns irrespective of their function. This may be because participants simply choose a word they recognised. In a study involving 59 secondary school-level and 29 university-level L1 Dutch students of English, Bogaards and Van der Kloot (2001) used a translation task to examine the usefulness of the systems used to present grammatical information in three MLDs (CIDE, COBUILD2, LDOCE3). There was no statistically significant difference in time taken to find the relevant information and correctness of the translations produced by participants using different dictionaries. In a further study, Bogaards and van der Kloot (2002) leave the question of findability aside concentrating instead on which type of grammatical information students used most and which type of information was most useful to them. The second study involved three groups of L1 Dutch speaking learners of English: the first comprised 33 secondary school students, the second 56 first-year university students; the third group contained 28 third-year university students. Students had to complete a translation task and underline the information they had used. Overall results suggest that grammatical codes relating to word classes are rarely used; examples are widely used especially, though not exclusively, by intermediate users. University users seemed to prefer less abstract grammatical information. Many users demonstrated a preference for using the definitions and examples even when there was more explicit COBUILD-style grammatical information available. Ideally then, definitions and examples should encode grammatical information in a natural way. CPA is possible means of achieving this. Renau and Battaner (2012) demonstrate the viability of CPA for the analysis of Spanish pronominal verbs and their representation in MLDs. However, in their work on a phraseological dictionary of science verbs G. Williams and Millon (2010) suggest that users could experience difficulty understanding CPA semantic classes at first.

The scant influence of the way in which verb syntax is encoded on the usefulness in translations might be related to the L1 of the participants under study. The results of Dziemianko's (2006) study involving 606 Polish learners of English at high school and university seem to run counter to Bogaards and Van der Kloot's (2002) findings. It seems that Polish learners of English tend to make more use of COBUILD type definitions than analytical ones when looking up information on verb syntax. However, overall participants made more use examples than either analytical or COBUILD style definitions. Since the Polish university participants were all attending linguistics and English grammar classes it is possible they were more familiar with the COBUILD encoding syntax than users working in other disciplines would be.

In the digital realm, Heid and Zimmermann (2012) report on two usability tests conducted in an attempt to find the most appropriate design for search interfaces for collocations in online dictionaries. In a test of German and Italian online dictionaries, Bank (2010) found that search interfaces which involved many steps had a negative impact on users' performance. Users preferred simple search procedures which returned a list of results. To investigate this further Heid & Zimmermann (2012) conducted usability tests comparing two mock-ups of collocational dictionaries; one profile-based and one search-engine based. They found that their advanced-level translator participants preferred the profile-based search for productive purposes.

3.5 Dictionaries and EAP

Learner dictionaries have always been closely linked the teaching and learning of English in academic contexts. The demands of teaching English in academic contexts were an important influence in the creation and development of learner dictionaries. For example

West's NMED was influenced by its editor's experiences as Principal of Dacca Teacher Training College (R. Smith, 2003). Likewise Hornby's work was undoubtedly influenced by his experiences in higher education first at the Tokyo University of Literature and Science, later at the Tokyo School of Foreign Languages and Tokyo Higher Normal school (Cowie, 2002).

In spite of their origins, the extent to which contemporary MLDs are apt for academic contexts has been questioned. D. Lea (2014a) argues that MLDs are only academic in the sense that they sometimes mark entries which pertain to AWL words and often include supplements dealing with academic writing. D. Lea (2014a) outlines creation process of a "genuinely academic learners' dictionary" (p. 181) the *Oxford Learner's Dictionary of Academic English* (OLDAE; OUP, 2014). The OLDAE adopts the frequency-based: general, sub-technical/academic, technical view of English vocabulary aiming for the middle ground of a general academic vocabulary. It is corpus-based, the text from which its decoding function was created comes from undergraduate higher education textbooks and the text which represents its encoding function comes from professionally written published academic journal articles. Echoing a common criticism of the use of professional writing as an example for learners or trainee writers, D. Lea (2014a) claims that such texts are not entirely appropriate for this group since examples derived from such texts may be overly complicated. This criticism is discussed in greater detail in the next chapter. The OLDAE's editors modify examples through process of "distillation of a number of different concordance lines, all of them very similar" (Lea, 2014a, p. 188) when the example taken directly from the concordance line is deemed too complicated. This process has interesting parallels with CPA; a procedure in which through an examination of corpus lines a semantic type is assigned to a lexical set, from which an implicature is then derived. This process effectively results in a general yet authentic example distilled from many corpus lines.

D. Lea (2014a) concedes that the conception of general academic vocabulary is somewhat problematic: "The student of economics […] is not well served by a mere description of a recession, when it is in fact a very precisely defined economic term."(p. 187). The OLDAE addresses this by providing a descriptive definition typical of MLDs followed by the strict discipline-specific definition. At first glance, Coffey's (2016) analysis of a sample of 100 OLDAE headwords, would seem to cast doubt on the academic nature of the dictionary's vocabulary since only one (*mass-spectrometer* from chemistry) is not included in the general OALD9. However, as Coffey points out, around a quarter of words in the sample demonstrate important differences in definitions and examples when compared those in the OALD9. Taking into account the importance of context in conditioning meaning, especially for productive use, the academic dictionary gives more thorough treatment to collocations than a general MLD. Around 700 "collocationally prolific words" (Lea, 2014, p. 184) are presented as separate entry lists in the style of a collocation dictionary.

The OLDAE is not the only academic learners' dictionary. *The Louvain EAP Dictionary* (LEAD; CECL, 2010) is a web-based English for Academic Purposes "dictionary-cum-writing aid" (Paquot, 2012, p. 163) for non-native writers. At the time of writing the LEAD dictionary was not widely available, access was restricted members of the Université Catholique de Louvain and the dictionary was undergoing beta testing. LEAD seems like a very promising resource, however in spite of three conference papers (Granger & Paquot, 2010a, 2010b, 2017), one book chapter (Paquot, 2012), and one article (Paquot & Granger, 2015) dealing with the dictionary, there is very little information available about its construction. Its creators list an emphasis on phraseology, as well as semasiological and

onomasiological, and semi-bilingualised access among its features. Users studying business, medicine, or linguistics are also given the choice of some discipline-specific customisation. However, it seems that this involves the provision of examples taken from texts from these disciplines rather than a systematic treatment and comparison of discipline-specific meaning.

A great deal of the research into the users and uses of learner dictionaries has been carried out in academic contexts. Many lexicographers work in universities and other higher education institutions. These institutions represent a rich source of participants for studies of dictionary use. The summary of research into dictionary use and users above reflects this insomuch as all but a handful of studies were carried out in an academic context or with university student participants. The majority of the findings of studies of the uses and users of learner dictionaries could equally apply to EAP lexical resources. In short, it is evident that dictionary access promotes vocabulary acquisition, that certain parts of microstructure are problematic, particularly those which deal with collocational and selectional restrictions, and that the effectiveness of examples depends on the purpose with which they are consulted. There is some evidence which suggests learners prefer sentence examples, and that some unsuccessful lookups are caused by the inclusion of examples of creative use rather than typical use.

There are several findings which are especially pertinent to EAP lexicography. Questionnaires dealing with user preferences indicate that advanced learners prefer MLDs. For this group, this type of dictionary has also been shown to promote more effective vocabulary retention than bilingual dictionaries. Assuming EAP users are primarily upper-intermediate or advanced learners these findings support the contention that EAP lexicography should focus primarily on monolingual resources. There is also an economic argument which supports this contention. In theory, monolingual dictionaries allow the concentration of resources on a handful of dictionaries while in the case of bilingual English dictionaries these resources would be diluted among various languages.

One of the reasons put forward for the lack of influence of dictionaries on comprehension scores is their failure to deal adequately with technical and scientific vocabulary. Similarly, defining scientific and technical concepts within a controlled vocabulary is also problematic. These findings support the need for multi-discipline lexicographical resources in EAP. The following chapter section outlines a number of approaches to discipline-specific EAP lexicography.

3.5.1 Discipline Specificity

The need for specialist dictionaries for EAP is clear. However, at present certain disciplines are better represented than others. For example, there are several monolingual English business dictionaries aimed at non-native speakers. This is not only because many non-native speakers have to use business English in their professional lives, but also, in an EAP context, because business studies and related disciplines have been by far the most popular subjects for international students (HESA, 2017). De Cock (2006) compares the treatment of seventeen key business English lexical-items in MLDs, with their treatment in specialist business MLDs (SBMLDs) she concludes that although the five MLDs cover between twelve and fifteen of the seventeen items, the *Oxford Business English Dictionary for Learners of English* (Parkinson & Noble, 2005) and *Longman Business English Dictionary* (Dignen, 2000) better meet the productive needs of the business English user. While the MLDs include many of the lexical items as sub-entries, in the SBMLDs fifteen of them are included as

carefully and systematically defined headwords supported by several usage examples in context. She also highlights the increased consideration of phraseology in SBMLDs.

In a comparison of the treatment of collocation in MLDs, collocational dictionaries, and SBMLDs, Walker (2009) examines the treatment of phraseology in business English in greater detail. He argues that the emphasis that both MLDs and SBMLDs place on frequency obscures small, yet significant differences in meaning between semantically related items. He also argues that general MLDs fail to explain "important semantic differences between items such as *run, head* and *manage*, or *aim*, *objective, target* and *goal" (p. 297)*. This could be remedied by focusing on these items in collocation. On a macrostructural level, he recommends grouping collocations by their semantic relations since the current practice of alphabetic organization encourages a view of the relations between collocations as arbitrary. He also calls the hegemony of nouns in collocational dictionary entries into question since there is there is no research indicating learners start the look up process with this part of speech.

In a methodologically innovative study involving 85 mixed English level, first-year university students of chemistry, Campoy Cubillo (2002) sought to discover this group's dictionary skills preferences. Participants created dictionaries containing around 100 words of their choice from the most important or useful texts they had studied during a first-year chemistry course. They were instructed to include as much information as they deemed necessary. They reported on the dictionary creation process in a short composition. The contents of dictionaries were then analysed in a bid to discover what chemistry students want in a dictionary. Analysis of the students' compositions revealed that most had employed bilingual dictionaries in the process of making their own dictionary. Campoy Cubillo regards the superficial nature of the examples provided, and the relatively few students who included other information, such as word sense, collocational information, and synonyms as evidence for a lack of dictionary skills on the part of the participants. The validity of this conclusion is questionable since there are many reasons why students might have omitted information. However the lexical content of the students' creations is more pertinent. Students' dictionaries contained both general English and specialised vocabulary. This suggests that they would like a general learners' dictionary which also contains vocabulary specific to their discipline.

In a different discipline, in two related studies which have wider implications than Campoy Cubillo (2002), de Cock (2006), and Walker (2009); G. Williams (2006, 2011) examines the treatment of words in the *Oxford Scientific*, a list of 265 word-forms aimed at helping users of the OALD understand scientific texts, in various dictionaries. G. Williams (2006) examines the treatment of *Oxford Scientific* words in the OALD, COBUILD, and MEDAL. G. Williams (2011) concentrates specifically the OALD. He argues that the treatment of specialist scientific vocabulary in MLDs is inconsistent and inadequate for an ESP/EAP audience. He demonstrates how a comparison of Oxford Scientific words in a general scientific corpus, the scientific section of the BNC, with a specialist corpus of parasitic plant biology articles (G. Williams, 1998) brings to light opportunities make slight changes to definitional prototypes to account for specialised usage. This would enable the creation of specialised production dictionaries for users who already have a substantial communicative ability in English but require help in production with that specialist vocabulary specific to their discipline in the same way that a general MLD provides help to the user in production in a general context. In the view of EAP adopted in this thesis, this description certainly applies to EAP users.

Although they do not identify their dictionary as an EAP dictionary, the work of Geoffrey Williams and colleagues on dictionaries for non-native English speaking scientists offers greater detail about the construction process and underlying rationale than the Louvain project. Building on research on collocational networks in a corpus of parasitic plant biology articles (G. Williams, 1998), G. Williams (2001, 2002) turns his attention to the use of collocational networks for headword extraction for specialised dictionaries. The crystallisation of this research can be seen in G. Williams and Millon's (2010, 2014) report on the creation of an *E-Advanced Learner's Dictionary of Verbs in Science*. Motivated by the lack of writing aids in MLDs, and the need to bridge the gap between general dictionaries and terminological resources, G. Williams and Millon (2014) outline a project to create an organic bottom-up dictionary based on the 33 million word BioMed corpus which comprises texts from biology and medicine journals. These texts include research articles, as well as letters, comments, and editorials. The 100 most frequently occurring verbs in the corpus are used as start nodes. These are linked to collocates extracted using statistical measures. These in turn link to other nodes in an iterative process, slowly building a complex network of lexicographical prototypes. By examining the evolution of node word using the OED and comparing collocational networks in general and specific corpora, light can be shed on etymological transfer and transfer between contexts. The outcome of the process is a semasiological and onomasiological dictionary which is organised alphabetically, as well as thematically around conceptual classes such as INVESTIGATE. A sub-class where a human does the examination is reproduced below (G. Williams & Millon, 2014; p. 51):

Investigate
To search or inquire into (1510)
X [We, author] investigate the effects (358) of (349)
Examine
Investigate by inspection (1330)
X [we, researcher, author, investigator, study] examine the effect(s) (535) of Y (518) on
Explore
To look closely into (*1592*)
X [We/study/analysis] explore the relationship(s)

The examples in italics are taken from the OED with the date they were first attested given in parenthesis. Numbers in boldface in parenthesis give frequency in the BioMed corpus.

3.6 Summary

This chapter has set out what is understood as an EAP user in the present study. As well providing a historical overview of the evolution of English pedagogical lexicography, it has reviewed research into dictionary use and users which will be helpful in achieving one of the aims of the study: The creation of guidelines for useful phraseological lexical resources aimed at EAP users. That said, it is important to bear in mind that dictionary research is context-bound and the variation in L1 background, proficiency-level, and academic setting limit the extent to which findings from a single study can be generalised for EAP users as a group. Similarly, pedagogical lexicography has undergone a great deal of change in the last 30 years, MLDs published today are different from those published 30 years ago, as a result the relevance of research carried out on older MLDs to MLDs today is sometimes questionable. The free agency of users represents yet another limitation to studies into the use of dictionaries and other lexicographical resources. Ultimately there is no means of

knowing how a user will use a lexicographical resource; if it will be used for its intended purpose or for some novel unforeseen function.

With these caveats in mind it is possible to make some generalisations about dictionaries and EAP. The first and most uncontroversial generalisation is that dictionaries have a positive influence on vocabulary acquisition. This lends weight to the aim of creating useful lexical resources for EAP users. Another generalisation suggested by observation-based and questionnaire-based research is that users' preference for and successful use of monolingual resources increases as a function of proficiency level. Taking a view of EAP users as predominantly upper-intermediate (B2) to advanced (C1/C2) users, this suggests that for the purposes of this study resources should be concentrated on guidelines for the creation of monolingual lexicographical resources. Research also suggests that the future of pedagogical lexicography is digital. It is clear that digital lexicographical resources offer many advantages in terms of customisation of access structures, size, and range of visual aids that can be included. It is also apparent from studies dealing with dictionaries and reading comprehension, defining style, and examples that the treatment of technical or discipline-specific vocabulary in lexicographical resources is a limiting factor in their usefulness to users. Failure to adequately encode phraseology, that is to say, selectional and collocational restrictions in an accessible manner is another limiting factor.

With this in mind, in the context of EAP, it is clear that more work is needed on sense discrimination between disciplines. Taking a phraseological view of language this naturally involves looking at patterns of collocation and how patterns might be usefully represented. The work on collocational resonance carried out by Geoffrey Williams and colleagues and Irene Renau and Paz Battaner's work on the representations of Spanish pronominal verbs in MLDs gives an indication of how such an aim might be achieved. However it is first necessary to examine differences in word meaning across academic disciplines.

This examination will be achieved through a corpus-based experiment comparing verbal collocations across the academic disciplines of history, management, and microbiology. The hypotheses tested in this experiment and the assumptions that underlie the study as a whole are laid out in the following chapter. Chapter 5 outlines the experimental procedure. Chapter 6 presents a statistical analysis of the results. A qualitative analysis including guidelines for and examples of how discipline specific phraseology might usefully be presented for EAP users is provided in Chapter 7. Conclusions are set out in Chapter 8.

4. HYPOTHESES AND ASSUMPTIONS

This dissertation is centred around a corpus-based experiment intended to examine differences in meaning and use of vocabulary across academic disciplines which are obfuscated on approaches currently adopted in studies of EAP lexis. In doing so it aims to demonstrate a means of phrase extraction which accounts for both semantic and syntactic concerns. The experiment and the analysis which follows it will test the following general hypotheses[9] which relate to the objectives of the present study.

The principal objective:

> To examine differences in meaning and use of vocabulary between academic disciplines that might be obfuscated by the distributional approaches to vocabulary selection habitually employed in English vocabulary studies.

Relates to the following hypothesis:

> The lexical-bundle approach obfuscates differences in meaning and use of vocabulary across academic disciplines.

The methodological sub-objective:

> To demonstrate a feasible means of phrase extraction that accounts for both semantic and syntactic concerns.

Relates to the following hypothesis:

> The method outlined in Chapter 5 provides a feasible means of vocabulary extraction that can deal with large quantities of data and both syntactic and semantic concerns.

The lexicographical sub-objective:

> To create guidelines for producing, and an illustrative examples of, a useful lexicographical resource for the EAP community.

Relates to the following hypothesis:

> CPA is a viable basis for the creation of an EAP lexicographical resource that provides a feasible response to many of the challenges faced by EFL lexicographers.

The testing of all the above hypothesis should make the achievement of the theoretical sub-objective possible:

> To examine how the behaviour of academic lexis relates to wider phraseological approaches.

The data resulting from the experiment will facilitate some objectives of the study. Namely, to create, illustrative examples of, and guidelines for producing, useful EAP vocabulary resources and to examine the extent to which findings about the behaviour of academic lexis relate to phraseological approaches to language in general. The hypotheses listed above, the corpus-based experiment which tests them, and the intention to create, illustrative examples of, and guidelines for producing, useful EAP vocabulary resources rely on several assumptions. The first group of assumptions relates to the nature of EAP

[9] These should not be confused with the formal hypotheses employed in statistical significance testing set out in Chapter 5 and reported in Chapter 6.

writing. The second assumption relates to the decision, in the present study, to concentrate on verbs and their collocates rather than collocations involving other parts of speech. Finally, retaking the thread laid out in the previous chapter, there are assumptions about the nature of EAP users and the market for EAP lexicographical resources.

4.1 Assumptions Related to Corpus Composition

The first assumption related to corpus composition concerns the mode of language the corpus contains. The hypotheses set out above are tested with written language. This is not because they apply only to this mode, but rather, a result of practical constraints such as the difficulty inherent in obtaining sufficient speech to compile a corpus of sufficient magnitude for a meaningful lexicographical study. Moreover, as suggested by the discussion of Linear Unit Grammar (Sinclair & Mauranen, 2006) in Chapter 1, the large quantity of time which would be needed to process this data. That said, a good deal of corpus research has been carried out on spoken academic language. In vocabulary studies, Simpson-Vlach and Ellis's (2010) *Academic Formulas List* is based in part on the Michigan Corpus of Academic Spoken English (Simpson, Briggs, Ovens, & Swales, 2002), while Biber and Barbieri (2007) compare lexical bundles in spoken and written academic registers. Particularly fruitful use of corpora of spoken academic English has been made, in the area of conversation analysis (see for example Barbieri, 2013; Morton, Walsh, & O'Keeffe, 2011; Walsh & Knight, 2016).

As is apparent from the review of literature on the teaching of and learning of vocabulary in Chapter 2, throughout ELT there is a widespread distinction made between students' productive and receptive needs. This distinction also holds in the debate surrounding the ideal composition of corpora in studies of EAP vocabulary. As far as receptive skills are concerned, it is generally accepted that both texts written by students and published academic writing (PAW) are useful to EAP students. The first generation of EAP wordlists (Campion & Elley, 1971; P. Ghadessy, 1979; Lynn, 1973; Praninskas, 1972) were created with receptive needs in mind and aimed primarily at prospective or first-year university students. They were compiled by examining the textbooks students would be expected to use in the course of their studies. Given that, in that pre-internet era, access to research articles (RAs) was not as simple as it is today, the decision to base lists on textbooks frequently consulted by students is understandable.

The implicit assumption of these word-list studies is that EAP users' receptive needs are limited to a requirement to understand textbooks. This is not the case. Firstly, it is premised on a narrow definition of EAP users as prospective students or undergraduate students which does not reflect the whole EAP-user population. Secondly, it assumes that that journal articles are not an important reference source for these students. In reference to the first point, as the discussion in the previous chapter suggests and as will be argued further in section 4.3 below, the EAP user population is more heterogeneous than this. As far as the second point is concerned, there is convincing evidence that journal articles have an important role at all levels of tertiary education. Weir, Hawkey, Green, Unaldi, & Devi, (2009) conducted a large-scale methodologically rigorous survey of the reading experiences of 766 students at the University of Bedfordshire. Although, students across all language backgrounds and academic levels regarded books as the most important information source on their course, 83.5 per cent of postgraduate participants and 70.7 per cent of the undergraduate participants definitely or mostly agreed that journal articles were important.

The corpus composition in those more recent EAP vocabulary studies which are not explicitly limited to receptive skills has provoked a great deal more debate particularly as far as writing is concerned. With the emergence of EAP writing instruction as a specialism in the late 1980s came an increased interest in how the teaching of disciplinary writing should proceed. At this stage, there were two contrasting paradigms. On the one hand, process writing drew strongly on the North-American tradition of teaching rhetoric and composition. Proponents of this approach such as Flower and Hayes (1977) placed great emphasis on individual inventiveness and use of rhetoric rather than the conventions of any particular discipline. In contrast, an approach based on genre analysis (Swales, 1990) advocated students critically analysing the genres of writing they needed to produce. This latter approach owes much to developments in systemic functional linguistics (Halliday, 1989) in the United Kingdom, and emerging multicultural and citizenship teaching practices in Australia (Cope & Kalantzis, 1993). Insomuch as it relies on published academic research articles, which it is assumed contain the lexis EAP writers need to produce, the present study falls within the genre analysis paradigm.

Several criticisms of the genre analysis approach to EAP writing instruction (EAPWI) have come to the fore. Firstly, genre-based EAPWI programmes have been criticised for their assumption that student writing is somehow intrinsically deficient and can be fixed in the same way a pathology might be cured, all the while failing to challenge the opaque and often contradictory literacy requirements placed on students by academic institutions (M. R. Lea & Street, 1998). In the same vein those working in critical English for academic purposes have criticised EAPWI at an epistemic level for its recourse to inappropriate textual exemplars which they claim reproduce official discourses while effectively silencing voices which challenge these discourses (Lillis, 2003; Turner, 2012). Finally, there are those who from an English as an academic lingua franca (ELFA) standpoint criticise EAPWI as requiring students "accommodate to a narrow assimilationist model of English" (Jenkins, 2011, p. 927). All these criticisms may well be valid from a philosophical or theoretical perspective. However, it is beyond the scope of this thesis with its primarily practical-applied objectives to address this question in any meaningful way. Indeed, as Tribble (2017) claims "despite the important insights into the workings of the international academy which these approaches offer, there is little evidence in the research literature of their having had a significant impact on classroom teaching as yet." (p. 32). In other words, in the EAP classroom a genre analysis approach to EAPWI is here to stay.

On the genre analysis approach to EAPWI a pertinent question centres on what form textual exemplars should take. In corpus-based studies, this debate focuses on the question of whether PAW such as RAs are suitable exemplars for EAP students. It is generally accepted that many advanced level EAP users such as professional researchers and academics will need to write RAs. The debate instead has tended to focus on the needs of pre-sessional, undergraduate, Masters, and PhD students. With reference to Nesi (2014). Timmis (2015) claims "student writing is important as it may present a more realistic and attainable target for learners" (p. 147). There are some who argue that academic texts written by professional researchers for publication have different linguistic requirements to those written by students for assessment. Hyland's (2008a) contrastive study of academic clusters, otherwise known as lexical bundles, in masters dissertations, PhD theses, and published research articles typifies the arguments against employing PAW in corpus studies of EAP writing. Hyland (2008a) suggests that difference in relative frequency of occurrence and functional types of four-item lexical bundles across corpora representing masters-level theses, PhD theses, and published RAs demonstrates that PAW is not a suitable model for student writing. To support this argument he turns to Swales's (1990) contention that

writing for publication represents a "norm developing" (p. 31) practice which is primarily concerned with creating knowledge through the peer review process, while student writing represents a "norm developed" (p. 31) practice through which the student displays his or her knowledge. However, in more recent work Hyland (2016a) concedes that, in terms of structure, humanities RAs are similar to student essays.

Some notable studies of EAP vocabulary which employ corpora consisting of texts written by students include Paquot's (2007) examination of the suitability of the AWL for productive purposes, Ädel and Erman's (2012) investigation the use of lexical bundles in undergraduate writing by native English and Swedish speaker linguistics students at a Swedish university, and Nesi and Gardner's (2012) examination of the types of texts students need to produce at British universities. An informal small-scale meta-analysis of 40 corpus-based studies dealing with EAP writing reveals no clear correlation between the stated purpose of study and the types of text which comprise the corpus (see Appendix A). Of the 27 studies loosely categorised as aimed at creating or evaluating wordlists or phraselists 14 employ corpora comprising RAs. The five pre-2000 wordlist studies are all based on textbook corpora. This suggests that, in practice, the availability of source texts plays an important role in determining corpus composition. Of the 13 remaining studies, 11 could be loosely classified as dealing with discourse analysis. In this set, four studies exclusively employ corpora of RAs, two make use of the learned and academic sections of general corpora, S. Gardner and Nesi (2012) employ the BAWE corpus which contains writing by undergraduates who are predominantly native-speakers of English, and Ädel and Erman (2012) employ a corpus of undergraduate writing by non-native speakers of English. The remaining two studies in the DA set compare RAs with student writing across various academic levels. The final two studies could be very loosely described as concerning phrase extraction. H. Yang (1986) employs textbooks, while G. Williams (1998) employs RAs. No effort was made to ensure that the sample of literature analysed in this swift procedure was representative of corpus-based studies of EAP lexis as a whole. However, it does indicate that in practice RAs, student writing, and other texts have been used fruitfully to investigate EAP vocabulary.

An argument which is commonly put forward to support the use of RAs in studies of EAP writing and lexis is that such texts represent the ideal "expert performance" (Bazerman, 1994, p. 131) that is to say, the kind of writing which students ideally want to achieve (Tribble, 2002). Many EAP vocabulary studies have adopted this approach. Recent examples include M. Yang's (2015) examination of frequently-used vocabulary in nursing RAs which she carried out in order to aid learners read and publish nursing articles in English, and Le and Harrington's (2015) examination of word clusters in the results sections of applied linguistics articles.

Arguments against the use of PAW in studies of EAP writing, may well be valid. However, contrary to claims of many authors the evidence from corpus-based studies does not necessarily always support these conclusions. For example, Hyland's (2008a) assertion that difference in relative frequency of occurrence and functional types of four-item lexical bundles across corpora representing masters-level theses, PhD theses, and published RAs demonstrates that PAW is not a suitable model for student writing. However, in addition the inherent inaccuracy of relying on frequency alone to draw conclusions about the behaviour of lexis, Hyland employs a relatively small corpus (3,400,000 tokens). Similarly, while the student writings in Hyland's corpus were written primarily by Cantonese L1 learners of English, the RAs were written, or subject to editorial oversight, by L1 or advanced L2 speakers of English. This gives rise to the possibility that the differences

highlighted by Hyland (2008a) and similar studies are not due to any fundamental differences between requirements of the genre but rather a difference in language competence. This should not be taken as support for the argument, frequently put forward proponents of ELFA, that EAPWI condescends to non-native speakers (Jenkins, 2014). Instead it might be more appropriate and pedagogically fruitful, not to see EAPWI in terms of native vs. non-native writers and instead adopt a view in terms of expert writers vs. apprentices. Indeed, there is experimental evidence to suggest that as regards proficiency in the use of academic collocations, experience with EAP is a better predictor than native or non-native speaker status (Frankenberg-Garcia, 2017).

The view of experience with genre as the determining factor in writing proficiency has consequences for the status of student writing in corpus studies of EAP vocabulary. For Granger and Tribble (1998) the difficulty of writing in a new genre is its unfamiliarity for the student. In the context of British universities, they argue that most EFL students are competent writers in their own academic cultures, but unfamiliar with the target performance in the British institutions. This argument could equally apply, not only to EFL students but, to all newly-arrived first-year students at these institutions. Consequently, it would be undesirable to use a corpus of texts written by students as a model. This is not to say that student writing has no place in studies of EAP vocabulary. Learner and L1 student corpora offer an excellent means of understanding the errors students produce, and analysing the type of texts students need to produce.

The idea that PAW is not a suitable model for apprentice writers is often put forward as argument against the use of PAW corpora in studies of EAP writing and vocabulary. This argument is condescending. It suggests that RAs present some unachievable objective for novice-EAP users, or that RA authors are somehow beyond reproach. Hyland (2008a) describes how in contrast to PAWs, masters students must "demonstrate a suitable degree of intellectual autonomy while recognising readers' greater experience in the field" while doctoral students must "present an understanding of disciplinary working through an appropriate exposition of research and argument" (Hyland, 2008a, p. 47). It is somewhat disingenuous to suggest that experienced academic writers do not share these obligations.

Another criticism of studies based on corpora of RAs which is particularly relevant to studies comparing disciplines is that such studies cannot reflect differences in knowledge dissemination practices across disciplines. Central to this notion is the idea of so-called hard vs. soft disciplines which in turn is based on the idea of a hierarchy of sciences with hard sciences such as physics at the pinnacle and social sciences at the bottom. The root of this hierarchy dates back at least as far to the writings of positivist philosopher Auguste Comte (1896). While a number of empirical studies have shown that the hard/soft hierarchy is widely and similarly perceived (see L. Smith, Best, Stubbs, Johnston, & Archibald, 2000), there is little agreement on the exact nature of the factors which underlie it. Factors proposed include the degree of consensus, which is assumed to be higher in hard sciences (see Hargens, 1988). Another explanation is provided by L. Smith, Best, Stubbs, Johnston, and Archibald (2000) who demonstrate a relationship between subjects' perception of hardness of science and the number of graphs or charts used in a RA. They show how this relationship also holds for sub-disciplines of psychology with behavioural psychology at the top of the hierarchy and counselling at the bottom and intermediate positions occupied by social psychology and developmental psychology among others. Elsewhere, divisions of academic disciplines into hard and soft categories have been proposed; with humanities and arts among the soft sciences and physical and life sciences among the hard (F. Huang, Finkelstein, & Rostan, 2013, p. 45 footnote). On this basis, a

tentative ordering of the disciplines involved in the present study on a continuum might include microbiology at the hard pole and history at the soft with management studies occupying an intermediate position.

Today the RA is of great importance in academic publication. For Hyland (2016a) it represents the "pre-eminent genre of the physical sciences, engineering, much of the social sciences and, increasingly, the humanities" (Location 2729 Kindle edition). Among the harder disciplines, computer science with its preference for peer-reviewed conference proceedings is a notable exception to this rule (Becher & Trowler, 2001), as is botany with the publication of genres known as *flora* and *treatments* (Swales, 1998). Maci (2015) claims that posters are a major source of knowledge dissemination in the hard sciences, yet neglected in the soft sciences. Traditionally, scholars working in the softer disciplines of social sciences and humanities have preferred to publish in books or monographs (Hicks, 2004; M. Huang & Chang, 2008). Savage (2010) regards the monograph as being held in particularly high prestige in History. Hyland (2016a) also stresses the importance of multi-contributor edited volumes in the humanities and social sciences. However, in recent years changes in academic practice particularly the importance of research evaluation systems have somewhat diminished the appeal of publishing in books, monographs and multi-contributor edited volumes provoking a shift in preference to RAs (D. Harvey, 2006; K. K. Ward et al., 2009; P. Williams, Stevenson, Nicholas, Watkinson, & Rowlands, 2009).

There is a trend towards homogenisation of knowledge dissemination practices across disciplines. Although differences in publication practices across disciplines exist, in EAP vocabulary studies they tend to be somewhat overstated. Sparks (2005) survey of knowledge dissemination practices amongst 750 academics working at British universities suggests that some forms of dissemination are more frequently in used in some fields than others. Respondents were asked how they disseminated their work. They could select as many dissemination options as they wished. The results are reproduced in Table 4.1. It is evident that monographs are more frequently used in arts and humanities, and languages (70.9 and 84 per cent of respondents respectively) than medical and biological sciences, and physical sciences (11.3 and 9.2 per cent of respondents respectively). Similarly, technical reports were far more frequently used in the physical sciences (30.9 per cent of respondents) than in other areas. However, journal articles are commonly used across all disciplines. Their use ranging from 87.9 per cent of respondents from physical sciences to 98 per cent of respondents from languages.

Table 4.1

Means of dissemination used across disciplinary areas (reproduced from Sparks, 2005, p. 38)

	Medical and biological sciences	Physical sciences	Social sciences	Languages and area studies	Arts and humanities
Pre-prints	21.3	36.7	40.8	34.0	29.1
Post-prints	40.0	46.4	38.6	40.0	33.0
Presentation	84.4	75.4	75.3	70.0	68.0
Journal article	91.3	87.9	96.9	98.0	95.1
Other periodical article	27.5	24.2	37.7	32.0	30.1
Monograph	11.3	9.2	30.0	84.0	70.9
Other book	20.6	14.5	34.1	64.0	57.3
Chapter in book	50.0	38.6	67.7	84.0	71.8
Peer reviewed conf. proceedings	62.5	76.8	65.9	64.0	62.1
Non-peer reviewed conf. proceedings	35.6	33.8	36.3	30.0	23.3
Third party reports – government/NGO	23.1	20.3	40.4	10.0	6.8
Third party reports – companies	9.4	21.3	15.2	4.0	1.9
Technical reports	7.5	30.9	9.4	4.0	0.0
Patent applications	9.4	9.7	0.4	0.0	1.0
Datasets	11.3	8.7	9.9	10.0	6.8
Software	5.0	25.6	4.9	4.0	3.9
Artefacts	0.6	1.9	0.9	0.0	1.9
Exhibitions	2.5	4.3	3.1	0.0	5.8
Performances	0.6	1.0	1.8	4.0	7.8

Note. Figures provided are percentages

There are of course some differences in the type, structure, and rhetoric of RAs across disciplines. My previous research has shown that microbiology RAs are on average around a third shorter than history or management RAs (Rees, 2013, 2016). This has implications for the validity of any comparison of vocabulary use across disciplines. The following chapter outlines how this is controlled for in the present study. In the hard sciences RAs are prominently empirical papers and follow the Introduction, Method, Results, Discussion structure. However, Tarone, Dwyer, Gillette, and Icke (1998) describe how astrophysics papers often take the form of logical arguments. Similarly, McCloskey (1994) shows how theoretical physics papers take the form of outlines for computer models. Kuteeva and McGrath (2013) demonstrate how the structure of theoretical papers in pure maths differs from that of empirical RAs in general. In terms of rhetoric, Hyland (2016a) claims that abstracts in hard disciplines emphasise novelty, while in the social sciences they stress importance of the work. He also highlights an enthusiasm for graphical abstracts in the hard sciences which is absent from the softer disciplines. In the softer disciplines RAs follow a general structure of a contextualising introduction, followed by a supporting body section, and finally conclusion which reviews the argument or offers a position. However, there some differences within and between soft disciplines. For example, in history research, Coffin (2006) makes a distinction between recording genres, explaining genres,

and arguing genres. None of these differences invalidate the argument for employing RA in a study of EAP lexis.

In phraseological studies, the type of corpus employed has consequences for the type of phraseology encountered. General academic vocabulary such as connectors of the type *however*, *therefore*, *similarly*, *on the one hand… on the other…*, are arguably among the easiest to master as such they are apt to be used, and often overused, in corpora of student writing (Granger & Tyson, 1996; Ha, 2016; Paquot, 2014). In contrast, the accurate use of discipline-specific collocations presents more of a challenge and as such their use may be limited to expert performances. This is another argument in favour of using RAs in the present corpus-based study of EAP phraseology.

Having justified the use of RA corpora in EAP vocabulary studies, the problem of how to select RAs for inclusion in a corpus arises. A review of previous corpus-based studies of academic language brings to light three approaches to the sampling of texts for inclusion. Firstly, there is a convenience sampling approach where the RAs used are those available to the researcher. Studies often employ open-access articles without access restrictions. This is the case in G. Williams's (1998) study of collocational networks in a corpus of plant biology research articles. In studies of single disciplines or contrastive studies with general language this approach is appropriate and effective, however the need in the present study for comparable corpora of RAs means that it is not feasible. Another approach involves expert judgements, where discipline specialists advise the researcher on which RAs to include. While compiling their pilot science wordlist, Coxhead and Hirsh (2007) asked lectures to provide material they expected their students to consult. While compiling his wordlist for foundation engineering undergraduates, J. Ward (2009) consulted lecturers at chemical, civil, electrical, industrial, and mechanical engineering faculties on the textbooks used on their undergraduate studies courses. This seems an appropriate approach, however in order to mitigate for the personal biases of the experts, a panel of experts is needed. In the case of the present study with its focus on history, management, and microbiology this approach is considered too logistically demanding. A final approach involves the use of journal citation indicators.

The most well-known measure of journal importance are perhaps the Journal Citation Reports (JCR) (Thomson Reuters, 2017). The JCR is based in part on the impact factor (IF) measure: a ratio of the number of current-year citations to the number of articles published in that journal during the previous two years. The IF measure has been criticised for a number of reasons. Firstly, it favours articles from hard sciences which have a longer half-life than those from softer disciplines. This means they are frequently cited for a longer period after publication (Cameron, 2005). The five-year impact-factor was created in an attempt to rectify this. Hyland (2016a) notes an unintended by-product: A cyclical effect whereby journals with high IFs attract better, or at least highly citable, RAs which in turn has a further positive effect on the journal's IF. Other problematic aspects include a small quantity of articles accounting for a disproportionate number of citations (Editorial, 2005; Seglen, 1997), bias towards disciplines which produce a large number of articles and have short peer-review times (Cameron, 2005), which have longer articles (Seglen, 1992), and which include a high proportion of review articles (Seglen, 1997). Elsewhere research has suggested that IF favours open-access journals (Murali et al., 2004), raised concerns over accuracy (Dong, Loh, & Mondry, 2005), and reported manipulation of IF through self-citation or apparent reciprocal citation agreements between journals (Van Noorden, 2012). Various alternative measures have been created in an attempt to address some of these shortcomings. The Scimago Journal Rank (SCImago, 2007) (SJR) takes the importance of

the journal in which citations appear into account. The h-index (Hirsch, 2005), originally envisaged as a measure of the productivity of individual scholars, has been applied as a measure of a journal importance. It takes into account the whole set of papers in a journal to correct for individual disproportionately highly cited papers (Harzing & Wal, 2009). However, these measures are not perfect. Many of the issues they present are related to differences between disciplines. In an increasingly interdisciplinary environment the question of which journal a discipline belongs in is an important one. Differences in citation practices across disciplines are also problematic; some disciplines cite more than others (Hyland, 1999), cell biology RAs tend to include a much higher numbers of citations than physics RAs (Editorial, 2005). There are also related geographical considerations, journals in the major North American indices tend to prioritise hard sciences (Hyland, 2016a), while the global scale of many indices cancels out important local differences. This is especially problematic in the humanities and social sciences. Despite these drawbacks, this approach has been used successfully in many RA-corpus-based studies of academic language. For example, Marco's (2000) study of medical RAs, Martinez and et al.'s (2009) study of agriculture RAs, Liu and Han's (2015) wordlist study of environmental studies RAs which also used expert judgements, and my previous research (Rees, 2013, 2016). Despite their shortcomings, given the previous successful use of citation indices to select RAs for inclusion in a corpus, and the problems the other approaches discussed represent in the context of the present study, it is assumed that citation indices represent a valid means by which to select articles for inclusion in a corpus.

In summary, this study of phraseology in academic language assumes that RAs represent the kind of language EAP users will encounter and strive to produce. From a receptive standpoint, although books are considered the most important source of knowledge, and academic level is an influencing factor, there can be little doubt that RAs comprise an important source of information for EAP users. From a productive standpoint, this assumption is also valid since in terms of EAPWI in a genre analysis paradigm RAs can be regarded as expert performances representing textual exemplars for EAP users. Characterisations of RAs as an unassailable goal for EAP users are condescending. Similarly, while it is true that PAW and student coursework have different goals these differences do not impede the effectiveness of RAs as a source of useful lexicographic data about collocation patterns in different academic disciplines.

4.2 Assumptions Regarding the Decision to Concentrate on Verbal Collocations

This chapter section deals with the assumptions underlying the decision to concentrate on collocations centred around verbs rather than those based on other parts of speech. At first glance, this many appear counterintuitive since the orthodoxy in lexicographical phraseological studies is to concentrate on nominal phrases.

The primacy of the noun is reflected in various lexicographical resources dealing with collocations. Walker (2009) finds a general assumption that users will first look up a noun form when using collocational dictionaries. Dictionary makers frequently employ psycholinguistic arguments to support this axiom. The following extract is taken from the font-matter of the *Oxford Collocations Dictionary for Students of English* (Runcie, 2002):

> When framing their ideas, people generally start from a noun. You might think of *rain* and want to know which adjective best describes rain when a lot falls in a short time. You would be unlikely to start with the adjective *heavy* and wonder what you could describe with it (*rain, breathing, damage, gunfire*?). Similarly, you might be looking for the verb to use when you do what you need to do in response to a *challenge*. But you would not choose *meet* and then choose what to meet (*a challenge, an acquaintance, your death, the expense*). (Runcie, 2002, p. IX.).

Such a view suggests a conservative conception of collocation, not as simple co-occurrences of words in text, but rather involving a dominant base and a collocate. For lexicographical purposes this is the view shared by Hausmann (1985) who suggests that in a dictionary collocates should be listed by their nominal base.

Intuitively, the contention that use first recourse of the dictionary user is to the nominal constituent when looking up a collocation seems plausible. However, in practice there is scant evidence to support this. In his experiment investigating learner look-up strategies in collocational dictionaries Bogaards (1999) found that users demonstrated a slight, but not statistically significant, preference for the noun. More generally, he found no clear preference for base or collocate and on many occasions the users who were Dutch learners of French simply chose a word they were familiar with.

Another argument frequently put forward to support the primacy of nominal collocations is that these are more frequent than collocations based on other parts of speech. "Nouns are more numerous than verbs, and noun tokens outnumber verb-tokens in a text. Consequently for any new word that one encounters the chances are that it will be a noun not a verb" (Taylor, 2002, p. 182). Thus far in this thesis the argument that frequency, whilst important, should come second place to saliency has become somewhat of a recurring motif. This argument is equally applicable here. More importantly, the argument that since noun tokens are more frequent than verb tokens lexicographers should concentrate their efforts on nouns, is fundamentally incompatible with a phraseological view of language. On phraseological approaches words in isolation do not have meaning, only meaning potential. Consider, for example, the Idiom Principle (Sinclair, 1991), Lexical Priming (Hoey, 2005), and TNE (Hanks, 2013) among others.

This chapter section has outlined some of the weaknesses in two commonly posited arguments in favour of concentrating on nominal collocations in lexicographic resources. The intention here is not to negate the importance of nominal collocations but rather to highlight the legitimacy of concentrating on verbal collocations. The remainder of this chapter section will present arguments in favour of this approach. Learner-centred or user-centred arguments have long been made to support the practice of listing collocations by nominal bases in collocational dictionaries:

> The most important class as far as bases are concerned is the noun because it is nouns which express the things and phenomena of this world about which we can talk. Adjectives and verbs can only be bases as far as they can be determined by adverbs. (Hausmann. 1985, p. 119, cited in Bogaards, 1999, p. 127)

Whilst plausible there is little empirical evidence to support this practice. Equally plausible learner-centred arguments can be made for studying collocations based on other parts of speech. These often vary according to language background. In his study of the productive

use of EFL dictionaries G. F. Huang (1985) describes how "as a rule, Chinese students have less trouble with the English noun than with the English verb" (p. 55). There is a great deal of anecdotal evidence that use of verbs in English is difficult for users of many proficiency levels and language backgrounds. A glance at Swan and Smith's (2001) survey of common difficulties faced by learners of English confirms this. Of the 22 languages or groups of languages discussed 18 list problems related to the use of verbs be they with complementation, post-modification, the use of non-finite verbs, or the use of phrasal verbs.

A final argument in favour of the assumption that verbal collocations represent a legitimate focus of research in the present study is concerns gaps in previous research. It is closely related to the evolution of theories of terminology. A great deal of studies dealing with specialist-lexis are rooted in the field of terminology. Until the relatively recent emergence of the Communicative Theory of Terminology (Cabré, 1999), terminological studies were restricted to nouns since under the General Theory of Terminology (Wüster, 1979) the noun is the only category to designate a concept. The practical consequence of this is that, for the most part, adjectives and other part-of-speech categories are not permitted as entries in terminological dictionaries. Likewise, the vast majority of terminological studies have tended to focus on nouns. As a result research on verbal collocations still lags behind. Although not a terminological study in any strict sense, the present study is a modest attempt to rectify this.

4.3 Assumptions about Ideal EAP Lexicographical Resources and their Market

The third assumption is that, in the guidelines for the treatment of phraseological EAP vocabulary which will result from this study, it is valid to concentrate on the creation of a monolingual lexicographical resource rather than a bilingual one. Underlying this assumption is the belief that this lexicographical resource will best serve higher proficiency level EAP users. It has been shown in the previous chapter how questionnaire-based studies show a tendency for user-preference for MLDs which increases as a function of proficiency level. Similarly test-based and observation-based research suggests that MLDs are more beneficial to higher-proficiency-level EAP users than those EAP users who are just starting out. Figure 4.1 shows the approximate relationship between EAP context, proficiency level, and dictionary-type preference as it is understood in the present study. The thick black line delimits the intended target market for the illustrative example of an EAP lexical resource envisaged as an eventual product of this dissertation.

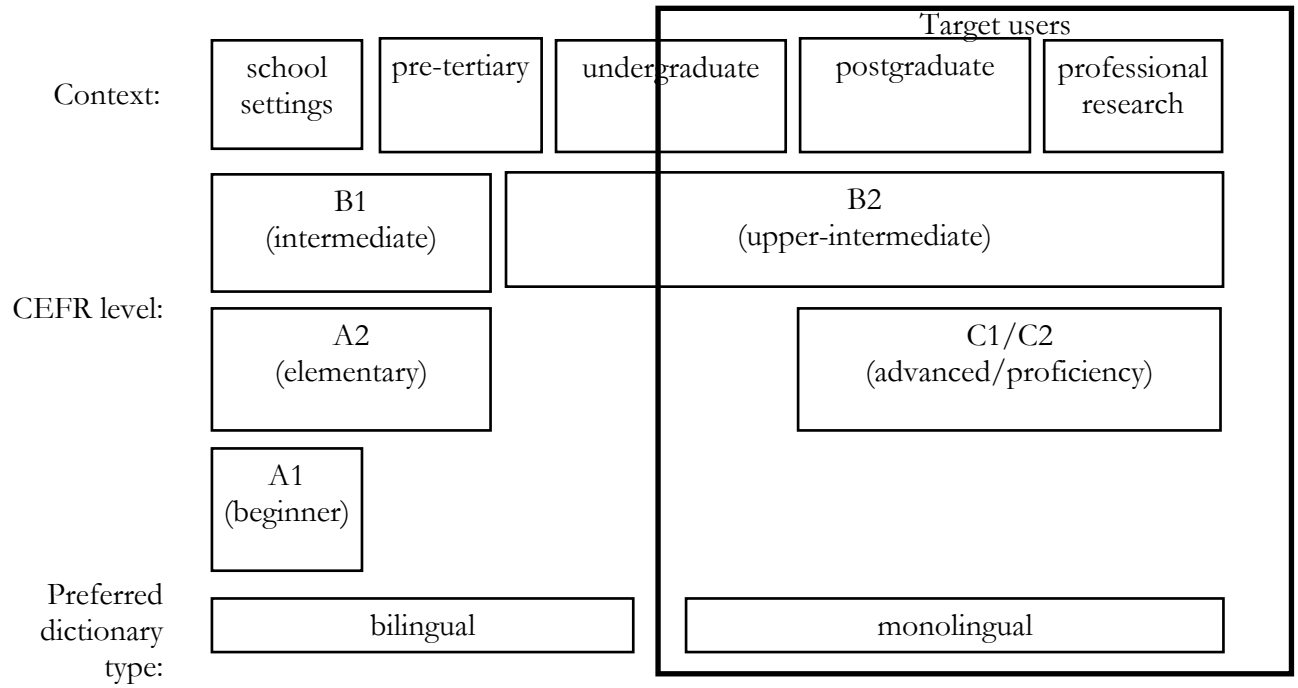

Figure. 4.1. EAP context, proficiency level, and target users and type of dictionary envisaged as product of the present study.

This projected target market is the result of careful consideration of the findings of previous research into dictionary use and users. However, dictionaries are often employed by users who bear scant resemblance to the audience envisaged by the designers. Moreover, users often employ dictionaries and other lexicographical resources in ways which their designers could not have contemplated.

The decision to concentrate on monolingual lexicographical resources is also supported by an economic and quality control argument. In his survey of attitudes towards dictionaries Lew (2004) suggests that the higher ratings given to MLDs in comparison to their bilingual Polish-English counterparts might be the result of a perception of higher quality. He further suggests that this perception is valid, reflects reality, and is due in part to the fact that international dictionary makers are better resourced than Polish dictionary publishers. Taking this argument one step further in a discipline-specific EAP context, it is possible to argue that monolingual resources are preferable to their bilingual counterparts as they allow a concentration of resources. Bilingual dictionary makers have to create a different dictionary for every L1 user market, while one MLD might cover the same target market as several bilingual dictionaries. In the case of commercial dictionaries staffing costs are an important consideration. In reference to monolingual lexicography Landau (2001) states: "Really top-flight dictionary editors are as rare as good pets" (p. 353). It follows that bilingual lexicographers are rarer still.

In summary, the decision to concentrate on monolingual lexicographical resources is motived by two factors. Firstly, based on dictionary-use studies it is apparent that such resources will better meet the needs of a greater proportion of the EAP-user population. Secondly, the creation of guidelines for monolingual lexicographical resources is practicable with the time and resources available for the present study, while the creation of such guidelines for bilingual resources is not.

4.4 Summary

This chapter has laid out the hypotheses which will be tested in the corpus-based experimental procedure detailed in the following chapter. It has also specified the assumptions underlying both this experiment and the wider objectives of the present study. The assumption that corpora of RAs are a legitimate reference for studies of EAP vocabulary has been shown to be shared by other researchers in the field. It has also been argued that EAP users need to understand RAs in their academic activity, and that RAs are an exemplar for the type of texts and by extension lexis that EAP users need to produce. Furthermore, a number of means of selecting RAs for inclusion in corpora have been examined and it has been argued that, although imperfect, in the context of the present study the use of citation indices represent an appropriate solution to this problem. The decision to concentrate on verbal collocations as opposed to, more frequent, nominal collocations has been shown to be valid in terms of user needs. It has also been argued that this decision responds to a gap in the literature. Finally, based on the research examined in the previous chapter, it has been argued that the needs of a greater proportion of EAP users can be feasibly and more adequately met by the creation of monolingual lexicographical resources than bilingual lexicographical resources. In a commercial lexicography context and in the context of the present study it has been argued that a monolingual approach is more practicable given the resources available.

The way in which citation indices will be employed to select journal articles remains to be specified. As does the precise way in which data about the collocations they contain will be analysed. The experimental procedure set out in the following chapter should address these issues. Chapter 6 provides a statistical comparison of the phraseology of verbs across the disciplines of history, management, and microbiology. Chapter 7 contains a qualitative analysis of the data and examples of how it might be represented in a lexicographical resource.

5. METHODOLOGY

The present study employs a corpus-based experiment to examine the phraseology, and by extension meaning of verbs, in the academic disciplines of history, microbiology, and management studies. This chapter describes the steps followed in this experiment. The experiment can be divided into three broad phrases. The first involves the creation of a corpus of academic journal articles. The second involves the processing and extraction of data from this corpus. The third is devoted to an experimental procedure intended to bring to light differences in verb collocation behaviour which might be relevant to the EAP user.

5.1 The AJACX2

This study tests the hypotheses set out in previous chapters with written language. This is not because these hypotheses only apply to this mode, but rather, a result of practical constraints such as the difficulty inherent in obtaining sufficient speech to compile a corpus, and the large quantity of time which would be needed to encode this data. The written mode, in contrast, is not subject to these constraints, electronic academic journals provide a convenient source of academic language, already classified by discipline, which can straightforwardly be made machine-readable. As argued in the previous chapter, they represent the type of language likely to be encountered by users of EAP. That said, as made clear in previous chapters, determining precisely what constitutes an academic discipline is a controversial issue. Similarly, delimiting the borders between disciplines is not a simple task. This is especially true given the greater emphasis which has been placed on interdisciplinary approaches in recent years. In addressing this issue, this study follows the strategy adopted by Cortes (2004) and Hyland (2008a, 2008b), both of which limit their comparisons of academic language to customarily unrelated fields. For example, Cortes (2004) examines differences between history and biology, while Hyland (2008a, 2008b) examines differences between electrical engineering, business studies, and applied linguistics. Following this example, the present study examines academic language across three sub-corpora: history, microbiology, and management studies. The examination of three sub-corpora is methodologically advantageous since it allows space triangulation (Burns, 2010). This permits more confidence in the results since, while it is possible that a difference in language between two sub-corpora could occur because of chance, it is less probable that this difference would occur by chance across all three disciplines.

This study makes use of an augmented version of the Academic Journal Article Corpus (AJAC). The original version was employed in previous research (Rees, 2013, 2016). The bespoke 8,088,429 token corpus is divided into three discipline-specific sub-corpora; history which contains 2,840,024 tokens, microbiology which contains 2,606,594 tokens, and management studies which contains 2,641,811 tokens. For brevity, the augmented corpus will be referred to as AJACX2[10]. A smaller-scale pilot study was carried out using the original AJAC corpus. The results of this pilot are referred to this chapter for the purposes of illustration.

[10] I am grateful to Dr Andrew Caines who suggested this abbreviation at the 2016 IVACS one-day symposium in Barcelona.

5.2 Materials

This section briefly outlines the materials employed in this experiment, setting out the motivation behind their selection.

5.2.1 Research Articles

The previous chapter argued that academic journal articles represent a suitable source of data for a study of the discipline-specific behaviour of verbal collocations in contexts relevant to EAP users. It briefly summarises some of the different forms in which knowledge is diffused in different educational contexts. Different disciplines, be they strictly academic or vocational, disseminate knowledge using different media. Despite this variation, RAs play an increasingly important role in knowledge dissemination in the great majority of disciplines (Sparks, 2005). This is the primary motivation for the use of RAs as the source of academic language in this study. In addition, the decision is partially motivated by practical concerns; while large quantities of research articles are easily available in fairly standardised electronic format, obtaining and digitising the necessary quantity of other text-types, for example textbooks, for a thoroughgoing analysis would be prohibitively expensive and time-consuming.

Great care must be taken in the selection of RAs for the corpus. A common criticism of corpus-based methods is the danger of researchers being influenced by confirmation bias while compiling corpora (Sinclair, 2005). There is a tension here between the need to build a corpus that is representative of the type and relative quantities of the language genuinely used by a group of users, and the possibility of data manipulation. Tognini-Bonelli (2001) posits a loose binary categorisation of corpora as either reference (Sinclair, 1996a), or specialised. With reference corpora a degree of representativity can be achieved by including huge quantities of texts from different, modes, registers, and a wide range of sources (Biber, 1994). However, in practice this is far from straightforward and ultimately true representativity is impossible (Sinclair, 2005). Instead, a judgement must be made about the degree of representativity acceptable. Achieving acceptable representativity is more problematic still in specialised corpora like the AJACX2 where "the aim of representativity by size is rendered impossible by the need to target disciplinary or thematic speciality" (G. Williams, 2002, p. 45).

Another important distinction in corpus type is temporal. In this regard, there are two principal approaches to data collection for the compilation of corpora. The monitor corpus (Sinclair, 1982) approach where a corpus expands, incorporating more data over time, and the balanced corpus approach where a corpus, which reflects language at a given point in time, is constructed according to a specific sampling frame. The present study adopts the latter approach, a sampling frame is adopted to ensure the corpus is reasonably balanced and representative of academic language across the three disciplines. For Sinclair (2005), as far as corpus construction is concerned, the terms balance and representativeness are frequently used yet vaguely defined. In the present study, representativeness is understood as pertaining to the degree to which the language contained within a corpus reflects that used by a given group of users. It has been argued thus far that a corpus comprising academic RAs is broadly representative of the language that EAP users read and write. According to Sinclair (2005) a balanced corpus contains proportions of different types of text which "correspond with informed and intuitive judgements" (Section 5, para. 1). The sampling criteria outlined below reflect such judgements.

The first criterion relates to the period under study. Academic research is not immune to the quirks of fashion; certain topics, methods, or approaches become fashionable only to disappear after a few years. These trends influence the language employed in the discipline, in this respect, they represent a deviation from the norm. In order that the range of language contained in the corpus is not adversely influenced by such deviations, this corpus samples journal articles over an eleven-year period: 2002 to 2012 inclusive.

It is not only temporal concerns which affect the range of language the corpus contains, the scope of the journal from which the articles are sampled is also important. With reference to Temmerman's notion of "umbrella categories" (2000, p. 96) such as biotechnology, G. Williams (2002) criticises the use of bibliographical categories in categorisation of texts for corpus construction. Such categories are motivated by administrative or organisational concerns rather than a true reflection the language use of a community of users. This is a valid criticism but the fact remains that in the construction of a specialist corpus, selection criteria are necessary. Judicious use of hierarchical categories does allow the creation of a corpus which is broadly representative of the language use of a fairly broad community of users. In the present study, this is achieved by controlling the scope of the journals included in the corpus. Journals representing very narrow sub-disciplines are excluded. This is reflected in the titles of the journals which meet the criteria (see Table 5.1 on p. 108), for example, *World History* is found yet *Water History* is not. The disciplines under study are history not the humanities in general, microbiology not biology, and management studies not the more general business studies. That is not to say that certain findings cannot be extrapolated form a narrower to a wider discipline. For example, certain findings from history might be extrapolated to the humanities in general. Similarly, certain findings from microbiology might be extrapolated to the sciences in general. Neither does recourse to hierarchical categories in text selection for corpus creation necessarily deny interdisciplinarity. It is precisely because no classification system is perfect that that there might be an article in the management sub-corpus dealing with the management of microbiological laboratories, or a history RA dealing with the evolution of a commercial organisation. Taking a broad view of the EAP community as comprising diverse users ranging from pre-sessional students to professional researchers, the scope criteria broadly reflect the language needs of EAP users.

In addition to these scope criteria, journal rankings and citation indexes provide another means of ensuring that the articles sampled are reasonably typical of those read and produced by EAP users in that discipline. Despite the geographical, disciplinary and cultural biases involved in the ranking systems for academic publishing set out in the previous chapter, it is assumed that those journals with a high average number of weighted citations contain language which is representative of their disciplines. The present study employs the SciMago (2007) Journal and Country Rank to categorise journals according to their discipline and popularity. The corpus contains only those articles from journals with a journal ranking indicator, a measure of a journal's impact, influence, or prestige, in the first quartile for their discipline for at least six of the eleven years sampled. Another factor which might affect the typicality of the language sampled is the dialect of English used in the journal, for this reason care is taken to include articles from journals published in the United States as well as journals published in the United Kingdom. However, it must be noted that academic publication is a global phenomenon and that as a result all the journals sampled accept contributions in both dialects of English. As outlined in the previous chapter, a possible objection to the use of journal articles in EAP vocabulary studies is that in many cases their authors are not native speakers of English. The validity of this objection is questionable. It might be argued that non-native authored RAs may not be the

most appropriate model for productive purposes because non-native writers have not mastered the required collocational repertoire (Yoiro, 1989). Not only is this argument condescending to non-native writers, there is some evidence that experience, not native-speaker status, is the influencing variable with mastery of academic collocations (Frankenberg-Garcia, 2017). From a receptive standpoint RAs, irrespective of the native language of the authors, represent an important source of information for students studying at UK universities (Weir et al., 2009). Moreover, editorial boards go some way to ensuring the language in journal articles is native-like (Flowerdew, 2001). Though as Rozycki and Johnson (2013) demonstrate, papers containing non-canonical grammar do occasionally get published and win awards.

Although the use of relative frequencies in inter-discipline comparison, mitigates the need for an equal quantity of RAs in each sub-corpora. The relative size of each sub-corpora could influence the sample of verbs selected for analysis following the procedure outlined in Section 5.3.2 below. In addition to ensuring representativeness, in the compilation of the original AJAC corpus it was assumed that sampling equally over an eleven-year span would ensure that sub-corpora represented approximately equal proportions of the overall corpus. A further assumption was that if each sub-corpus contained, firstly, an equal number of journal articles from each year; secondly, an equal number of articles from three journals from the same discipline; and thirdly that the articles from the sampled journal were selected at random, then the sub-corpora would be more representative than would have been the case had only one journal been sampled. This sampling frame also helps mitigate the risk of confirmation bias.

In spite of these attempts to ensure balance and representativeness, it became apparent in previous research (Rees, 2016) and during the pilot study phase of the present study that the original version of AJAC corpus still has some limitations as a source of data for the study of verbs and their collocates. Firstly, in common with many other corpus studies there was the problem of sparseness of data, some of the examples of discipline-specific behaviour of AWL words found in Rees (2016) were based on a small number of corpus-derived examples. The risk here is that such examples are anomalies and therefore not representative of the typical language of a discipline but instead the idiosyncrasies of a particular author or editor. Another corpus related limitation, specific to the AJAC, concerns the size of the microbiology sub-corpus. The assumption that sampling journal articles equally over an extended period would ensure that sub-corpora represented approximately equal proportions of the overall corpus proved to be false since microbiology articles are approximately 3000 tokens shorter on average than those from history and management studies.

These limitations are somewhat mitigated in the present study by increasing the corpus size. The AJACX2 is more than double the size of the original AJAC corpus (AJAC: 3,264,011 tokens; AJACX2: 8,088,429 tokens). However, it is worth noting that word frequencies are Zipfian in nature, around half the word types in a corpus occur only once, this means that doubling the corpus size will not necessary lead to a proportional increase in the occurrence of a given verb.

The problem of the undersized microbiology sub-corpora is solved by including a third more articles in this sub-corpus. In total the microbiology sub-corpus contains 352 articles from eight microbiology journals compared with 264 articles from six journals as the case in history and management Studies sub-corpora. The resulting AJACX2 contains 8,088,429 tokens in three discipline-specific sub-corpora of approximately equal size; history which

contains 2,840,024 tokens, microbiology which contains 2,606,594 tokens, and management studies which contains 2,641,811 tokens. Table 5.1 below provides a summary of the composition of the AJACX2. Bibliographic references for the articles it contains can be found in Appendix B.

5.2.2 Sketch Engine

After the text-processing procedures discussed in Section 5.3.1, the Sketch Engine (SE) (Kilgarriff, Rychlý, Smrž, & Tugwell, 2004) is employed to compile the journal articles into a corpus. SE is an online corpus query system. It is employed in the compilation of the AJAC and AJACX2 because it provides a convenient means of lemmatisation. It contains an implementation of the English Penn TreeTagger Part-of-Speech Tagger (Schmid ,1994), which has a reported tag accuracy of 96.36 per cent. The sketch grammar used is English PennTB-TreeTagger 2.5 developed by Marcus, Santorini, and Marcinkiewicz (1993).

In addition to corpus compilation, SE also facilitates the exploitation of corpora. Its wordlist tool provides a practical means of extracting the occurrences of the verbs under study in the AJAXC2. The system's Sketch Difference (WSDiff) tool allows the visual comparison of words in different sub-corpora according to their salient collocational context. An example is given in Figure 5.1; the greener a word is, the more closely it collocates with *assemble* in the History sub-corpora. The redder a word is, the more closely it relates to *assemble* in the Microbiology sub-corpora.

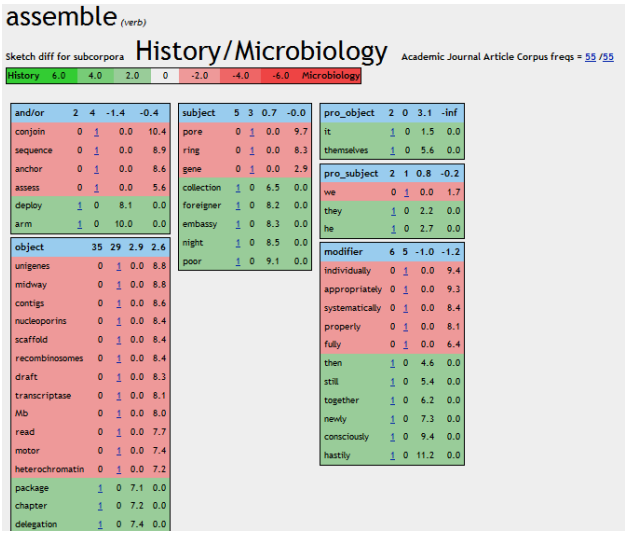

Figure 5.1. Partial screenshot of a WSDiff for *assemble* across History/Microbiology sub-corpora in the original AJAC

The WSDiff tool arranges collocates in sets according to their syntactic role. The user can see the keyword and collocate(s) in context by clicking on the figure to the right of the collocate. Previous research (Rees, 2013, 2016) has employed the WSDiff tool as a shortcut in the CPA process. It was used to draw inferences about the predominant semantic types of verb collocates in each sub-corpora of the original AJAC. The semantic types found in each sub-corpora were then compared. Effectively, this process permitted an approximate CPA without the need for manual annotation of corpus lines. The findings gave an initial indication of the difference in verb phraseology across disciplines. While, the WSDiff may still prove useful in the analysis process. The procedure adopted in the adopted in the present study offers a far greater degree of precision.

At the commencement of this project there was no information about the direct annotation of corpus lines using CPA semantic types in SE. However, Baisa, El Maarouf, Rychlý, and Rambousek (2015) report on a suite of tools which extend SE for the purposes of CPA annotation. Access to these tools is restricted as they are in the development stage. Therefore, another approach to corpus annotation is adopted (see Section 5.3.3). Ultimately, the presence of the WSDiff and wordlist tools along with considerations of ease-of-use, data security, and development time were the principal reasons for employing SE rather than a specially developed tool which may have offered greater levels of customisation.

5.2.3 The Pattern Dictionary of English Verbs

The Pattern Dictionary of English Verbs is intended to be the digital repository of information on the collocational patterns of verbs in general English. The patterns it contains are the result of CPA, a lexicographical technique for mapping meaning onto text. As demonstrated towards the end of Chapter 1, TNE provides the theoretical framework underlying CPA. In CPA, in order to document norms, alternations, and exploitations of syntagmatic verb patterns and by extension verb meaning, corpus lines are examined. In this examination, the SPOCA clause roles or 'slots' of systemic grammar represent the argument structure of the verb. Take, for example, the following adapted corpus line[11] from (Cinková & Hanks, 2010):

Subject	**Predicator (verb)**	**Object**	**Adverbial (also known as Adjunct)**
He	grasped	the handle	of the door

The lexical items occupying these slots must be labelled with the appropriate semantic type. In TNE semantic types represent the prototypical properties shared by the lexical items over which they generalise. They are arranged in a shallow ontology (CPASO). By comparing a number of corpus lines, the analyst can decide on the appropriate level in the CPASO necessary to obtain a pattern which generalises over many corpus lines. Lines (1a) to (1c) below were taken from the BNC50[12] sample used to create the PDEV entry for *appeal*. The semantic types assigned in the example perfectly reflect the semantics of each lexical item. However, when taken as a set it is necessary to 'climb the CPASO' to find the semantic type which adequately encapsulates the meaning of the lexical items as a set. In the case of the example sentences below these semantic types are shown in the pattern: [[Human]] *appeals* {to} [[Abstract Entity]]. Occasionally, contextual roles are employed to account for distinctions in verb meaning between certain contexts, for example: "[[Human = Film Director]] *shoot"* and "[[Human = Sports Player]] *shoot*" (El Maarouf, Baisa, Bradbury, & Hanks, 2014 p. 1002). Similarly, lexical sets are sometimes included to represent set phrases, for example: [[Human]] or [[Institution]] *denies* {responsibility or liability}.

[11] C - Complement is not present in this example. For Cinková and Hanks, (2010, p. 2) it is "a phrase that is co-referential either with the subject of the sentence, as in He is happy; he is the President or with the direct object, as in They elected him President; it made him happy"

[12] A 50-million word sample of the BNC

1.

a. [[Human]] appeals {to} [[Concept]]

Example: we cannot explain more complex needs by **appealing** to a concept of a universal individual

b. [[Human]] appeals {to} [[Proposition]]

Example: Wyclif was not breaking new ground but **appealing** to ideas already long current in certain

c. [[Human]] appeals {to} [[Information]]

Example: they **appealed** to empirical evidence to support their

The corpus-based nature of CPA semantic types is another factor motivating the use of PDEV in the present study. Other more widely used resources for word sense disambiguation have been criticised for their top-down approach. Hanks and Pustejovsky (2010) praise WordNet (Fellbaum, 1998) for its completeness, but criticise the inclusion of invented items, which are not naturally found in English, in order to complete the semantic hierarchy. As Fontanelle (2012) states: "Hierarchies offered by the tool are sometimes little more than a figment of the compiler's imagination and do not correspond to anything that is empirically observable" (p. 439). The use of glosses from the pre-corpus LDOCE1 only compounds this problem. FrameNet (Atkins et al*.,* 2003, Fillmore et al*.,* 2003, Ruppenhofer et al*.,* 2005) is similarly critiqued, since although annotated corpus examples of populated frames are provided, the frames themselves are the result of annotator intuitions. Annotator intuition is less of an issue in the PDEV as the CPA semantic types employed cover large lexical sets such as [[Human]] or [[Artifact]] and are derived inductively from corpus lines. This bottom-up approach to semantic labelling is beneficial to the study of domain specific language since, if corpus evidence exists, it permits the minting of new semantic types. In the pilot study this was the case with the semantic type [[Microbiological Entity]] and *specify* in the microbiology sub-corpus.

Annotation is not a straightforward process, a summary of some of the problems annotators must address is provided in Section 5.3.3 below. Further difficulties will be discussed in the analysis and conclusion chapters.

The result of this annotation process are patterns[13], each of which represents a typical use of particular sense of the verb. The PDEV is the online repository of these patterns. Its principal access structure is semasiologic: The user accesses a verb entry by searching for it using the search function or clicking on that verb in the index. It also offers onomasiologic functions, for example, by selecting a semantic type in the CPASO, the user can see a list of those noun collocates the semantic type generalises over, the argument slot these collocates occur in, and a list of verbs which form patterns types with collocates of this semantic type. See Figure 5.2 for an example with [[Information]].

[13] In the comparative application of CPA concieved in the present study, the term 'pattern type' is preferred to the original CPA term 'pattern'. The addition of 'type' may seem redundant as patterns are already abstract generalisations over corpus texts. However, the need to draw comparisons between relative quantities of distinct patterns necessitates another level of abstraction. On this level of abstraction, patterns in the original CPA sense represent tokens.

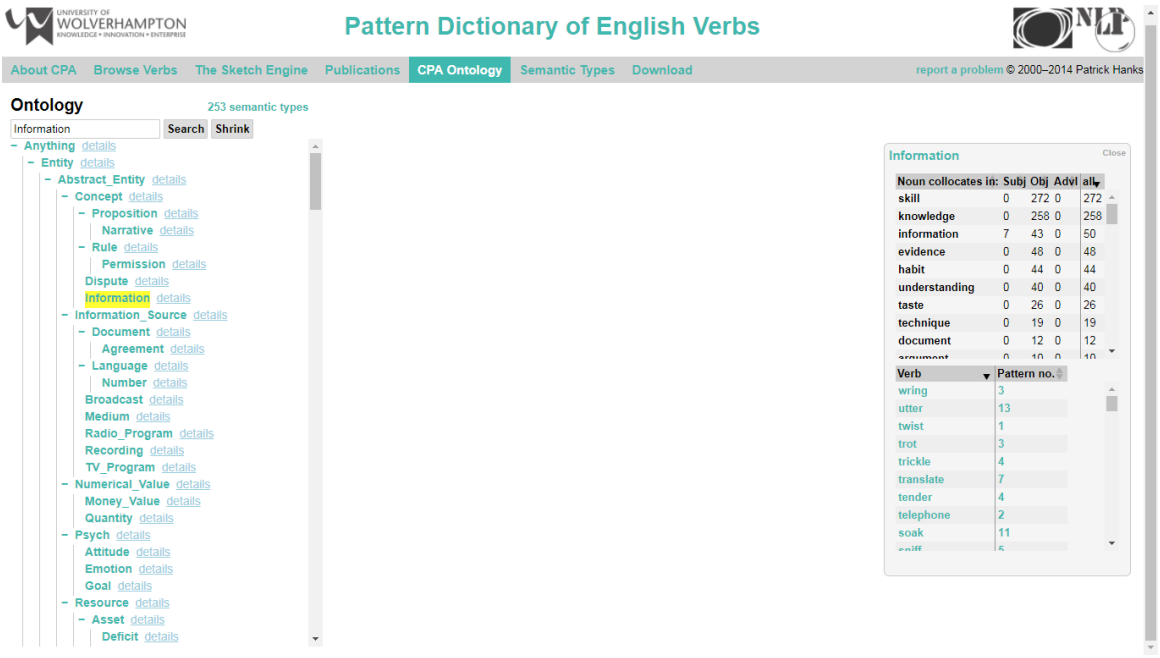

Figure 5.2. Screenshot of noun collocates for [[Information]] in PDEV

In addition to the paradigmatic pattern types for each sense of the verb, PDEV entries comprise primary implicatures which give the meaning of the pattern along with labels indicating usage information such as register (e.g. Formal) and domain (e.g. Law) (see Figure 5.3). Corpus-derived examples of text which maps to the syntagmatic pattern type are also provided. Alongside each pattern, a percentage is provided. This indicates the proportion of the sample a given pattern type represents. This is calculated by dividing the number of lines which correspond to a pattern type by the total number of lines in the sample.

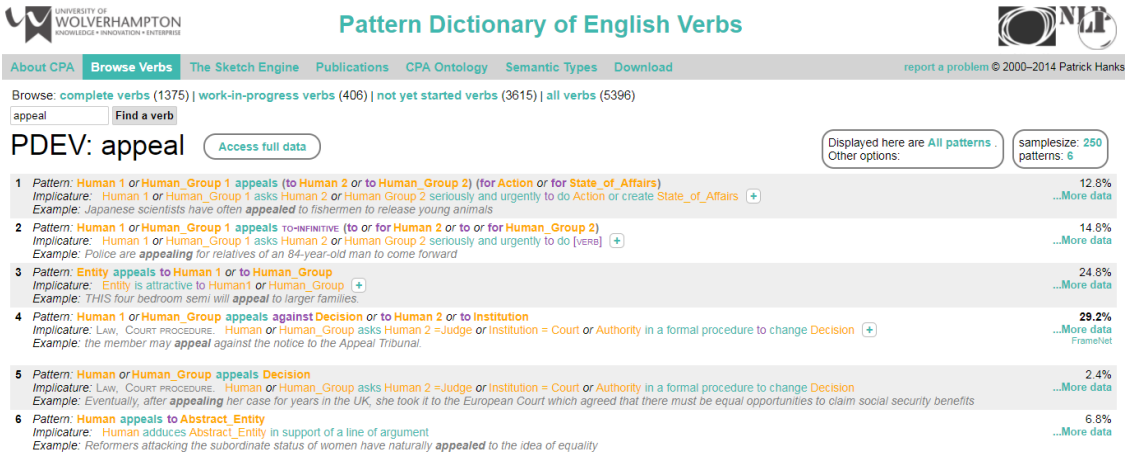

Figure 5.3. Screenshot of the PDEV entry for *appeal*.

A great advantage of digital lexicography is that it allows lexicographers to easily show the evidence on which entries are based. By clicking on 'More Data' users can see a sample of the corpus lines from the BNC50 which a given pattern type generalises over. Clicking 'Access Full Data' reveals all the corpus lines in the sample used to build the entry for that verb. In the literature dealing with PDEV, it is reported that samples usually contain around 250 lines (Bradbury & El Maarouf, 2013; Cinková & Hanks, 2010; El Maarouf et

al., 2014). In other CPA applications, the sample is reported as containing as many as 500 lines (Hanks, 2008). In practice, around half of the 1375 completed verb entries in the PDEV at the time of writing are based on less than 70 corpus lines. The screenshot in Figure 5.4 shows some of the corpus lines in used to create the PDEV entry for *approve*. The numbers in green boxes beside the key word indicate which pattern type generalises over the corpus line. Additional labels indicating variation in use are also found '.a' for anomalous collocate, '.e' for exploitation, '.f' for figurative use, and '.s' for syntactically anomalous.

element in his package of proposals to be	approved	1	without agreement on all the others. The
last Friday the European Foreign Ministers	approved	1.a	a package of sanctions. Work is now proceeding
whether the proposed company statute can be	approved	1.a	by a majority vote (as the commission would
King, he said, could only do so in terms	approved	1.a	by his ministers. The King then said: `
announced on Nov. 13. The new Cabinet was	approved	1.a	by the Dáil on the same day. Vincent
sought. All German military exports must be	approved	1.a	by the Federal Security Council.
to which he referred and which have been	approved	1.a	so far, only 23 were named in reorganisation
, Fumaroli insists, is a pose. The State	approves	1.a	the artistic tastes of a fashionable avant-garde
Italian parliament after protracted debate	approved	1.s	on Jan. 16 (by 382 to 201 in the Chamber
Italian and Belgian aircraft [see below], NATO	approved	1.s	on Jan. 17 the strengthening of its naval
economic restructuring programmes led it to	approve	1.s	on July 14, 1989, a financial package of
conferred by means of a warrant which must be	approved	2	(and is normally signed) by him personally
1989 (but that only 821 applications were	approved	2	). This compared with 16,726 in 1988, itself
note, invoice received, invoice checked and	approved	2	, and statement reconciled and account paid
to the Committee. 2. MMRC minutes, once	approved	2	, are circulated to all members of MCC.

Figure 5.4. Partial screenshot of the BNC50 corpus lines for *approve* in PDEV.

In the pilot-study, it became apparent that the following corpus line was marked '.e' for exploitation in the PDEV indicating a semantic-type exploitation, however, in the microbiology sub-corpus this exploitation predominated and therefore represented the norm in this discipline. Indicating a microbiology-specific use of *specify*.

a stretch of DNA that **specifies** [1.e] a single type of protein is a gene.

In the present dissertation, the PDEV serves two purposes. It provides a useful template for the semantic annotation of verbs and their collocates. It also provides a means of ensuring a reasonable degree of validity and accuracy in annotation. In January 2016, the PDEV contained 1286 verbs with their CPA patterns. Each entry is derived by a single trained analyst. Cinková, Holub, and Kríž (2012) have demonstrated a high degree of inter-annotator-agreement in CPA annotation, and furthermore, that generally, where discrepancies do exist they are due to a lack of annotation guidelines or confusion as to where to delimit the border between norms and exploitations rather than more fundamental flaws in the CPA system or the theory underlying it. Access to the sample of corpus lines provides an indication of how other annotators have classified lexical items in certain contexts. Similarly, the onomasiological function allows the annotator to see the nouns that have been generalised over by a given semantic type. It must be noted that annotation must take place bottom-up from corpus line to semantic type, not top-down from semantic type to corpus line. Nonetheless, a summary of the nouns commonly generalised over by a given sematic type is beneficial in cases of doubt. The PDEV also has a secondary role allowing the comparison of patterns in the discipline-specific corpora in the AJACX2 with those of the general English BNC50. Although this is not the principal aim of this study, having this reference resource may be advantageous in analysis and discussion.

5.2.4 ATLAS.ti

The use of computer software to analyse data has traditionally been reserved for quantitative statistical analysis. This changed with the widespread emergence Computer Assisted Qualitative Data Analysis (CAQDAS) software and techniques in the early 1990s. Originally developed to aid those working in Grounded Theory (Glasser & Strauss, 1967). CAQDAS software typically combines tools for database and text management, coding and annotation, data retrieval, and representing relations within and between data sets. ATLAS.ti ("ATLAS.ti," 2008) is among the most widely used CAQDAS packages available (Paulus, Lester, & Dempster, 2014; Silver & Lewins, 2014)

In linguistics in general researchers have eschewed CAQDAS software such as ATLAS.ti aimed at the social sciences in general in favour of software designed for a specific task. For example, coding gestural behaviour and sign language is often done with ELAN (Lausberg & Sloetjes, 2009), and coding prosodic and phonetic transcription with Praat (Boersma & Weenink, 2017). Working in discourse analysis, Paulus and Lester (2015) demonstrate how ATLAS.ti can be used to realise these tasks. A similar situation exists in corpus linguistics where studies tend to employ corpus query tools such as WordSmith Tools (M. Scott, 2017), AntConc (Anthony, 2017), Wmatrix (Rayson, 2008), or SE. This is because although ATLAS.ti works well as a concordancer and can be used to group clouds of closely associated words, it lacks many features which are particularly useful in corpus linguistics research. For example, automatic part-of-speech tagging, lemmatisation, and extraction of statistically salient collocations or keywords. However, the utility of ATLAS.to corpus linguistics and related research has been recognised. In his corpus-driven research on argumentation in reading groups O'Halloran (2011) combines automatic part-of-speech tagging conducted with Wmatrix with manual annotation of discourse functions conducted using ATLAS.ti. In lexicography, G. Williams (2017) highlights the advantages of ATLAS.ti in mapping Henri Basnage de Beauval's expanded edition of the *Dictionnaire Universel* (Furetière, 1701). The strategy of combining specialist corpus tools with ATLAS.ti is followed here. Pilot analyses carried out in preparation for the present study revealed ATLAS.ti and SE to be practical and robust options, while several corpus annotation tools proved unstable and laborious when used for annotation. A review of the literature revealed no other ESP lexicography studies employing ATLAS.ti. In this respect the present study is breaking new ground. The relevance of the advantages of specific features of ATLAS.ti to the present study are exemplified in section 5.3.3 below.

5.3 Procedures

Previous sections have outlined the rationale underlying this corpus-based study and introduced the principal resources it employs. This section sets out, in detail, the experimental procedure. First the procedure for the compilation of the AJACX2 is detailed. Next the procedure employed for the selection of the sample of verbs under study is presented along with the underlying rationale. Then the procedure analysing these verbs is stipulated and, finally there is a discussion of the statistical measures employed in the study.

5.3.1 Compilation of Corpus

The journals containing the articles sampled in the corpus are selected according to the criteria set out in section 5.1 above. In total the corpus comprises 880 articles taken from twenty journals. In order to ensure adequate range, one article from every n^{th} issue is

selected. Each article in the issue is assigned a number then a random number generator (a simple Python script) is used to select the article to be sampled. The precise frequency with which articles are sampled (summarised in Table 5.1 below) depends on the publication schedule of each journal.

Having selected the articles, the next step involves the extraction and formatting of text. At the time both AJACs were compiled, none of the academic journals selected were open-access. They are only accessible via web browsers connecting through academic networks. Employing corpora of RAs from open-access journals is one possible means of avoiding these access issues. The number of open-access journals is increasing rapidly (Laakso et al., 2011). From a productive standpoint this bodes well for their inclusion in the AJACX2 since it might suggest that more EAP users are producing open-access RAs. From a receptive standpoint, the picture is not so clear, while there is some evidence to suggest they are widely downloaded and widely cited (Davis & Walters, 2011; Eysenbach, 2006; Lawrence, 2001; Metcalfe, 2006), other studies suggest that this is due to an initial peak in citations (Davis & Walters, 2011) and the initial harvesting of open-access RAs for reading later (Outsell, 2009 cited in Hyland, 2016a). Overall, over the longer-term, evidence suggests that restricted-access journals are cited with greater frequency (Björk & Solomon, 2012). This would seem to be borne out by the SciMago index (SCImago, 2007) where the great majority of the top-ranking journals for the disciplines under study are not open access. It must be noted that this trend in not so pronounced in microbiology.

As a result of the decision to use journals behind a paywall, mass data extraction tools cannot be easily employed. Similarly, differences in formatting between journals make batch-processing tools impractical. Instead, where HTML versions of articles exist, in the majority of the articles included, text is cut and pasted into a text editor, abstracts, headings, mathematical formulas in-line with the body of text, tables, and footnotes are removed. A series of regular expression macros is applied to remove extraneous whitespace and numeration. Each article is then saved as a plain text file and given a filename code indicating the name, issue and number of the source journal which is then cross referenced with author information. A similar process is applied for earlier (predominantly pre-2006) issues where only Portable Document Format (pdf) articles are available, however, in this case text was extracted from the pdf files using ABBYY Fine Reader 11 OCR (ABBYY, 2011) software. Finally, files are manually checked for errors such as missing white-space, and encoding errors. If authors have written more than one article included in the corpus, these articles are flagged at this point. The resulting files are arranged in subfolders according to discipline. These folders are then compressed in .zip format and uploaded to SE where the corpus is compiled and tagged using the TreeTagger Part-of-Speech Tagger (Schmid, 1994). The sketch grammar used is the English PennTB-TreeTagger 2.5. Figure 5.5 illustrates this procedure.

Table 5.1

Journals, sampling periodicity, and articles in the AJAXC2

Journal Sampled	Issues Published per Year	Sample Periodicity	# RAs	Length in tokens
Historical Journal	4	1 RA every issue	44	10480
Journal of World History	4	1 RA every issue	44	10692
Past and Present	4 (plus supplementals)	1 RA every issue	44	10965
The Journal of Modern History	4	1 RA every issue	44	13430
Journal of Contemporary History	4	1 RA every issue	44	8886
Journal of Social History	4	1 RA every issue	44	10094
Total history			264	10758 (avg.)
FEMS Microbiology Reviews	6	1 RA from 4 issues a year	44	13136
Trends in Microbiology	12	1 RA every 4th issue	44	3471
Current Opinions in Microbiology	6	1 RA from 4 issues a year	44	1829
Molecular Microbiology	24	1 RA per 1st iss. of vol.	44	7924
Annual Review of Microbiology	1	4 RA per issue	44	7384
Cellular Microbiology	12	1 RA every 4th issue	44	7817
Journal of Applied Microbiology	12 (plus supplementals)	1 RA every 4th issue	44	4434
Clinical Microbiology Reviews	4	1 RA every issue	44	13247
Total microbiology			352	7405 (avg.)
Journal of Management	8	1 RA every 2nd issue	44	10149
Organization Science	6	1 RA from 4 issues a year	44	10722
Strategic Management Journal	13	1 RA from 4 issues a year	44	10454
Management Science	10	1 RA from 4 issues a year	44	9053
Organization Studies	6	1 RA from 4 issues a year	44	9874
The Journal of Management Studies	8	1 RA every 2nd issue	44	9790
Total management			264	10007 (avg.)
Overall total			880	9390 (avg.)

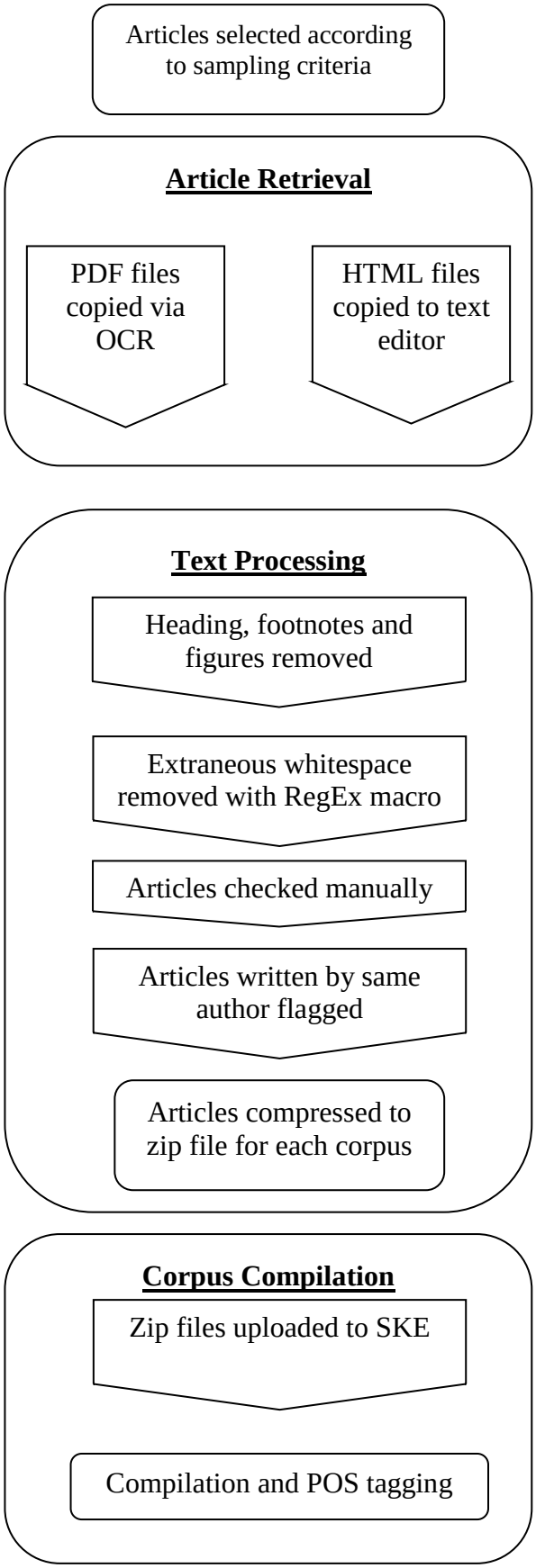

Figure 5.5. Flow diagram illustrating corpus compilation procedure

5.3.2 Selection of Verbs

The present study takes a purposive approach to sampling. Only the 1286 verbs completed in the PDEV in January 2016 are considered for analysis. In order to be included in the sample a verb must occur twenty times or more in each sub-corpus and in at least ten different journal articles written by distinct authors. These restrictions are not arbitrary. On the contrary, according to Sinclair (2005), the twenty occurrence threshold represents the point at which a trained lexicographer is able to make judgments about meaning. The ten-article by distinct author threshold is an attempt to ensure than the verbs are truly representative of the discipline rather than idiosyncrasies of a handful of authors. Similarly, the high degree of inter-annotator agreement regarding PDEV patterns reported (Cinková & Hanks, 2010) makes them a useful support the annotation of sentences in the study. While this sampling regime does not account for the entire range of language that EAP users are likely to encounter and produce, it does allow more robust and reliable analysis than would otherwise be possible.

In common with corpus compilation process, the selection of verbs to study creates a potential for confirmation bias. As Hanks (2013) explains, there is a tendency to regard "corpora as 'fish ponds' in which to angle for fish that will fit independently conceived hypotheses and theories. Fish that do not fit the theory are thrown back into the pond." (Hanks, 2013; p. 6). In order to counter this threat, Leech (1992) advocates an approach of 'total accountability' which avoids conscious selection of data. The verb selection criteria adopted in no way constitute the conscious selection of data. Returning to Hanks's fishpond metaphor there is no attempt to fish for data to fit hypothesises about discipline-specific language.

The first stage of the verb selection process involves the SE wordlist tool. A wordlist was extracted from each sub-corpus using a whitelist filter containing all 1286 PDEV verbs completed in January 2016. Those verbs which did not meet the twenty-occurrence threshold in each sub-corpus are rejected. This left 139 candidate verbs. The wordlist tool was employed once more, this time to count the number of documents in which each candidate verb occurred. Three were no verbs which occurred under the threshold.

The wordlist tool searches for the various inflected forms of the word listed in the whitelist. It does not filter forms by parts of speech. For this reason, the second stage of the sampling procedure involves searching each sub-corpus using SE's corpus query language for instances of the 138 lemmas tagged as verbal forms. The number of articles containing each verbal form is also noted. This left 108 verbs in history, 103 in management, and 104 in microbiology which met the twenty-occurrence threshold in ten different articles.

In the process of selecting journal articles for inclusion in the corpus no attempt was made to control for the number of articles included written by a given author or authors. The rationale behind this choice being that the number of journals selected per discipline, the frequency with which their issues are included in the corpus, and the random selection of articles within those issues would result in a corpus which reflects the language EAP users produce and encounter. However, there is a danger of reaching conclusions based on authors' idiosyncratic uses of certain verbs. To counter this only verbs occurring in ten or more articles written by different authors are included in the sample. The third stage of the verb selection procedure ensures that the verbs occur in at least ten articles written by different authors. A simple Python script counts articles which are written by the same

authors. It returns a count of such authors for each verb in each sub-corpus. Although there are a handful of articles written by the same authors in the AJACX2, the application of this author filter did not affect the number of verbs which met the ten articles by different author threshold. Figure 5.6 below illustrates the various stages of verb selection process and the number of verbs remaining in the sample at each stage.

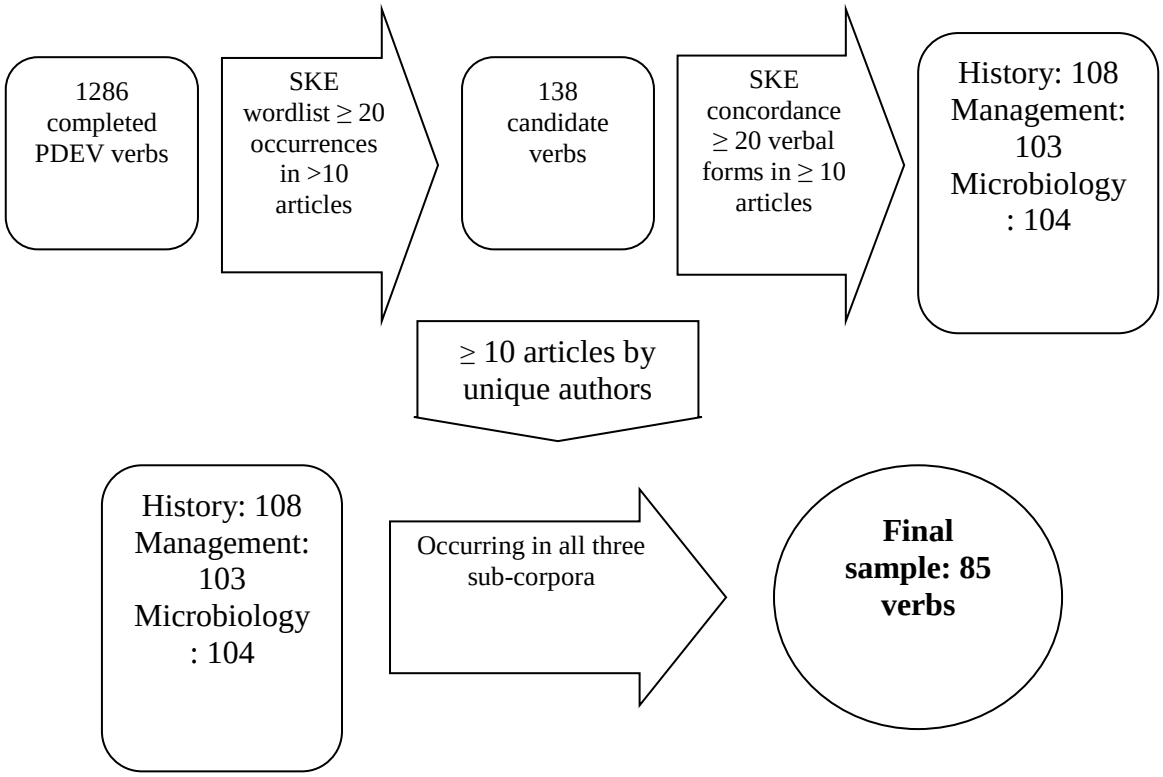

Figure. 5.6. Flow diagram of verb selection procedure

Table 5.2 shows the result of the sampling procedure. Overall there are 85 PDEV verbs which occurred over 20 times in every sub-corpus in at least 10 articles written by different authors. CPA is time intensive process. For example, at the time of writing the PDEV has been in development for over five years and 1388 of the proposed 5396 entries have been completed. It would not be feasible to annotate all 85 verbs. Instead, each verb is assigned a number and 30 verbs are selected at random for annotation.

Table 5.2

Verbs which meet the sampling criteria and their frequency of occurrence

Verb	Hist	Man	MB	Verb	Hist	Man	MB	Verb	Hist	Man	MB
absorb*	62	63	26	appear*	1367	803	1267	export	91	23	54
accelerate*	44	43	45	apply*	356	654	314	face	421	545	48
accept*	579	226	65	appreciate	99	51	27	fail*	482	311	250
accommodate*	95	70	37	approve	117	52	33	follow*	1195	1358	1360
accompany*	204	71	95	argue	1461	1358	95	handle*	74	84	21
accomplish*	73	138	67	arise	194	436	161	impede	24	36	23
account*	111	367	100	arrange	111	33	37	last	89	45	27
accumulate*	42	111	277	ascribe	68	42	24	lead*	1298	1564	1510
achieve	345	725	288	ask	618	582	37	maintain*	709	566	459
acquire	274	704	359	assemble	91	56	157	manage*	296	747	22
act*	421	426	518	avoid	327	379	167	mask	25	30	22
adapt	354	276	215	build	452	693	72	need*	809	1267	511
add	274	448	615	call*	1202	450	355	note*	773	849	413
adhere	53	27	89	choose	399	753	89	operate	221	779	109
adjust	25	166	63	claim	864	137	22	overcome*	138	193	112
administer	103	60	71	classify	73	155	156	point	456	258	109
admit	254	34	26	conduct*	167	572	229	prescribe*	26	41	22
adopt	385	578	57	construct	233	443	299	preserve	226	112	46
advance	142	147	20	consume	29	61	40	propose*	279	817	628
affect	246	1269	918	continue	1043	374	185	replicate*	30	144	277
afford	101	79	22	cross	46	47	58	see*	2514	2224	1545
agree	419	233	43	cultivate	35	24	54	shed	79	93	111
aid	85	38	78	decline	126	82	57	specify*	70	325	57
aim	227	181	72	die	439	20	137	stress	264	116	49
align	36	182	54	distinguish	238	308	175	submit	119	58	42
alter	110	169	289	distribute	125	321	116	treat*	399	234	527
analyse	70	187	539	engage	370	835	48	trigger	44	116	361
answer*	145	89	41	explain*	876	1104	431	yield	103	294	244
anticipate	79	144	27								

Note. * indicates verb randomly selected for analysis

The decision to limit the study to this sample is essentially a trade-off between representativeness on one hand, and practicality and reliability of analysis on the other. A larger sample which included verbs not in the PDEV, as well as verbs which occurred below frequency and distinct-authorship thresholds, might have better represented the language which EAP users encounter and strive to produce. However, the large amount of data such an approach would produce would make it impossible to conduct a thoroughgoing analysis. Concentrating on a smaller sample of verbs allows more in-depth analysis of collocation patterns in the disciplines under study.

Even before proceeding to the analysis stage the sample of verbs above indicates a clear deficiency of traditional wordlist approaches to EAP vocabulary. Several of the verbs

sampled are polysemous. For example, the MEDAL1 lists three senses for *admit* in general English: Firstly, 'to agree that something is true'; secondly, 'to take someone into hospital for medical treatment' and finally, 'to allow someone to enter a place'. Polysemy such as this, which could certainly be problematic for a L2 English learner, is concealed by traditional wordlist approaches.

Polysemy also raises a methodological problem. There are several sampled words which are heterosemous. For example, *act* could be a verb or a noun. Although the accuracy of POS tagging has improved a great deal in recent years, heterosemous forms continue to represent a significant challenge. It is likely that a substantial number of the heterosemous forms tagged as verbs are in fact nouns. Such cases are noted in the annotation process and omitted. If the number of tagging errors is so great that the verb falls below the 20 occurrences in 10 articles written by distinct authors then that verb is removed from the sample entirely.

The relationship between the sample and other representations of academic language is far from straightforward. Only 25 of the 85 verbs in the sample appear in the AWL. On one account, this may be because the majority of verbs in the sample are representative of general English and not academic in the least. However, scholars including Biber (2006) have highlighted the role sampled verbs such as *aim*, *explain*, *face*, and *propose,* none of which are included in the AWL*,* play as discourse markers in academic writing. Adopting a different perspective; 60 of the verbs sampled are not included in the AWL. Given that the sampling framework set out above ensures that the verbs sampled have a widespread and repeated use in academic journal articles, this casts doubt on Coxhead's (2000) claim that the AWL represents academic English.

5.3.3 Annotation of Corpus Lines

Annotation proceeds on a verb-by-sub-corpus basis. A text file containing the sample of corpus lines for a given verb in a given sub-corpus is loaded into ATLAS.ti where it is known as a 'primary document'. ATLAS.ti allows strings of text to be highlighted and assigned a code. When these strings of text are linked with a code they are known as 'quotations'. In the present study the code assigned is normally a CPA semantic type such as [[Human]]. However, on occasion, when a specific lexical item predominates in an argument slot the 'in-vivo coding' function is used to assign a code of the same name. For example highlighting the lexical item *places* in the text "students apply for places at university", and using the 'in-vivo coding' function would create a code called *places*. In the PDEV [[Human]] *applies* (for {place}) is listed for the verb *apply*. Unlike some of the linguistic annotation packages trialled in the pilot study, ATLAS.ti does not require that codes be specified a priori. This is advantageous since CPA semantic types are not a closed class and thus it is possible that new CPA semantic types emerge from the corpus. Take for example, the emergence of [[Microbiological Entity]] from the *specify* microbiology sample in the pilot experiment.

The sample size customarily reported in CPA-based studies is 250 lines (Bradbury & El Maarouf, 2013; Cinková & Hanks, 2010; El Maarouf et al., 2014), however, annotation of samples of as large as 500 lines has been reported (Hanks, 2008). In practice, such as in the creation of the PDEV, annotators make the best use they can of the data available. In the present study, an upper limit of 150 lines is set. In those cases where A and B samples are taken, they are combined in the interdisciplinary comparison to create a 300 line sample. In

cases where there are not 150 lines available all available lines are annotated. Table 5.3 gives sample sizes for the random selection of 30 verbs annotated.

Table 5.3

Size of the sample annotated for the 30 randomly selected words

Verb	Hist. Sample Size	Man. Sample Size	Micro. Sample Size	Verb	Hist. Sample Size	Man. Sample Size	Micro. Sample Size
absorb	62	63	26	fail	150	150	150
accelerate	44	43	45	follow	300	300	300
accept	300	157	65	handle	74	81	21
accommodate	95	70	37	lead	300	300	300
accompany	150	71	95	maintain	300	300	300
accomplish	73	138	67	manage	150	300	22
account	111	300	100	need	300	300	300
accumulate	42	111	145	note	300	300	300
act	300	300	300	overcome	138	150	112
answer	145	89	41	prescribe	26	41	22
appear	300	300	300	propose	150	150	150
apply	300	300	300	replicate	30	144	150
call	300	300	300	see	300	300	300
conduct	150	300	150	specify	70	300	54
explain	300	300	300	treat	150	150	150

Patten types emerge in the annotation process. Once every line in a sample has been annotated, it is necessary to record how many patterns of each type have occurred. For intransitive patterns, for example, [[Human]] *manages,* this could be straightforwardly achieved by generating a report table summarising the quantity of each code (usually a semantic type) per primary document. However, the existence of patterns containing two or more semantic types complicates the process. The most time efficient way to proceed is to count patterns manually and then check this count using the report functions. Figure 5.7 shows part of the quotation report table for the verbs annotated in this experimental procedure. The quantity of each pattern type found in each sample is recorded in preparation for the statistical analysis stage. The corpus lines annotated in this process can be found on the CD accompanying this thesis where they are sorted by pattern type, verb, and discipline.

	?	ABSTRACT	ACTION	ACTIVITY	ADJ	ANIMAL	ANYTHING
P24: accumulateHIST.txt	0	26	0	0	0	0	0
P25: accumulateMAN.txt	0	91	0	1	0	0	0
P26: accumulateMB.txt	0	11	0	1	0	2	0
P27: actMG.txt	0	12	1	0	1	0	38
P28: actHIST.txt	0	5	0	1	1	0	47
P29: actMB.txt	0	0	0	0	0	2	10
P30: appearMB.txt	0	2	0	0	13	0	0
P31: appearMG.txt	0	12	0	0	27	0	0
P32: appearHIST.txt	0	11	0	0	21	0	0
P33: applyMB.txt	0	0	0	0	0	0	0
P34: applyMG.txt	0	3	0	0	0	0	0
P35: applyHIST.txt	0	2	0	2	0	0	0
P36: callHIST.txt	0	0	9	2	0	0	100
P37: callMG.txt	0	0	15	5	0	0	107
P38: callMB.txt	0	0	2	2	0	0	141
P40: callMB_B.txt	0	0	2	0	0	0	143
P41: callMG_B.txt	0	3	11	1	0	0	109
P42: callHIST_B.txt	0	0	12	2	0	0	86

Figure. 5.7. Partial screenshot of quotation report table

In addition to the coding and report function, ATLAS.ti offers a number of other features which are particularly useful to CPA. Chief amongst these is the ability to comment quotations, strings of text assigned a code, using memos. Memos often include reasoning about why a lexical item maps to a specific semantic type, intuitions about emerging patterns, and speculation about exploitations and the mechanisms underlying them. Figure 5.8 shows a memo being used to speculate about interdisciplinarity in the AJACX2. In many cases, further analysis shows that tentative hypotheses and reflections recorded in memos are unfounded. Even so, they serve as useful reflection points.

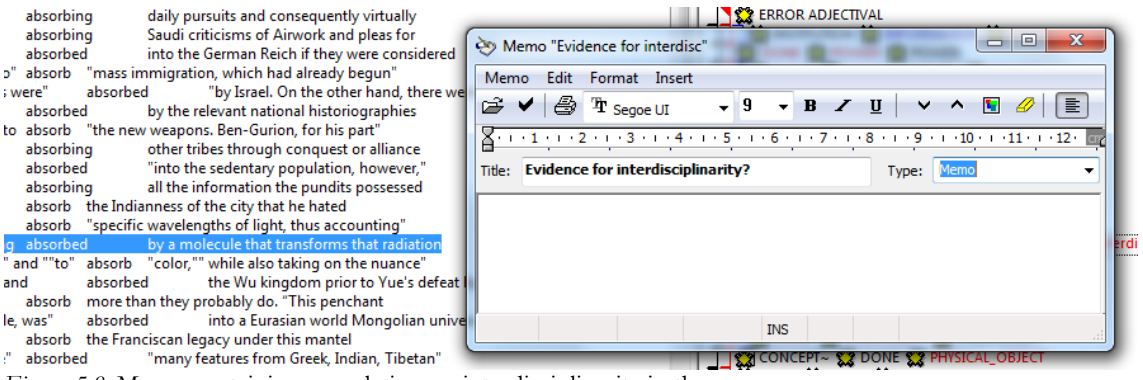

Figure. 5.8. Memo containing speculation on interdisciplinarity in the corpus

The ability to assign codes (usually semantic types) to corpus lines is fundamental to CPA. The ability to work in the other direction is also advantageous. The code manager function of ATLAS.ti allows the user to see, in context, all the stings of texts, in this case lexical items, which have been mapped to a given code. This helps ensure that semantic types are being assigned to lexical items in a consistent manner. It is particularly helpful in cases of doubt when a quick overview of the contexts in which a particular semantic type has been used previously often reveals the most appropriate choice. Similarly, it can also be used to test the validity of novel semantic types. Figure 5.9 juxtaposes the quotations for [[Physical Object]] and [[MB Entity]]. When seen together as a set it seems clear that that the lexical

items in each window denote things with quite different characteristics. This casts doubt on the PDEV practice of grouping them both under [[Physical Object]].

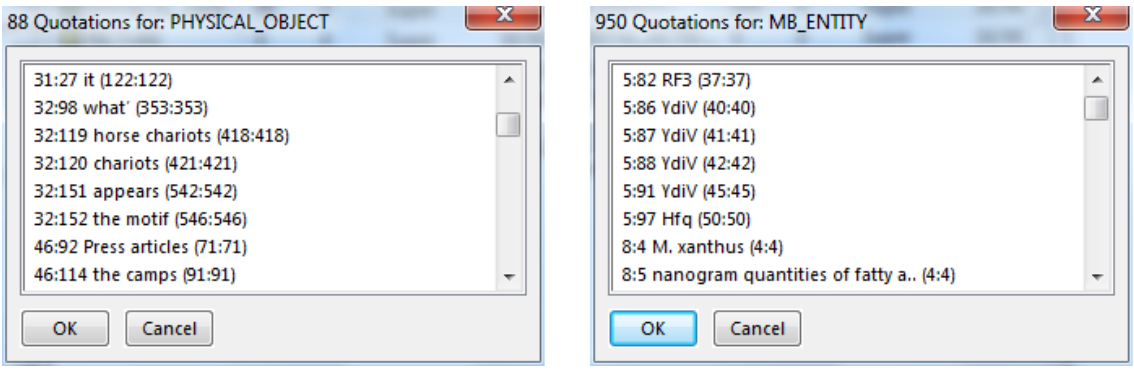

Figure. 5.9. Juxtaposition of quotation windows showing lexical items labelled as [[Physical Object]] and [[MB Entity]]

While not unique to ATLAS.ti, the ability to search for stings of text across documents is certainly useful. If a salient collocation is detected, for example an anomalous collocate, the search function represents a straightforward manner of estimating how widespread this is in the samples of the AJACX2 under study.

A final advantage of ATLAS.ti over the majority of annotation and corpus tools trialled in the pilot study is its robustness. Many of the tools tested proved to be extremely unstable when used with even moderate quantities of data. This lead to frequent data loss. Fortunately, crashes are extremely infrequent with ATLAS.ti and frequent auto-saves and redundant backups mean large-scale data loss is uncommon.

Annotation of the corpus is technologically a fairly straightforward process. However, some contexts provide annotation challenges. The remainder of this section outlines some of these challenges using examples taken from the BNC50 corpus lines used to create the entries in the PDEV.

Syntactic alternations pose a challenge because, at first glance, they may not appear to conform to prototypical patterns. This is true of diathesis alternations where the agent is not always immediately apparent. Take for example the following corpus line used to derive the PDEV entry for *specify*:

> the antero-posterior and dorso-ventral are **specified**, in a manner which is not yet understood

The line is marked as the first pattern type: [[Human]] or [[Document]] or [[Institution]] *specifies* however the agent is not apparent here from the corpus line and must be inferred by the annotator.

Causative and inchoative alternations pose similar problems. Examples (2a) to (2c) are corpus lines used to derive the PDEV entry for *accumulate*.

> 2a. [Edward]discovered a new currant species and **accumulated** a collection which includes the type specimen
> 2b. within the company itself. Thus costs are **accumulated**[…] but no revenue

2c. Considerable evidence has **accumulated** that there is a genetic component to bipolar

The causative alternation in (2a) and the passive in (2b) pose little problem in annotation, both fit the dominant pattern type: [[Human]] or [[Institution]] *accumulates* [[Entity]]. However, care must be taken not to attribute the same pattern to the inchoative in (2c). There is a temptation to assign the dominant pattern [[Human]] or [[Institution]] *accumulates* [[Entity]] when in fact there is no indication of the agent or the cause of the accumulation. A more adequate pattern is: [[Abstract Entity]] *accumulates*.

Indirect object alternations occur in English with verbs of giving. At the time of writing such verbs have yet to be completed in the PDEV; those that have been completed have BNC50 corpus lines which contain no examples of this behaviour. However, the question of how this alternation should be annotated, should it be encountered, is an important one consequently the invented sentences below are provided:

3a. The teacher gave the student a book.
3b. The teacher gave the book to the student.

On the one hand, since there is no semantic or pragmatic difference between the two sentences both could fit the pattern [[Human 1]] *gives* [[Object]] to [[Human 2]]. On the other hand, and crucially for the purposes of the present study, authors from a given discipline might demonstrate a preference for either recipient position. For this reason it is important to be aware, and take note of this type of alternation during the annotation process.

Syntactic alternations involving reciprocal constructions also pose a problem in annotation. Care must be taken to distinguish between patterns which represent corpus lines (3a) and (3b) below:

3a. while I get the house together,' he said, **embracing** her gently before taking us into the house
3b. affection between the two was obvious. Then they **embraced**. Just a peck on the cheek.

Line (3a) maps to the pattern type: [[Human 1]] *embraces* [[Human 2]]. However, the reciprocal construction in (3b) may map to the pattern [[Human = Plural]] *embraces*.
Ellipsis alternations also represent a potential difficulty for the annotator. Some ellipsis alternations are more challenging than others. Take for example 4 below:

4.
The older horses would immediately thunder off in the opposite direction and the youngsters would immediately **follow**. In this way, youngsters learn what is

The verb *follow* has no direct object, if it were not for the context given in the text preceding the verb it would not be possible to derive the pattern [[Animal 1]] *follows* [[Animal 2]]. This null-object alternation could prove somewhat challenging for an annotator. However, the extent of this challenge is debatable as: "A writer can reasonably expect that a reader will proceed sequentially through a text; such a reader can predict what the expected direct object is, which means that it does not need to be stated explicitly." (Hanks, 2013, p. 198)

Clausal ellipsis poses a greater problem in annotation. For example the canonical pattern type for *urge* is: [[Human 1]] or [[Institution 1]] *urges* [[Human 2]] or [[Institution]] to-infinitive. However ellipsis of the infinitive clause and the object makes occasionally makes this pattern difficult to detect.

> 5a. in a speech delivered at Chongup, Rhee **urged** a separate government for south Korea.
> 5b. Thatcher and Mr Lamont's former Treasury boss, **urged** a tax increase to cut the soaring public

In (5) the pattern type: [[Human 1]] or [[Institution 1]] *urges* [[Human 2]] or [[Institution 2]] [to-infinitive]. The infinitive clause, would not be apparent were it not for the context given. The implied meaning in (5) could be (5a): Rhee *urged* authorities to create a separate government; and (5b): The former Treasury boss *urged* authorizes to enact a tax increase. Care must be taken to recognise such ellipsis in annotation.

Care must also be taken to recognise adverbial ellipsis such as the ellipsis of an adjectival from-clause. This can lead to a significant change in sense. For example, in (6a) *abstain* has the canonical sense of 'voluntarily avoiding something unhealthy', in this case risk-taking. However the absence of a from-clause in (6b) points to a sense akin to 'deliberately refraining from casting one's vote'.

> 6a . For compulsive gamblers it is advisable to **abstain** from all activities that involve risk-taking
> 6b. allowed to leave Iraq.) China was **abstaining** , Qian Qichen said, because of its belief

Conative alternations also provide interesting annotation challenges. In conative constructions the object of a verb in transitive construction is realised as the object of a prepositional at-phrase (Levin 1993, p. 42). Semantically conative constructions reflect 'attempted actions' and as such and aspectual change from the transitive connotation.

> 7a. paradise which offers beautiful views as you **sip** your Tequila Sunrise at sunset.
> 7b. We can't take our drink -- we can't just **sip** at it, we have to get blotto. We had to

The implicature of (7a) is:

> [[Human]] drinks [[Beverage]] in very small mouthfuls.

The addition of the preposition *at,* which could be easily missed in the annotation process, gives a quite different implicature:

> [[Human]] drinks [[Beverage]] in a few occasional very small mouthfuls without consuming all of Beverage.

Resultative constructions present a similar annotation challenge. Such constructions focus on the result of the action rather than the action itself.

> 8a. effect that black has on colours -- Poussin **painted** on black canvas, and so did Braque
> 8b. I just ripped up the carpet and **painted** the walls white!

While in (8a) focus is placed on the action of painting in (8b) emphasis is placed on the colour of the walls. This important distinction could be easily missed amongst a mass of corpus lines.

Chapter 1 outlining different approaches to phraseology and Chapter 2 discussing vocabulary selection for language classes highlight the challenge phrasal verbs present for distributional approaches such as the lexical-bundle approach. Such verbs also present an annotation challenge for the CPA approach for reasons similar to conative alterations and resultative constructions.

> 9.
> The reasons for this are complex and to **explain** the situation all **away** in terms of social

It is entirely possible that, in a sea of corpus lines, an annotator might fail to notice the particle *away* thus radically misrepresenting the meaning of the sentence, derive the pattern:

> [[Human]] or [[Institution]] or [[Document]] or [[Proposition]] or [[Eventuality 1]] *explains* [[Eventuality 2]]
>
> Implicature: [[Human]] or [[Institution]] formulates (in conversation with [[Human 2]] or in a [[Document]] intended for [[Human 2]]to read) a [[Proposition]] concerning the cause or effect of an [[Eventuality 2]].

The more appropriate pattern would be:

> [[Human 1]] *explains* [[Eventuality]] *away*,
> Implicature: [[Human 1]] formulates [[Proposition]] as a way of accounting for [[Eventuality]] = Inconvenient.

The way in which that-clauses are annotated in the PDEV is also challenging. In the PDEV patterns, clauses beginning with *that* abound. Take for example the second pattern type listed for *specify*: [[Human]] or [[Document]] or [[Institution]] *specifies* that-clause. An examination of the corpus lines used to derive this pattern type reveals:

the amendment invalid. The amendment **specified** That "No law varying the compensation
company house-style. An organisation will **specify** that all its letters must be typewritten
existed when they wrote. Thus Saint-Pierre **specified** that existing frontiers were not to be
way of Yahweh. Line B, however, in **specifying** that the "clearing"; is a matter
hold Tessas. The Tessa legislation **specifies** that the account must not be held on behalf
up with a dozen possible ideas. But I've **specified** that the address must be within a half-hour
Article 2(2) of the Order.' It is **specified** that the explanatory note is not part of
Wheldrake Ings account book for 1868-1934, it is **specified** that the meadows be mown
the subtour involving towns 1, 3 and 5 by **specifying** that the tour must go from town 1 to 2,
interpretative or hermeneutic setting, we need to **specify** that there is meaning both in `the
behaviour

In these cases, the that-clause could be subsumed by the semantic type [[Rule]]. From the point of view of theoretical eloquence this is preferable as it permits a more accurate description of language without an increase in the complexity of the theory. From the point of the user of the dictionary for productive purposes it might be argued that it is preferable to precisely describe what type of information the that-clause could contain. However, in annotation, patterns types with that-clauses and to-infinitives are treated as separate pattern types even though in many cases their implicature is the same as a pattern type with another syntagmatic arrangement. This decision is motivated by a concern for the needs of the user; in the pilot study undertaken in preparation for the present study it was evident that syntagmatic arrangements with that-clauses or to-infinitives tend to occur considerably more frequently in certain disciplines. This information would be beneficial to the user working in these disciplines. This might also explain the presence of separate to-infinitive and that-clause patterns in the PDEV.

5.3.4 Data Analysis

The data-analysis stage of the study comprises two phases. The first quantitative phase involves the 30 of the 85 verbs in the verb sample randomly selected for analysis. In this phase the composition of each sample of corpus lines is compared. In the second qualitative phase a random selection 10 of the verbs from the verb sample are examined in more detail.

The comparison of collocational behaviour involves the comparison of the relative frequencies of the CPA patterns derived from a sample of corpus lines from each discipline. A key principal of CPA of verbs is that each pattern type activates a distinct verb meaning. It follows that differences in pattern type frequencies across disciplines represent differences in verb meaning across disciplines. Descriptive statistics can be employed to facilitate comparison across sub-corpora. On this approach, the percentage of corpus lines a pattern type represents is calculated for each verb sample in each sub-corpora. A common criticism of many lexical studies and of language research in general is the absence of quantitative methods and statistical techniques to support qualitative analysis (Cantos Gómez, 2013). The appropriateness of such statistical techniques is somewhat limited in the present study since there is no discrete number of verbs exhibiting discipline-specific behaviour which indicates that academic language is discipline specific. However, statistical analysis is employed ensure that any differences in pattern type frequency between corpus line samples are statistically significant. This use of statistical hypothesis testing represents an original application of CPA.

In the present study the statistical analysis comprises two broad stages. The first stage involves an interdisciplinary comparison in which the log-likelihood (LL) and approximate Bayes Factor (BIC) statistics are calculated in three pairwise comparisons for each pattern type found in annotation: History/management studies, management studies/microbiology, and microbiology/history. The occurrence of a given pattern type at significantly different frequency in any of these comparisons could indicate that a verb is being used with a different meaning. However, there is always the possibility that any differences found may be due to the inherent variation between texts and not to some more systemic variation in the patterns used in different disciplines. As Rayson and Garside (2000) state "given the non-random nature of words in a text, we are always likely to find frequencies of words which differ across any two texts" (p. 2). In the second stage of statistical analysis, an intradisciplinary comparison is intended to control for this possibility. In this intradisciplinary comparison a selection of samples containing 300 annotated lines are randomly split into two 150-line samples (samples A and B). The frequency of occurrence of each pattern type is then compared by calculating the LL and BIC statistics. The assumption being that no significant difference between the frequency of a given pattern type between samples A and B indicates that any significant difference found between disciplines is the result of a variable other than inherent variation in texts. Given the composition of the AJACX2 it is reasonable to assume that this variable is academic discipline.

Of the 30 verbs randomly selected for annotation 11 occur in 300 or more corpus lines in each sub-corpora. All these verbs are included in the intra-discipline comparison. However, there is the possibility that these frequently occurring verbs are atypical; that their high frequency is somehow reflected in a uniformity in collocational behaviour in each discipline. For this reason a further five verbs which occurred 300 times or more in only one sub-corpora were included. This brings the number of verbs included in the intradisciplinary comparison to 16 in total. Table 5.4 shows the verbs and disciplines involved in the interdisciplinary comparison (N = 150 for both the A and B samples).

Table 5.4

Verbs and disciplines involved in intradisciplinary comparison

Verb	History	Management	Microbiology
accept	X		
account		X	
act	X	X	X
appear	X	X	X
apply	X	X	X
call	X	X	X
conduct		X	
explain	X	X	X
follow	X	X	X
lead	X	X	X
maintain	X	X	X
manage		X	
need	X	X	X
note	X	X	X
see	X	X	X
specify		X	

Note. X indicates intradisciplinary comparison

In the comparisons described above, the frequency of occurrence of a given pattern type represents the dependent variable. This is an interval variable (ratio). In the inter-discipline comparison, the academic discipline under study represents the independent variable. In the intra-discipline comparison the randomised sample from which a line derives is the independent variable. Both these variables are categorical. The test recommended for variables with these characteristics in social science and health sciences textbooks, and several linguistics textbooks dealing with statistics is the chi-square test (Cantos Gómez, 2013; Oakes, 1998). However, chi-square is a non-parametric test and as such its use is based on the assumption that data are normally distributed. Dunning (1993) demonstrates that this assumption is not valid in the study of natural language texts smaller than billions of tokens in length. He shows that the LL statistic, which assumes binomial or multinomial distributions, represents a more reliable alternative for smaller texts. The LL test is also more reliable than the chi-square test with sparse data. The latter test tends to overestimate significance when the expected frequency is less than five (Rayson & Garside, 2000). This is the case for certain pattern types in the AJACX2.

The LL test has been applied to in a number of corpus studies (A. Wilson, 2013). It is the default option in many corpus analysis programs such as AntConc (Anthony, 2017), WordSmith Tools (M. Scott, 2017), and Wmatrix (Rayson, 2008). Importantly, in the context of this study it has been used successfully in studies involving semantic tagging (Rayson & Garside, 2000). In the interdisciplinary stage of the present study the LL statistic

is calculated in three pairwise comparisons for each pattern type found: History/management studies, management studies/microbiology, and microbiology/history. In the intradisciplinary comparison samples A and B are compared for each discipline. These comparisons are achieved by constructing contingency tables like Table 5.5 below.

Table 5.5

Contingency table for pairwise comparison using LL statistic

	Sample One	Sample Two	Total
Freq. of Pattern	A	B	a+b
Freq. of All Other Patterns	c-a	d-b	c+d-a-b

In the table above c and d correspond to the N values for a given pattern type. The observed values (O) for a given pattern type are represented by a and b. The following formula is used to calculate the expected values (E):

$$E_i = \frac{N_i \sum_i O_i}{\sum_i N_i}$$

There is no need to normalise the figures to account for different sample sizes as the formula takes this into account. Having calculated the expected frequency the LL statistic is given:

$$-2\ln \lambda = 2\sum_i O_i \ln\left(\frac{O_i}{E_i}\right)$$

Since calculating the LL statistic is computationally undemanding, the use of statistical packages such as SPSS offers little advantage in practical terms or in terms of reliability. Instead, Rayson's (2016) implementation in Microsoft Excel is employed.

The LL test indicates whether a difference is significant or not at a given p-value. Contrary to widespread belief, it provides no information about the magnitude of difference (Gries, 2005; A. Wilson, 2013). The same is true of all test statistics which are measured against a p-value; chi-squared, fishers-exact, t-test etc. A related problem is that all these significance test statistics conflate the effect of an independent variable, in the case of the present study the origin of the sample in which pattern types occur, with the size of the sample. What is needed, then, is a measure of effect size.

Bayes Factors have been increasingly employed to measure effect size in many fields especially genetics (Sawcer, 2010). More recently they have been fruitfully employed in studies of the effect of bilingual dictionary use on writing tests amongst Polish secondary school children (Lew, 2016). The Bayesian Information Criterion (BIC) can be used alongside the log-likelihood statistic as an approximate Bayes Factor in order to measure effect size (Kass & Raftery, 1995). A. Wilson (2013) demonstrates how approximate Bayes Factors (BIC) can be used to compare lexical items, part-of-speech categories, or semantic fields across corpora. Approximate Bayes Factors (BIC) can be given straightforwardly for the LL statistic:

$\text{BIC} \approx \text{LL} - \log(\text{N})$

Here N = total number of patterns in both samples. The approximate Bayes Factor given represents a degree of evidence against the null hypothesis under test. A. Wilson (2013) provides Table 5.6 based on Kass and Rafery (1995, p. 777):

Table 5.6
Degrees of evidence against the null-hypothesis (H_0)

Approximate Bayes Factor	Degree of evidence against the null-hypothesis (H_0)
0-2	not worth more than a bare mention
2-6	positive evidence against H_0
6-10	strong evidence against H_0
>10	very strong evidence against H_0

A number of corpus linguists have highlighted the need for effect size measures in corpus studies (Gries, 2005; Kilgarriff, 2001). To address this, following the procedure set out above, Approximate Bayes Factors (BIC) are employed in the present study. The use of Bayesian approach is commonplace in other disciplines and there is also precedent for its successful application in corpus linguistics (A. Wilson, 2013) and lexicography (Lew, 2016). As Table 5.6 above indicates BIC is uncomplicated to interpret. Furthermore, its calculation is undemanding and is implemented in Rayson's (2016) Microsoft Excel LL calculator.

In both interdisciplinary and intradisciplinary comparisons an approximate BIC value indicating at least positive evidence against H_0 is set as the minimum threshold level for significant difference between pattern-type occurrences. The significance level adopted is $p = < .05$. In the interdisciplinary comparison, the following null hypothesis is tested:

> H_0: there is no difference in the frequency of occurrence of a given pattern type across sub-corpora

The alternative hypothesis is as follows:

> H_1: there is a difference in the frequency of occurrence of a given pattern type across sub-corpora

In the intradisciplinary comparison the hypotheses tested are reversed:

> H_0: there is a difference in the frequency of occurrence of a given pattern type across sub-corpora

The alternative hypothesis is as follows:

> H_1: there is no difference in the frequency of occurrence of a given pattern type across sub-corpora

The statistical analysis reported in the next chapter is followed by a more qualitative analysis and discussion in Chapter 7. This qualitative stage examines some of the discipline-

specific differences found in the quantitative stage in greater detail and discusses how they might be usefully included in a lexicographical resource for the EAP community.

6. STATISTICAL ANALYSIS

This chapter summarises the results of the experimental procedure outlined in the previous chapter. Section 6.1 reports on the interdisciplinary comparison. It treats each of the 30 verbs annotated individually. Section 6.2 presents a general summary of the intradisciplinary comparison. Finally, in Section 6.3, some provisional conclusions and reflections on the implications of the finding for EAP pedagogy and lexicography are presented.

The previous chapter contained a discussion of the problems inherent in annotating corpus lines with a view to performing a statistically valid comparison between disciplines. The principal problems are deciding, firstly, on the requisite degree of granularity in annotation —deciding if patterns with fine-grained distinctions in semantic type represent separate pattern types; and secondly, whether to widen patterns to include a semantic type which typically only occurs in one discipline, or alternatively treat the pattern containing the anomalous semantic type as a separate pattern type in its own right. The notion of implicature is fundamental to the CPA response to these problems. In the case of the first problem, this means climbing the CPASO in order that the semantic type recorded for statistical analysis is the highest possible in the CPASO without admitting lexical items which would change the sense of the verb. Similarly, in the case of the second problem, for the purposes of statistical comparison, patterns are widened to include anomalous semantic types on condition that these changes do not alter the core sense of the implicature.

1.

> [[Human]] or [[Abstract_Entity]] **absorbs** [[Asset]]
> <Human or Abstract_Entity uses Asset - This is not typically considered a good use of Asset>
> The past nine years have […], *absorbed* countless years of time in two opposing local authorities

Example 1, taken for the PDEV, provides an illustration of how the question of granularity in annotation is addressed. The NP in the subject slot: *the past nine years,* could have been represented by [[Time Period]]. However, bearing in mind the context of the other corpus lines for *absorb* the annotator has climbed the CPASO effectively subsuming [[Time Period]] under its parent node [[Abstract Entity]].

2.

> [[Location]] or [[Building]] or [[Machine]] **accommodates** [[Physical Object]] or [[MB Entity]] or [[Activity]]
> <Location, in case of MB Entity part of cell, or Building or Machine or provides enough space for Physical Object or MB Entity to be located or for Activity to take place>
> Ex. History:
> Lecture rooms 'with sufficient space to *accommodate* an audience' in the houses of advocates
> Extra galleries were opened to *accommodate* the huge range of his photographic record
> Ex. Management:
> Ultimately, websites began *accommodating* financial transactions.
> Ex. Microbiology:
> antigen-binding pocket that is designed to *accommodate* the hydrocarbon chains of lipids
> the conventional clathrin cage cannot *accommodate* microorganisms the size of bacteria, even

Example 2, taken from the AJACX2, provides an illustration of the widening of the scope of a pattern to include a semantic type only present in one discipline. Examples from

history and management include the semantic types [[Physical Object]] or [[Activity]] in the object slot while the lexical items in microbiology could be generalised over with [[MB Entity]]. Widening the pattern to include this anomalous semantic type results in little significant change to the implicature.

This approach has methodological and practical implications. It ensures that comparisons across disciplines are in fact like-for-like. It avoids the possibility that a use of a semantic type which occurs in only one discipline creates the unwarranted impression of a difference in pattern-type frequency across sub-corpora. Additionally, it may produce data of benefit to the user to the extent that it allows valid comparison with general English through the PDEV where the anomalous semantic type may well have occurred only to be counted as an exploitation of a more prevalent pattern. The idea here is not to conduct a thorough comparison with general English, but to indicate in a general manner the extent to which academic uses are distinct from general English. From the perspective of the user of any eventual lexicographical resource climbing the CPASO and widening patterns to include semantic types that only occur in one discipline results in the loss of information that is of potential use. A similar argument applies to distinct syntagmatic arrangements which involve no great difference in implicature, yet occur reasonably frequently in one discipline and rarely in another. In the pilot study conducted in preparation for the present study the pattern: [[Human]] or [[Institution]] *accepts* [that clause] occurred frequently across all three disciplines. However, in microbiology the syntactic alternation: {it} is ({generally} {well} {widely}) *accepted* {that} accounted for 100 per cent of occurrences. This information would surely be of use to a student of microbiology.

The approach taken in the statistical analysis which follows is the same as that adopted in the PDEV. Patterns generalize over corpus lines to the maximum extent possible without provoking a change in implicature. There is however one notable exception to this rule as regards syntagmatic arrangements which is also present in the PDEV: Patterns types with that-clauses and to-infinitives are treated as separate pattern types even though in many cases their implicature is the same as a pattern type with another syntagmatic arrangement. This decision is motivated by a concern for the needs of the user; in the pilot study undertaken in preparation for the present study it was evident that syntagmatic arrangements with that-clauses or to-infinitives tend to occur considerably more frequently in certain disciplines. This information would be beneficial to the user working in these disciplines. This might also explain the presence of separate to-infinitive and that-clause patterns in the PDEV.

In the present chapter a rigorous statistical analysis is undertaken with the aim of testing the general hypothesis: Verbs exhibit different meanings in different academic disciplines and analysis of the semantic and syntagmatic patterns reveals differences elided by distributional approaches to vocabulary extraction. This analysis will provide an initial indication of the type of differences that occur across disciplines. The following qualitative-analysis chapter will further illustrate the nature of this difference while bringing to light characteristics which have remained hidden during the statistical analysis phase. The practical outcome envisaged is to posit a means of representing the discipline-specific behaviour of verbs in lexicographical resources that is useful to the EAP community.

Following the annotation criteria set out above, after annotating 16,420 corpus lines of text a total of 241 pattern types were found.

6.1 Interdisciplinary Comparison

As set out in the previous chapter, in order to test the hypothesis of the study the log-likelihood (LL) statistic is employed with an assumed significance level of at least $p < .05$. For the interdisciplinary comparison, the following null hypothesis is tested:

> H_0: there is no difference in the frequency of occurrence of a given pattern type across sub-corpora

The alternative hypothesis is as follows:

> H_1: there is a difference in the frequency of occurrence of a given pattern type across sub-corpora

Contrary to widespread belief, rejection or otherwise of the H_0 reveals nothing about the actual size of the difference between categories (A. Wilson, 2013). For this reason, the present study adopts the effect size measure the Bayes Information Criterion (BIC) to evaluate the amount of evidence provided against the H_0 by the log-likelihood test. Kass and Raftery (1995) suggest that in conjunction with a significant log-likelihood statistic a BIC score of ≥ 2.00 represents positive evidence against the H_0. These are the threshold criteria adopted for discipline-specific behaviour of a pattern in the following statistical analysis. For ease of reading full patterns, LL scores, raw and percentage frequency figures are not provided in-line with the text. This information can be found in Appendix C. Line charts are employed to compare the proportion of a sample that a pattern type represents across sub-corpora. Henceforth in this chapter pattern types are referred to using the abbreviation PT followed by a number. Summary tables are given as a reference for full pattern types and to indicate those patterns meeting the criteria for discipline-specific behaviour. Sample sizes are represented using the standard nomenclature of *N*, in the present study *N* represents the number of individual patterns (pattern tokens), or in other words corpus lines, for a given verb in a given discipline.

6.1.1 *Absorb*

An examination of the corpus revealed 13 patterns types with *absorb* two of which, [[Stuff]] *absorbs* {strongly} (PT12), and [[Liquid 1]] *is absorbed* with [[Liquid 2]] (PT13), are not present in general English as represented by the PDEV. There are no PTs which occur across all three sub-corpora. Neither are there PTs which are exclusive to management (*N* = 63). Three PTs occur exclusively in microbiology (*N* = 26). These include patterns 12, 13, and three. Two PTs occur exclusively in history (*N* = 62): patterns seven and 11. Three occur in both history and microbiology: patterns types one, two, and four. Five occur in both history and management: pattern types five, six, eight, nine, and 10.

An examination of the distribution of patterns across disciplines gives an initial indication of the collocational behaviour of verbs in different disciplines. An inspection of frequency of occurrence, expressed as a proportion of each sample in Figure. 6.1 below, reveals a more nuanced situation with many PTs appearing just once or twice. These include five PTs in history and the two new PTs, not present in general English, which are found in microbiology. In spite of the low number of occurrences, there is some indication of *absorb* demonstrating collocational behaviour in microbiology which is distinct to that displayed in history and management.

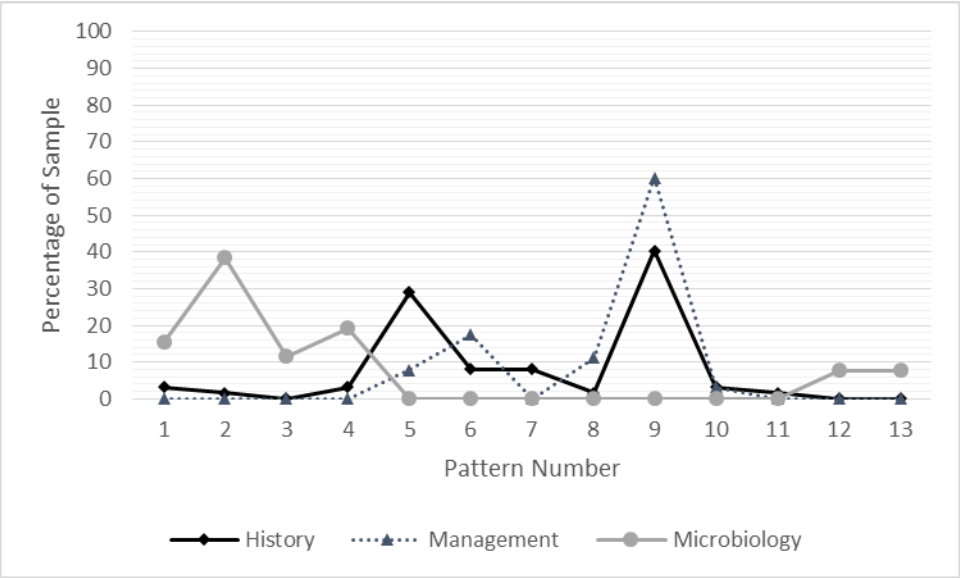

Figure 6.1. Interdisciplinary comparison for *absorb*.

This slight tendency towards a split between history and management on one hand, and microbiology on the other is reflected in a cross-sub-corpora comparison of the frequency of occurrence of each pattern. A comparison of management and microbiology (summarised in Table 6.1 below) reveals six patterns occurring at significantly different frequencies and with BIC scores indicating at least positive evidence against H_0, a comparison of microbiology and history reveals three, while in a comparison between management and history there are two patterns which meet these criteria. Additionally, it should be noted that two of the patterns which occurred exclusively in microbiology (PTs 12 & 13) reached the significance threshold in a comparison of microbiology and history, however, the low frequency of occurrence (2 hits each) is reflected in an effect size which does not reach the threshold for consideration as evidence against the H_0.

Table 6.1

Result of cross-corpora comparison for absorb

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Physical_Object 1]] absorbs [[Stuff]] or [[Physical_Object 2]] <Stuff or Physical_Object 2 becomes part of Physical Object 1 by means of a natural process>		X	
2	[[Human]] or [[Animal]] or [[Plant]] or [[Body_Part]] or [[MB_Entity]] absorbs [[Stuff]] <Human or Animal or Plant or Body_Part takes Stuff = Nutrient or Liquid into the body so that Stuff or Liquid becomes an integral part of Human or Animal or Plant or Cell or Cell part>		X	X
3	[[Physical_Object]] absorbs [[Radiation]] <Physical_Object receives Radiation in large quantities so that Radiation becomes part of Physical Object>		X	

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

Table 6.1 (cont.)
Result of cross-corpora comparison for absorb

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
4	[[Physical_Object]] or [[Stuff]] absorbs [[Energy]] <Physical_Object receives and utilizes or stores Energy>		X	
5	[[Human_Group 1]] or [[Institution 1]] absorbs [[Human_Group 2]] or [[Institution 2]] <Human_Group 2 or Institution 2 becomes a part of Human Group 1 or Institution 1>	X		X
6	[[Eventuality 1]] or [[Institution]] absorbs [[Eventuality 2]] or [[Institution]] <Eventuality 2 becomes a part of Eventuality 1 or Institution>		X	
7	[[Human]] or [[Abstract_Entity]] absorbs [[Asset]] <Human or Abstract_Entity uses Asset - This is not typically considered a good use of [[Asset]]>	X		
8	[[Institution]] or [[Human]] absorbs [[Deficit]] <Business | Economics. Jargon. Institution or Human = Business Person is robust or wealthy enough to bear the adverse effects of Deficit>			
9	[[Human]] or [[Institution]] absorbs [[Anything]] or [[Attitude]] <Human learns about, understands, and accepts Anything = Concept or Attitude>		X	X
10	[[Human]] is or become absorbed in or by [[Activity]] or in or by [[Anything]] <Human is greatly interested in Activity or Anything - This is typically to the exclusion of any other [[Activity | Anything]]>			
11	[[Abstract_Entity]] absorbs [[property]] <Abstract_Entity is influenced by property of another abstract entity>			
12	[[Stuff]] absorbs {strongly} <stuff = liquid or other molecule enters MB entity via a microbiological process>			
13	[[Liquid 1]] is absorbed with [[Liquid 2]] <Human adds Liquid 2 to Liquid 2 = chemical reagent, to test if reaction occurs>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.2 *Accelerate*

An examination of the corpus sample for *accelerate* reveals three different PTs, all of which appear in general English as represented by the PDEV. Two PTs, [[Anything]] *accelerates* [[Process]] or [[Pace]] (PT1), and [[Process]] or [[Pace]] *accelerates* (PT2), occur at reasonably high frequencies across all three disciplines. The third PT, [[Human]] *accelerates* [[Proposition]], is present solely in the history sub-corpus where it occurs only once.

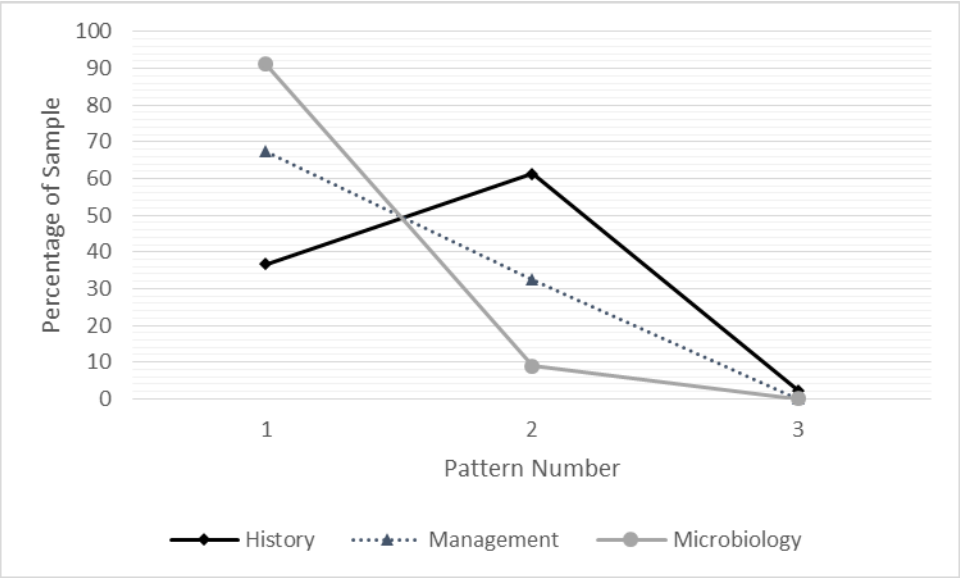

Figure 6.2. Interdisciplinary comparison for *accelerate*.

At first glance, an inspection of frequency of occurrence expressed as a proportion of each sample (Figure 6.2) suggests that although all PTs occur across the three disciplines, there are important differences in the frequency at which they occur. This is not entirely confirmed by significance testing (Table 6.2). Although, comparison using LL score shows significant differences in frequency of the first PT between history (*N* = 44) and management (*N*= 43) and between all disciplines for the second PT, the effect sizes suggest that the only difference which represents evidence against the H_0, in this case to a very strong degree, is the difference between microbiology (*N* = 45) and history for PT two.

Table 6.2

Result of Cross-corpora Comparison for accelerate

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Anything]] accelerates [[Process]] or [[Pace]] <Anything causes Pace of Process to proceed more quickly>			
2	[[Process]] or [[Pace]] accelerates <Pace of Process proceeds more and more quickly>			X
3	[[Human]] accelerates [[Proposition]] <Human expresses Speech Act with greater urgency>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.3 *Accept*

The verb *accept* demonstrates prolific collocational behaviour. It occurs in 17 PTs in the AJACX2, nine of which are PTs found in general English as represented by the PDEV.

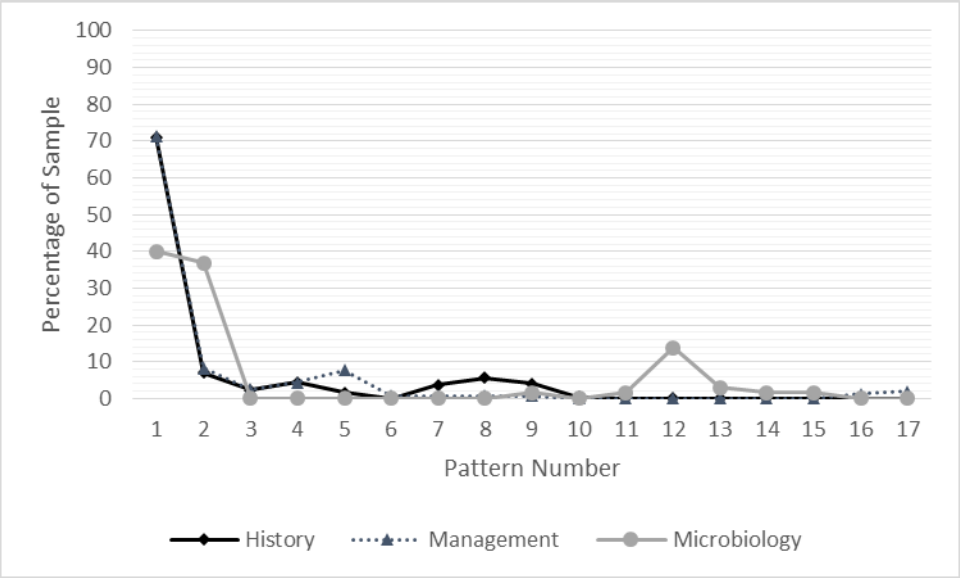

Figure 6.3. Interdisciplinary comparison for *accept*.

As illustrated by Figure 6.3 above, three PTs are found across all sub-corpora. These include PTs one and two which occur frequently across all disciplines, and pattern nine which is limited to a single occurrence in management and microbiology. Five PTs, three, five, seven, and eight, occur in both history (N = 300) and management (N = 157) although in the case of management PTs seven and eight are limited to single occurrences. There is one instance of the PT ten: [[Location]] *accepts* [[Human Group]] which occurs in history. Similarly the three patterns unique to management, PTs six, 16, and 17, found exclusively in microbiology, occur at very low frequencies. The same is true of all but one of the five PTs, patterns 11 – 15, in microbiology (N = 65). Given the large differences in sample sizes, an examination of significance and effect size is required for a more accurate picture of *accept* across disciplines.

A cross-sub-corpora comparison of the frequency of occurrence of each pattern (summarised in Table 6.3 below) reveals some evidence for a split between history and management on one hand, and microbiology on the other. A comparison of management and microbiology reveals four PTs occurring at significantly different frequencies with BIC scores indicating at least positive evidence against H_0, similarly a comparison of microbiology and history also reveals four PTs which meet these criteria. In contrast, a comparison between management and history reveals two PTs which meet these criteria.

Table 6.3

Result of cross-corpora comparison for accept

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human]] or [[Institution]] accepts [[Proposition]] or [[Concept]] or [[Eventuality]] <Human or Institution agrees that Proposition or Concept or Eventuality is correct and does not need to be contested>		X	X
2	[[Human]] or [[Institution]] accepts <Human or Institution agrees that that [clause] is true or correct>		X	X
3	[[Human]] or [[Institution]] accepts {responsibility} (for Eventuality or for Entity) <Human or Institution agrees that he, she, or it is or will be responsible for Eventuality or Entity>			
4	[[Human 1]] accepts [[Entity]] or [[Money]] <Human 1 consents to receive Entity = Valuable or Money as a gift (or bribe) or as part of business transaction from Human 2>			
5	[[Human]] accepts [[Offer]] <Human 1 agrees to act on the Offer of Human 2>	X	X	
6	[[Human 1]] accepts {resignation} <Human 1 = Authority receives and consents to resignation of Human 2 = Employee>			
7	[[Human]] accepts [[Human Role]] <Human agrees to undertake the work specified as required to fulfil appointment or job>			
8	[[Human 1]] or [[Institution 1]] accepts [[Human 2]] or [[Institution 2]] as [[Human_Role]] or as [[Institution_Role]] <Human 1 or Institution 1 acknowledges that Human 2 or Institution 2 is suitable for Human_Role or Institution_Role>	X		
9	[[Human_Group 1]] or [[Human 1]] accepts [[Human 2]] or [[Human_Group 2]] <Human_Group 1 or Human 1 agrees to allow Human 2 or Human Group 2 to become a member of or be respected by Human Group 1>			
10	[[Location]] accepts [[Human Group]] <Location is suitable for Human Group>			
11	[[MB Entity]] accepts {task} <MB Entity= part of cell undertakes task in cell>			
12	[[MB Entity1]] accepts [[MB Entity 2]] <MB Entity 1 takes MB Entity 2 which becomes an integral part of MB Entity 1>		X	X
13	[[MB_Entity 1]] accepts [[MB_Entity 2] as [[MB_Role]] <MB_Entity 2 plays a role in the normal functioning of MB_Entity 1>			X

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

Table 6.3 (cont.)

Result of cross-corpora comparison for accept

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
14	[[Device]] accepts [[Stuff]] <Device is capable of processing Stuff = biological sample>			
15	[[MB Entity]] accepts {topology} <MB Entity=gene occurs in a given spatial configuration>			
16	[[Concept 1]] accepts [[Concept 2]] or [[Eventuality]] <Concept 1 is compatible with Concept 2 or Eventuality>			
17	[[Human]] or [[Institution]] accepts [[Document]] <Human or Institution agrees that Document meets the standards required for publication>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.4 *Accommodate*

An examination of the collocational behaviour of *accommodate* in the corpus reveals eight PTs in total, six of which occur in general English as represented by the PDEV. An analysis of their frequency of occurrence, expressed as a percentage of the sample (Figure 6.4), across sub-corpora reveals three core PTs which represent a good proportion of the sample in two of the three disciplines. These are PT two, three, and eight. There are three PTs which occur, at very low frequencies, in both history (N = 95) and management (N = 70) yet are absent in Microbiology. These include PTs one, four, and five. PT seven is unique to history in the AJACX2 where it occurs three times. In the AJACX2, PT six is unique to the microbiology (N = 37) sub-corpus where it occurs six times.

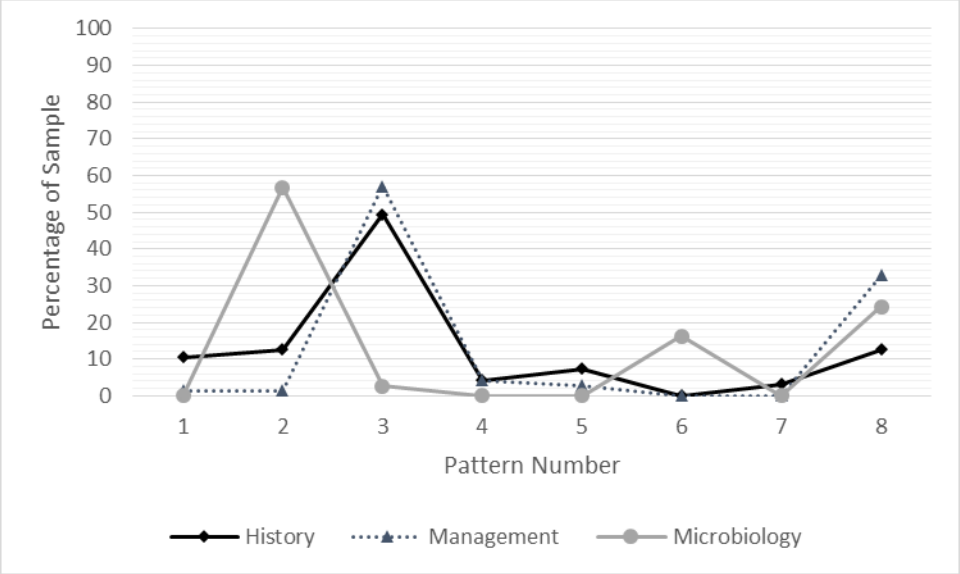

Figure 6.4. Interdisciplinary comparison for *accommodate*.

As for the significance of these differences (summarised in Table. 6.4), in a comparison of history and management PTs two and eight meet the criteria for meaningful difference. In a comparison of management and microbiology there are four PTs which meet the criteria, and applying the same procedure to microbiology and history reveals three significant differences between PT occurrence.

Table 6.4

Result of cross-corpora comparison for accommodate

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human 1]] or [[Institution]] or [[Location]] accommodates [[Human 2]] or [[Human_Group]] or [[Animal_Group]] <Human 1 or Institution provides suitable space in Location for Human 2 or Human_Group or Animal_Group to live or work in>			
2	[[Location]] or [[Building]] or [[Machine]] accommodates [[Physical_Object]] or [[MB Entity]] or [[Activity]] <Location, in case of MB Entity part of cell, or Building or Machine or provides enough space for Physical_Object or MB_Entity to be located or for Activity to take place>	X	X	X
3	[[Human]] or [[Institution]] accommodates [[Eventuality]] <Human or Institution deliberately adapts the way they think or behave to take account of Eventuality>		X	X
4	[[Human 1]] accommodates [[Human 2]] <Human 1 makes a special effort meet the needs or wishes of Human 2>			
5	[[Human]] or [[Institution]] accommodates [[Self]]to [[Abstract_Entity]] or to [[Eventuality]] <Human or Institution deliberately adapts the way they think or behave to take account of Abstract_Entity or Eventuality>			
6	[[Animate]] or [[Plant]] or [[MB Entity]] accommodates (Eventuality) <Animate or Plant or MB Entity makes adjustments in behaviour or structure in order to cope with Eventuality>		X	X
7	[[Human]] or [[Activity]] accommodates [[Eventuality 1]] to [[Eventuality 2]] <Human or Activity adjusts Eventuality 1 in order to cope with Eventuality 2>			
8	[[Concept]] or [[State of affairs]] accommodates [[Eventuality]] <Concept or state of affairs provides the conditions necessary for Eventuality to be realised>	X	X	

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.5 *Accompany*

There are five PTs found in an examination of the collocational behaviour of *accompany* in the AJACX2. Four of these, PTs one to four, are found in general English as represented by the PDEV. Figure 6.5 shows the distribution of patterns across disciplines and the proportion of each sample they represent. Patterns one to three occur across all disciplines, while pattern four is exclusive to history (*N* = 150) and pattern five exclusive to management studies (*N* = 71). In contrast to the other verbs addressed thus far in the study there are no patterns exclusive to microbiology (*N*= 95).

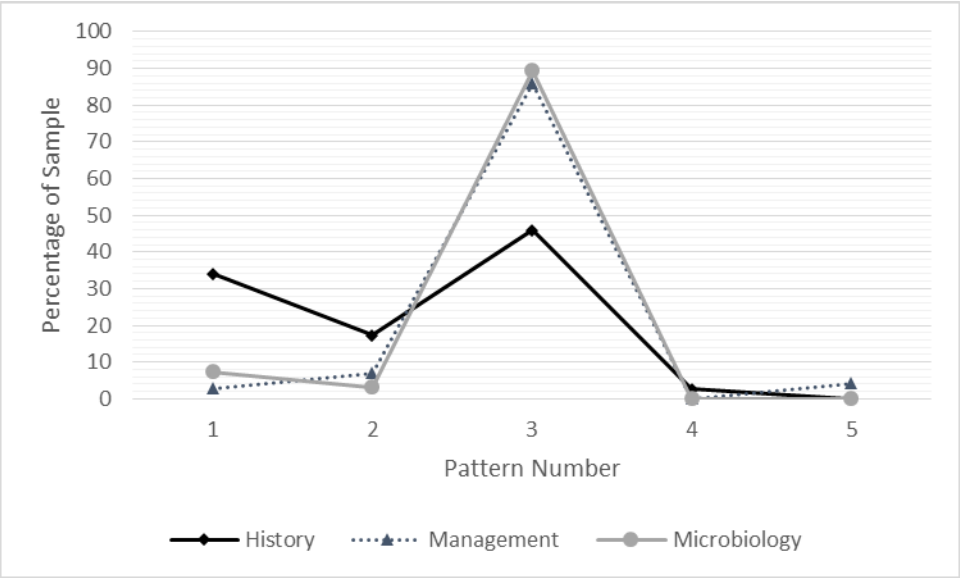

Figure 6.5. Interdisciplinary comparison for *accompany*.

Figure 6.5 above illustrates the limitations of relying on a binary present-or-not model to compare verbal collocation patterns across academic disciplines. Although PTs one to three occur across all disciplines, there is a notable difference in the proportion of the sample they represent in history on the one hand, and microbiology and management on the other. An examination of the significance of these differences (summarised in Table 6.5 below) partially confirms this impression with differences between PTs one to three meeting the established thresholds in the case of microbiology and history, and PTs one and three meeting these thresholds in the case of history and management. In the case of the latter comparison a low effect size score for pattern two places the significance of the difference in doubt. The collocational behaviour of this verb seems to run contrary to the humanities/science division hinted at as thus far.

Table 6.5

Result of cross-corpora comparison for accompany

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human 1]] or [[Vehicle 1]] or [[Human_Group 1]] or [[MB Entity1]] accompanies [[Human 2]] or [[Animal]] or [[Vehicle 2]] or [[Human_Group 2]] or [[MB Entity 2]] (Direction) <Human 1 or Human_Group 1 or Vehicle 1 or MB_Entity 2 goes with Human 2 or Human Group 2 or Animal or Vehicle 2 or MB Entity 2, either when travelling Direction or (if Human 1, 2) also at a social function>	X		X
2	[[Document 1]] accompanies [[Document 2]] or [[Artifact]] <Document 1 is transmitted or published at the same time as Document 2 or Artifact = Valuable, in order to explain, verify, or supplement it>			X
3	[[Eventuality 1]] accompanies [[Eventuality 2]] <Eventuality 1 occurs at the same time as Eventuality 2 and may be associated with it in some way>	X		X

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

Table 6.5 (cont.)

Result of cross-corpora comparison for accompany

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
4	[[Human 1]] or [[Musical_Performance 1]] accompanies [[Human 2]] or [[Musical_Performance 2]] <Human 1 = Musician plays Musical_Performance 1 at the same time as Musical Performance 2 is played by Human 2 = Musician>			
5	[[Abstract Entity 1]] accompanies [[Abstract Entity 2]] <Abstract Entity 1 forms an important part of Abstract Entity 2>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.6 *Accomplish*

The analysis of the corpus revealed a single PT for *accomplish*: [[Human]] or [[Institution]] or [[MB Entity]] or [[Process]] *accomplishes* [[Activity]] or [[Eventuality]]. This PT is wider than that found in general English as represented by the PDEV. The general pattern does not contain [[MB Entity]] or [[Process]]. A close examination of the corpus lines used to create the PDEV entry for this verb revealed no processes which could not be attributed to human agency and no lexical items which might be classed as [[MB Entity]]. However, in the microbiology sub-corpus (N = 67), these semantic types occurred in the majority of lines (n = 45). This points to the existence of a discipline-specific alternation in which [[MB Entity]] or [[Process]] alternates with [[Human]] or [[Institution]] with *accomplish* in microbiology.

6.1.7 *Account*

The verb *account* exhibits five PTs in the AJACX2, the first four of which occur in general English as represented by the PDEV. The fifth PT occurs exclusively in microbiology (N = 100). The first two PTs occur at relatively high frequencies across all disciplines. The third pattern type occurs frequently in history (N = 111) and management (N = 300) while the fourth pattern is limited to a handful of occurrences in history and microbiology. Figure 6.6 illustrates this distribution as a percentage of the sample from each discipline.

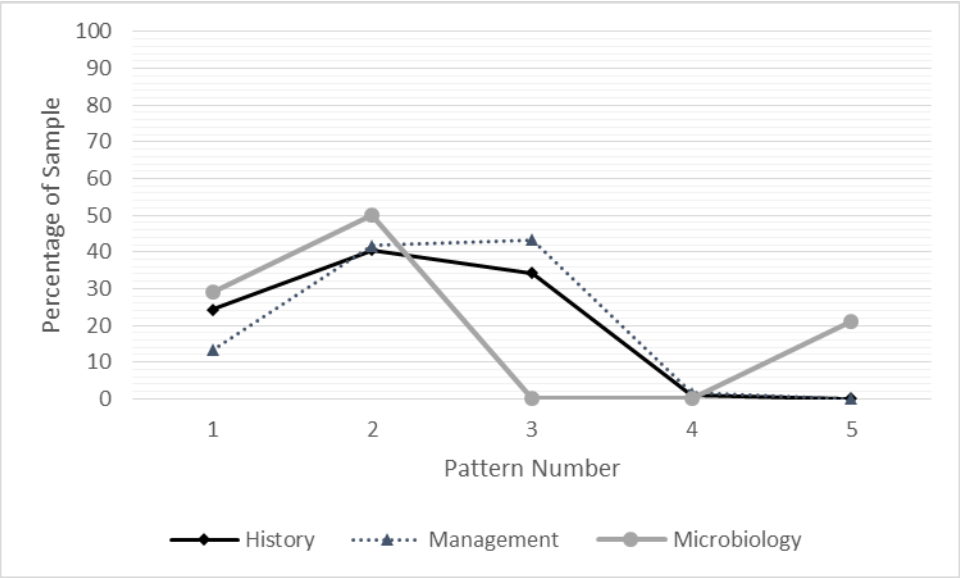

Figure 6.6. Interdisciplinary comparison for *account*.

A statistical comparison between history and management (Table 6.6), reveals no significant difference in frequency of occurrence for each PT. In a comparison of management and microbiology, PTs one, three, and five meet the criteria for significant difference. In a comparison of microbiology and history PTs three and five meet the criteria. Here again, there is evidence for a humanities/science split.

Table 6.6
Result of cross-corpora comparison for account

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Entity 1]] or [[Eventuality 1]] accounts for [[Numerical_Value]] of [[Entity 2]] or of [[Eventuality 2]] <Entity 1 or Eventuality 1 makes up a certain Numerical_Value = Percentage of the total number of Entity 2 or of Eventuality 2>		X	
2	[[Eventuality 1]] accounts for [[Eventuality 2]] <Eventuality 1 was a factor in causing Eventuality 2 to happen>			
3	[[Human]]or [[Concept]] accounts for [[Eventuality]] <Human or Concept explains or justifies Eventuality>		X	X
4	[[Human]] accounts for [[Entity]] <Human is able to explain where Entity is or how it was used>			
5	[[MB Entity]] accounts for [[Eventuality]] <the presence of MB_Entity explains Eventuality>		X	X

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.8 *Accumulate*

An analysis of *accumulate* in the AJACX2 brings to light four PTs, all of which have equivalents in general English as represented by the PDEV. As illustrated in Figure 6.7, PTs one and four occur quite frequently across all sub-corpora. The other two PTs are limited to microbiology (N = 145). PT two: [[Body_Part]] accumulates [[Stuff]] is derived

from a single corpus line, PT three: [[Stuff]] accumulates in [[Location]] or on [[Location]] accounts for slightly under half (n = 72) of all the pattern occurrences in microbiology.

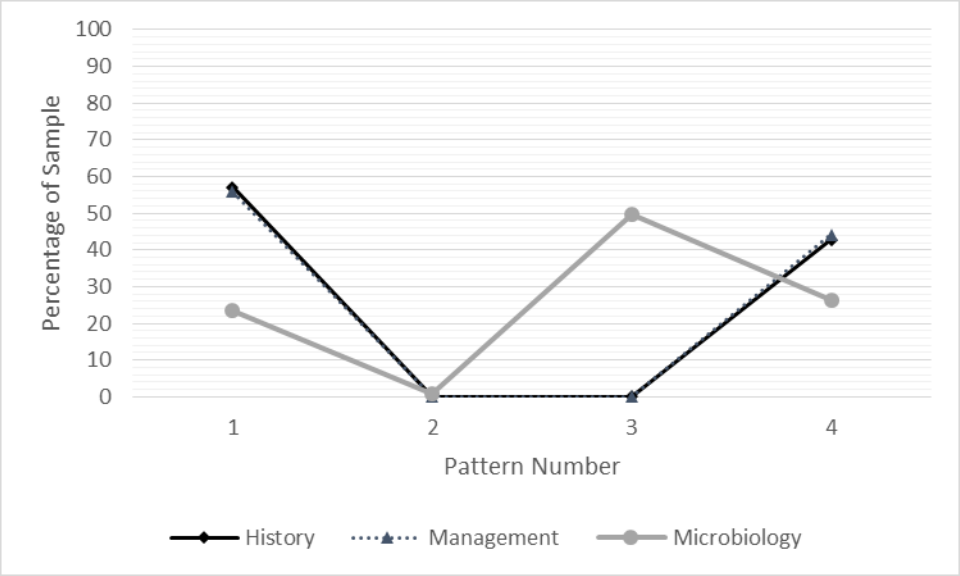

Figure 6.7 Interdisciplinary comparison for *accumulate*.

Statistical analysis reveals significant differences in PT frequency of occurrence between disciplines (summarised in Table 6.7). The single occurrence of PT two prohibits valid comparison. However, the PTs one and three meet the significance criteria in comparisons of management (N = 111) and microbiology, and microbiology and history. There are no significant differences in frequency of occurrence revealed in a comparison of history (N = 42) and management. This suggests, once again, a division with history and management on one side and microbiology on the other.

Table 6.7

Result of cross-corpora comparison for accumulate

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human]] or [[Institution]] or [[MB Entity]] accumulates [[Entity]] <Human or Institution acquires or experiences an amount of Entity over a period of time>		X	X
2	[[Body_Part]] accumulates [[Stuff]] <Stuff collects and builds up in Body_Part - Typically, this is undesirable>			
3	[[Stuff]] accumulates in [[Location]] or on [[Location]] <A quantity of Stuff collects and builds up in Location or on Location>		X	X
4	[[Abstract Entity]] or [[Document]] or [[MB Entity]] accumulates <The amount of Abstract_Entity available or experienced increases>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.9 *Act*

Analysis of the AJACX2 reveals 13 PTs for *act,* 11 of which occur in the general English as represented by the PDEV. As Figure 6.8 indicates, in the majority of its manifestations the verb occurs in the first three PTs in all disciplines (*N* = 300 in each discipline). Other notable PTs include pattern eight for all disciplines, and pattern six for history and management. The remaining PTs represent one or two occurrences.

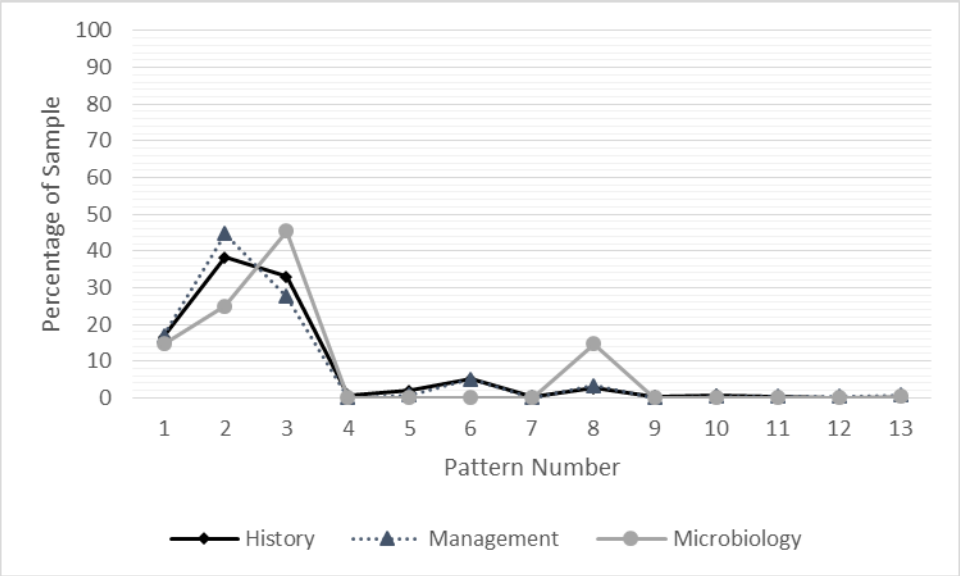

Figure 6.8 Interdisciplinary comparison for *act*

Significance testing once again reveals information about verb usage across academic disciplines which might be useful for the EAP student (Table 6.8). PT one occurs at a fairly consistent level across all three disciplines. Pairwise comparisons of both history and management with microbiology reveal that PTs two, three, six, and eight meet the criteria for significant difference. The low frequencies of the remaining patterns make reliable hypothesis testing unfeasible this is reflected in the effect-size scores obtained.

Table 6.8
Result of cross-corpora comparison for act

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human]] or [[Institution]] or [[Animal]] or [[Machine]] or [[MB unit]] acts <Human or Institution or Animal or Machine = Agent performs a motivated Action>			
2	[[Human]] or [[Institution]] or [[MB unit]] acts [[Manner]] <Human or Institution behaves in the manner specified>		X	X
3	[[Entity 1]] acts as or like [[Anything]] <In a particular situation, Entity 1 performs the role or function specified>		X	X
4	[[Human 1]] or [[Institution 1]] acts for [[Human 2]] or for [[Institution 2]] <Human 1 = Lawyer or Agent or Institution 1 is employed to perform certain tasks for Human 2 or Institution 2>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

Table 6.8 (cont.)
Result of cross-corpora comparison for act

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
5	[[Human 1]] or [[Institution 1]] acts on behalf of [[Human 2]] or on behalf of [[Institution 2]] <Human 1 or Institution 1 performs a motivated Action for the benefit and/or at the request of Human 2 or Institution 2>			
6	[[Human]] or [[Institution]] acts on [[Eventuality]] <Human or Institution behaves in accordance with Eventuality=Motivation>		X	X
7	[[Human]] or [[Institution]] acts under [[Rule]] or under [[Command]] <Human or Institution behaves in accordance with Rule or Command>			
8	[[Entity 1]] acts on or upon [[Entity 2]] <Entity1 has a particular effect on Entity 2>		X	X
9	[[Human]] acts (Role) (in Performance) <Human plays Role = Theatrical (in Performance)>			
10	Phrasal verb. [[Human]] acts [[Event]] or [[Human_Role]] or [[Emotion]] out <Human performs Role, not necessarily sincerely, or behaves as if feeling Emotion>			
11	[[Human]] acts [[ADJ]] <Human behaves in the manner specified by ADJ>			
12	[[Drug]] acts <Drug has an effect>			
13	[[Process]] acts <Process exerts an influence>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.10 *Answer*

An examination of the collocational behaviour of *answer* in the AJACX2 uncovers 12 PTs, all which are found in general English as represented by the PDEV. However, this overall figure gives a misleading impression of wide-ranging collocational behaviour, when in fact the situation varies a great deal across disciplines (see Figure 6.9). For example, in microbiology ($N = 41$) PT two accounts for all occurrences. In management ($N = 89$) only PTs one, two, four, and seven are found although, all but the second PT are limited to a handful of occurrences. Only PT four is absent from the history sub-corpus ($N = 145$) although, many of the patterns occur as single instances.

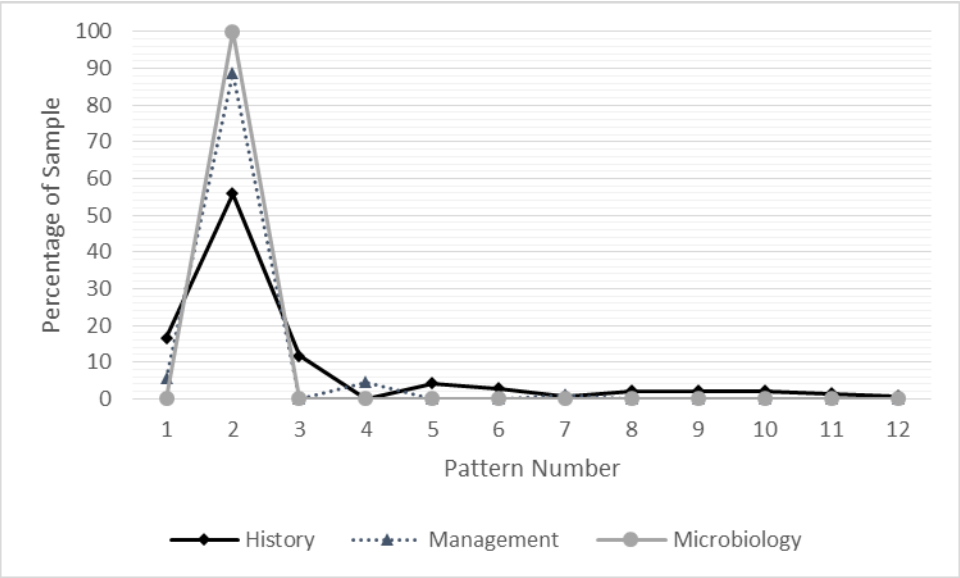

Figure 6.9 Interdisciplinary comparison for *answer*

The relatively low occurrences of most PTs might explain the results of the statistical comparison. To a certain extent these run counter to those considered thus far in the statistical analysis. While a comparison of microbiology and history reveals that PTs one, two, and three meet the threshold levels for significant difference, there are no PTs which meet the threshold level in a comparison of management and microbiology. In contrast to the impression given by previously analysed verbs PTs two, three, and four exhibit significant difference between history and management.

Table 6.9

Result of cross-corpora comparison for answer

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human]] answers quote or that [clause] <Human says that [clause] or [quote] in response to a question or statement by someone else>			X
2	[[Human 1]] or [[Institution]] or [[Document]] answers [[Question]] or [[Human 2]] or [[Speech_Act]] < Human 1 or Institution or Document says or writes Proposition that is intended to provide relevant information in response to Question or Speech_Act that has been asked by Human 2>	X		X
3	[[Human]] or [[Institution]] or [[Document]] answers [[Speech_Act]] <Human or Institution says something or writes something in Document intended to refute Speech_Act = Accusation>	X		X
4	[[Human]] answers {telephone} or {call} <Human accepts incoming call and speaks into telephone after it rings>	X		
5	[[Human]] answers [[Document]] <Human 1 writes an email or letter in response to Document = Mail from human 2>			
6	[[Human 1]] or [[Institution 1]] answers to [[Deity]] or to [[Human 2]] or to [[Institution 2]] <Human 1 has an obligation to account responsibly for his/her actions to Deity or Human 2 or Institution 2>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

Table 6.9 (cont.)

Result of cross-corpora comparison for answer

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
7	[[Human 1]] answers ([[Human 2]]) back Human 1 says something rude, impertinent, or forthright and unexpected in response to something that (Human 2) has said			
8	[[Artifact]] or [[Proposition]] answers need or purpose <Artifact or Proposition provides what is necessary for some purpose>			
9	[[Deity]] or [[Eventuality]] answers {prayer} <Eventuality desired by Human happens>			
10	[[Human]] answers for [[Eventuality]] <Human takes responsibility for Eventuality>			
11	[[Human 1]] answers {the description of} [[Human 2]] <Human 1 has the features listed in the description of Human 2>			
12	[[Human]] answers {calling} <Human undertakes activity in response to impulse or inclination>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.11 *Appear*

The verb *appear* is found highly frequently (N = 300 across all three sub-corpora). An examination of its collocational behaviour reveals 13 PTs, 11 of which occur in general English as represented by the PDEV. *Appear* exhibits its greatest collocational diversity in history with PTs six, 11, 12, and 13 occurring exclusively in this discipline. PT four occurs in history and microbiology and the remaining PTs occur across all three disciplines although as Figure 6.10 illustrates this is often at low frequency.

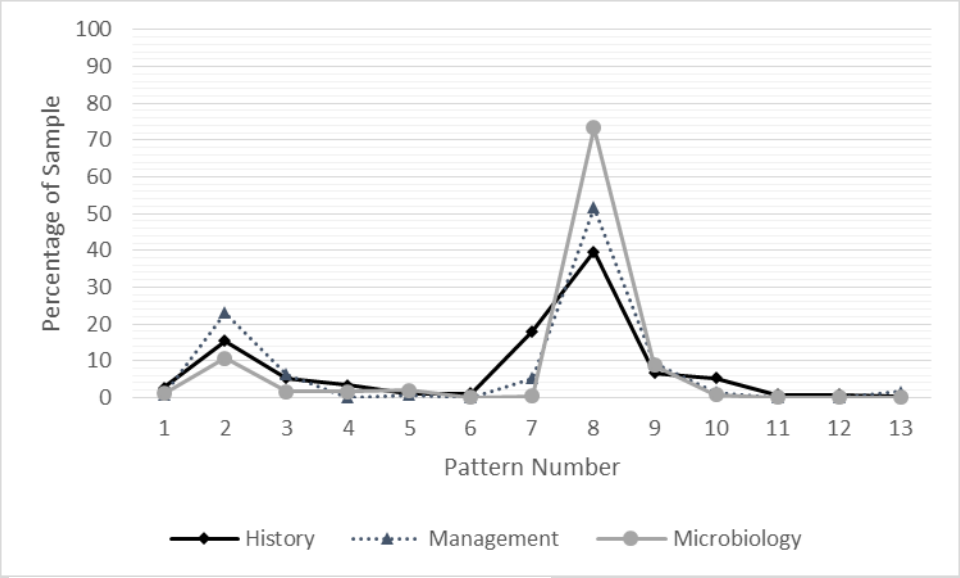

Figure 6.10 Interdisciplinary comparison for *appear*.

A statistical analysis (summarised in Table 6.10) reveals that PT seven occurs at significantly different levels in every discipline. Given its constituent semantic types,

[[Document]], [[Document Part]], and [[Image]], this may reflect the varying degree of intertextuality present in each discipline with history representing an extremely intertextual discipline (Buehl, 2017). In addition to PT seven, PTs two, three, and eight exhibit significant differences in a comparison of management and microbiology and eight and ten in a comparison of microbiology and history.

Table 6.10

Result of cross-corpora comparison for appear

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Physical Object] appears <Physical_Object becomes visible>			
2	[[Anything]] appears [ADJ] <Anything seems to be ADJ>		X	
3	[[Abstract Entity]] or [[State of Affairs]] appears <Abstract_Entity or State_of_Affairs becomes noticeable or comes into existence>		X	
4	[[Stuff]] OR [[Physical Object]] appears [[Time Period]] <Stuff or Physical_Object comes into use or becomes available Time_Period>			
5	[[Human]] OR [[Animal]] or [[MB Entity]] appears ([[Location]]) <Human or Animal arrives at Location>			
6	[[Human]] appears in [[Performance]] on [[TV Programme]] on [[Radio Programme]] or on [[Stage]] <Human takes part in Performance>			
7	[[Document 1]] OR [[Image]] appears (in [[Document 2]] or [[Document Part]]) <Document 1 or Image is published (in Document 2 or Document Part)>	X	X	X
8	[[Anything]] appears [To-Infinitive] <Anything seems to/inf [verb]>		X	X
9	{It} appears [[That Clause]] <that-[CLAUSE] is probably true>			
10	[[Entity 1]] or [[Eventuality 1]] appears (as) [[Entity 2]] or(as) [[Eventuality 2]] <Entity 1 or Eventuality 1 is perceived as or assumed to be Entity 2 or Eventuality 2>			X
11	[[Human]] appears in or before [[Institution]] or before [[Human]] <Law: court procedure. Human presents Self formally in Court or before Judge as a Defendant to be tried in respect of Accusation>			
12	[[Abstract Entity]] appear in [[Document]] <Document mentions abstract entity>			
13	[[Human]] appears as {an MP} <Human undertakes the role of MP>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.12 *Apply*

Analysis of the AJACX2 brings to light eight collocational PTs for the verb *apply* (*N* = 300 across all sub-corpora), five of which appear in general English as represented by the PDEV. Across all sub-corpora, *apply* occurs predominantly in the first PT. This is closely followed the second. PTs five and six occur across all disciplines albeit at fairly low frequencies. PT four, another notable PT, occurs only in microbiology and history. PT seven occurs only in history where it is found once. The sole instance of PT eight is found in microbiology.

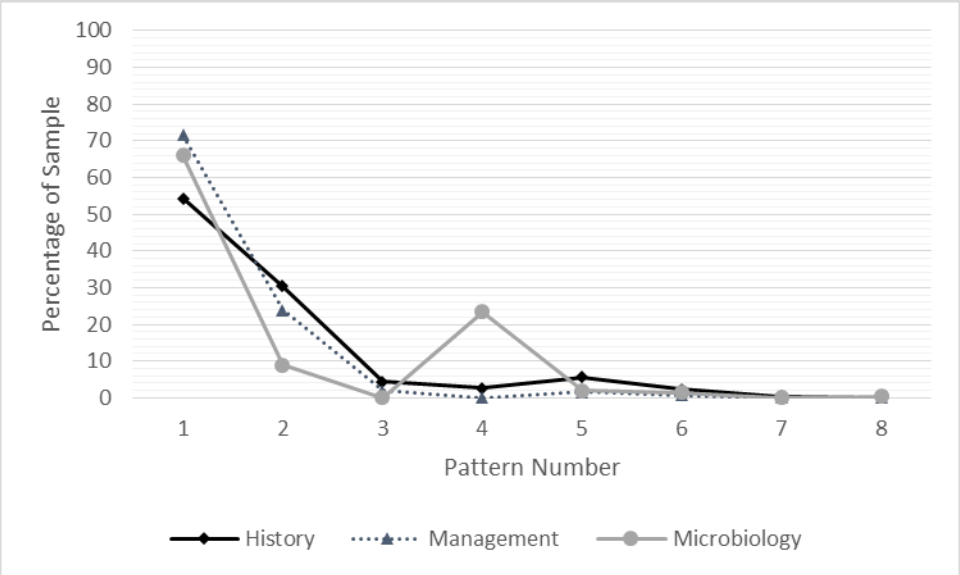

Figure 6.11 Interdisciplinary comparison for *apply*.

As far as significance testing is concerned (Table 6.11), no significant difference is apparent between disciplines for the frequently occurring PT one. This suggests that this pattern represents a general academic use of *apply* across all disciplines. For PT two, all comparisons with microbiology reveal significant difference. This again suggesting a split between the humanities on the one hand and science on the other. There is a significant difference between microbiology and history for PT three. For PT four which does not occur in management, all three pairwise comparisons result in significant differences. PT five which represents around two per cent of the sample in management and microbiology and around six per cent in history demonstrates no significant difference.

Table 6.11

Result of cross-corpora comparison for apply

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human]] or [[Institution]] applies [[Concept]] or [[Process]] (to [[State of Affairs]]) <Human or Institution makes use of Concept or Process in a particular situation or State_of_Affairs>			
2	[[Concept]] or [[Process]] applies (to [[State of Affairs]] or [[Activity]]) <Concept or Process is relevant to State_of_Affairs or Activity>		X	X

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

Table 6.11 (cont.)

Result of cross-corpora comparison for apply

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
3	[[Human]] applies for [[Benefit]] benefit or injunction or extension or admission or readmission or {[NP] to be [V]} <Human 1 formally asks Human 2 or Institution to give Benefit to Human 1>			X
4	[[Human]] OR [[Device]] applies [Fluid]] or [[Stuff]] to [[Surface]] <Human (uses Device to) spread Fluid or Stuff on Surface, typically in order to heal, improve, protect Surface or conduct test>	X	X	X
5	[[Human]] applies [[Word]] to [[Anything]] <Human uses Word to suitably refer to Anything>			
6	[[Human]] or [[Institution]] applies {pressure} <Human tries to influence decision or action of another human or institution>			
7	[[Human]] applies [[Self]] to [[Activity]] <Human makes great effort to conduct activity>			
8	[[Human]] applies {caution} <Human take a cautious approach to [[Anything]]>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.13 *Call*

An examination of the corpus reveals 20 PTs for the verb *call* (N = 300 in all disciplines) making it the most collocationally prolific item examined thus far, 18 of these PTs are found in general English as represented by the PDEV. As Figure 6.12 illustrates, once again, the greatest variety of PTs can be found in history, with all 19 PTs occurring. Microbiology is the least diverse discipline in terms of PTs with six types occurring. There are ten PTs found in management. In history and management, patterns one, two, and four predominate, with pattern ten also occurring regularly, while in microbiology the vast majority of instances of *call* appear in PT one.

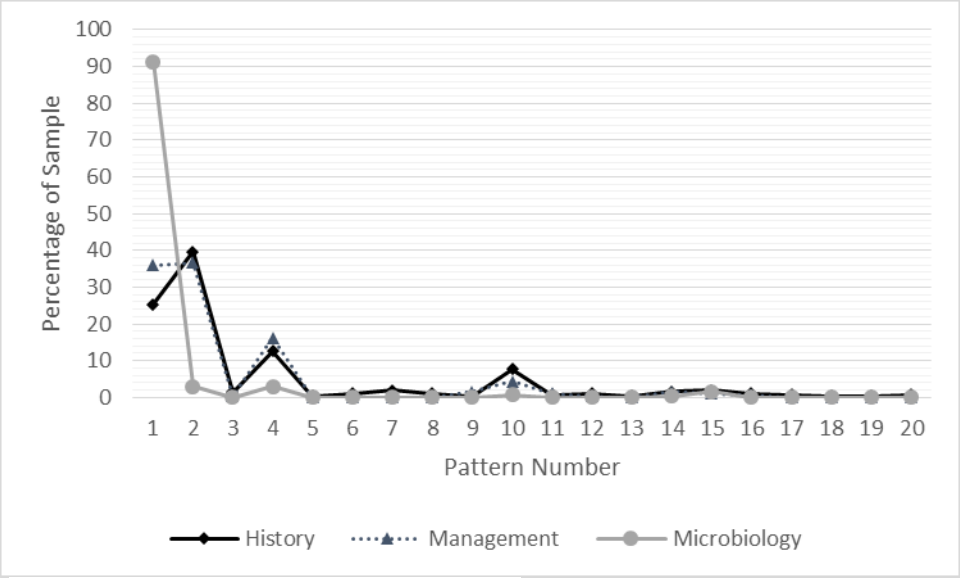

Figure 6.12 Interdisciplinary comparison for *call*.

The impression of a division with history and management on one side, and microbiology on the other is confirmed by the statistical comparison summarised in Table 6.12. A significant difference is apparent for PTs one, two, four, and ten in comparisons with microbiology.

Table 6.12
Result of cross-corpora comparison for call

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Anything]] is called [[Name]] <Anything is referred to as Name>		X	X
2	[[Human]] OR [[Institution]] calls [[Anything]] [[Name]] <Human or Institution invents or uses Name to refer to Anything>		X	X
3	[[Human]] or [[Institution]] calls [[Event]] <Human or Institution instructs people to cause Event = Meeting or Action to happen immediately>			
4	[[Human]] or [[Institution]] or [[Document]] calls {for} [[Action]] or {for} [[State of Affairs]] <Human or Institution or Document says that other people should do Action or create [State_of_Affairs]>		X	X
5	[[Human]] or [[Institution]] calls [[Human 2]]{for} [[Activity]] <Human 1 or Institution officially invites or instructs Human 2 to take part in Activity>			
6	[[Human 1]] calls ((in or round) (on [[Human 2]]) at [[Location]]) <Human 1 goes to Location in order to meet Human 2>			
7	[[Human 1]] calls [[Human 2]] <Human 1 shouts to Human 2, typically in order to ask Human 2 to come to them>			
8	[[Human]] calls [[Speech_Act]] (out) <Human says Speech_Act or QUOTE in a loud, clear voice>			
9	[[Human 1]] calls [[Human 2]] or [[Institution]] or [[Number]] <Human 1 contacts Human 2 or Institution on the telephone by selecting Number = Telephone Number>			
10	[[Human 1]] or [[Institution 2]] or [[Document]] calls {on} or {upon} [[Human 2]] OR [Institution 2]] [TO-INFINITIVE] <Human 1 or Institution 1 or Document asks Human 2 or Institution 2 to [verb]>		X	X
11	[[Human 1]] or [[Institution 1]] calls [[Human 2]] or [[Institution 2]] {in} <Human 1 or Institution 1 asks Human 2 or Institution 2 to come and give help or advice>			
12	[[Human]] calls [[Activity]] {off} <Human cancels Activity>			
13	[[Human]] calls [[Information]] or [[Image]] {up} <Human causes Information or Image to be displayed on a computer screen>			
14	[[Human]] calls {attention to} [[Event]] <Human asks people to notice Event>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

Table 6.12 (cont.)
Result of cross-corpora comparison for call

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
15	[[Human]] or [[Anything] calls [[Proposition]] {into question} <Human or Anything provides a reason for doubting whether Proposition is correct>			
16	[[Human 1]] calls {upon} [[Human 2]] <Human 1 asks Human 2 for assistance>			
17	[[Human 1]] calls [[Anything]] {down upon} [[Human 2]] <Human evokes Anything to affect Human 2>			
18	[[Human]] is called to {celibacy} <Human chooses to remain celibate>			
19	[[Concept]] called into {being} <Human creates concept to explain or deal with Eventuality>			
20	[[Institution]] calls [[Human]] {up} <Institution = Government requires Human to undertake military service>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.14 *Conduct*

An examination of the collocational behaviour of *conduct* in the AJACX2 reveals six PTs, four of which have equivalents in general English as represented by the PDEV. However, *conduct* is only behind *accomplish* in terms of its conservative collocational behaviour. PT one: [[Human]] or [[Institution]] *conducts* [[Activity]], accounts for 94 per cent of instances in history (N = 150), 99.33 per cent in management (N = 300), and 98 per cent in microbiology (N = 150). The next most prevalent PT is PT three: [[Human]] *conducts* [[Self]] [Manner] which occurs five times in history and twice in management. For all PTs significance testing failed to reject the H_0 of no significant difference between frequencies of occurrence across disciplines.

6.1.15 *Explain*

An analysis of the AJACX2 reveals five PTs for the verb *explain*, all of which are found in general English as represented by the PDEV. The first PT is the most prevalent across all three disciplines (N = 300) closely followed by the second. PT five is limited to history and microbiology all other PTs occur across all disciplines albeit representing different proportions of the sample in each.

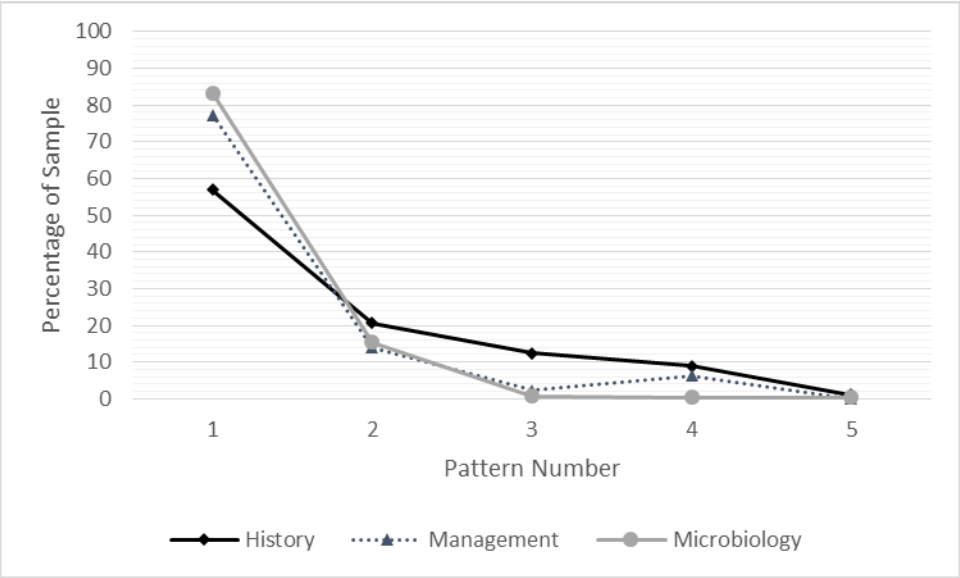

Figure 6.13. Interdisciplinary comparison for *explain*.

In contrast to most of the verbs studied thus far the chart in Figure 6.13 indicates a greater distinction between history on the one hand, and microbiology and management on the other. This distinction is supported by significance testing (summarised in Table 6.13) with PTs one and three meeting the threshold for significant difference in both comparisons with history. PT four demonstrates significant difference between microbiology and the other two disciplines.

Table 6.13

Result of cross-corpora comparison for explain

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human]] or [[Institution]] or [[Document]] or [[Proposition]] or [[Eventuality 1]] explains [[Eventuality 2]] (in terms of [[Eventuality 3]]) (to [[Human 2]]) <Human or Institution formulates (in conversation with Human 2 or in a Document intended for Human 2 to read) a Proposition concerning the cause or effect of an Eventuality 2>	X		X
2	[[Human]] or [[Institution]] or [[Document]] or [[Proposition]] explains (to [[Human 2]]) <Human or Institution formulates Proposition in Document concerning the cause or effect of an Eventuality that is expressed in wh- [clause]>			
3	[[Human]] or [[Institution]] explains [[That-Clause]] (to [[Human 2]]) <Human 1 or Document tells Human 2) that [clause], as a way of accounting for Eventuality (often implied rather than stated explicitly)>	X		X
4	[[Human]] or [[Institution]] explains [[Quote]] <Human 1 or Document tells Human 2) that [clause], as a way of accounting for Eventuality (often implied rather than stated explicitly)>		X	X
5	[[Human]] or [[Institution]] explains [[Eventuality]] {away} <Human 1 formulates Proposition as a way of accounting for Eventuality = Inconvenient>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.16 *Fail*

Nine PTs were uncovered for *fail* in the analysis of the corpus (*N* = 150 in all disciplines), all of which have equivalents in general English as represented by the PDEV. At first glance, patterns one to four represent major PTs overall while pattern six is particularly prevalent in microbiology.

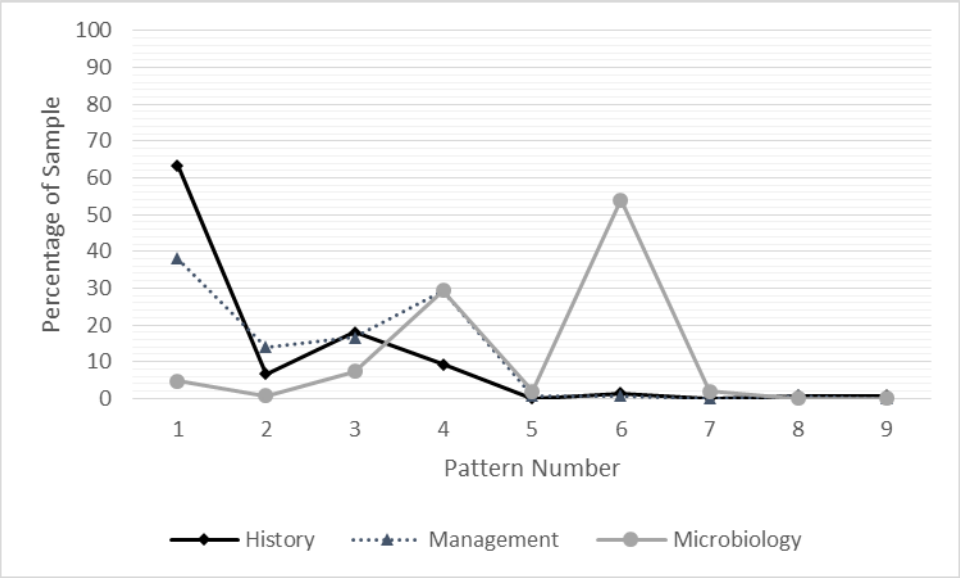

Figure 6.14. Interdisciplinary comparison for *fail.*

Significance testing reveals that there are significant differences in the frequency of occurrence of PT one across all three disciplines (Table 6.14). PT two occurs significantly less frequently in microbiology than the other disciplines. PT four is significantly less frequent in history than management and microbiology. PT six is clearly significantly more frequent in microbiology than in the other disciplines.

Table 6.14
Result of cross-corpora comparison for fail

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human]] or [[Institution]] fails [[to-infinitive]] <Human or Institution does not do something that they intended to do or were expected to do>	X	X	X
2	[[Human]] or [[Institution]] fails (in[[Activity]]) <Human or Institution attempts to do Activity but does not produce the expected or intended result>		X	X
3	[[Activity]] or [[System]] fails <Activity or System does not produce the expected or intended result>			
4	[[Activity]] or [[System]] or [[Abstract_Entity]] fails [[to-infinitive]] <Activity or System or Abstract_Entity does not produce the expected or intended result>	X		X
5	[[Artifact]] or [[Plant]] or [[Body_Part]] or [[MB Unit]] fails <Artifact or Body_Part or Plant does not function effectively>			

Table 6.14 (cont.)
Result of cross-corpora comparison for fail

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
6	[[Artifact]] or [[Body_Part]] or [[MB Entity]] or [[Plant]] fails [[to-infinitive]] <Artifact or Body_Part or Plant does not work or perform in the way that it is intended or expected to>		X	X
7	[[Human 1]] fails [[Activity]] <Human does not meet the standard required to be successful in Activity = Test>			
8	[[Human 1]] or [[System]] or [[Institution 1]] fails [[Human 2]] or [[Institution 2]] <Human 1 or System or Institution 1 fails Human 2 or Institution 2>			
9	[[Weather_Event]] fails <Weather_Event = Desirable does not occur in the way that is expected or hoped for>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.17 *Follow*

The verb *follow* is prolific in its collocational behaviour. An examination of the corpus (*N* = 300 across all disciplines) revealed 19 PTs, 16 of which are found in general English as represented by the PDEV. That said, as indicated by Figure 6.15 below, six PTs: two, four, nine, 12 and 14 are particularly prevalent.

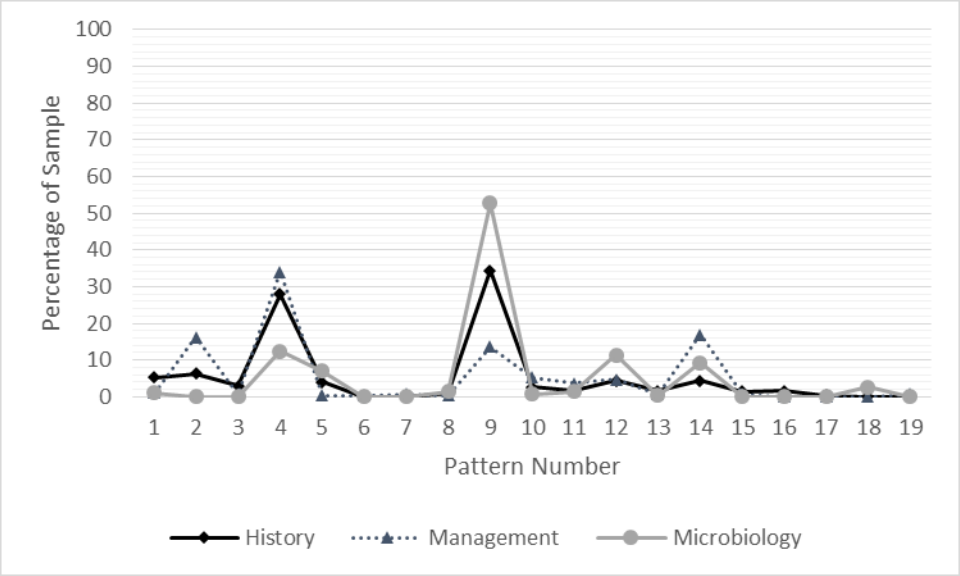

Figure 6.15 Interdisciplinary comparison for *follow*

The differences in PTs two and nine for all disciplines are significant (Table 6.15). PTs, four, 12 and 18 demonstrate significant difference in comparisons with microbiology. Although, it must be pointed out that PT 18: [[MB Entity]] follows [[Route]] is unique to microbiology. PT one occurs significantly more frequently in history than in other disciplines. PT five occurs significantly less frequently in management than in other disciplines. PTs three and 15 occur at a significantly higher frequency in history than microbiology. PT ten occurs at a significantly higher frequency in management than microbiology.

Table 6.15

Result of cross-corpora comparison for follow

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human 1]] or [[Animal 1]] or [[Vehicle 1]] or [[MB Entity 1]] follows [[Human 2]] or [[Animal 2]] or [[Vehicle 2]] or [[MB Entity 2]] <Human 1 or Animal 1 or Vehicle 1 moves in the same direction as that selected by Human 2 or Animal 2 or the driver of Vehicle 2>	X		X
2	[[Human 1]] follows [[Human 2]] <Human 1 = Student or Disciple studies and is influenced by or tries to practice the teachings of Human 2 = Teacher>	X	X	X
3	[[Human_Group 1]] or [[Institution 1]] follows [[Human_Group 2]] or [[Institution 2]] <Human_Group 1 or Institution 1 is influenced by and thus copies the behaviour of Human Group 2 or Institution 2>			X
4	[[Human]] or [[Institution]] or [[Concept]] or [[Process]]follows [[Command]] or [[Rule]] or [[Plan]] or [[Document]] <Human or Institution acts in accordance with Command or Rule or Plan (expressed in Document)>		X	X
5	[[Human]] or [[Institution]] follows [[Event]] <Human or Institution takes an interest in Event = Unfolding>	X	X	
6	[[Human]] follows [[Proposition]] or [[wh-clause]] <Human is able to understand Proposition or wh-clause>			
7	[[Human]] or [[Institution]] follows {lead} or {line of enquiry} <Human or Institution pursues an investigation based on specific information>			
8	[[Human]] or [[Institution]] or [[Event]] follows {trend} <Human or Institution performs in line with the expectations dictated by trend >			
9	[[Event 1]] follows (Event 2) <Event 1 happens after and typically as a consequence of Event 2>	X	X	X
10	[[Eventuality 1]] follows from [[Eventuality 2]] <Eventuality 1 is a necessary consequence of Eventuality 2>		X	
11	{it} follows [[that-clause]] (from Eventuality 1) <Eventuality 2 expressed by [CLAUSE] is a necessary consequence of Eventuality 1 (either expressed in a from-phrase or one that was previously stated)>			
12	[[Anything 1]] follows [[Anything 2]] <Anything 1 = Item in List comes after Anything 2 = Item in List in a list or sequence>		X	X
13	[[Human]] follows {up} [[Eventuality]] <Human monitors Eventuality and finds out more about it>			
14	as follows <as will now be stated here>	X		
15	[[Human 1]] or [[Institution 1]] follows {suit} <Human 1 or Institution 1 does the same as Human 2 or Institution 2 has done>			X
16	[[Human 1]] follows in footsteps {of [[Human 2]]} <Human 1 engages in similar activity to one previously engaged in by Human 2>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

Table 6.15 (cont.)
Result of cross-corpora comparison for follow

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
17	[[Human]] follows up with [[Speech Act]] <Human adds additional details to information already given>			
18	[[MB Entity]] follows [[Route]] <MB Entity follows an established route or conduit>		X	X
19	[[Human]] follows {through on commitment}<human fulfils a commitment>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.18 *Handle*

An analysis of the collocational behaviour of *handle* reveals seven PTs in the corpus, six of which are present in general English as represented by the PDEV. However, as Figure 6.16 shows that only the first three PTs found occur across all disciplines. PT four does not occur in microbiology (N = 21). PTs five and six are exclusive to history (N = 74) where they occur at extremely low frequencies. There is only one instance of pattern seven this occurs in microbiology.

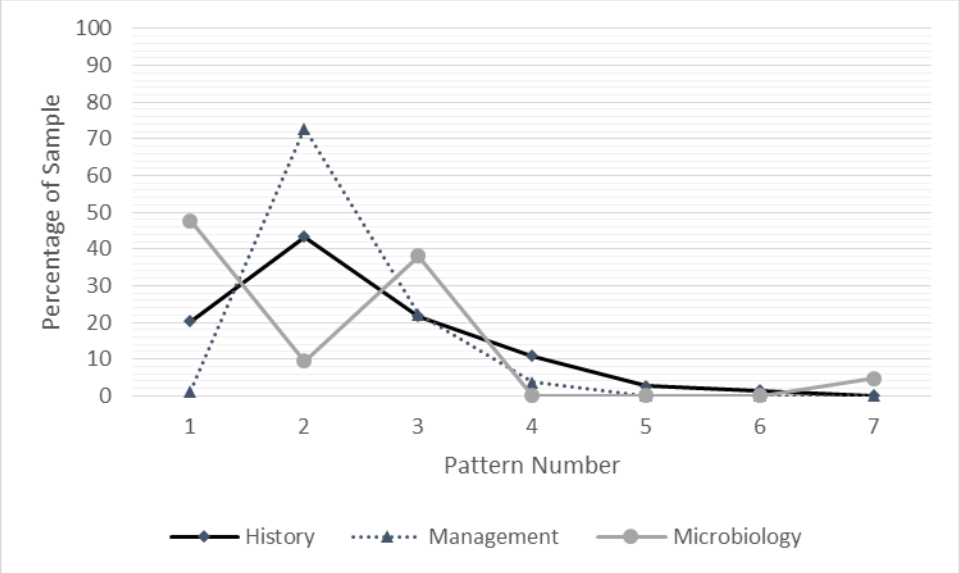

Figure 6.16 Interdisciplinary comparison for *handle*.

The significance testing summarised in Table 6.16 reveals difference between history and management (N = 81) and management and microbiology for PT one and between all disciplines for PT two.

Table 6.16

Result of cross-corpora comparison for handle

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human]] handles [[Physical_Object]] <Human uses his/her hands to hold, manipulate, and/or use Physical_Object>	X	X	
2	[[Human]] or [[Institution]] or [[Computer]] handles [[Eventuality]] (Manner) <Human or Institution or Computer uses intelligence or an established procedure to deal with Eventuality as required or in a particular manner>	X	X	X
3	[[Human]] or [[Institution]] or[[Device]] or [[MB Entity]] handles [[Entity]] or [[Stuff]] <Human or Institution or Device processes or deals with Entity or Stuff>			
4	[[Human 1]] handles [[Human 2]] or [[Human_Group]] (Manner) <Human 1 manages or copes with Human 2 or Human_Group (in a particular manner)>			
5	[[Human]] handles [[Artifact]] <Human receives Artifact = that should not legally have in his/her possession>			
6	[[Human]] handles [[Self]] (manner) <Human behaves in a particular manner>			
7	[[Human]] handles [[Animal]] <Human cares for and commands animal>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.19 *Lead*

An examination of the corpus brings to light 11 PTs for *lead* (*N* = 300 across all disciplines), all of which are found in general English as represented by the PDEV. While all 11 are found in history, only PTs one, three, and five are found in microbiology. PT one is the by far the most prevalent across all disciplines.

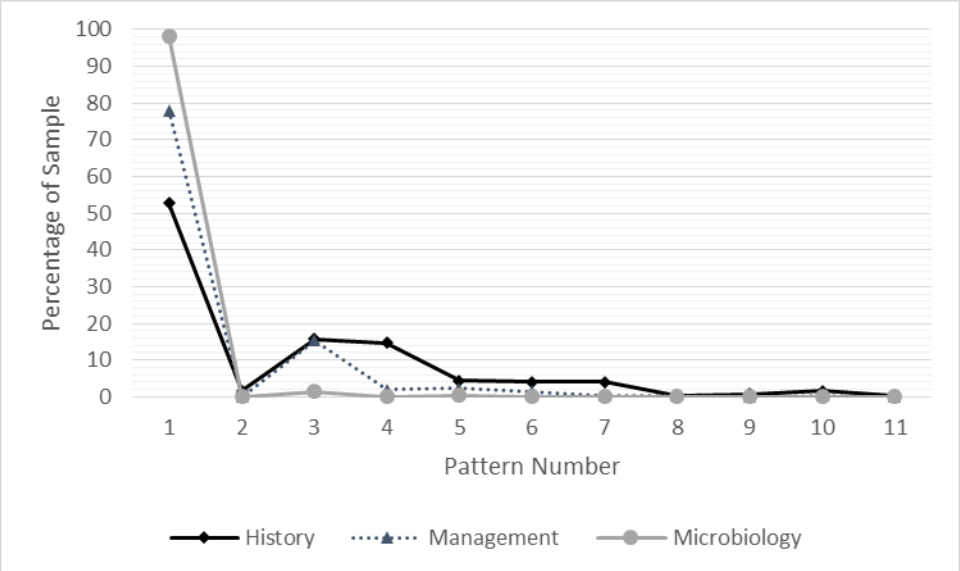

Figure 6.17 Interdisciplinary comparison for *lead.*

The absence of instances of most PTs in microbiology is reflected in significance testing with six PTs, types one and three to seven, reaching the threshold level for significant difference in a comparison of microbiology and history. PTs one, four, and seven also meet the significance threshold in a comparison of history and management. While a comparison of management and microbiology reveals a significant difference for PT three.

Table 6.17

Result of cross-corpora comparison for lead

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Eventuality 1]] leads to [[Eventuality 2]] <Eventuality 1 is the cause of Eventuality 2>	X		X
2	[[Eventuality 1]] leads {up to} [[Eventuality 2]] <Eventuality 1 = Plural precedes but may not be the cause of Eventuality 2>			
3	[[Eventuality]] leads [[Human]] or [[Institution]] [[to-infinitive]] <Eventuality causes, enables, or encourages Human or Institution to/INF [verb]>		X	X
4	[[Human]] or [[Institution 1]] leads [[Human_Group]] or [[Institution 2]] <Human or Institution 1 organizes, directs, or provides a model for Activity of Human_Group or Institution 2>	X		X
5	[[Human]] or [[Institution]] leads [[Activity]] <Human or Institution organizes, directs, or provides a model for the Activity of Human_Group>			X
6	[[Human 1]] leads [[Human 2]] or [[Human_Group]] (Direction to Location) <Human 1 organizes and directs the movement of Human 2 or Human_Group to Location by accompanying Human 2 or Human Group and showing the way>			X
7	[[Route]] leads [[Direction]] to [[Location]] <Route is a path or road Direction or to Location>	X		X
8	[[Human]] or [[Human_Group]] leads (Activity) <Human = Competitor or Human_Group = Competitor is in first place in Activity = Competition or Race)>			
9	[[Human]] or [[Institution]] leads the way(in-ING) <Human or Institution is the best at doing something and shows how it can best be done>			
10	[[Human]] or [[Animate]] leads MOD {life} or {existence} <Human or Institution is the best at doing something and shows how it can best be done>			
11	[[Human 1]] or [[Eventuality]] leads [[Human 2]] on <Human 1 or Eventuality encourages Human 2, typically to do or believe something stupid or reprehensible>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.20 *Maintain*

The corpus pattern analysis revealed six PTs for *maintain* (*N* = 300 across all disciplines), all PTs, except PT six, are present in general English as represented by the PDEV. The first PT represents a high proportion of the sample across all disciplines.

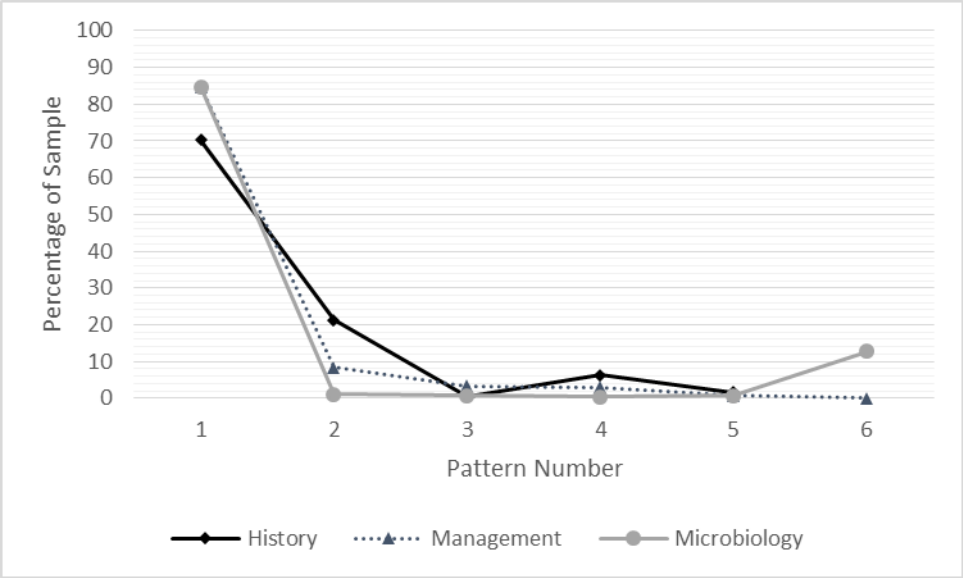

Figure 6.18 Interdisciplinary comparison for *maintain*.

The results of significance summarised in Table 6.18 reveal that there is no significant difference in frequency of occurrence between disciplines for PT one. All three pairwise comparisons indicate significant difference for PT two. PT three occurs significantly more frequently in management than in history. PT four occurs significantly more frequently in history than microbiology. The sematic type [[MB Entity]] found in subject position in PT six suggests that this pattern is exclusive to this discipline. Testing confirms that the occurrence of PT six in microbiology and its absence from the other disciplines is significant from statistical standpoint.

Table 6.18

Result of cross-corpora comparison for maintain

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human]] or [[Process]] maintains [[State_of_Affairs]] <Human takes action to ensure or Process has the effect that State_of_Affairs remains unchanged>			
2	[[Human]] or [[Institution]] or [[Document]] maintains [[that-clause]] <Human or Institution or Document asserts strongly and consistently over time that [clause] or [Proposition]>	X	X	X
3	[[Human]] or [[Institution]] maintains [[Document]] <Human or Institution regularly and systematically updates Document>	X		
4	[[Human]] maintains [[Building]] or [[Vehicle]] or [[Artifact]] <Human takes action to ensure that Building or Vehicle or Artifact is kept in good working order>			X

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

Table 6.18 (cont.)

Result of cross-corpora comparison for maintain

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
5	[[[Human 1]] maintains [[Self]] or [[Human 2]] <Human 1 provides food and drink for Self or Human 2, enabling Self or Human 2 to stay alive>			
6	[[MB_Entity]] is maintained in [[Location]] or on [[Surface]] <MB Entity is place in Location or on Surface for the purpose of a test>		X	X

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.21 *Manage*

Examination of the corpus reveals nine PTs for *manage*, eight of which have equivalents in general English as represented by the PDEV. As Figure 6.19 suggests, the prevalent PT varies considerably across disciplines. PT one represents the prevalent PT for history (*N* = 150), type six for management (*N* = 300) and type nine for microbiology (*N* = 22).

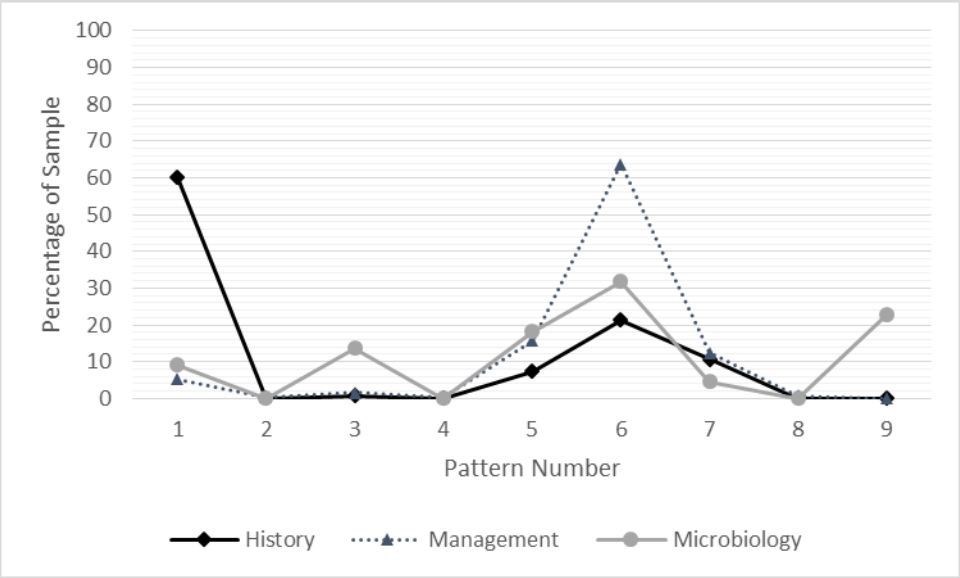

Figure 6.19 Interdisciplinary comparison for *manage*.

The results of significance testing (summarised in Table 6.19) support this initial impression. All comparisons with history demonstrate significant differences for PT one. All comparisons with microbiology demonstrate significant differences for PT nine. There is also significant difference between history and microbiology for PT three and between history and management for PT six.

Table 6.19

Result of cross-corpora comparison for manage

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human]] or [[Institution]] or [[MB Entity]] manages [to-infinitive] <Human or Institution is able to/inf [verb], despite difficult circumstances>	X		X
2	[[Human]] manages <Human is able to deal with difficult circumstances successfully>			
3	[[Human]] or [[Institution]] [[MB Entity]] manages [[Action]] or [[State_of_Affairs]] <Human or Institution is able to perform Action = Desirable or achieve State_of_Affairs = Desirable, despite difficult circumstances>			X
4	[[Human]] manages [[Artifact]] <Human is able to lift, carry, or use Artifact without difficulty>			
5	[[Human]] or [[Institution 1]] manages [[Human 2]] or [[Institution 2]] <Human or Institution 1 is responsible for and controls Institution 2>			
6	[[Human]] or [[Institution]] or [[MB Entity]] manages [[System]] or [[Activity]] <Human or Institution is responsible for and controls System or Activity = Ongoing>	X		
7	[[Human]] or [[Institution]] manages [[Resource]] <Human or Institution is responsible for Resource>			
8	[[Human]] or [[Institution]] manages [[Location]] <Human or Institution is responsible for the use and maintenance of Location>			
9	[[Human]] manages [[Disease]] <Human alleviates symptoms of Disease with drugs or Therapy>		X	X

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.22 *Need*

Corpus pattern analysis reveals five PTs for *need* (*N* = 300 across all disciplines), all of which are found in general English as represented by the PDEV. As Figure 6.20 illustrates, with the exception of PT five which is quite infrequent and does not occur in management, these patterns occur at reasonably high frequency across all sub-corpora.

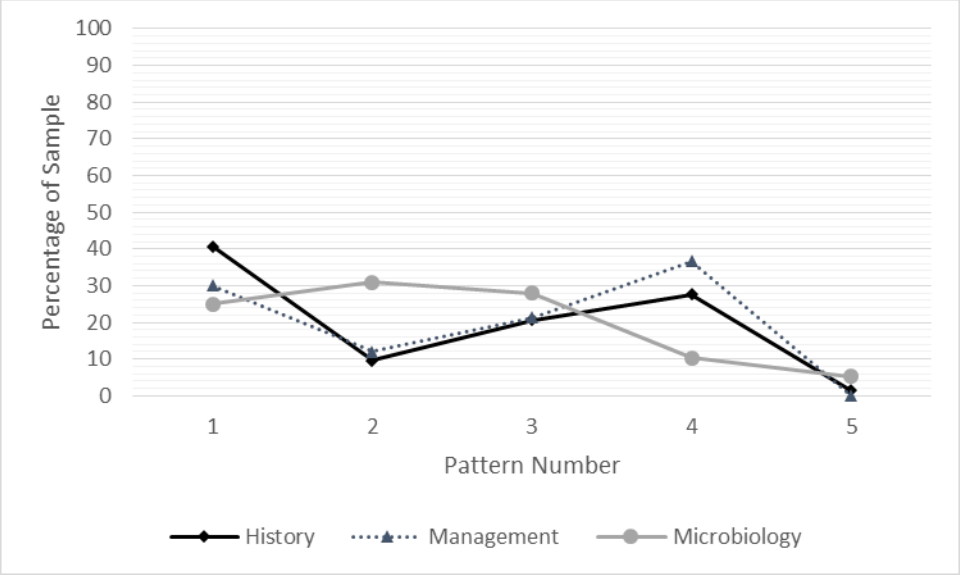

Figure 6.20 Interdisciplinary comparison for *need.*

Once again significance testing (summarised in Table 6.20) supports the impression of a division between microbiology and the other disciplines. All comparisons with microbiology show significant difference for PTs two and four. PT one is also significantly more frequent in history than microbiology and PT five significantly less frequent in management than microbiology.

Table 6.20

Result of cross-corpora comparison for need

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human]] or [[Institution]] needs [[Eventuality]] or [[Entity]] <Human or Institution requires that Eventuality or Entity must be realized or available, in order to accomplish some goal>			X
2	[[Entity 1]] or [[Eventuality 1]] needs [[Entity 2]] or [[Eventuality 2]] <Entity 1 or Eventuality 1 is an essential precondition for or attribute of Entity 2 or Eventuality 2>		X	X
3	[[Entity]] or [[Eventuality]] needs [[to-infinitive]] <An essential precondition for the realization of Eventuality is that [verb] must be realized typically to be>			
4	[[Human]] or needs [[to-infinitive]] <Human must do [verb]>		X	X
5	[[Plant]] or [[Animate]] or [[MB Entity]] needs [[Eventuality]] or [[Stuff]] <Plant or Animate must have Eventuality or Stuff in order to survive and flourish>		X	

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.23 *Note*

A corpus pattern analysis for *note* reveals four PTs (*N* = 300 in all disciplines), all except PT4 are found in general English as represented by the PDEV. Figure 6.21 shows PTs as a proportion of the sample from each discipline. The first two PTs occur at reasonably high frequencies across all three sub-corpora. The third is much less frequent particularly in microbiology. The final pattern is exclusive to microbiology where it occurs only five times.

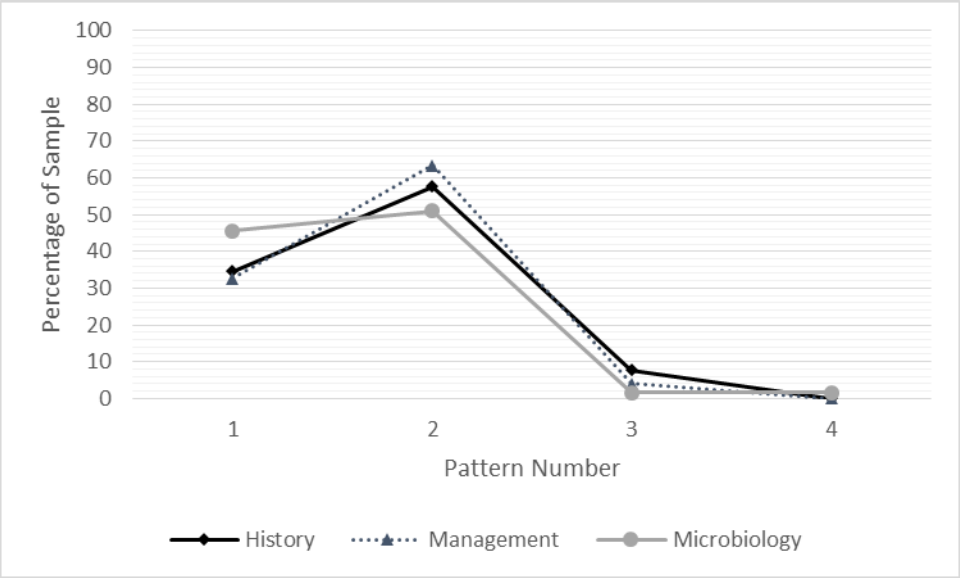

Figure 6.21 Interdisciplinary comparison for *note.*

As indicated by Table 6.21, the only difference in frequency of occurrence across sub-corpora which met the threshold level occurs in a comparison between microbiology and history for PT three.

Table 6.21

Result of cross-corpora comparison for note

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human]] notes [[Information]] or [[Eventuality]] or [[Visible_Feature]] <Human notices and pays particular attention to Information about Eventuality or Visible_Feature>			
2	[[Human]] notes [[quote]] or [[that-clause]] or [[wh-clause]] <Human notices and pays particular attention to Eventuality or Visible_Feature>			
3	[[Human]] notes [[Information]] (on or in [[Document]]) (under [[Document_Part]]) <Human makes a written note of Information (on | in Document) (under Document_Part = Heading)>			X
4	[[MB Entity]] is noted [[Name]] <MB Entity is referred to as Name>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.24 *Overcome*

In terms of collocational behaviour, of the verbs examined thus far only *conduct* is more conservative than *overcome*. An analysis of the AJACX2 reveals only three patterns the latter two of which are absent from management (N = 150 total patterns) and microbiology (N = 112 total patterns), occurring exclusively in history (N = 138 total patterns).

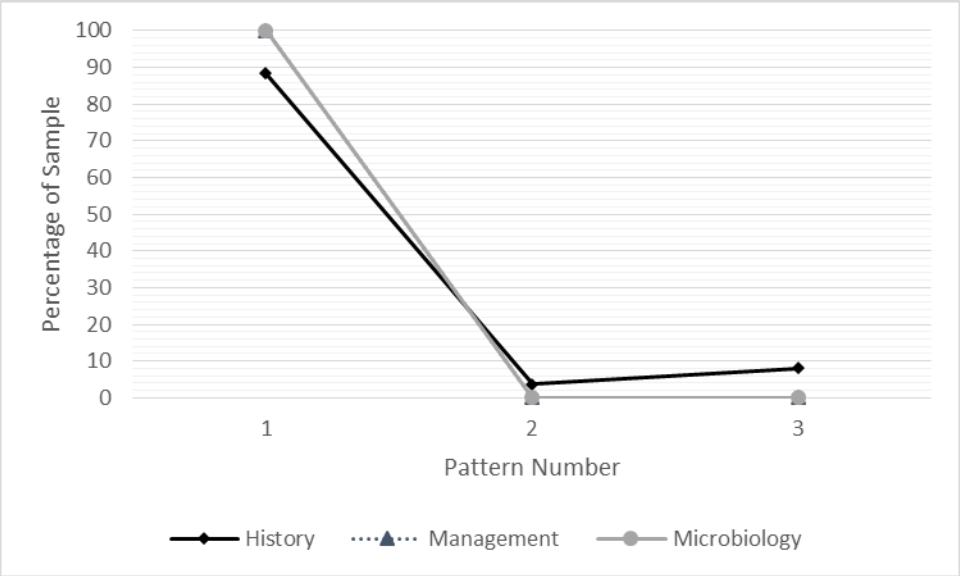

Figure 6.22 Interdisciplinary comparison for *overcome*.

As shown by Table 6.22, the only significant difference occurs in the case of PT three. The presence of this PT in history and its absence in management proves significant from a statistical standpoint.

Table 6.22
Result of cross-corpora comparison for overcome

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human]] or [[Institution]] or [[Activity]] or [[MB Entity]] overcomes [[Eventuality]] <Human or Institution or Activity successfully deals with Eventuality = Problem>			
2	[[Eventuality]] overcomes [[Human]] <Human fails to deal successfully with Eventuality = Problem>			
3	[[Human 1]] or [[Human_Group 1]] overcomes [[Human 2]] or [[Human_Group 2]] <Human 1 or Human_Group 1 defeats Human 2 or Human Group 2 in a fight, war, or contest>	X		

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.25 *Prescribe*

The verb *prescribe* is even more conservative in its collocational behaviour than *overcome*. Overall its frequency occurrence is low: in history N = 26, management N = 41, and in microbiology N = 22 only slightly above the cut-off threshold of 20 required for inclusion in the sample of verbs examined in this experiment. A corpus pattern analysis reveals only two PTs both of which are also present in PDEV. The first [[Human]] *prescribes* [[Drug]] or [[Activity]] ((for Human 2) (for Illness)) a represents reasonably high proportion of occurrences across all sub-corpora. The second pattern [[Human]] or [[Institution]] or [[Rule]] *prescribes* [[State of Affairs]] or [[Activity]] is far more prevalent in history and management than microbiology.

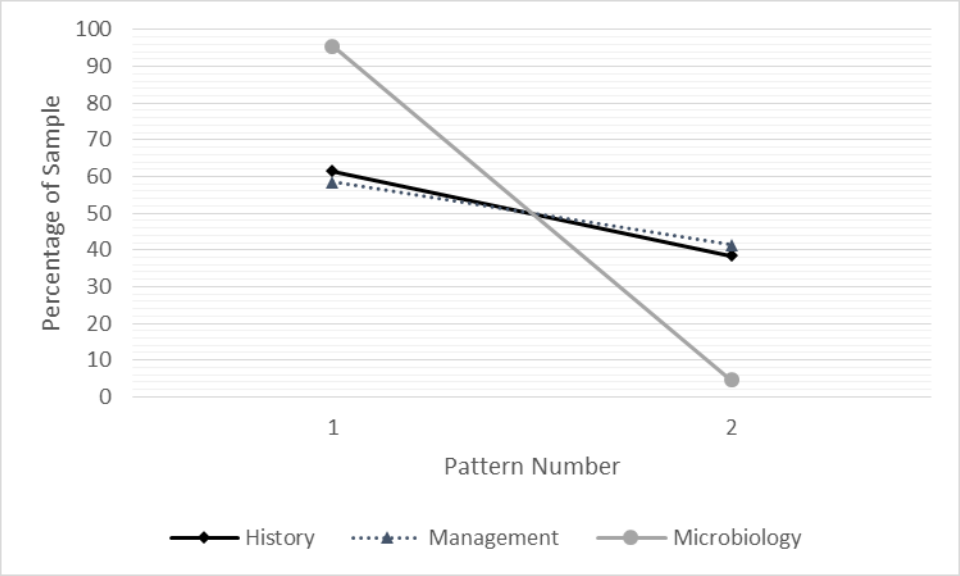

Figure 6.23 Interdisciplinary comparison for *prescribe*.

As indicated by Table 6.23, testing confirms that the difference of the proportion of sample represented by PT2 in comparison with microbiology is significant.

Table 6.23

Result of cross-corpora comparison for prescribe

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human]] prescribes [[Drug]] or [[Activity]] ((for Human 2)(for Illness)) <Human = Medical Professional advises the use of Drug or Activity = Treatment to treat Human 2 = Patient) (for Illness))>			
2	[[Human]] or [[Institution]] or [[Rule]] prescribes [[State_of_Affairs]] or [[Activity]] <Formal. Human or Institution or Rule authoritatively sets out the terms and conditions necessary for State_of_Affairs or Activity to occur>		X	X

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.26 *Propose*

An examination of the corpus reveals six PTs for *propose* (N = 150 in all sub-corpora), all of which can be found in general English as represented by the PDEV. As Figure 6.24 shows only PTs one, three, and four occur across all disciplinary sub-corpora. The remaining two are restricted to history where they are limited to a handful of instances.

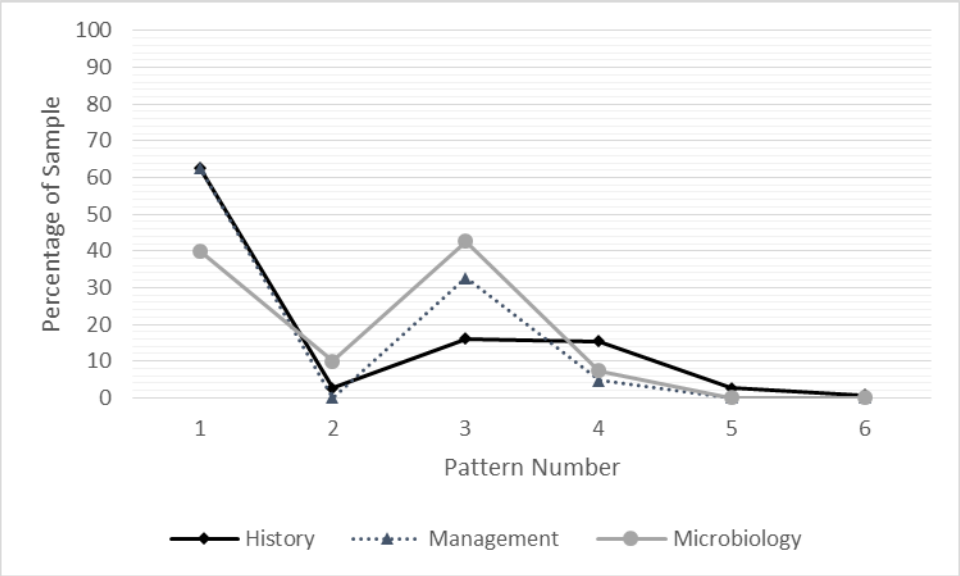

Figure 6.24 Interdisciplinary comparison for *propose*.

There is significant difference between management and microbiology for PT two, in both comparisons with history for PT three, and between history and management for PT four.

Table 6.24

Result of cross-corpora comparison for propose

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human]] or [[Institution]] or [[Document]] proposes [[Action]] or [[Plan]] or [[Proposition]] <Human or Institution suggests, for consideration by others, Plan or Proposition as a basis for Action>			
2	[[Human]] or [[Institution]] proposes [[Entity]] <Human or Institution suggests, as a plan for consideration by others, that Entity should be created>		X	
3	[[Human]] or [[Document]] proposes [[that-clause]] <Human suggests [CLAUSE] as a proposition for consideration by others>	X		X
4	[[Human]] or [[Institution]] or [[Document]] proposes [[to-infinitive]] <Human or Institution intends to [verb]>	X		
5	[[Human 1]] proposes [[Human 2]] for or as [[Human_Role]] <Human 1 formally suggests that Human 2 should be elected or appointed to undertake Human_Role>			
6	[[Human 1]] proposes {marriage} <Human 1 asks Human 2 to marry him or her>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.27 *Replicate*

Six PTs were found in an analysis of the collocational behaviour of *replicate* in the AJACX2 corpus. With the exception of PT six, all of these can be found in general English as represented by the PDEV. As shown by Figure 6.25, no single PT predominates in all disciplines. Notable PTs include: two which is prevalent in history (N = 30) and management (N = 144), three which is prevalent in management, and four which dominates microbiology (N = 150).

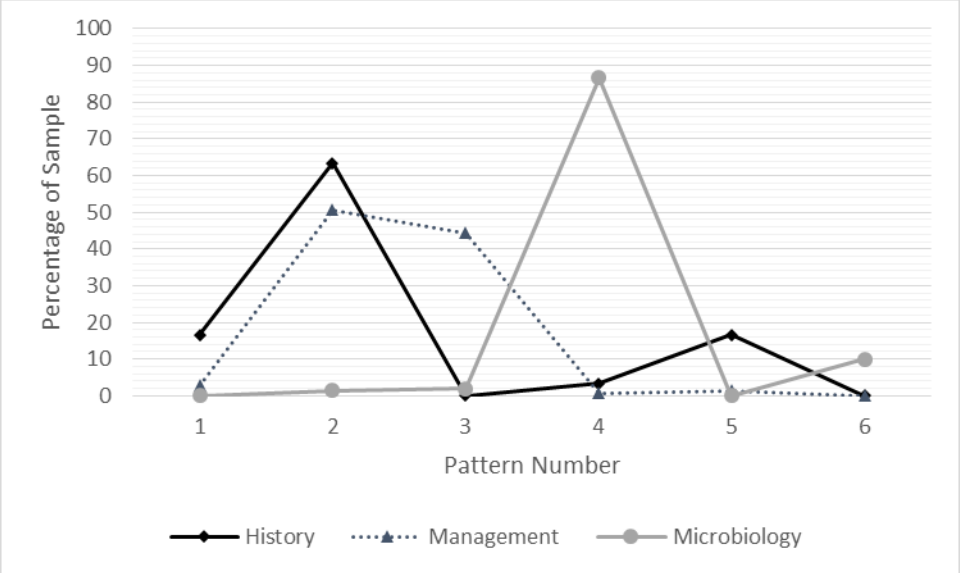

Figure 6.25 Interdisciplinary comparison for *replicate*.

The results of significance testing summarised in Table 6.25 provide further information. All comparisons with microbiology reveal significant differences for PTs two and four. PT three represents a significantly greater proportion of the sample in management than it does in other disciplines. PT five occurs significantly more frequently in history than it does in the other disciplines. There is also significant difference between microbiology and history for PT one and microbiology and management for PT six.

Table 6.25

Result of cross-corpora comparison for replicate

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human]] replicates [[Artifact]] <Human creates a copy of Artifact>			X
2	[[Human]] or [[Institution]] or [[Eventuality 1]] replicates [[Eventuality 2]] <Human or Institution or Eventuality 1 recreates Eventuality 2>		X	X
3	[[Human]] or [[Activity]] replicates [[Eventuality]] or {results findings} <Science. Human = Scientist or Activity = Study repeats Eventuality = Experiment in order to investigate whether this produces the same results or findings>	X	X	

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

Table 6.25 (cont.)

Result of cross-corpora comparison for replicate

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
4	[[MB_Entity]] replicates (Self) <MB Entity = DNA or Cell produces an exact copy of Self>		X	X
5	[[Concept]] or [[Institution]] replicates ([[Self]]) <Concept reoccurs in different context>	X		X
6	[[Human]] [[MB_Entity 1]] or [[Process]] replicates [[MB_Entity 2]] <[[MB_Entity 1]] or [[Process]] produces copy of [[MB_Entity 2]]>		X	

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.28 *See*

See is a frequently occurring verb across all disciplines (N = 300 in all sub-corpora). Corpus pattern analysis reveals 15 PTs in the AJACX2, of which all but pattern 15: [[Human1]] sees {the writing on the wall} are found in the PDEV. However, as Figure 6.26 shows, only the first eight occur at reasonably high frequencies and then not always across all disciplines.

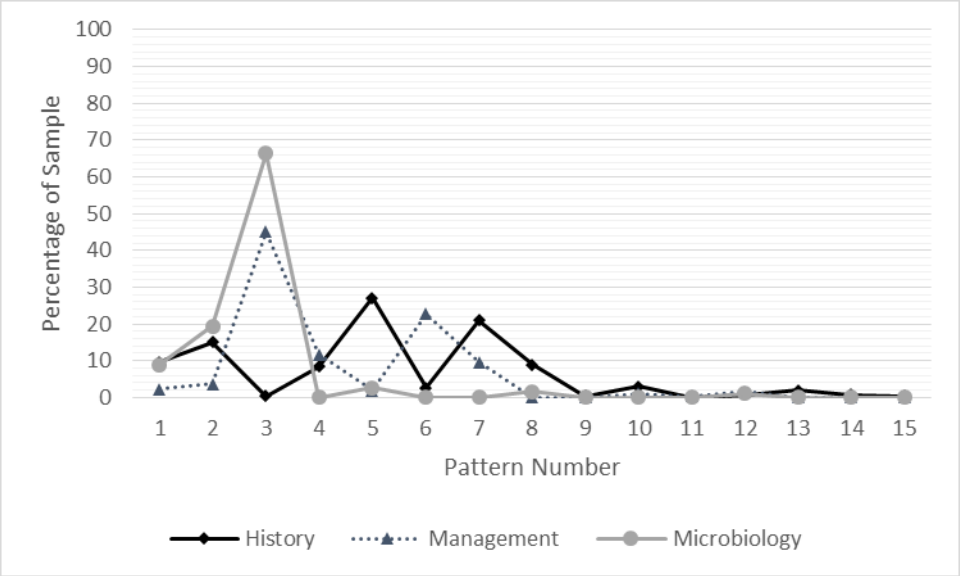

Figure 6.26 Interdisciplinary comparison for *see*.

Significance tests (summarised in Table 6.26) on PTs three, six and seven indicate significant difference in all pairwise comparisons. PTs one and two are significantly less frequent in management. PT four is significantly less frequent in microbiology. PTs five and eight are significantly more frequent in history. Finally, the absence of PT ten in microbiology is significant when compared with history.

Table 6.26

Result of cross-corpora comparison for see

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human]] or [[Animal]] sees [[Physical_Object]] or [[Stuff]] <Human or Animal perceives or observes Physical_Object or Stuff with his or her eyes>	X	X	
2	[[Human]] sees [[Event]] <Human is a witness to Event [NO ADVL]>	X	X	
3	[[Human]] sees [[Document]] or [[Document_Part]] <Human refers to Document or Document_Part for further information>	X	X	X
4	[[Human]] sees [[Location]] <Human refers to a Document_Part as a reference>		X	X
5	[[Human]] sees [[Proposition]] or [[Concept]] <Human achieves an understanding of Proposition or Concept>	X		X
6	[[Human]] sees [[Anything]] [[Manner]] <Human regard Anything Manner, i.e., in a particular way>	X	X	X
7	[[Human]] sees [[Eventuality]] or [[Physical_Object]] as [[Anything]] <Human classifies Eventuality or Physical_Object as being Anything>	X	X	X
8	[[Human]] sees [[that-clause]] or [[wh-clause]] <Human achieves an understanding of Proposition or Concept embodied in that [clause] or wh- [clause]>	X		X
9	[[Human]] or[[Institution]] sees [[Emotion]] or [[Attitude]] or [[Responsibility]] or [[Obligation]] <Human or Institution experiences an Emotion or Attitude or Obligation or Responsibility>			
10	[[Human 1]] sees [[Human 2]] <Human 1 goes to and spends some time with Human 2, typically for social reasons>			X
11	[[Human 1]] sees [[Human 2]] <Human 1 consults Human 2 = Expert or Doctor in order to get expert advice or treatment>			
12	[[Action]]is seen [[to-infinitive]] <Action is considered to/inf [verb], where the clause identifies an effect of the Action>			
13	[[Time_Period]] or [[Time_Point]] sees [[Event]] <Time_Period or Time_Point is the time when Event occurs>			
14	[[Human]] sees {fit} <Human considers it appropriate to/inf [verb]>			
15	[[Human1]] sees {the writing on the wall} <Human realises that Event=negative will happen soon>			

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.29 *Specify*

Only three PTs reveal themselves for *specify*, all of which are present in general English as represented by the PDEV. As indicated by Figure 6.27, the first occurs reasonably frequently in all samples, the second is limited to history (*N* = 70) and management (*N* = 300), and the third is exclusive to microbiology (*N* = 54) in the AJACX2.

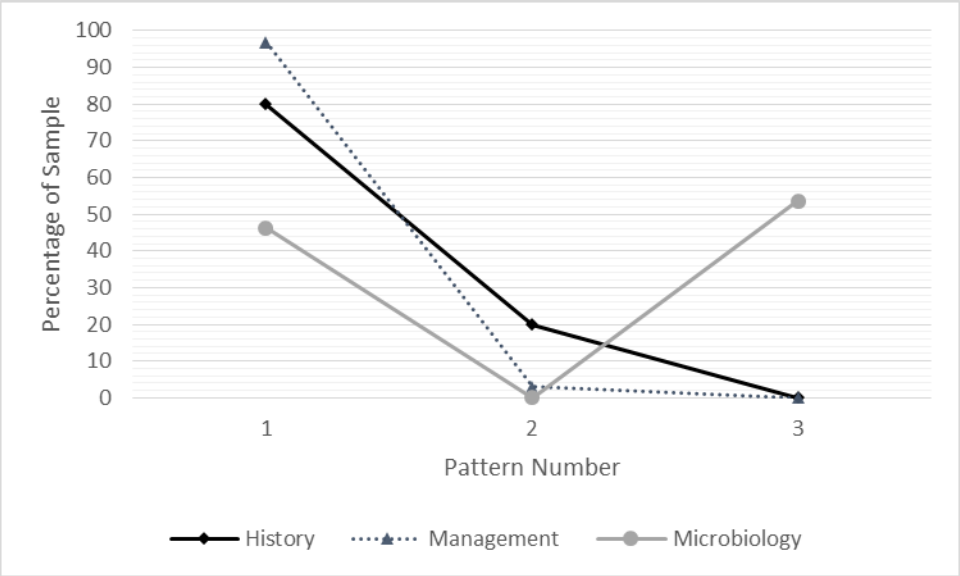

Figure 6.27 Interdisciplinary comparison for *specify*.

The results of significance testing (summarised in Table 6.27) reveal that there is significant difference in the frequency of occurrence of PT one between management, where it occurs most frequently, and microbiology where it occurs least frequently. PT two occurs significantly more frequently in history than in the other disciplines. PT three is only present in microbiology where significance testing suggests that difference is not a result of chance.

Table 6.27

Result of cross-corpora comparison for specify

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human]] or [[Institution]] or [[Concept]] or [[Document]] specifies [[Anything]] <Human or Institution or Document states clearly and precisely that Anything is a necessary condition for some activity, process, or concept>		X	
2	[[Human]] or [[Institution]] or [[Concept]] or [[Document]] specifies [[that]] <Human or Institution or Document states clearly and precisely the conditions implied by [that [CLAUSE] as being necessary for some activity, process, or concept>	X		X
3	[[MB Entity 1]] specifies [[MB Entity 2]] <MB Entity 1 contains the instructions necessary for the creation of MB Entity 2>		X	X

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.1.30 *Treat*

A corpus pattern analysis reveals seven PTs for *treat*, five of which are present in general English as represented in the PDEV. The proportion of each discipline sample that each PT represents is represented in Figure 6.28. As might be expected, given the focus of the discipline, the collocational behaviour of this verb in microbiology appears distinct from the other disciplines. This is particularly evident in PT one where history and management represent a greater percentage of the sample than microbiology, and patterns two and seven where the reverse is true.

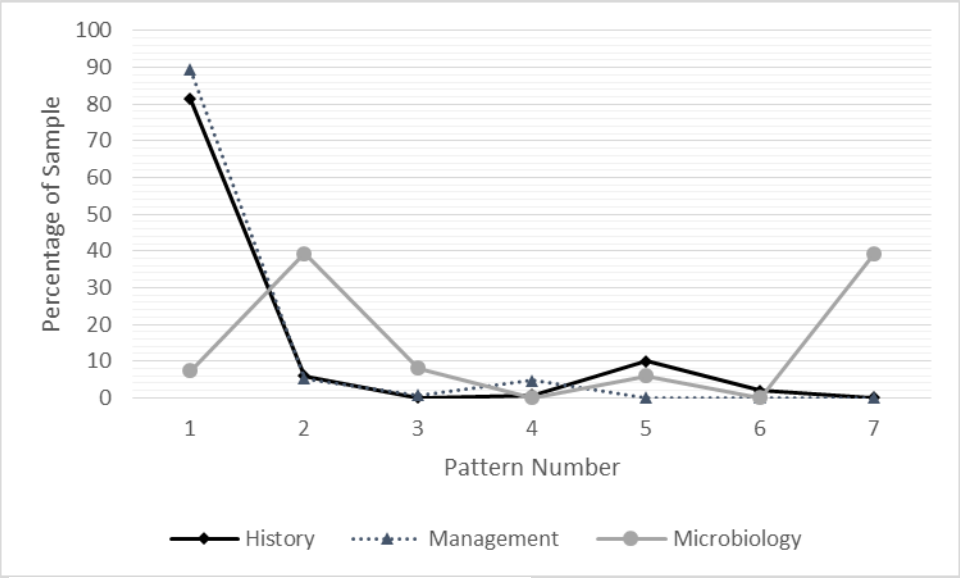

Figure 6.28 Interdisciplinary comparison for *treat*.

Significance testing (summarised in Table 6.28) confirms this impression with PTs one, two, and seven reaching the difference threshold in all comparisons with microbiology. PT three represents a significantly higher proportion of the sample for microbiology than history. PT five represents a significantly smaller proportion of the management sample than the other two disciplines.

Table 6.28
Result of cross-corpora comparison for treat

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
1	[[Human 1]] or [[Institution 1]] or [[Animal 1]] treats [[Human 2]] or [[Animal 2]] or [[Entity]] or [[Event]] Manner <Human 1 or Institution 1 or Animal 1 behaves toward Human 2 or Animal 2 or Entity or Event in the Manner specified>		X	X
2	[[Human 1]] or [[Process]] or [[Drug]] treats [[Human 2]] or [[Animal]] or [[Disease]] or [[Injury]] <Human 1 = Health Professional applies a Drug or Process = Medical to Human 2 =Patient for the purpose of curing the patient`s Disease or Injury>		X	X

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

Table 6.28 (cont.)
Result of cross-corpora comparison for treat

PT #	Pattern <Implicature>	Hist/ Man	Man/ Micro	Micro/ Hist
3	[[Human]] or [[Device]] treats [[Inanimate]] (with [[Stuff]] or by [[Process]]) <The chemical or other properties of Inanimate are improved or otherwise changed by Process or the application of Stuff>			X
4	[[Human 1]] treats [[Human 2]] or [[Self]] to [[Eventuality]] <Human 1 gives or pays for Eventuality = Good as a benefit for Human 2 or Self>			
5	[[Human]] treats [[Anything]] <Human discusses Anything = Topic>	X		X
6	[[Human 2]] treats with [[Human 2]] <Human 1 reaches a peace agreement with Human 2>			
7	[[Human]] treats [[MB entity 1]] or [[Animal]] with [[MB entity 2]] or [[Drug]] <Human introduces drug or MB entity 2 to MB Entity 1=cell or Animal for experimental purposes>		X	X

Note. X = significant at ≥ $p < .05$ and a BIC value signifying ≥ positive evidence against H_0

6.2 Intradisciplinary Comparison

It is possible that the differences in collocational behaviour outlined above are the result of factors other than academic discipline. It is widely accepted that the correlation of variables does not imply causality. An increasingly common criticism of corpus-based vocabulary studies which compare frequency measures across corpora is that the differences found are often not due to any of the variables under investigation, but instead differences that naturally occur any large body of text. Dunning (1993) has demonstrated how widely respected studies in corpus linguistics have failed to account for this possibility.

In an attempt to ensure that the differences in frequency of occurrence of PTs found between disciplines are not solely the result of the non-random nature of texts, the present study conducts a further round of significance testing in which verbs in the sample with a number of occurrences greater than 300 are randomly divided into two samples, A and B, of 150 lines. As in the inter-disciplinary comparison reported above, the LL test is applied and the BIC is employed to measure effect size. However in this instance the hypothesis tested are reversed:

> H_0: there is a difference in the frequency of occurrence of a given PT across sub-corpora

The alternative hypothesis is as follows:

> H_1: there is no difference in the frequency of occurrence of a given PT across sub-corpora

To ensure reliable comparison, ideally samples A and B should contain at least 150 lines from each discipline. This is the case for 11 of the verbs in the intradisciplinary

comparison. However, limiting the intradisciplinary samples to high frequency verbs allows the possibility that high frequency is the explanatory factor behind the verbs' behaviour. In other words, it is possible that there is a group of 'general' academic verbs which occur across all disciplines. To mitigate this possibility five verbs which only occur 300 or more times in one discipline are also included in the intradisciplinary comparison giving a total of 16.

The intradisciplinary comparison does not reveal a single PT which occurred at significantly different frequency in the A and B samples. Since these samples contain randomly selected corpus lines for each verb, this suggests that the differences frequency of occurrence of PTs found in the interdisciplinary comparison are not due to the non-random nature of the text itself, but rather some other explanatory variable. It seems reasonable to assume that this variable is academic discipline. It should be noted that two of the verbs in the intradisciplinary comparison, *accomplish* and *conduct*, do not demonstrate significant difference in the occurrence of their PTs across disciplines in the interdisciplinary comparison. The complete results of the intradisciplinary comparison can be found in Appendix D.

6.3 Summary

This statistical analysis of the collocation patterns of verbs in academic journal articles has provided clear evidence of verbs exhibiting distinct meanings in different academic disciplines. Each PT represents a sense, in many cases a given PT predominates in a discipline. In addition, there are often significant differences in the frequency with which a PT in used across disciplines. At first glance this may seem trivial, it seems obvious that different disciplines with their different focus should use verbs in different ways. However, in spite of its apparent obviousness, it runs counter to the rationale underlying discrete-item general academic wordlists.

The impact this analysis could have on EAP lexicographic practice, and by extension EAP users, is also far from trivial. It is apparent from the comparison of patterns found in the AJACX2 with those found in the PDEV that EAP users cannot rely on collocational dictionaries created using corpora of general English. More importantly, from a receptive standpoint while a user with a general definition of a verb and the benefit of context may be able to deduce the meaning of a discipline-specific use, in many instances this is unlikely be the case.

Going beyond the conclusion that verbs exhibit different meanings in different disciplines, the statistical analysis also reveals some important trends in the behaviour of verbs. There are PTs which occur at a reasonably high frequency across all disciplines these may be considered 'general' academic vocabulary which should be included in any EAP lexicographic resource. There is a suggestion of a humanities/science split with history and management on one side and microbiology on the other. Given the cost implications of producing subject specific dictionaries this is a valuable finding. These findings are discussed in further detail in the qualitative analysis reported in the following chapter.

7. QUALITATIVE ANALYSIS AND GUIDELINES FOR LEXICOGRAPHICAL REPRESENTATION

The previous chapter reported on a statistical analysis which demonstrated that patterns representing different verb meanings account for significantly different proportions of corpus samples from different academic disciplines. This casts many of the assumptions underlying discrete-item general academic wordlists into doubt. The present chapter examines interdisciplinary difference in more detail, reporting on a qualitative analysis of ten of the verbs included in the statistical analysis. These verbs were selected at random. Many of the observations made in this chapter are not statistically significant in the AJACX2 but they are indicative of the type of discipline-specific phraseological behaviour encountered during the annotation process and certainly merit further investigation in their own right. At the end of the chapter guidelines for how these findings might be usefully represented in EAP lexicographical resources are put forward along with example dictionary entries.

7.1 Qualitative Analysis

7.1.1 *Accept*

As is apparent from the statistical analysis reported in the previous chapter the verb *accept* demonstrates prolific collocational behaviour. It occurs in 17 pattern types in the AJACX2 nine of which are pattern types found in general English as represented by the PDEV. However, in the AJACX2 only three pattern types occur across all disciplines. These include pattern types one and two which occur frequently across all disciplines, and pattern nine which is limited to a single occurrence in management and microbiology. This analysis concentrates on the first two patterns.

> Pattern one: [[Human]] or [[Institution]] **accepts** [[Proposition]] or [[Concept]] or [[Eventuality]]
> Implicature: Human or Institution agrees that Proposition or Concept or Eventuality is correct and does not need to be contested

Pattern type one accounts for 71 per cent of the sample in history (N = 300), 71.34 per cent in management (N = 157), and 40 per cent of the sample in microbiology (N = 65). The statistical testing reported in the previous chapter demonstrated significant difference in the frequency of occurrence of this pattern type between the latter discipline and the former disciplines. A finer-grained analysis of the corpus lines brings to light further differences which the purely frequency-based comparison obfuscates.

An immediately obvious difference is the apparent semantic-type alternation which occurs frequently in history where [[Human]] or [[Institution]] alternates with [[Location]] in the subject slot. For example: "*Towns and territories accepted the reformation*". This kind of alternation is extremely widespread in language in general. Given the subject matter of history, this synecdochic relation between location and humans or institutions in a position of authority, is unsurprising. It is unlikely to cause the user any great problem. It produces no great change in the meaning of the event, although there is a slight change in focus to emphasise the geographical location of the [[Human]] or [[Institution]] doing the accepting.

An examination of semantic prosody reveals a further difference between the instances of pattern type one found in history and those found in management. On the occasions where the semantic type [[Eventuality]] appears in the object slot in management it frequently imbues the verb with a negative aura. For example, "*employees are more accepting of an unfavourable outcome"*. This would not be problematic from a receptive standpoint, however for a novice EAP writer working in this discipline an explanatory note on this phenomenon could prove useful. It is possible that this trend is an artefact of corpus composition, perhaps the management sub-corpus contained an unrepresentatively large proportion of articles dealing with institutional problems. However, the sampling procedures adopted mitigate this possibility somewhat.

> Pattern type two: [[Human]] or [[Institution]] **accepts** [that-clause]
> Implicature: Human or Institution agrees that [that-clause] is true or correct

A key difference in the collocation behaviour of *accept* is apparent in the comparison of microbiology on one hand, and history and management on the other. In microbiology (*N* = 65), 36.92 per cent of the instances of the verb *accept* occur in pattern type two; in management (*N* = 157), the figure is 7.28 per cent; and in history (*N* = 300) the figure is 7 per cent. The tests reported in the previous chapter confirm the significance of these differences in frequency. An in-depth examination of those corpus lines annotated as pattern type two reveals a further discipline-specific difference in the phraseology of *accept*. In microbiology all 26 instances of pattern type two exhibit the following syntactic construction:

> {it} is ({generally}{well}{widely}) **accepted** {that}

In history only three of the 21 instances of this pattern (14.29 per cent) exhibited this alternation. In management the alternation occurred in four out of 13 instances (30.77 per cent).

The distinction between so-called hard and soft disciplines provides a tentative explanation for the marked difference in frequency of this alternation across disciplines. Harder disciplines place greater emphasis on certainty (Hargens, 1988). It is hardly surprising, then, that in microbiology, the hardest discipline under study, variations of the phrase *it is accepted that* abound to the extent that they represent a domain-specific norm. In contrast, in history, the softest discipline under study, this alternation occurs less frequently because there is a much lesser degree of certainty. The different functions of RAs across disciplines, discussed in Chapter 4, provide a related explanation: The reason this syntactic alternation is not treated as an exploitation is that it does not alter the basic meaning of the verb. Instead, the alternation of [[Human]] or [[Institution]] in subject slot, with the dummy subject {it} merely shifts the focus from the agent to the [that-clause] being accepted. With this in mind, the contrasting frequencies of the alternation can be seen to reflect the contrasting functions or purposes of RAs in each discipline. RAs in harder disciplines like microbiology tend to report procedures and findings (Hyland, 2016a). This explains the focus on what has been accepted rather than who has accepted it. In contrast, RAs in history tend to describe events or argue for particular interpretations of such events; placing greater emphasis on human agents involved (Coffin, 2006). Taken as a set, management RAs represent a middle ground with articles both reporting on research procedures, and describing events and arguing for particular concepts or interpretations.

The treatment of this alternation in lexicographical resources raises important questions. This alternation involves no change in the basic meaning of the verb and as such is not likely to pose a great problem to the user from a receptive standpoint. However, given its ubiquity in microbiology, it would make sense to make the users working in this discipline aware that this is the prototypical form.

7.1.2 *Accomplish*

In history (N = 73) and management (N = 138) the phraseology of *accomplish* is uniform. The verb denotes a human agent successfully completing a task or bringing about an eventuality. This use is represented in the following pattern:

> Pattern: [[Human]] or [[Institution]] **accomplishes** [[Activity]] or [[Eventuality]]
> Implicature: Human successfully completes Activity = Task or brings about Eventuality

Corpus lines from microbiology (N = 67) also map to this pattern. This is exemplified the illustrative selection of lines from microbiology in (a) to (e) below. The lexical items {detection}, {mapping}, {demonstration}, and {identification} might all reasonably have semantic type [[Activity]]. While {ECL protocol}, {model systems}, and {determining susceptibility} necessarily involve human cognition.

a. Detection was **accomplished** using the ECL protocol
b. mapping of regions of the chromosome has been **accomplished** in several model systems using two
c. demonstration of the microbial activities was not **accomplished** before. Using a set of molecular and
d. techniques are not used, identification can be **accomplished** by determining susceptibility to optochin
e. policy for nearly 20 years. This was not **accomplished** by 2000 or 2005 and does not appear likely

The situation is different in other corpus lines from microbiology. The lexical items {copper binding}, {incorporation}, and {interactions} can feasibly be grouped as [[Activity]] or [[Eventuality]]. However, {conserved methionine residues}, {interactions with HIV-1 Gag}, {production of degradative enzymes}, {urease accessory proteins}, and {secreted molecules} do not involve of human agency. These lexical items might be better classed as [[MB entity]] or [[Process = microbiological]].

f. indicated that copper binding in CusF is **accomplished** in part by two conserved methionine residues
g. Incorporation of LysRS into the virion is **accomplished** via specific interactions with HIV-1 Gag
h. acquiring nutrients. This is thought to be **accomplished** by the production of degradative enzymes
i. the nickel ions into the metallocenter are **accomplished** by the urease accessory proteins UreD chaperone
j. environment. These interactions are often **accomplished** by the distal effects of secreted molecules

These new semantic types can be incorporated in the pattern set out above causing only slight changes to the verb's meaning. They can be regarded as discipline-specific alternations. The alternation of [[MB Entity]] with [[Human]] was found to be widespread

in the AJACX2. In some instances this caused a change in the core meaning of the verb and resulted in the creation of a new pattern, in other instances frequently occurring alternations had little effect on verb meaning. In the latter case, the original pattern was simply widened to include the new semantic types. Notable exceptions to this procedure include listing those patterns with [that-clauses] separately. The rationale underlying this decision is set out at the start of the previous chapter. The latter approach was adopted with *accomplish* for the purposes of the statistical analysis reported in the previous chapter.

> Pattern: [[Human]] or [[Institution]] or [[MB entity]] or [[Process = microbiological]] **accomplishes** [[Activity]] or [[Eventuality]]
> Implicature: Human or Microbiological entity successfully completes Activity or brings about Eventuality

Widening the pattern to include the new semantic type ensured that the statistical analysis compared patterns that truly represented different word meanings rather than patterns that were different only because they reflected the different subject focus of each discipline. In other words, [[Human 1]] *treats* [[Human 2]], and [[Human 1]] *treats* [[MB Entity]] are included as separate patterns because they activate different senses of the verb *treat*, rather than as a reflection of the necessity of microbiologists to talk about [[MB Entities]].

This conservative approach to the creation of new patterns ensures an empirically sound comparison of their frequency of occurrence across disciplines. However, it is not always conducive to the representation of information in a manner that would be useful to an EAP user working in a particular discipline. The behaviour of *accomplish* in microbiology provides a clear example of this. In history and management, a mixture of active and passive constructions involving *accomplish* can be found. In contrast, in microbiology the predominant construction is passive and all of the instances of *accomplish* collocating with [[MB Entity]] or [[Process = biological]] take this form. For this reason, it is suggested that any entry in an EAP lexicographical resource aimed at microbiologists reflect the following pattern:

> [[Activity]] or [[Eventuality]] is **accomplished** {by} [[MB Entity]] or [[Process = biological]]

The inclusion of separate patterns in lexical resources for all active and passive constructions would not be feasible. However, on the basis of the evidence from the AJACX2 the passive form represented by the pattern above predominates in microbiology. It is reasonable to assume therefore that for users working in this discipline this is the form that should be represented. It could be argued that such considerations belong in the domain of syntax and are beyond the scope of lexicography. However, on a phraseological approach to language, lexicographical resources should reflect the predominant patterns in a discipline in the form in which they occur.

7.1.3 *Apply*

An analysis of the corpus brought to light eight collocational pattern types for the verb *apply* (N = 300 in all sub-corpora). This qualitative analysis focuses on pattern types one and four. Pattern type one is found at high frequencies across all sub-corpora. It accounts for 54.33 per cent of the history sample, 71.67 per cent of management, and 64 per cent of microbiology. Statistical testing found no significant difference in the relative frequency of occurrence of this pattern across the three disciplines. In the AJACX2, pattern type four

occurs only in history where it accounts for 2.67 per cent of occurrences, and microbiology, where it accounts for 23.33 per cent of occurrences. Significant difference in relative frequency was noted in all three pairwise comparisons for this pattern.

> Pattern type one: [[Human]] or [[Institution]] **applies** [[Concept]] or [[Process]] (to [[State of Affairs]])
> Implicature: Human or Institution makes use of Concept or Process in a particular situation or State of Affairs
> Pattern type four: [[Human]] or [[Device]] **applies** [[Fluid]] or [[Stuff]] to [[Surface]]
> Implicature: Human (uses Device to) spread Fluid or Stuff on Surface, typically in order to heal, improve, protect Surface or conduct test

Although no significant difference in frequency was found for pattern type one across the sub-corpora, the annotation did reveal phraseology which merits further comment. A number of lines in the history sub-corpus proved a challenge in the annotation process:

> (1)
> a. no balm whatsoever would be **applied** to tender consciences by his magistrate
> b. why was Zumárraga led down the path of **applying** the heavy cudgel of Inquisition to the indigenous neophytes
> c. he signally failed to **apply** the blade to conventional wisdom on gender

In all the corpus lines in (1) metaphor is being used to exploit a norm. At first glance, the temptation might be to map all these sentences to pattern type four: [[Human]] *applies* [[Fluid]] or [[Stuff]] to [[Surface]]. This is especially so in the case of (1a). However, on reflection it is clear that the {balm} in (a), {cudgel} in (b), and {blade} in (c) have been chosen to achieve a pragmatic effect which is achieved through a semantic-type coercion. In the case of (a) {balm} is coerced from [[Fluid]] or [[Stuff]] to [[Process]]. At the same time the coercion exploits pattern ambiguity, evoking a strong resonance with the sense of healing or improving conveyed by pattern type four. In a CG or RCG framework, this might be explained by the novel pattern which represents an expression which lacks unit status, being meaningful as a result of its similarity to the established pattern which represents an expression with unit status. In the case of (b) and (c) the semantic type [[Weapon]] is coerced to [[Process]] altering the semantic prosody of the verb and imbuing it with an aura of violence or abruptness.

> (2) and half killed them with blows, and he **applied** fire to my foster-daughters, having stripped

The corpus line in (2) should not be confused with an exploitation through metaphor of pattern one or four. It is in fact an archaic use of *apply* which in collocation with *fire* has the sense of the verb *burn*. For the purposes of statistical comparison it was classified as an exploitation of pattern type four. Statistically this isolated occurrence is of little significance however it does illustrate the problems archaic language might pose to an EAP user working in history.

The situation in microbiology is far more straightforward. No examples of archaic uses were found for any of the verbs analysed. More significantly, all instances of corpus lines mapping to pattern type four were entirely literal typically referring to applying a solution containing a sample of a biological material to be tested to some kind of test medium. It could be argued that instead of lumping this meaning under the very general implicature of: *Human (uses Device to) spread Fluid or Stuff on Surface, typically in order to heal, improve, protect*

Surface or conduct test this pattern should be treated as a domain-specific norm in its own right. A discipline-specific realisation of pattern type four for microbiology would be:

> Pattern: [[Human = scientist]] **applies** [[Stuff = sample or culture]] to [[Surface = test medium]]
> Implicature: Human (typically a scientist) applies a sample to surface (typically some kind of test medium) in order to or conduct a microbiological test
> Passive alternation: [[Stuff = sample or culture]] is **applied** to [[Surface = test medium]] (by [[Human = scientist]])

Further, evidence of this procedural function can be seen in a passive alternation of pattern type four. In microbiology these patterns generalise over corpus lines which are overwhelmingly passive. This is not the case in the other two disciplines under study. In history and to a lesser extent management, the focus of the sentence in on the actor, in contrast microbiology RAs are primarily concerned with the events taking place in the procedure.

The exploitation of pattern type one through metaphor occurs in three different history RAs. It was not apparent in any other sub-corpora. While it represents an interesting, cognitively salient, exploitation which is rhetorically effective and easy to recall, it is not statistically salient since it is limited to three history RAs written by distinct authors. An additional search in the BNC for various lexical items with the semantic type [[Weapon]] in collocation with *apply* revealed no results. This suggests that the exploitation may well be an idiosyncrasy of these authors and thus, on balance should not be included in a lexicographical resource for EAP users. In contrast, the passive alternation of pattern type four occurs frequently in microbiology where it represents a discipline-specific norm. For this reason it should certainly be included in any EAP lexicographical resource aimed at users from this discipline. Its predominantly passive configuration raises the question of what form the canonical pattern should take. As Hanks (2013, p. 189) states, it would be ill-advised to include both active and passive patterns in a general pattern dictionary; nevertheless, it is certainly worth considering whether the active construction truly represents the canonical form.

7.1.4 *Conduct*

In statistical analysis only *conduct* and *accomplish* displayed no statistically significant differences in collocational behaviour across disciplines. Pattern type one accounts for 94 per cent of instances in history (N = 150), 99.33 per cent in management (N = 300), and 98 per cent in microbiology (N = 150):

> Pattern type one: [[Human]] or [[Institution]] **conducts** [[Activity]]
> Implicature: Human or Institution carries out Activity

A more detailed examination of the corpus lines reveals nothing to counter this impression of uniformity. There is a generalised active/passive alternation but this is widespread in English in general and since it seems to occur fairly frequently across all sub-corpora with *conduct* it does not merit any special mention in an EAP lexicographical resource.

Pattern type six occurs only twice in the AJACX2. As the semantic type in subject position suggests, it is exclusive to microbiology:

> Pattern type six: [[MB Entity]] **conducts** [[Process]]
> Implicature: MB entity carries out Process = microbiological

The two lines which map to this pattern are:

and the two ATG4 peptidases are thought to **conduct** the pivotal steps in the formation of autophagous in this strand. The complex **conducts** the cycling (loading and unloading) of

Although represented as a pattern in its own right, this pattern could equally be treated as an alternation of the predominant pattern type one. In the CPASO [[Process]] is a hyponym of [[Activity]]. The alternation of [[MB Entity]] with [[Human]] in subject slot is widespread throughout the AJACX2. It may have come about as pioneers in microbiology, seeking to effectively communicate their work, drew analogies between observations in their work and lived human experience. In cognitive terms it has become entrenched with repeated frequent use. For the purposes of statistical analysis, where such alternations occur frequently the more general pattern is widened to include the new semantic type. This ensures that the statistical analysis compares patterns that truly activate different word meanings rather than patterns that are different only because they reflect the different subject focus of each discipline. In the case of *conduct* pattern type one was not widened because the phraseology represented by pattern type six was extremely infrequent. The low frequency of occurrence meant the decision to list it as a separate pattern in its own right had a negligible effect on the statistical analysis.

This analysis of the collocational behaviour *conduct* in the AJACX2 has implications for the representation of the verb in EAP lexicographical resources and materials. Pattern type one reflects a meaning of a verb which is frequently employed in all the disciplines addressed in the present study. It should therefore be included in any lexicographical resource aimed at a general EAP audience. The inclusion in a lexicographical resource of the other patterns listed is not justified by data from the AJACX2 alone. However, this does not exclude the possibility of their inclusion after further investigation in other sources.

7.1.5 *Explain*

The verb *explain* occurs at high frequency across all the sub-corpora (history: 876 occurrences; management: 1104 occurrences; microbiology: 431 occurrences). A statistical analysis of its phraseology (N = 300 in all sub-corpora) revealed that it does not reflect the general tendency for a marked difference in frequency between history and management on the one hand, and microbiology on the other. The first indications of this can be seen in the frequency of occurrence of pattern type one. In management and microbiology this pattern accounts for 77.33 per cent and 83.33 per cent of the sample respectively, in history the figure is significantly lower at 57 per cent.

> Pattern type one: [[Human]] or [[Institution]] or [[Document]] or [[Proposition]] or [[Eventuality 1]] **explains** [[Eventuality 2]] (in terms of [[Eventuality 3]]) (to [[Human 2]])
> Implicature: Human or Institution formulates (in conversation with Human 2 or in a Document intended for Human 2 to read) a Proposition concerning the cause or

effect of an Eventuality 2 - typically, using a hypothesis or theory that has already satisfactorily explained [[Eventuality 3]]

As the implicature suggests, typical uses of this pattern involve a human explaining an eventuality using a hypothesis that has previously been used to explain something else. Since the history RAs in the corpus contain fewer explicit references to theories, and historical analysis typically relies on hypothesis testing to a lesser extent than analysis in microbiology and management (N. Wilson, 1999), this marked difference in frequency is hardly surprising.

The distinct functions of RAs across disciplines also explain differences in frequencies of other patterns. Pattern type two accounts for 20.67 per cent of the occurrences of *explain* in history and 14 per cent of the occurrences in management. Log-likelihood testing comparing the frequency of occurrence of the across both sub-corpora give a score slightly over the critical value. However the effect size score indicated positive evidence in favour of H_0 rather than against.

Pattern type two: [[Human]] or [[Institution]] or [[Document]] or [[Proposition]] explains (to [[Human 2]])
Implicature: Human or Institution formulates Proposition in Document concerning the cause or effect of an Eventuality that is expressed in wh- [clause]

Although this difference is not significant, it could be explained by the relatively greater use of *explain* with a reporting function in history. Compare lines (a) to (e) from the history sub-corpus with line (f) to (j) from the microbiology sub-corpus:

a. A few years later, Guizot **explained** how the various passions inherited from
b. intervenes in the narrative at this point to **explain** what is going on, noting how after death
c. scrutiny. 23 Gadi Algazi, in particular, has **explained** how Brunner's historical scholarship on
d. climatic theories of health. This, Paul **explains**, is why the north produces people so prolifically
e. an invisible narrator described its 'practical modern appointments', **explaining** how the new labour-saving devices worked

f. this mechanism **explains** how aRNAs are produced and are specifically
g. the coamplification model can **explain** how gene clusters are maintained
h. the loop **explains** how the C-signal rises continuously from
i. these uncertainties certainly **explain** why some thermophilic enzymes were found
j. the structural efflux pump **explains** why they are regulated in the same way

In (a) to (e) the semantic type of the lexical item in subject slot is [[Human]]; in lines (f) to (j) it is [[Proposition]]. The alternation of these semantic types is fairly widespread since the formulation of a proposition requires human cognition. In most cases no change in core meaning is activated by this alternation. However, in the case of (a) to (e) above *explain* exhibits a meaning similar to *report*. In contrast in lines (f) to (j) a [[Proposition]] activates a meaning of *explain* pertaining to the cause of an eventuality.

If it were the case that the differing phraseology of *explain* in history is due to its reporting function, this difference might also be reflected in the relative occurrence of pattern type four. This pattern generalises a use of *explain* which involves reporting quotes:

Pattern type four: [[Human]] or [[Institution]] **explains** [[Quote]]
Implicature: Human says [QUOTE] as a way of accounting for Eventuality (often implied rather than stated explicitly)

Pattern type four does occur significantly more frequently history where it accounts for 9 per cent of the sample, than microbiology where its single instance accounts for 0.33 per cent of the sample. There is no significant difference between history, and management (6.33 per cent of the sample). Several articles in management report the response of participants in field interviews. This is generally achieved through the use of literal quotes enclosed in quotation marks. Reporting practices in history are more varied. On many occasions authors employ a paraphrase using a [that clause]:

Pattern type three: [[Human]] or [[Institution]] **explains** [That-Clause] (to [[Human 2]])
Implicature: Human 1 or Document tells (Human 2) that [clause], as a way of accounting for Eventuality (often implied rather than stated explicitly)

This difference in reporting practices explains why pattern type three accounts for a significantly higher proportion of the sample in history (12.33 per cent) than it does in management (2.33 per cent) and microbiology (0.67 per cent).

The objective underlying conducting a statistical comparison of relative pattern frequency across discipline-specific sub-corpora is to examine whether the prototypical meaning potentials activated for a verb differ across academic disciplines. A key assumption underlying the statistical comparison is that each pattern represents a distinct core meaning of the verb. The analyses undertaken above suggest that this may not be the case for pattern types two, three, and four, which could equally be accounted for by the following more general pattern:

Pattern type 2a: [[Human]] or [[Institution]] or [[Document]] or [[Proposition]] **explains** [wh-clause], [that-clause] or [[Quote]] (to [[Human 2]])
Implicature: Human or Institution formulates Proposition in Document or [[Quote]] or that-[clause] or [wh-clause] concerning the cause or effect of an Eventuality

A statistical reanalysis of *explain* reveals that the broader pattern accounts for a significantly higher portion of the sample in history, than it does in management and microbiology at $P < 0.0001$ and effect size indicating very strong evidence against the null hypothesis of no significant difference. Full details are provided in an additional table in Appendix C.

In TNE there is no strict line which delimits norms, alternations, and exploitations. Ultimately, different speech communities will perceive the relationship differently. With this in mind, with the aim of providing users with information relevant to their disciplines, the more delicate distinctions could prove useful. EAP writers working in management might find it useful to know that in their discipline *explain,* when used for reporting, is followed by a quote rather than by a paraphrase inside a that-clause. In contrast, pattern type one should be included in a general EAP lexicographical resource since, despite significant differences in frequencies, it is highly frequent in every discipline in the AJACX2.

7.1.6 *Manage*

Before examining a single line of any sub-corpus what is immediately apparent is the marked difference in the frequency of this verb across disciplines. *Manage* occurs 747 times in management placing it among the most frequently occurring of the verbs studied. This is hardly surprising given the focus and the name of the discipline. *Manage* also occurs frequently in history with 296 occurrences, slightly under the threshold for intradisciplinary comparison. In contrast, in microbiology it occurs 22 times. Pattern type one predominates in the history sample (*N* = 150) accounting for 60 per cent of occurrences, in management (*N* = 300) it accounts for 5.33 per cent of the sample, while the two instances of this pattern in microbiology (*N* = 22) account for 9.09 per cent of the sample.

> Pattern type one: [[Human]] or [[Institution]] or [[MB Entity]] **manages** [to-infinitive]
> Implicature: Human or Institution is able to/inf [verb], despite difficult circumstances

A tentative hypothesis for the vast difference in overall frequency of occurrence of the verb between disciplines may rest upon the distinction between hard and soft disciplines in terms of the need to express conation. Hard disciplines are characterised by certainty, goals are either realised or not realised, however, the use of *manage*, particularly with the meaning activated by pattern type one suggests a lesser degree of certainty. Writers in history seek to report whether a human or institution is able to achieve goal and the difficulty of the circumstances in which this was done. In microbiology the difficulty of the circumstances takes second place to reporting the procedure.

Pattern type six also occurs frequently representing 21.33 per cent of the history sample, 63.67 per cent of management, and 31.82 per cent of microbiology.

> Pattern type six: [[Human]] or [[Institution]] or [[MB Entity]] **manages** [[System]] or [[Activity]]
> Implicature: Human or Institution or MB Entity is responsible for and controls System or Activity = Ongoing

Once again the focus of the discipline explains the predominance of pattern six in management. For an EAP user working in management studies the default interpretation of *manage* would denote responsibility for an ongoing process or system. More surprising, perhaps, is the relatively high frequency of this pattern type in microbiology. Here, [[MB Entity]] predominates in subject slot and is typically responsible for controlling an ongoing microbiological process. The presence of [[MB Entity]] in microbiology in a position occupied by [[Human]] in general English is widespread, as will become apparent in analysis of other verbs, and points to an exploitation. This is clearly a not literal use of *manage*. An [[MB Entity]] e.g. a *cell* or a *gene* does not have the cognitive capacity to manage a process. It is possible that, faced with the necessity to describe new concepts and discoveries writers in microbiology exploited the existing pattern. Over time this exploitation of pattern type six has become conventionalised in microbiology and at present could be considered a discipline specific norm. In the terminology of cognitive linguistics it has become entrenched.

Pattern type nine is limited to microbiology in the AJACX2. It has five occurrences and accounts for 22.73 per cent of the sample. It represents another example of discipline-specific collocational behaviour in microbiology:

> Pattern type nine: [[Human]] **manages** [[Disease]]
> Implicature: Human alleviates symptoms of Disease with Drugs or therapy

While it could be argued that [[Human]] *manages* [[Disease]] could be regarded as an alternation of pattern six, in fact, the meaning activated is much narrower. It is closer in meaning to *alleviate* than the sense activated by pattern type six which implies *controlling* or taking responsibility for something. The exclusive and relatively widespread occurrence of this pattern microbiology in the AJACX2 and its usefulness for students and researchers working in medicine and social care, suggests these users would also be well served by this distinction.

All of these differences have implications for EAP lexicography. In general the difference in frequency of *manage* between history and management on one hand, and microbiology on the other hint at EAP users' needs being best served by separate dictionaries; one aimed at hard disciplines the other at softer disciplines. This impression is reinforced by the prevalence of pattern type one in history with its emphasis on achievement in difficult circumstances. Pattern type six illustrates the challenges diachronic aspects of meaning pose to lexicography. This is especially so in scientific and technical domains. Outside microbiology, lexical items of the semantic type [[MB entity]] are seldom present in the subject slot of this pattern type. With this in mind, it is reasonable to hypothesise that this alternation may have come about through what was originally an exploitation. A now conventionalised metaphor was originally used to express a novel concept in a target domain, microbiology, in terms of an established concept from the source domain of organisational structure. The alternation between [[MB Entity]] and [[Human]] or [[Institution]] is fairly transparent and unlikely to cause the user any great problem. However, the widespread nature of this alternation is problematic from a lexicographical point of view. There is a tension between the desire to comprehensively report the facts of the language on one hand, and not to overwhelm the user with superfluous information on the other. A possible solution might involve a note explaining that this kind of alternation is prevalent in microbiology while not formalising it in every single pattern entered in the dictionary. Pattern type nine represents a similar problem. It might be argued that the microbiological sense expressed by pattern nine is a subset of pattern six and could be subsumed by this sense in any lexicographical resource. Ultimately, different members of a speech community and different analysts will have different perspectives on this relationship. However, the alternation between [[Disease]], and [[System]] or [[Activity]] in the object slot is far less transparent than that between [[Human]] and [[MB Entity]] in the subject slot of pattern type six. Moreover, as one might expect, in microbiological contexts nouns denoting diseases are much more common in object position. This makes the alternation particularly salient in this domain and suggests that it should be included as a pattern type in its own right in a resource for microbiologists.

7.1.7 *Note*

Note (*N* = 300 in all sub-corpora) appears in four pattern types in the AJACX2, the first two denote the sense of noticing or paying attention, while the latter two denote a sense of recording information. The only significant difference in frequency revealed by the statistical analysis reported in the previous chapter occurred between history and

microbiology for pattern type three. However, a finer-grained analysis of the corpus lines reveals differences in other patterns.

> Pattern type two: [[Human]] **notes** [[quote]] or [that-clause] or [wh-clause]
> Implicature: Human notices and pays particular attention to Eventuality or Visible Feature

For pattern type two, which accounts for 63.33 per cent of occurrences in management and 51 per cent of occurrences in microbiology, the object slot is typically filled by a [that-clause], however in history, which accounts for 57.67 per cent of occurrences, quotations predominate in this slot. As has been seen in the analysis of *explain*, this is hardly surprising given that a good deal of historians' work involves reporting speech and documents. The importance of documentation in history and, to a lesser (not statistically significant) extent management, is also reflected in the comparison of the occurrence of pattern type three with microbiology.

> Pattern type three: [[Human]] **notes** [[Information]] (on or in [[Document]]) (under [[Document_Part]])
> Implicature: Human makes a written note of Information (on | in Document) (under Document_Part = Heading)
> Pattern type four: [[MB Entity]] is **noted** [[Name]]
> Implicature: MB Entity is referred to as Name

Pattern types three and four, which deal with the sense of recording information, occur infrequently. Pattern type three represents 7.67 per cent of the history sample, 4 per cent of the management sample, and 1.67 per cent of the microbiology sample. Pattern type four occurs only in microbiology, where it is found five times, and represents 1.67 per cent of the sample. Although these frequencies are relatively low they reveal an interesting perspective on the evolution of domain-specific and, in the context of the present study, discipline-specific meaning.

The corpus lines which map to these patterns differ both syntactically and semantically. Both these differences bring about slight changes in meaning. Pattern type three occurs in both active and passive forms. In most cases the agent is expressed. In the case of the passive this is achieved via a prepositional phrase with *by*. The five instances of pattern type four occur exclusively in microbiology and in a passive construction. In all instances the auxiliary *be* is elided:

described that specific bacterial strains **noted** probiotics, and specific nondigestible
pathogenesis-related (PR) proteins was described. Those **noted** PR-2 proteins have a -1 3-glucanase activity
catalytic domain of the toxin into the cytosol (**noted** CNF1-CD). This phenomenon is probably triggered
expression of genes coding for specific proteins **noted** pathogenesis-related (PR) proteins was
anti-Gbs1477 antiserum. The last two lanes (**noted** IP) represent immunoadsorptions performed

An active form of pattern type four could be represented as follows:

> [[Human]] **notes** [[MB Entity]] [[Name]]

Here [[Name]] can be regarded as a hyponym of [[Information]]. With this is mind it is possible that pattern type four is in fact an exploitation of pattern three. The [[Human]] *noting* the [[Information = name of MB entity]] being elided perhaps to place the focus on

the name of the [[MB Entity]] rather than the scientist doing the naming. Having said that, the generalisability of this example must be questioned due to the low frequency of occurrence and limited number of source RAs in which it is found in the AJACX2. An examination of 12-million token sub-corpus of journal articles from health sciences in the Oxford Corpus of Academic English (D. Lea, 2014a) confirms this impression revealing only four lines which mapped to this pattern.

An in-depth examination of those corpus lines containing *note* in the AJACX2 reveals few findings with serious implications for the representation of its phraseology in EAP lexicographical resources. The relatively high number of quotes in object slot for pattern type two is a result of the subject focus of history as a discipline. In the analysis of *explain* the specification of three separate patterns for [what-clause], [that-clause], and [quote] in object slot allowed tentative hypothesises about the quotation practices in each discipline to be formulated. Information about such differences may prove useful to writers working in each discipline. In the examination of the corpus lines for *note* there is no evidence to suggest the lines should map to distinct pattern types, although it is suggested that the slightly, though not significantly, greater prevalence of pattern type two in history is due to an increased tendency report quotations using *note* in this discipline. Further investigation is needed, but on balance it would be prudent to include an example of this use of *note* in any general academic lexicographical resource. Pattern type four may be more problematic for the user. The semantic and syntactic differences it demonstrates in relation to the more general pattern type three make deriving its meaning far from straightforward. However, because this use is extremely infrequent accounting for only 1.67 per cent of the microbiology sub-corpus, and only found in three source RAs, it would be unwise to include it any lexicographical resources based on evidence from the AJACX2 alone.

7.1.8 *Replicate*

Before examining corpus lines an immediately obvious difference concerns the frequency of occurrence of *replicate* in each sub-corpus. The verb occurs only 30 times in history, 146 times in management, and 211 times in microbiology (N = 150). Once again, an a priori explanation might attribute this to the hardness of the discipline with microbiology, and to a lesser extent management, RAs placing an emphasis on replication of results that is not generally required, expected, or even possible in history (Munslow, 1997). However, an analysis of corpus lines reveals that pattern type three with the meaning of repeating an experiment in order to investigate whether this produces the same results or findings occurs most frequently in management where it accounts for 44.44 per cent of occurrences, not in microbiology where it accounts for 2 per cent:

> Pattern type three [[Human]] or [[Activity]] **replicates** [[Eventuality]] or {results, findings}
> Implicature: Science. Human = Scientist or Activity = Study repeats Eventuality = Experiment in order to investigate whether this produces the same results or findings

The confounding factor here is the presence of two pattern types which are closely related to microbiology. These are pattern type four which accounts for 3.33 per cent of history, 0.69 per cent of management, and 86.67 per cent of microbiology; and pattern type six which accounts for 10 per cent of microbiology, yet is absent in other disciplines:

Pattern type four: [[MB_Entity]] **replicates** (Self)
Implicature: MB Entity = DNA or Cell produces an exact copy of Self
Pattern type six: [[MB_Entity 1]] or [[Process]] **replicates** [[MB_Entity 2]]
MB Entity 1 or Process produces copy of MB Entity 2

These pattern types represent a pre-eminently discipline-specific use of *replicate* which has a clear specialist meaning in microbiology. This skews the sample somewhat, since taken together pattern types four and six generalise over 96.67 per cent of the occurrences of *replicate* annotated. If a discipline contains a conventionalised technical use of a given verb, it is possible that writers in this discipline will avoid other uses of this verb in order to prevent confusion with the technical term. This tendency for prevalent use of a pattern with a clearly specialist technical meaning is reflected in statistical comparisons of verb use across disciplines.

A related consideration concerns the adequacy of CPA semantic types for describing technical and scientific language. The semantic type [[MB Entity]] does not figure in the CPASO listed in the PDEV. The PDEV represents the meaning of the lines mapped to pattern type four using the pattern: [[Physical Object = DNA or Cells]] *replicates* (Self). In the PDEV noun collocates found with [[Physical Object]] include {stone, window, floor} the only lexical item found with scientific connotations was {ovum}. In contrast, a global view of the microbiology sub-corpus makes it clear that [[Physical Object]] does not adequately describe the kind of lexical items found in the sub-corpus, for example, {cell, DNA, gene, genome, strand}. Specifying these as contextual roles would have been an inelegant and time-consuming solution which did not faithfully reflect the corpus data. For this reason, and for the purposes of the present study, a semantic type representing microbiological entity: [[MB Entity]] was added. This is clearly permitted within the framework of CPA.

The idea of self-replication generalised by pattern type four is not unique to microbiology. It is also seen in history and management. The following corpus lines are found in the history sub-corpus of the AJACX2:

a. and political-economic constructs would **replicate** themselves across Africa—particularly
b. because they're good for us but because they **replicate** effectively. Dennett is well
c. for the carrier, but in some situations it **replicates** when the carrier's act inspires
d. carrier's act inspires others. But most memes **replicate** not just with our blessings, but
e. some metaphors clump together nicely and **replicate** successfully, but I haven't found

The following two corpus lines occur in the management sub-corpus:

f. innovations contain information to enable them to **replicate** themselves in a similar way to genes
g. which opens the possibility that firms **replicate** and adapt simultaneously, and I find that

In these lines the lexical item in subject slot is clearly not of the semantic type [[MB Entity]] or [[Physical Object]]. The lexical item {constructs} in a, the {memes} and {metaphors} referenced in (b) to (e), and {information} in (d) might best be generalised over by [[Concept]]. The lexical item {firm} in pertains to the semantic type [[Institution]]. In the annotation process it was decided that these apparently anomalous semantic types were evidence of a further pattern:

Pattern type five: [[Concept]] or [[Institution]] **replicates** ([[Self]])
Implicature: Concept or Institutions reoccur in different contexts

Pattern type five represents 16.67 per cent of the sample in history and 1.39 per cent in management; it is absent from the microbiology sub-corpus. The pattern might represent an exploitation of the microbiology-specific norm represented by pattern type four. An analogy is made between the physical replication of cells, DNA, or genes in microbiology and the more abstract denotation of the replication of concepts. This is made particularly clear by the simile present in (f) which can be paraphrased: *Innovations replicate themselves like genes*. In CG terms the expressions represented by these pattern types are meaningful because one novel expression which lacks unit status is similar to the other established expression with unit status. In this case it is not immediately obvious which is which.

In retrospect since they occur in such a limited context, four lines from two RAs in history and two lines from management, there is little evidence that the use shown in pattern type five has become conventionalised. It might have been more prudent, for the purposes of statistical comparison, to count these uses of *replicate* as exploitations of pattern type four. In any event, recalculating LL and BIC scores did not result in any meaningful difference. A comparison between history and management still reveals no significant difference. The comparisons with microbiology reveal difference at $P < 0.0001$ and very strong evidence against the null hypothesis of no significant difference.

From an EAP lexicography standpoint the limited context in which pattern type five is found means that it would not be advisable to include it in any EAP pattern dictionary. In contrast, the wide range of microbiological contexts in which pattern types four and six are found means that their inclusion in an EAP lexicographical resource aimed at those working in microbiology is essential. The marked difference in frequency of occurrence of these patterns representing distinct senses of the verb *replicate* provides further evidence against a general-academic approach to the selection of vocabulary for EAP courses and materials.

7.1.9 *Specify*

Although the CPA annotation revealed only three pattern types in the AJACX2 for *specify*, a more detailed examination of its collocational patterns suggests that representing its phraseology in a way that is useful to EAP users in the three disciplines is far from straightforward. Pattern type one represents 80 per cent of the sample in history ($N = 70$), 97 per cent in management ($N = 300$), and 46.30 per cent in microbiology ($N = 54$):

> Pattern type one: [[Human]] or [[Institution]] or [[Concept]] or [[Document]]
> **specifies** [[Anything]]
> Implicature: Human or Institution or Document or Concept states clearly and precisely that Anything is a necessary condition for some activity, Process, or Concept

An immediately obvious difference is the apparent semantic-type alternation which occurs frequently in management where many lexical items labelled as [[Human]] or [[Institution]] or [[Document]] could just as well be labelled as [[Concept]]. For example, in the line *a symmetric pure-strategy equilibrium specifies a pair of functions* the noun phrase in the object slot {a symmetric pure-strategy equilibrium} could legitimately be labelled as a [[Concept]]. This possible alternation is widespread in management where at least 40 out of 300 lines, some 13.33 per cent, could legitimately be labelled this way. A common lexical item in this slot is {model} and or {theory}. In contrast, none of the lines in the history or microbiology sub-

corpora display this alternation. An examination of the corpus lines in the sample used to derive the pattern in the PDEV reveals only two lines which could be legitimately labelled with [[Concept]]. During annotation these lines were not counted as separate patterns because concepts necessarily involve human cognition. However, give the relatively high frequency of occurrence of this alternation in management, any EAP lexicographical resource aimed at users working in management studies should list this discipline-specific alternation separately.

> Pattern type two: [[Human]] or [[Institution]] or [[Concept]] or [[Document]] **specifies** [[that]]
> Implicature: Human or Institution or Document states clearly and precisely the conditions implied by that [clause] as being necessary for some activity, process, or concept

Another salient finding from the statistical analysis of the collocational patterns of *specify* is the marked difference in frequency of occurrence of pattern type two across disciplines. In history this this pattern type accounts for 20 per cent of the sample, in management 3 per cent, while in microbiology there are no lines which map to this pattern type. Given the subject matter of history, it is hardly surprising that writers in this discipline report on the actions of humans and institutions and the contents of documents using this construction. At the opposite extreme is microbiology, where RAs focus on reporting experimental procedures and results rather what other humans have said. Once again management represents a middle ground since it contains both research reports, theoretical papers, and case studies. Here again, then, there is tentative evidence of a split between harder and softer disciplines. However, as was the case with *replicate*, the situation is complicated by the presence of a pre-eminently microbiological use of *specify* which represents 53.70 per cent of the sample in microbiology:

> Pattern type three: [[MB Entity 1]] **specifies** [[MB Entity 2]]
> Implicature: MB Entity 1 contains the instructions necessary for the creation of MB Entity 2

As the presence of the semantic type [[MB Entity]] in both subject and object slots suggests, in the AJACX2 pattern type three occurred exclusively in microbiology. It might be argued that this pattern type could be subsumed by pattern type one. An examination of the corpus lines used to derive the equivalent of pattern type one in the PDEV reveals one line which could map to pattern type three: *a stretch of DNA which specifies a single type of protein*. This line is marked as an exploitation of pattern type one. Given its low frequency of occurrence in the PDEV sample this decision seems justifiable. However, in the microbiology sub-corpus where it accounts for 53 per cent of lines with *specify* it no longer represents an exploitation but rather the norm. More importantly, there is a difference in sense:

a.
this step only the unidirectional relations **specified** in Figure 1 were examined. To assess the
additional study will be required to more clearly **specify** the role of specific emotions with respect
variables. Therefore, our nine hypotheses, **specified** in Table 5, propose that each independent
particular period. The planner may prefer to **specify** an internal service level. In order to
the analyst and expert to work together to **specify** several different partitions and explicitly

b.
presence of a region encompassing the genes **specifying** the oligopeptide transport system Opp [
encodes more proteins: while the IL1403 genome **specifies** 2270 predicted proteins larger than 60
1998). Gene expression levels are usually **specified** by particular promoter sequences and their
REase, whereas two genes have been shown to **specify** the LlaJI restriction activity. The MTases
its permeability to protons; hdeA, which **specifies** a periplasmic chaperon hypothesized to

The lines in (a) map to pattern type one. They were randomly selected from the management sample. The lines in (b) are judged to map to pattern type three. They were randomly selected from the microbiology sample. The semantic type of the lexical items in subject slot in (a) is [[Concept]] or [[Document]] or [[Human]]. There is no prototypical lexical set in object position therefore the semantic type [[Anything]] is assigned. This results in the rather general implicature: *Human or Institution or Document or Concept states clearly and precisely that Anything is a necessary condition for some activity, process, or concept*. In (b) the lexical items in both subject and object positions denote genome, genes, parts of genes and other parts of cells therefore the semantic type [[MB Entity]] was assigned. This results in the implicature *MB Entity 1 contains the instructions necessary for the creation of MB Entity 2* which implies a far more specific sense than the implicature of (a). Given that lines which map to pattern type three abound in microbiology, it stands to reason that this pattern should be listed in any lexicographical resource aimed at people working in microbiology. Having said that, pattern types one and three are clearly related. It is possible that pattern type three is the conventionalised form of what was originally an exploitation of pattern one. This exploitation may have been prompted by the need to express a novel concept, namely genes or genetic instructions, in a way that could be easily understood. By exploiting pattern one, an analogy is drawn between a human or document giving instructions to achieve a task, and a gene or other unit of genetic material containing instructions for the creation of a cell or other microbiological entity. Originally, this would have been rhetorically effective; today it has become conventionalised or entrenched.

In summary, a closer examination of the corpus lines containing *specify* in the AJACX2 reveals findings with implications for the representation of phraseology in an EAP lexicographical resources. In regards to pattern type one the alternation of [[Human]] or [[Institution]] with [[Concept]] is particularly salient for those working in management. The prevalence of pattern type two in history and its relative absence in the other disciplines under study suggests that this sense of *specify* would be particularly important for those working in disciplines where reporting speech and the content of documents is essential. Neither of these patterns are likely to prove problematic from a receptive standpoint. The same cannot be said for pattern type three which represents a pre-eminently discipline-specific sense of the verb *specify* which might prove problematic for users accustomed to the more general sense of this verb. This sense would have to be included in any lexicographical resource aimed at users working in microbiology. Pattern three also calls into question the validity of a statistical comparison of the frequency of patterns when it is apparent that in one discipline the verb has a clear specialist technical meaning. In cases like this it could be argued that writers in this discipline might eschew uses of *specify* with a more general meaning to avoid confusion.

7.1.10 *Treat*

Like many of the verbs analysed, the statistical analysis of *treat* revealed a marked contrast in phraseology between history and management on one hand, and microbiology on the

other. In the case of the first two pattern types, this can be explained by the semantic type of the recipient of the action and the subject focus of each discipline.

> Pattern type one: [[Human 1]] or [[Institution 1]] or [[Animal 1]] **treats** [[Human 2]] or [[Animal 2]] or [[Entity]] or [[Event]] Manner
> Implicature: Human 1 or Institution 1 or Animal 1 behaves toward Human 2 or Animal 2 or Entity or Event in the Manner specified
> Pattern type two: [[Human 1]] or [[Process]] or [[Drug]] **treats** [[Human 2]] or [[Animal]] or [[Disease]] or [[Injury]]
> Implicature: Human 1 = Health Professional applies a Drug or Process = Medical to Human 2 =Patient for the purpose of curing the patient`s Disease or Injury

Pattern type one with its emphasis on treating someone or something in a certain manner represents a high proportion of the sample in history where it accounts for 81.33 per cent of occurrences of the verb, and in management where it accounts for 90.67 per cent. In microbiology it represents only 7.33 per cent of the sample (N = 150 in all sub-corpora). Pattern type two which implies a health professional applying a drug to treat a patient is far more frequent in microbiology where it accounts for 39.33 per cent of the sample compared with 6 per cent in history, and 5.33 per cent in management. Pattern type three occurs predominantly in microbiology with 10 occurrences representing 6.67 per cent of the sample. There are three other instances of this pattern type which occur in management representing 2 per cent of the sample. The latter all come from the same source RA. Nothing about the semantic types of this pattern suggests anything inherently microbiological:

> Pattern type three: [[Human]] or [[Device]] **treats** [[Inanimate]] (with [[Stuff]] or by [[Process]])
> Implicature: The chemical or other properties of Inanimate are improved or otherwise changed by Process or the application of Stuff

However, CPA is an inductive rather than a deductive process. Meaning is mapped bottom-up from the corpus to the pattern rather than top-down from pattern to corpus. The following corpus lines are taken from the microbiology sub corpus:

a. was isolated from a pilot plant bioreactor **treating** arsenopyrite/pyrite in Kazakstan . It oxidizes
b. estimates of the total quantity of copper ore **treated** by microbially assisted processes are difficult
c. proved to be robust and have successfully **treated** gold-bearing concentrate for extended periods
d. bacteria was found in pilot-scale bioreactors **treating** a nickel pentlandite-pyrrhotite ore. In
e. altered the microbial landscape. A soil **treated** with sludge containing either high or low
f. in the nitrifying reactor systems used to **treat** water in the large saltwater aquaria at
g. The photo mask was then **treated** with an antiadhesion pellicle coating (
h. the marine environment is only partially **treated** and is not disinfected (examples of partially
i. wastewater treatment plants collect and **treat** community effluent, evaluation of viruses
j. processed via a washer and the effluent is **treated** in a nitrification/denitrification system

Aside from some passive alternations which place emphasis on the treatment process rather than the person or device administering the treatment these lines map to pattern type three. When they are expressed, arguments in the subject slot have the sematic type [[Device]] while those in the object slot typically have the semantic type [[Inanimate]]. Any additional arguments usually involve chemicals or chemical processes. Lines (k) to (l) below are slightly more problematic; while lexical items in the other argument slot {listwise deletion} and {incorporating…} can feasibly be labelled as [[Process]], the object slot is

filled by {missing values} and {potential autocorrelation}. These lexical items would be better grouped as [[Abstract Entity]] than [[Inanimate]].

k. full model (Model 7). Missing values were **treated** with listwise deletion, which accounts
l. observations after listwise deletion was applied to **treat** missing values. Missing values occurred
m. than interfirm variation in performance. I **treated** potential autocorrelation by incorporating

This exploitation of pattern type three reflects the human tendency to draw analogies between abstract concepts and physical experience. This is important in many of the cognitive theories of language outlined in Chapter 1. In this case the abstract denotation of {missing values} resonates with the physical experience of processing materials. It would unwise to make generalisation about the prevalence of this kind analogy on the basis of a handful of corpus lines taken from a single source RA. However, for many speakers the idea of data as a material which can be mined is very much conventionalised. For example, the term *data mining* is often used synonymously with *data processing*. Data is talked about in similar terms to an ore or mineral to be exploited. With this in mind, the fact that lexical items denoting data in (k) to (m) occur in a similar context to lexical items denoting mineral ores occur in (a) to (d) is perhaps not coincidental.

Instances of *treat* with the meaning reflected in pattern type five are found in all sub-corpora. In history they represent 10 per cent of the sample, in management and microbiology they represent 2 per cent of the sample.

> Pattern type five: [[Human]] **treats** [[Anything]]
> Implicature: Human discusses Anything = Topic

This pattern has two clear functions in the AJACX2. It is used to reference relevant work from other sources for example "*[…] the history of the English church have been treated at length by others*", and "*these topics are covered in excellent recent review articles […] and will not be treated in detail here*". It also serves as a means of organising discourse and alerting the reader to the structure and content of the RA: *"In what follows I treat viewing and responding to the Sputnik replicas"*. History is an extremely intertextual discipline (Coffin, 2006), historians cite sources with great frequency, for this reason the relatively high proportion of the history sample represented by pattern type five is not surprising. Previous research has found that discourse organising patterns are particularly prevalent in the abstract sections of RAs (Hyland, 2004). Since abstracts were not included in the AJACX2 such patterns might be underrepresented in the corpus.

> Pattern type six: [[Human 1]] **treats** with [[Human 2]]
> Implicature: Human 1 reaches a peace agreement with Human 2

Pattern type six is extremely infrequent in the AJACX2 its three instances occur in history representing two percent of the sample. However, this intransitive use of a normally transitive verb is salient. The absence of this pattern from the PDEV and the other AJACX2 sub-corpora suggest that this pattern represents a discipline-specific norm. In the online OED the oldest citation for the verb *treat* fits this pattern ("treat, v," 2017). It is not a use listed in any of the latest editions of the big five MLDs.

> Pattern type two: [[Human 1]] or [[Process]] or [[Drug]] treats [[Human 2]] or [[Animal]] or [[Disease]] or [[Injury]]
> Implicature: Human 1 = Health Professional applies a Drug or Process = Medical to Human 2 =Patient for the purpose of curing the patient`s Disease or Injury

Pattern type seven: [[Human]] **treats** [[MB entity 1]] or [[Animal]] with [[MB entity 2]] or [[Drug]]
Implicature: Human introduces drug or MB entity 2 to MB Entity 1=cell or Animal for experimental purposes

The statistical testing reported in the previous chapter demonstrates that both pattern types two and seven account for a significantly larger proportion of the microbiology sample than they do in the other disciplines under study. Pattern type seven occurs exclusively in microbiology where it represents 43.33 percent of the sample. Pattern type two accounts for six per cent of the history sample, 5.33 per cent of the management sample, and 39.33 per cent of the microbiology sample. Formally, the patterns are similar, however, while pattern type two activates a clearly medical meaning of *treat* and is often followed by (with [[Drug]] or with [[Process]]), in pattern seven the argument: *with [[MB Entity 2]] or [[Drug]]* is obligatory and activates a distinct verb meaning related to experimentation.

The phraseological behaviour of *treat* has clear implications for its representation in EAP lexicographical resources. Pattern type three should certainly be included in lexicographical resources for disciplines such as physical sciences and engineering. Further study is needed to establish whether the analogy with data and physical materials in the management sub-corpus is widespread and justifies broadening the pattern and including it in resources aimed at other disciplines. Previous research suggests that pattern type five: [[Human]] *treats* [[Anything]], plays an important role in organising discourse and reporting previous findings in academic writing in general. Evidence from the present study suggests that this role is particularly important in history. Pattern type six reflects a use of *treat* that was once prototypical in general usage. Comparison with current general dictionaries and MLDs suggests that this is no longer the case. This intransitive use of a normally transitive verb could certainly prove problematic from both productive and receptive perspectives. For this reason it should be included in an EAP lexicographical resource aimed at users working in history. Despite the formal similarity of their phraseology, lexicographically patterns two and seven should be treated differently. Although pattern two represents a medical use of the verb *treat* its occurrence in other corpora besides microbiology suggests that this use is generalised enough to merit inclusion in a general-academic lexical resource. In contrast, pattern type seven occurs exclusively in microbiology and implies a sense of carrying out experiments on living organisms or parts of living organisms, this suggests that it should be included in any lexicographical resource aimed at those working in the biological sciences. However, its relevance for EAP users working in other areas can be questioned.

7.2 Summary of the Analysis

Both statistical and qualitative analyses of the AJAXC2 point to a clear division in the collocational behaviour of verbs: In terms of their collocational patterns, verbs in the history and management studies sub-corpora tend to behave similarly, and in marked contrast to verbs in microbiology. The verb *explain* is a notable exception to this trend. Common factors which might explain this split include differences in the degree of consensus required in each discipline. This is evidenced by the behaviour of *accept*. The varying need to express conation is another factor. This is apparent in different prototypical meanings of *manage* across disciplines. Variation in disciplinary discourse conventions also plays a role. This can be seen in the uses of *explain* and *note* to report

quotations. Similarly, differences in disciplinary move structure are hinted at by some uses of *treat*.

The principal aim of this study is not to account for differences in the collocational behaviour of verbs across academic disciplines, but to demonstrate how this varies across disciplines. It has a secondary aim of investigating how these differences might be represented in lexicographical resources in a way that is useful to EAP users. Auguste Comte's (1896) notion of a hierarchy of sciences is helpful in this task. The split between history and management on one hand, and microbiology on the other could be extrapolated to a general division between hard and soft sciences. These could be termed harder disciplines and softer disciplines or, in more user friendly terms, even *arts and humanities* and *sciences*. Irrespective of labels, these categories offer a feasible way of grouping EAP vocabulary. The creation of individual resources for every discipline would, in many cases not be a viable option. Respecting this apparent division offers a practical solution. Discipline-specific EAP resources might take the form of separate resources for hard and soft disciplines, or a single resources with information on verb usage in hard and soft disciplines clearly labelled as such.

A discipline-specific EAP lexicographical resource does not entail the exclusion of general academic vocabulary. Patterns which represent verb meanings which are frequently used in all disciplines irrespective of significant differences in frequencies should be included in both general EAP resources, and multi-discipline resources. For example, the pattern: [[Human]] or [[Institution]] *accepts* [that-clause], represents a significantly higher proportion of the sample in microbiology than it does in the history or management. Nonetheless, it is still used frequently in all the disciplines studied. Conversely, patterns which predominate in one discipline yet appear at extremely limited frequency in the others should only be included in a resource aimed at users working in the dominant discipline, or be clearly marked in a general EAP resource. For example, the pattern [[MB Entity]] *replicates* (Self), accounts for 130 out of 150 occurrences of the verb in microbiology (86.67 per cent), but only one occurrence in history (3.33 per cent) and one in management (0.69 per cent). The limitations of using a solely frequency-based approach for vocabulary selection are once again apparent here. An examination of the corpus lines reveals that the instances of this pattern which are not from the microbiology sub-corpus are, in fact, describing microbiological processes. This is also evidence of interdisciplinarity and the limitations of bibliographical classification of RAs.

The analysis process also revealed several alternations which are more common in some disciplines than others. Chief among these are semantic-type alternations. Although many of these alternations are widespread in general English, some of them tend to occur far more frequently in some disciplines than others. This is the case with alternation of [[Human]] with [[Location]] in collocation with *accept* in history, or [[Concept]] with [[Human]] with many verbs including *specify* in management. Many of the alternations can be described as transparent in so much as the rationale which underlies them can be straightforwardly discerned. The [[Human]] / [[Location]] *accepts* alternation is an attempt to emphasise the geographical location where the event took place. The [[Human]] / [[Concept]] is understandable since concepts are the result of human cognition. Other alternations are less transparent, for example the alternation of [[MB Entity]] with [[Human]] in collocation with *accomplish* has little relation to cognition. This alternation occurs uniquely in microbiology and as such represents a truly discipline-specific alternation. It can, along with many similar alternations, perhaps best be explained by the need of microbiologists to explain once novel concepts in familiar terms. In TNE terms a

pattern activating a normal meaning of the verb was exploited to express a new concept. Over time this exploitation has become conventionalised. Taking a broad overview of academic language, this conventionalised exploitation represents an alternation since it only occurs in one discipline. For the purposes of this study it is labelled a discipline-specific pattern.

The lexicographic representation of all transparent and widespread semantic-type alternations would be unfeasible and would lead to a great deal of redundant representation. However, brief explanatory notes on these alternations might prove useful for EAP writers. True discipline-specific semantic-type alternations deserve fuller treatment and should be described in any discipline-specific EAP vocabulary resource.

In addition to semantic-type alternations, analysis also revealed syntactic alternations. Key amongst these was the hegemony of the alternation represented by the pattern: {it} is ({generally} {well} {widely}) *accepted* {that} in microbiology. The prevalence of passive alternations with certain verb patterns in certain disciplines, for example *apply* in microbiology, also has implications for lexicographical representation. It would not feasible to represent every syntactic alternation for every verb, however, data from the AJACX2 suggests that the current practice of representing patterns in active and declarative form is not always justified. In a lexicographical resource aimed at microbiologists the prototypical sense of *accept* should be represented by the pattern: {it} is ({generally} {well} {widely}) *accepted* {that}. Similarly, the sense of *apply* in microbiology represented by: [[Human = scientist]] *applies* [[Stuff = sample or culture]] to [[Surface = test medium]] might better be represented in a passive construction.

Assigning a function to patterns allows them to be grouped in ways which may be useful to the EAP community. For example certain patterns with *explain, note*, or *specify* might be grouped under the title *reporting what has been said or written*.

The analysis of the AJACX2 revealed examples of exploitations such as the exploitation of *apply* where [[Weapon]] is coerced to [[Process]] in the history sub-corpus discussed above. These exploitations suggest a good deal about how normal uses of verbs are exploited for pragmatic and rhetorical effect, however, they are not yet conventionalised and may never become conventionalised and as such should not be included in an EAP lexicographical resource.

7.3 Guidelines for Lexicographical Representation

The construction of a fully-fledged lexicographical resource was never envisaged as an immediate outcome of the present study. The principal aim has always been to empirically investigate the phraseological behaviour of vocabulary across academic disciplines. A secondary, though nonetheless important, aim is to demonstrate how insights from such an investigation might be applied to the creation of examples of, and guidelines for producing, lexicographical resources which are useful for EAP users. The AJACX2 data alone are not the basis upon which to build a lexicographical resource. While they are adequate for a comparison of verb phraseology and by extension meaning across academic disciplines, the construction of lexicographical resource would require a larger corpus, and other supplementary sources of data. The adequacy of the AJACX2 is discussed in further detail in the final chapter of this dissertation.

The lack of statistical testing in the majority of the finer-grained analysis reported in this chapter also raises questions of methodological adequacy. Most of the differences observed in the qualitative analysis were widespread and systematic, although, the observations they are based on were not subject to any statistical testing. However, considered independently the statistical analysis reported in the previous chapter was adequate to the aim of demonstrating that patterns activating different verb meanings occur at significantly different frequencies in distinct academic disciplines. The qualitative analysis reported in the present chapter provides a fuller picture indicating of the kind of differences a useful EAP lexicographic resource needs to reflect.

Thus far in this dissertation, the vague term 'lexicographical resource' has been purposefully employed. This is because the digitisation of lexicography has opened means of accessing lexicographical information far more varied than those which might traditionally be thought of as dictionaries. To use the term dictionary might unnecessarily and unhelpfully create a limited impression of the ways in which the findings of this study might be applied. At the same time, a set of nebulous abstract guidelines would be of little practical use. With these considerations in mind the guidelines and examples set out in the remaining sections of this chapter refer to a corpus-based online pattern dictionary of EAP verbs. In line with most pattern and collocational dictionaries, it is envisaged that its principal function is encoding. However, it may also prove useful in decoding. The decision to concentrate on this kind of resource is a logical consequence of the CPA methodology employed in the experiment at the heart of the present study. An online pattern dictionary also offers a middle ground between traditional paper-based dictionaries and more novel lexicographical resources such as those integrated with text editors, web-browsers, e-readers, and pedagogical activity generators. Insights from an idealised pattern dictionary could be applied, with minimal changes, to both novel and traditional contexts.

Irrespective of the form a dictionary takes an important first step must involve the definition of the audience, and, in the case of commercial resources, the market. The present study adopts the traditional view of EAP as English for teaching, learning, and research that takes place at universities and other similar higher education institutions. EAP users are conceived of as predominantly upper-intermediate (B2) to advanced (C1/C2) level users who need to produce and understand English in academic contexts. It is possible to create guidelines for, and examples of a dictionary aimed at EAP users working in a particular discipline, or those working in harder or softer disciplines. However, a multi-discipline dictionary would meet the needs of a larger audience and present greater opportunities to apply the results of the analysis above. It would also prove particularly useful in the current academic context of increasingly interdisciplinary research. A multi-discipline dictionary should not be understood as referring to a dictionary comprising general-academic or semitechnical vocabulary in the same way general-academic wordlists and certain MLDs do. Instead, it should be understood as striving to include entries for those patterns which are used across many academic disciplines, as well as discipline-specific patterns in an appropriately marked manner.

The selection of items for inclusion in the word list of the lexicographical resource presents a greater challenge than defining the potential audience. The approach to verb selection adopted in the corpus-based experiment at the centre of the present study took the 1286 PDEV verbs which had been completed at the time of sampling as a starting point. In order to be considered for inclusion in the sample for analysis these verbs had to occur at least 20 times; the point at which a trained lexicographer can begin to make judgements about verb meaning (Sinclair, 2005). A threshold of verb appearance in ten different articles

written by distinct authors was also adopted to ensure the idiosyncrasies of a particular author were not overly represented. Limiting the study to the completed PDEV verbs ensured a useful reference for difficult cases in annotation and thus increased annotation reliability. However, although this procedure was essential to a reliable comparison of the phraseology of verbs across disciplines, it limited the range of language which could be studied and has consequences for the representativeness of the verbs under study. The verb list in the PDEV is based on frequency of occurrence measure in the BNC50. This is problematic for two reasons: Firstly, since the BNC is a general corpus it does not necessarily reflect frequencies of occurrence of verbs in academic language; secondly it is over twenty years old. This is particularly problematic for the representation of meaning in fast-moving science and technology disciplines. For these reasons, it is not the foundation on which to build an EAP lexicographical resource which truly reflects the range of language which users encounter and have to produce.

The analysis above has illustrated the benefits of relative frequency as a means of discriminating between general academic patterns such as [[Human]] or [[Institution]] *accepts* [that-clause], and discipline specific patterns such as [[MB Entity]] *replicates* (Self). Frequency also offers a means of populating the wordlist. In the *E-Advanced Learner's Dictionary of Verbs in Science* (G Williams & Millon, 2010, 2014; G Williams, Millon, & Alonso, 2012) collocational networks are used to select headwords. Starting from the 100 most frequent verbs in a corpus of biomedical texts statistical procedures are used to find most the salient collocates. These collocates are, in turn, networked to other verbs leading to growing list of headwords populated bottom-up from the corpus, rather than top-down from the lexicographers' intuition about users' needs. Although there are a number of issues to perfect with this approach, such as which verbs to include in a stop-list, and the point at which to stop including new words in the lexical resource, this procedure offers a principled approach to headword selection for specialist dictionaries.

The use of corpora to select headwords for specialist dictionaries is certainly more valid than relying on intuition. However, since no corpus could attest to language use in its entirely, there exists the possibility of silence in the absence of data. Feedback from the EAP community, subject experts, and dictionary use studies is fundamental to noticing these silences. More generally, although the guidelines and examples presented below are based on a thorough review of dictionary use studies, research on the presentation of phraseology and academic vocabulary, and the results of the analysis above; feedback and user studies present the only way of evaluating their effectiveness for the creation lexicographical resources. It should be clear from the discussion in Chapter 3 that meaningful dictionary use studies are complex and time-consuming undertakings. To undertake such a study in addition to the corpus-based comparison reported above is beyond the scope of this dissertation. However, it does offer an interesting avenue for future research.

Access structure is another important consideration in the creation of lexicographical resources. One of the principal advantages of digital lexicographical resources is that they are not constrained by traditional alphabetical ordering. Search functionality offers many possible means of access to an entry. Research suggests that users prefer simple one-step search functionality (Bank, 2010; Heid & Zimmermann, 2012). For this reason, in a dictionary of verbal collocations, users should be able to search for any form of the verb and be taken directly to the relevant entry. This is consistent with the principle of using as few metalinguistic terms as possible which is apparent from several dictionary use studies (Béjoint, 1981; Bogaards & Van der Kloot, 2002; Nesi & Haill, 2002; Neubach & Cohen,

1988). It should be clear from the literature review and analysis that CPA offers the possibility of arranging verbs in functional or semantic classes. This facilitates the creation of a lexicographical resource that is both semasiological and onomasiological. It is apparent from the annotation process, and has been suggested elsewhere that CPA semantic types are not immediately understandable to the user (G Williams & Millon, 2010). For this reason alternative labels must be considered. Hunston and Francis's (2000) "notional groups" (p. 109) represent, perhaps, a more user-friendly means of classifying verb patterns. It is easy to see how descriptors such as *describing agreement* might be used to index patterns representing this meaning of *accept*: [[Human]] or [[Institution]] *accepts* [[Proposition]] or [[Concept]] or [[Eventuality]], [[Human]] or [[Institution]] *accepts* [that-clause], as well as the *it is accepted that* alternation which represents the norm in microbiology. Similarly, *reporting what has been said or written* could group those patterns for the verbs *explain*, *note*, *specify* judged to have a reporting function in the analysis above. This approach could also work for discipline-specific patterns. The descriptor *talking about the creation of cells or DNA* could group the following microbiological patterns:

> *Replicate* pattern type six: [[MB_Entity 1]] or [[Process]] replicates [[MB_Entity 2]]
> Implicature: MB Entity 1 or Process produces copy of MB Entity 2
>
> *Specify* pattern type three: [[MB Entity 1]] specifies [[MB Entity 2]]
> Implicature: MB Entity 1 contains the instructions necessary for the creation of MB Entity

Ideally users should be able to search for these descriptors in natural language without having to resort to any special syntax. The conventions of CPA are not always conducive to this aim. CPA involves mapping the various specific meanings expressed by corpus lines to general patterns, this is an adequate means of achieving the semasiological function of distinguishing verb meaning, however, it is difficult for writers to work back from the general to the specific. For this reason, it might be advisable that patterns list typical members of a lexical set rather than the CPA semantic types they belong to. This would have the additional advantage of allowing users to search for patterns by common collocate. This function would also be particularly helpful in decoding. On this approach, the pattern [[Human]] or [[Institution]] *accepts* [[Proposition]] or [[Concept]] or [[Eventuality]] might be better represented:

> a person OR organisation ***accepts*** something is true OR a theory or concept or model explains something OR something has happened

Representing patterns in this way fulfils many of the criteria a user-friendly defining style that Hanks (1987) describes as "dictionary as prose" (p. 118). This style is prose-like and avoids metalanguage and the use of parenthesis; both of which are factors found to be particularly problematic by several dictionary use studies discussed in Chapter 3 (Nesi & Haill, 2002; Nesi & Meara, 1994). It takes only minor adjustment to convert the user-friendly pattern above to a COBUILD style definition like that for the second sense for *accept* listed in the learner section of Collins Dictionary Online:

> If you **accept** an idea, statement, or fact, you believe that it is true or valid.

This is also the definition style which fared most positively in the user studies on defining style discussed in Chapter 3 (Cumming et al., 1994; Nesi & Meara, 1994). It is hardly surprising given the dictionaries shared heritage and corpus-based methodology that

PDEV-type implicatures can be easily converted to COBUILD style definitions. For example, compare the COBUILD style definition above with the PDEV pattern below:

> PDEV pattern one: [[Human]] or [[Institution]] accepts [[Proposition]] or [[Concept]] or [[Eventuality]]
> Implicature: Human or Institution agrees that Proposition or Concept or Eventuality is correct and does not need to be contested

There is some evidence to suggest that users perceive definitions written with a restricted defining style as more intelligible (MacFarquhar & Richards, 1983). However, not only do they have the potential to give a misleading impression of word meaning (Hanks, 1987), but defining scientific and technical terms with a restricted vocabulary has been shown to be particularly challenging (Herbst, 1986). For these reasons a restricted defining vocabulary is avoided here.

Lew and Dziemianko's (2012) contention that folk-defining style definitions do not contain the information necessary for Polish speaking users to determine the part-of-speech of the headword, is not relevant in a pattern dictionary of verbs. More pertinent is Teubert's (2001) claim that well-selected citations can be of greater benefit to the user than definitions in encoding dictionaries. However, the free agency of the user makes a secondary decoding role a real possibility. The combination of user-friendly patterns and well-chosen examples should suffice in allowing users to grasp meaning. An additional benefit of not converting patterns to definitions is that it avoids problems of using *if you* statements with morally reprehensible acts, for example, *if you murder someone, you...* and other problematic culturally-bound situations (Hanks, 1987). In any event, research suggests that examples, not definitions, are the first port of all for most non-linguist users (Bogaards & Van der Kloot, 2002; Dziemanko, 2006, 2012).

The need to distinguish between examples intended to help users with encoding and examples intended to help students with decoding is gaining greater acceptance. Humblé (2001) highlights the need for examples to contain contextual clues in order to help user with decoding, and patterns of syntax and collocation to help users with encoding. The results of Frankenberg-Garcia's (2012, 2014, 2015) studies lend support to this argument.

In a corpus-based pattern dictionary, patterns distilled from corpus lines bridge the gap between traditional definitions and examples. In many cases, the CPA patterns from the AJACX2 provide the type of information which helps users with encoding. Take, for example, the many patterns with [to-infinitives] or [that-clauses] or the complementation information in pattern type three for *note*:

> [[Human]] **notes** [[Information]] (on or in [[Document]]) (under [[Document Part]])
> User-friendly pattern: a human notes information SOMETIMES on OR in a document OR SOMETIMES under a heading

The availability of several corpus-based examples for each pattern make the use of parenthesis to indicate optionality unnecessary. The possibility of linking to an abundance of corpus lines should, in most cases, provide sufficient contextual clues to facilitate comprehension. If further information is needed there is, in theory, the possibility of linking from corpus lines to the wider context of the source document.

7.3 Example Entries

The entries below provide an illustrative example of how the guidelines above might be put into practice.

7.3.1 Entry for the Verb *Apply*

Apply

1. a person OR organisation ***applies*** theory OR model OR process SOMETIMES to address a situation

reason-of-state literature, may have been **applied** to practical politics. But the crucial [HISTORY]
This has made it difficult to effectively **apply** and combine theoretical lenses, to delineate [MANAGEMENT]
discouraged, bacteriologists did not attempt to **apply** the Biological Species Concept (or any [MICROBIOLOGY]

2. a theory OR model OR process ***applies*** SOMETIMES to a situation OR activity

"certain ""natural"" principles of morality" **apply** even without government or religion to
diversification also does not **apply** to a single-business CEO. Accordingly, [MANAGEMENT]
CLSI screening and confirmatory tests **apply** only to Escherichia coli, Klebsiella spp [MICROBIOLOGY]

3. a person ***applies*** for benefit OR an injunction OR an extension OR an admission OR readmission

of supporters and economic resources in **applying** for exemptions. The revised law would not [HISTORY]
upon further promotion. In 1809 Liverpool **applied** to the prince regent for a grant of £1,000 [HISTORY]
list. Firms included on the WM list must **apply** for recognition. This raises the possibility [MANAGEMENT]

4. a sample OR culture is ***applied*** to a surface OFTEN a test medium SOMETIMES by a scientist [MICROBIOLOGY]

The dialysed sample was **applied** to a hydroxyapatite column
100 µl of culture was **applied** to the face of the wedge array channel
700 µl of the final concentrated sample was **applied** to DNA extraction

5. a person ***applies*** a word OR term to anything

"Solidarity" was a term that could and did **apply** to different, overlapping, and sometimes [HISTORY]
Menu is used here as a metaphor that can be **applied** to a wide range of service interactions [MANAGEMENT]
The term Trojan horse has been **applied** to many biological packages, such as mobile [MICROBIOLOGY]

6. a person OR an organisation ***applies*** pressure

the crown initially **applied** pressure to compel religious orthodoxy [HISTORY]
viral mutagens **apply** less selective pressure to a viral population [MANAGEMENT]
efficient. For example, many companies have **applied** accountability pressures on managers to [MANAGEMENT]

Figure. 7.1. An illustrative example of an entry for verb *apply*

In the user friendly patterns in the illustrative entry shown in Figure 7.1, operators such as OR, SOMETIMES, or OFTEN are preferred to the more traditional means of indicating optionality via parenthesis. This has the advantage of being more prose-like. It also provides a user-friendly means of conveying the approximate frequency of certain alternations using SOMETIMES or OFTEN. The discipline labels, in this case, [HISTORY], [MANAGEMENT], or [MICROBIOLOGY] indicate the discipline from which a corpus-derived example is taken. Discipline-specific patterns such as (4) above are indicated via a label under the headword. In patterns such as (3) which only occurs in two disciplines, the majority of examples come from the discipline where the pattern occurs

most frequently. In the case of (3) this is history. In any case, it is envisaged that the user is able to click on any example to view more corpus lines over which the pattern generalises, sorted by discipline. There is also the possibility of further context in the form of the source text from which the line is taken. The presentation of sense labels and salient collocates in white writing on a blue background is an attempt to imitate the signposting strategy in the LDOCE5. In an eye-tracking study, Dziemianko (2016) found this style to be the most effective at speeding up recognition and retention. However, there is a great deal more research needed in this area.

7.3.2 Example Entries for Functional Groups

Function: *reporting what has been said or written*

Explain

a person OR organisation ***explains*** that… SOMETIMES to another person

[Note for Historians]

Presenting the draft to Scott, Lindsay **explained** that its intention was to inspire respect [HIST]
objections against the Xiongnu marriage code, he **explains** that their purpose is to safeguard the [HIST]
experience. The cover sheet for the case **explained** that the task was to diagnose the causes [MANAGEMENT]

a person OR organisation ***explains*** "something that is said or written using quotation marks"

[Note for Management Students]

"domestic agriculture was of a purely subsistence nature" he **explained** to a local intendant" [HIST]
George Hulme **explained:** "I think what started me thinking about this was when…" [HIST]
as one shop manager **explained**: "If you look after my client, when I'm away…" [MANAGEMENT]

Note

a person ***notes*** "something that is said or written using quotation marks" OR that… OR what happens or is said or written, when, where, why something happens or is written

the sister recalling this story for the chronicle **noted**: "no more widows in our convent" [HIST]
participate in this study. It should be **noted** that research suggests that 71% of firms [MANAGEMENT]
of 50 samples by RT-PCR . These authors **noted** that while no correlation could be established [MICROBIOLOGY]

Specify

a person OR organisation OR theory OR model OR document ***specifies*** that…

[Not found in the microbiology sources used to make this dictionary]

labour'. The Vagrancy Act went further, **specifying** that the 'labour' to be undertaken by slaves [HISTORY]
he **specified** that the "cowries come from the Maldives" [HISTORY]
group from another" (Hofstede, 1984: 21). It **specifies** how things are to be evaluated and what [MANAGEMENT]

Figure. 7.2. An illustrative example of an entry for the functional group *reporting what has been said or written*.

The conventions for displaying patterns adopted for onomasiological entries are much the same as those for semasiological entries. A key difference is the order in which patterns appear. In Figure 7.1 the overall frequency of occurrence of a pattern type in the samples analysed was used to order the patterns. However, this may not be appropriate when an entry includes different verbs. In Figure 7.2 alphabetical order is used.

The illustrative entry in 7.2 contains an example of the type of explanatory note mentioned in the analysis above. On clicking [Note for Historians] the user would see:

> In the history articles used to make this dictionary there was a slight tendency to report information using *explain* and a paraphrase rather than a direct quote

While clicking the [Note for Management Students] would reveal:

> In the management studies articles used to make this dictionary this pattern was often used to report participant responses in interviews

Functional groups can also be discipline specific as illustrated by Figure 7.3 below.

Function: *talking about the creation of cells or DNA*

Replicate

a cell OR part of a cell OR piece of DNA ***replicates*** another cell OR part of a cell OR piece of DNA

reporter (R) and quencher (Q) dyes. As DNA pol **replicates** the template strand, hydrolysis of the [MICROBIOLOGY]
organisms are able to efficiently and faithfully **replicate** their DNA are of critical importance in [MICROBIOLOGY]
P proteins to transcribe and eventually **replicate** the HPIV genome . The P protein of HPIV [MICROBIOLOGY]

Specify

a genome OR cell OR part of a cell OR piece of DNA ***specifies*** another cell OR part of a cell OR piece of DNA

the genes busAA-AB, **specifying** a high-affinity betaine uptake system [MICROBIOLOGY]
arcABD1C1C2TD2 in which arcD1 and arcD2 **specify** putative arginine ornithine (1:1) antiporter [MICROBIOLOGY]
Gene expression levels are usually **specified** by particular promoter sequences and their [MICROBIOLOGY]

Figure. 7.3. An illustrative example of an entry for functional group *talking about the creation of cells or DNA.*

The qualitative analysis in this chapter complements the statistical analysis reported in the previous chapter. The findings support the need for a discipline-specific phraseological approach to the selection of vocabulary for EAP courses and materials and show the limitations of a single-word general-academic wordlist approach. In doing so, they highlight the limitations of primarily frequency-based approaches to the study of phraseology such as the lexical bundle approach. They also demonstrate the feasibility of CPA in the study of academic lexis across disciplines. With reference to the review of dictionary use studies undertaken in Chapter 3, the present chapter has also posited some guidelines for the useful presentation of these findings in EAP lexicographical resources, as well as, brief illustrative examples of how these guidelines might be applied in an online EAP verb pattern dictionary. Along the way, a number of limitations of the present study have been suggested along with opportunities for further research. These will be discussed in more detail in the final conclusions chapter of this dissertation.

8. CONTRIBUTIONS, CHALLENGES, AND FUTURE DIRECTIONS

This chapter aims to summarise the contribution made by the present study and examine possibilities for further research. It begins by summarising the findings related to the principal objective before working back summarising further contributions made in the broad sections outlined in the introduction. Section 8.1 first addresses Chapters 4 and 5 dealing with the corpus-based experiment at the heart of the present study, and then Chapters 1, 2, and 3 which present the key concepts which underlie it. Section 8.2 addresses the challenges encountered in achieving these contributions. Section 8.3 discusses possible avenues for future research.

The principal objective of the present study was: To examine differences in meaning and use of vocabulary across academic disciplines which might be obfuscated by the distributional approaches to vocabulary selection habitually employed in English vocabulary studies. The findings reported in Chapters 6 and 7 clearly indicate that such differences exist.

Chapter 6 reported on a statistical comparison of the collocational behaviour of verbs, and by extension verb meanings, across the three sub-corpora of the AJACX2. It demonstrated that verb pattern types often account for significantly different proportions of samples in sub-corpora representing each discipline in the AJACX2. This finding runs counter to the rationale underlying a generalist approach to the selection of vocabulary for EAP courses and materials. The analysis also pointed to a split in collocational behaviour between history and management on one hand, and microbiology on the other. The chapter also reported on an intradisciplinary comparison which sought to control for the possibility that the differences observed between samples in the interdisciplinary comparison were due to the non-random nature of words in text. This second comparison involved two random samples for a given verb from the same discipline. None of the verbs included in this second comparison showed significant difference in collocational behaviour. This suggests that the differences in pattern type frequencies observed in the interdisciplinary comparison are due to a factor other than the non-random nature of language. Given the composition of the corpus it is reasonable to assume that the explanatory variable is academic discipline.

In reporting the findings of a qualitative analysis Chapter 7 went further than Chapter 6, it involved a finer-grained analysis of the corpus-lines of a random selection of ten of the 30 annotated verbs. This revealed semantic-type alternations which are more prevalent in some disciplines than others. For example, the synecdochical alternation of [[Location]] with [[Human]] in the subject slot of *accept* in history. It also revealed discipline-specific tendencies involving semantic prosody. For example, in management: [[Human]] or [[Institution]] *accepts* [[Eventuality = negative]]. These types of differences can often be explained by the focus of the discipline. It was argued that other types of differences such as the dominance of variations of *it is accepted that* in microbiology, and the presence of patterns indicating different conventions for reporting speech, can be linked to the prototypical function of research articles in the disciplines studied: In broad terms history articles argue, while microbiology articles report. Management studies RAs fulfil both functions.

Chapter 7 also reported corpus insights into metaphor as well as exploitation of pattern ambiguity. This can be novel and limited in scope; as is the case with the example of

magistrates applying balm to tender consciences, which exploits the prototypical pattern [[Human]] *applies* [[Concept]] or [[Process]] (to [[State of Affairs]), and resonates with medical uses of *apply* implying healing. These kinds of exploitations are employed to achieve a pragmatic effect. There is also evidence of the widespread use of what might have once been exploitation becoming conventionalised. For example, the widespread occurrence of [[MB Entity]] in slots prototypically occupied by [[Human]] with verbs involving human cognition. This may have come about as pioneer microbiologists sought to explain their new discoveries in terms of their lived experience.

Distributional approaches to vocabulary selection, such as the lexical-bundle approach, adopted in many other studies have strengths, however, they would have elided many of these insights. Given the increasing importance of standardised tests of academic English to academic and professional lives of EAP users, and the large sums of time and money invested in academic English courses and materials, it is essential that methods of vocabulary selection can truly capture the language that the EAP community encounters and strives to use. With this in mind the present study makes an important contribution to the study of EAP lexis.

8.1 Further Contributions

In the process of arriving at the findings summarised above, the present study has made a number of further contributions to the field. These include methodological contributions arising from the design of the corpus-based experimental procedure at the heart of the study. To prepare the ground Chapter 4 set out hypothesises to be tested in the corpus-based experimental procedure aimed at examining differences in the meaning and use of vocabulary across academic disciplines. With reference to previous corpus-based studies of academic lexis, and research on the language needs of users in EAP contexts, it argued that a corpus of research articles from academic journals could be considered reasonably representative of the language EAP users encounter and strive to produce. It also argued that on a phraseological approach to language the primacy of the noun in lexicographical studies can be called into doubt.

Chapter 5 detailed the corpus-based experimental procedure itself. The methodology adopted represents a novel application of the lexicographical technique of corpus pattern analysis. This technique was employed along with statistical hypothesis testing to compare the collocational behaviour, and by extension the meaning, of verb patterns across three sub-corpora which represent the academic disciplines of history, microbiology, and management studies. It is not only the decision to use statistical measures to compare CPA pattern types across disciplines which breaks new ground, the *type* of statistical measures employed is also somewhat innovative. Bayesian statistics have been successfully employed as effect-size measures in a small number of studies in the wider discipline of applied linguistics (A. Wilson, 2013), and are frequently used in microbiology (Sawcer, 2010), however the present study may have been the first to employ such measures in an examination of EAP lexis. In addition to this adaptation of CPA for comparative purposes, the present study was also methodologically innovative in its use of CAQDAS. Traditionally, ATLAS.ti has not been employed in corpus linguistics. However, in the wider discipline of applied linguistics, a small number of studies have proved the utility of ATLAS.ti in conversation analysis (O'Halloran, 2011; Paulus & Lester, 2015). It has also been employed in the examination of historic dictionaries (G Williams, 2017). The present study may have been the first to exploit ATLAS.ti in a study of EAP lexis.

The findings reported in Chapters 6 and 7 are testament to the adequacy of this CPA-based procedure as a feasible means of phrase extraction which accounts for both semantic and syntactic concerns. Demonstrating such a means of vocabulary selection was a secondary objective of the present study. This contribution is all the more significant still since the procedures set out in Chapter 5 could, with minor alterations, straightforwardly be applied to several other comparative applications in corpus-linguistics.

In addition to the methodological contributions outlined above, the present study has made some contributions which could be classified as pedagogical or lexicographical. Chapter 2 sought to situate vocabulary acquisition in EAP in the wider context of vocabulary acquisition in general. Through a review of literature on second language vocabulary acquisition, it sought to establish what knowing a word or phrase entails and thus uncover the rationale behind the productive/receptive dichotomy prevalent in language teaching. This was followed by a discussion about which words or phrases should be taught and learned. It was argued that the traditional recourse to the criterion of frequency should be tempered by the contextual needs of the user. In EAP contexts, an effect of the pre-eminence of frequency in approaches to vocabulary selection has been the dominance of wordlists such as the AWL. In addition to not reflecting differences in vocabulary needs across disciplines, traditionally, such lists have not taken account of phraseological concerns. In this regard, academic phrase-lists, be they general (Simpson-Vlach & Ellis, 2010) or discipline-specific (Hsu, 2014), present a major advance in the selection of vocabulary. However, these phrase lists are selected on the basis of variants of a frequency-based n-gram approach widely known as the lexical-bundle approach. On this approach strings of text, habitually of three or four words in length, occurring above a given frequency threshold are extracted. A central hypothesis of the present study is that the positive development represented by the phraseological turn in EAP vocabulary selection is limited by this form or distribution-based method of phrase extraction. A further question addressed in this review of the literature was how vocabulary can be most effectively taught and learned. In this respect, it was demonstrated that dictionaries and other lexicographical resources have a role to play in a great deal of vocabulary acquisition, procedures, strategies, and activities.

Chapter 3 summarised research on lexicographical resources in what could reasonably be considered pertinent contexts for EAP users. In light of a review of questionnaire-based, observation-based, and test-based research into dictionary use and users, it has argued that a broad EAP community's lexicographical needs would be best met by a monolingual lexicographical resource. In addition to providing an appraisal of existing EAP lexicographical resources, it also reviewed research into specific aspects of dictionary use such as defining and example style, visual elements, and the treatment of phraseology in monolingual learners' dictionaries and phraseological dictionaries.

The product of the insights from the review carried out in Chapters 2 and 3 and the findings reported in Chapters 6 and 7 fulfils a secondary objective of the present study: To create guidelines for producing, and an illustrative example of, a useful lexicographical resource for the EAP community. Guidelines for the creation of an online pattern dictionary of EAP verbs and the illustrative examples of its entries can be found in Sections 7.3 and 7.4 respectively. This marks a minor contribution of the present study to the field of EAP lexicography.

The methodological and lexicographical contributions discussed thus far could be classified as practical or applicable in nature. To the extent that they chime with phraseological

accounts of language, some of the findings of the present study might also be considered contributions to theory. Chapter 1 sought to outline, in broad terms, some major currents of linguistic theory which, in contrast to generative approaches, envisage a central and organising role for the phrase in language. It also sought to present a practical technique by means of which many of the insights gained from this this review of theory could be applied to the study of academic lexis. In this regard it has demonstrated that Hanks's (2013) theory of norms and exploitations and its practical counterpart CPA rely heavily, not only on insights gained from other corpus linguists, for example, Sinclair's (1991) Idiom and Open-choice Principles and Hoey's (2005) notion of Lexical Priming, but also on concepts from other linguistic currents. For example, TNE reflects a usage-based model of the lexicon (Bybee, 2006) compatible with those posited in Cognitive Grammar (Langacker, 2008), and Radical Construction Grammar (Croft, 2001, 2005). The hierarchical ontology of semantic types at the heart of CPA owes much to Pustejovsky's (1995) Generative Lexicon Theory. Additionally, as the SPOCA clause roles fundamental to CPA suggest, in TNE the notion of argument structure is broadly compatible with that of functional grammar (Haliday, 1994).

In light of the insights into phraseological approaches to language gained from the review in Chapter 1 and the results of the CPA analysis of the AJAXC2 reported in Chapters 6 and 7, the present study has examined how the behaviour of academic lexis relates to wider approaches to phraseology. In addition to providing evidence further supporting corpus linguistic concepts such as semantic prosody and alternation, this constitutes one of the present study's secondary objectives. Concepts from cognitive linguistics can plausibly account for facts uncovered by CPA analysis. Take, for example, the idea that novel expressions which lack unit status are meaningful if they are sufficiently similar to existing expressions which have unit status. This provides a plausible account of how humans were first able to understand patterns with *manage,* a verb which prototypically implies human cognition, with non-human arguments. For example, it is plausible that [[MB Entity]] *manages* [[Activity]] as a result of the once novel pattern's similarity to the established [[Human]] *manages* [[Activity]]. With repeated usage the newer pattern has become entrenched. More generally, the systematic alternation of more abstract semantic types, in this case [[MB Entity]], with more concreate ones, such as [[HUMAN]], might suggest the existence of a Lakovian conceptual metaphor along the lines of CELLS ARE HUMAN BEINGS. However, to truly investigate the lexicalisation of metaphors more fully, a diachronic approach is needed.

8.2 Challenges

In order to make the contributions outlined in the previous section, a number of challenges had to be overcome. These include factors related to the corpus, annotation, sampling, the methodological adequacy of CPA, and the lexicographical representation of the findings.

8.2.1 Corpus Size

Several aspects of corpus construction proved to be somewhat challenging. Size is a perpetual concern in corpus construction. A limitation of the original AJAC, evident in previous research, was that many of the AWL words it contained occurred at extremely low frequencies. This was especially problematic in the microbiology sub-corpus which contained articles which were, on average, one third shorter than those in sub-corpora representing history and management. In an attempt to mitigate this the AJACX2 is more

than twice the size of the original. It also contains a third more articles in microbiology than the original corpus. It was not envisaged that this increase in corpus size would lead to a proportional increase in the frequency of individual words since word frequency is Zipfian in nature. However, it was hoped that there would be a perceivable drop in findings that were based on only a handful of occurrences. In the AJAC, in microbiology, *access* occurs only 8 times labelled as a verb whereas in the AJACX2 it occurs 21 times. This indicates that the strategy of expanding the corpus was successful. Ultimately, the smallest verb sample annotated in the present study was that of *prescribe* in microbiology which contained 22 occurrences shared between two pattern types.

Overall, taken as a set, verbs occurred at sufficiently high frequencies for a valid cross-discipline comparison of verbal collocational behaviour. However, it might be argued that in cases of individual verbs which have a wide variety of collocational pattern types, yet occur at just over the frequency threshold of 20 occurrences, this was not the case. The statistical measures adopted in the present study mitigate the possibility of Type I errors. The LL statistic has been demonstrated to be more reliable than chi-square with low expected frequencies. Similarly, the effect-size threshold adopted ensures that *P* values indicating significant difference which were calculated based on extremely low frequencies of occurrence were discarded.

Low frequency of occurrence is perhaps more problematic in the qualitative analysis stage of the present study where some of the observations made were based on a small quantity of occurrences. For example, speculation about the possible alternation of [[Inanimate]] with [[Abstract Entity]] in object slot of pattern type three for *treat* in order to draw analogies between the treatment of data and the treatment or raw materials such as minerals, is made on the basis of two corpus lines. In spite of this limitation, the observations made are at least indicative of the type of phraseological behaviour encountered during the annotation phase and they are clearly advantageous to the creation of useful lexicographical resources for the EAP community.

The methodological design of the present study has some features which mitigate the minor problems brought about by the limited size of the corpus. The concept of data triangulation, adopting multiple perspectives on data, allows more confidence in findings than would have otherwise been possible. Space triangulation, in this case comparing academic language across three sub-corpora representing three disciplines allows more confidence in the existence of difference than would be possible with a two-discipline comparison. Methodological triangulation, in this case, combining a statistical analysis stage with more qualitative techniques, allows greater confidence than would be possible had only one approach been adopted.

In the period which has elapsed since the commencement of the present project, large-scale corpora of written academic language have become more widely accessible. For example the Oxford Corpus of Academic English (D. Lea, 2014b) has recently been made available, on request, via SE. It contains around 85-million words from 27 different disciplines, predominantly from academic journals, but also from textbooks and other scholarly literature. Although the AJACX2 was adequate to the task it was designed for, repeating the procedures of the present study in a larger corpus covering a greater number of disciplines and with a wider variety of text types may result in further insights into the phraseology of academic lexis across disciplines.

8.2.2 Corpus Composition

In addition to obtaining an adequate size, some decisions about corpus composition also proved challenging. One such decision; not to include abstracts in the AJAXC2, had consequences for corpus size. The decision to remove headings, tables, graphics and those equations not in-line with the text of the RAs is justifiable since these elements contain relatively little lexical content. In contrast, if abstracts are typically around 300 words in length, their removal represents a loss of around 264,000 tokens. This represents around 3.33 per cent of the total corpus size. More significantly, since abstracts have their own discourse and lexical conventions (see G.Hu & Cao, 2011), the decision to discard them has made the AJACX2 less representative of the lexis used by the EAP community than it might have been had abstracts been included. This is particularly true from a receptive standpoint, since abstracts serve as a summary of the contents of an article aimed at helping the reader decide whether it is worth reading the whole article, it stands to reason that they are more widely read than the other parts of the article. Although the exclusion of abstracts does affect corpus size and representativity, it in no way invalidates the findings of the study. The repetition of the present study with a corpus including RA abstracts would be beneficial, as would the application of its methods for a comparison of lexis used in distinct RA sections.

The decision to construct a corpus of research articles will be seen by some as a major limitation of the present study. It is true that the decision to concentrate on the written mode means that examination of spoken and sign language in academic contexts is not possible. It is, perhaps, true that these modes are relatively underrepresented in EAP vocabulary resources. Aside from these concessions, the criticism of the construction of a corpus of RAs for the study of academic lexis across disciplines is not valid. In Chapters 4 and 5, it was argued that academic journals are widely read by all levels of the EAP community. Furthermore, obtaining a sufficient quantity and range of textbooks for a meaningful examination would be fraught with logistical difficulties. As far as writing is concerned, the argument that RAs are not suitable textual exemplars for students is at best questionable and at worst condescending. This should not be taken to mean that there is no place for corpora of student writing in studies of EAP lexis and lexicography. Error analysis in corpora, of student writing, written by both native and non-native writers, provides an invaluable source of data which allows teachers, curriculum designers, dictionary writers, and other material creators to fine tune their practice to the needs of their public. In fact, the comparison of the findings about the verbs examined in the present study with their behaviour in corpora of student writing presents an interesting avenue for future investigation. Similarly, much of what is written about what comprises the ideal textual exemplar for novice writers is based on speculation or language-internal characteristics, for example, difference in frequency of occurrence of lexical bundles across text types. There is a pressing need for more research about what writers themselves think. Obtaining useful findings would involve more than simply asking established academic writers: How did you learn to write? Or what examples of writing have proved useful to you? However a well-designed study could provide invaluable insights into what constitutes a suitable example for novice academic writers.

The language background of the authors' of many of the articles included in the AJACX2 might be seen by some as a limitation of the present study. In spite of growing acceptance of English as a lingua franca in academic contexts, for some, the presence of articles written by non-native speakers will invalidate the AJACX2 as a source of data on lexis for the EAP community. Aside from the fact that ELFA is a reality and reflects much of the

language that EAP users experience, there are two principal arguments against this view: Firstly, with some notable exceptions (Rozycki & Johnson, 2013), editorial boards and the peer review process ensure that RAs contain language that is reasonably native-like. Secondly, there is some research, at least as far as academic collocations are concerned, to suggest that it is experience rather than native-speaker status which is the defining factor in accurate production (Frankenberg-Garcia, 2017). In future research into EAP lexis using corpora of RA, researchers may wish to control for the language background of authors. However, it has been argued in the present study that, not only is this technically challenging, it would also negatively impact the extent to which the corpus truly represented the language needs of the EAP community.

A further corpus-construction challenge concerns the selection of RAs for inclusion in the AJACX2. The use of journal citation indices in this regard was a fairly novel approach, having only been adopted by a handful of corpus investigations of academic language. G. Williams (2002) criticises the use bibliographical categorisations in the construction of specialised corpora. He argues that such classifications are motivated by administrative or organisational concerns rather than a true reflection the language use of a community of users. However, no means of categorising knowledge is perfect, the disciplinary categories from the journal citation ranking employed in the present study are broadly representative of the divisions in language use of a fairly broad community of EAP users. A further argument against the use of discrete bibliographic categories is that they cannot truly reflect interdisciplinarity. However, in the representation of the language EAP users encounter and strive to produce; if it is assumed that bibliographic categories are imperfect, and that much academic research is interdisciplinary, then the descriptive inadequacy of bibliographic categories becomes an advantage. In the AJAXC2, this is evidenced by the presence of articles in the history sub-corpus proposing medical explanations for the behaviour of historical figures, articles in the management sub-corpus dealing with the management of healthcare, and articles in the microbiology sub-corpus dealing with the management of health and safety law. In addition to the general problems involved in categorising knowledge, citation indices and measures also have inherent geographical, disciplinary and cultural biases. However, the fact remains that in the creation of specialised corpora selection criteria are essential. In order to obtain a fuller impression of the representativity of the RAs included in the AJACX2, a future study could seek the opinion of expert informants in the academic disciplines it comprises.

8.2.3 Annotation

The question of how to ensure validity and reliability in annotation presents a challenge to all corpus annotation projects. Ideally validity and reliability are achieved through the collaborative creation and constant revision of annotation guidelines and use of inter-annotator reliability measures to monitor consistency between annotators. In the context of the present study, the use of multiple annotators was not possible. Another means of ensuring valid and reliable annotation was needed. To overcome this challenge, the PDEV was employed as a reference in cases of doubt in the annotation process.

The need for an annotation reference had consequences for the range of language which could be examined in the present study. In conjunction with the corpus compilation criteria, the frequency and authorship thresholds adopted in the sampling of verbs for examination were an attempt to ensure that the verbs examined were representative of those which EAP users encounter and strive to use. Nonetheless, the potential sample was limited to those verbs which had been completed in the PDEV. Furthermore, the verb list

in the PDEV is based on their frequency of occurrence in the BNC50. This is problematic for two reasons: Firstly, since the BNC is a general corpus it does not necessarily reflect frequencies of occurrence of verbs in academic language; secondly it is over twenty years old and contains texts dating back to 1964 (Oxford Text Archive, 2015). This is particularly problematic for the representation of meaning in fast-moving science and technology disciplines. To mitigate these problems future studies should select verbs on the basis of frequency in the corpus under study and use multiple annotators and inter-annotator reliability measures in an attempt to ensure reliability in annotation.

The need for an annotation reference is particularly exigent in the case of CPA, as although, a number of short papers reporting CPA-based studies have been published, at present there are no official detailed guidelines for annotation. In this regard it is hoped that the procedures set out in Chapters 5, and notes and reflections in Chapter 6 and 7 might be of use to future CPA-based studies. Explanations and examples of what this study has termed 'widening patterns' and 'climbing the CPASO', along with the many examples of annotated corpus lines should be of use in future CPA studies. The decision to differentiate between patterns and pattern types in manner a similar to the distinction made between, words (tokens) and types in corpus linguistics in general might also be of value. This may appear trivial. However, the need to differentiate between a single instance of a pattern and many instances which map to a pattern is vital when using CPA to compare corpora. Using conventional CPA terms to do this proved burdensome.

8.2.4 Sampling

Decisions about which verbs to study also proved somewhat of a challenge. It has been established that the practical need for an annotation reference limited the range of language which could be studied to those 1286 verbs in the PDEV which had been completed at the time the sample was taken in January 2016. Additionally, the threshold of 20 occurrences was selected because it is said to represent the point at which a trained lexicographer is able to make judgments about meaning (Sinclair, 2005). The ten-occurrences in RAs of distinct authorship threshold was adopted to ensure that no single author's idiosyncratic use of a verb was overly represented in the sample. Although this approach has been successful, further insight might be gained from repeating the procedure with various threshold levels.

8.2.5 CPA

As the discussion in above indicates, CPA has proved to be a viable means of examining differences in the semantic and syntactic behaviour of verbs across corpora. Many of the differences it revealed would have been elided on distributional approaches to phrase extraction such as the lexical-bundle approach. However, the manual classification of verb arguments with semantic labels is an endeavour which requires a great deal of time. In contrast, the extraction of stings of a certain length which occur above a given frequency threshold takes minutes and has been employed in studies which have provided important insights into EAP vocabulary. For example, Biber and Barbieri's (2007) examination of lexical bundles in spoken and written language in academic contexts, Simpson-Vlach and Ellis's (2010) general academic phrase list, and Hsu's (2014) phrase list for prospective EFL business postgraduates, come to mind, among many others. Since form and distributional approaches to vocabulary selection such as the lexical-bundle approach require no prior manual annotation of corpus lines, initial phrase extraction can be achieved significantly more quickly than is possible with CPA. In contrast, the need for manual annotation of

corpus lines with semantic types presents a challenge to the use of CPA in the selection of MWUs for EAP. Automatic semantic tagging represents a potential response to this challenge. A supervised automatic semantic tagger based on the CPASO is currently in development however, in common with other semantic tagging applications, accuracy rates reported are low (66% in micro-average F1 across a sample of 25 frequent verbs) (El Maarouf & Baisa, 2013). An alternative, CPA-based, semi-automatic means of making broad comparisons across corpora has been demonstrated in previous research (Rees, 2013, 2016). On this approach, the statistically salient collocates for a given verb in two corpora are displayed using SE's WSDiff tool. By examining the salient collocates in syntactic positions, an analyst is able to assign a prototypical semantic type to the verb arguments of verbs in each corpora. Although it offers significant efficiencies in annotation time, this ad hoc CPA necessarily involves a significant loss in precision.

At first glance, the nature of the semantic labels employed in TNE and CPA also presents a challenge to the use of CPA in the description of technical or scientific terms. CPA is designed to examine general language. It employs an ontology of semantic types which represent the prototypical properties shared by the lexical items over which they generalise. This is perfectly adequate for general language. In certain contexts *mayor*, *teacher*, *lexicographer* could all be generalised over by the semantic type [[Human]], likewise *the government, the university, the home office* could be generalised over by the semantic type [[Institution]]. Scientific or technical uses sometimes present a greater challenge. In the PDEV lexical items such as *cell, DNA, gene, genome, strand* were labelled as [[Physical Object]] along with *stone, wood, floor*. It has been argued that these groups have different prototypical properties. In the context of the microbiology sub-corpus it became apparent a new semantic type, [[Microbiological Entity]], was needed. From a theoretical standpoint, CPA semantic types are not a closed class. As a corpus-based theory in which meaning is projected bottom-up from the corpus line rather than top-down from the pattern, TNE permits the creation of novel semantic types in light of corpus evidence.

The presence of pattern types with a clear technical or discipline-specific meaning also has implications for the statistical comparison of patterns across disciplines. It is possible that writers in this discipline will avoid other uses of this verb in order to prevent confusion with the technical term. This tendency for prevalent use of a pattern with a clearly specialist technical meaning is reflected in statistical comparisons of verb use across disciplines.

8.2.6 Guidelines and Illustrative Examples of Lexicographical Representation

The creation of a fully-fledged lexicographical resources was never envisaged as an immediate outcome of the present study. This is not possible on the basis of the AJACX2 data alone. The principal aim has always been to empirically investigate the phraseological behaviour of vocabulary across academic disciplines. However, even on these terms, the secondary aim of creating guidelines for, and illustrative examples of, how the insights gleaned from the present study might be usefully presented in lexicographical resources for the EAP community proved challenging.

The digitalisation of lexicography has significantly broadened the scope of what can be considered a lexicographical resource. Lexicographical resources might be considered on a scale of innovation. For example traditional printed dictionaries might be found at one pole and text-editor integrated writing tools at the other. In the present study, the challenge was presenting the insights gained in a manner that, with some changes, could straightforwardly

apply to this whole range of possibilities without seeming so nebulous and general that conceiving of any practical application required a good deal of imagination. With these concerns in mind, the guidelines in Chapter 7 were limited to the creation of an online multi-disciplinary EAP verb pattern dictionary.

In addition to the challenge of what type of dictionary to present in Chapter 7, there was the challenge of defining its scope or intended audience. Possible options included creating guidelines and examples for a dictionary aimed at EAP users working in one of the three disciplines examined, creating guidelines and examples for a multi-disciplinary dictionary including all three disciplines studied, or attempting to reflect the observed trend of difference between hard and soft disciplines. Ultimately, it was decided that guidelines and examples for a multi-disciplinary online verb pattern dictionary for the disciplines of history, management, and microbiology were the most appropriate means of faithfully exploiting the data obtained in the present study.

A further challenge involved how to adequately visualise data in the illustrative examples of entries posited. It might be argued that research on visual elements in lexicographical resources has not yet caught up with the digital revolution in lexicography. While there have been studies of user perceptions of multimedia elements such as animations in dictionaries (Lew & Doroszewska, 2009), the majority of research has tended to focus on visual elements which are found in paper dictionaries for example, the use of parenthesis, signposts, and the colour of sense divisions. The illustrative examples in Chapter 7 represent an attempt to apply this research. However, there are many insights from the field of data-visualisation which might fruitfully be applied to EAP lexicography. Some work has already begun in this regard (Roberts et al., 2017), however there is still a great deal more to achieve.

8.3 Future Directions

Bearing in mind the discussion above the most obvious course for future praxis points towards the creation of a fully-fledged multi-discipline EAP resource. This would involve preparatory investigations about the ideal corpus composition, specific lexical difficulties typically faced by EAP users, and the wider lexicographical needs of the EAP community. The next step would involve the collection of lexicographical data and drawing up an initial wordlist. A research phase could provide a good opportunity to further investigate some of the insights about academic lexis found the present study. The production stage would represent an opportunity to put the latest research on lexicographical visualisation into practice. Finally, an evaluation stage would allow an opportunity to test the resource produced and put the evaluation techniques discussed in Chapter 3 into practice.

8.3.1 Preparatory Steps

A preparatory step for the creation of any corpus aimed at fulfilling the productive component of this idealised multi-discipline resource would be to conduct fieldwork investigating what established academic writers regard as most adequate textual exemplars for less experienced writers. Similarly, error analysis in corpora of native and non-native student writing would provide a helpful means of focusing attention on those words academic writers find particularly troublesome.

Once market research had established which disciplines it would be financially viable to include in the resource, the views of expert informants on which text types and sources best represent the needs of EAP users in each of these disciplines could be sought. These views could then be crossed referenced with other selection criteria such as journal citation measures.

8.3.2 Data Collection

It was argued in the guidelines for the creation of useful lexicographical resources in Section 7.3 that data from the AJACX2 is not sufficient for the creation of a complete multidisciplinary EAP lexicographical resource. There is a clear need for data sources from other academic disciplines. This would permit further investigation of the proposed split between hard and soft disciplines. Even without the views of expert informants it is clear that to accurately reflect the range of academic language that users experience, a spoken academic language component is required. If expert informants suggested that RAs were included in the corpus, as research on academic publication outlined in Chapters 4 and 5 suggests they would, then abstracts should also be included. Although, as suggested above it might first be advantageous to conduct research about the lexis of different RA sections. A well balanced and representative corpus would provide an adequate source from which to draw a headword list. Any new corpus constructed should be complimented with data from existing corpora, both academic and general. Ideally, in order to obtain a better view of semantic transfer over time, such as the evolution of metaphors, the corpus should contain a diachronic element. G. Williams and Millon (2014) report on a procedure which employs the OED as a corpus in order examine the evolution of the word *probe*. This seems like a viable approach.

8.3.3 Research

In conjunction with these other sources, this corpus would provide the ideal basis for further research into some of the phenomena that have been noted in the present study. This includes further systematic study of discipline-specific alternations of all types. Given the difference in pattern types used for reporting speech noted in the present study, the conventions for making intertextual references across disciplines presents a potentially rewarding avenue for future research. The annotation necessary to conduct this research would ideally be done collaboratively with the annotation criteria under constant review. Statistical measures of inter-annotator reliability could be employed to measure consistency in annotation. In the case of inconsistency between annotators the criteria could be revised in to provide more precise guidance on the difficult cases.

8.3.4 Production

As in the present study, the data gleaned in the production of the idealised multi-discipline dictionary would be applicable to a wide range of lexicographical resources. However, in recent years there has been a trend towards the seamless integration lexicographical data with software. From a productive perspective this involves greater integration with writing environments such as word-processing software. As far as academic phraseology is concerned, work has recently begun in this regard (Frankenberg-Garcia et al., 2017). From a receptive perspective, this might involve the integration of lexicographical data with web-browsers or other means of reading electronic documents. These avenues offer an ideal

testing ground to experiment with the application of visualisation techniques from the discipline of computer science to lexicography (Roberts et al., 2017).

8.3.5 Evaluation

The primary objective of the present study was to examine differences in meaning and use of academic lexis which might be obfuscated by distributional or form-based approaches to the selection of vocabulary for EAP courses and materials. One of its secondary objectives was to create guidelines for, and illustrative examples of, the useful representation of phraseology in lexical resources for EAP users. As far as the examples produced in the present study are concerned, it is assumed that, although they do not constitute a complete lexicographical resource, they are illustrative of the type of the resource that would be useful for an EAP user working in one of the disciplines examined. There are two arguments in favour of this assumption. Firstly, the guidelines and examples created are based on data obtained from carefully designed procedures based on a corpus of language which is representative of the language EAP users encounter and strive to produce. Secondly, they are based on principles and findings obtained from a well-founded careful review of research on EAP lexicography.

Ideally it would have been advantageous to test these assumptions more directly. However, it is clear from the review in Chapter 3 that meaningful dictionary use research is not a simple undertaking. In general terms, it should be large scale; many respected studies of dictionary use have involved hundreds of participants. It should ideally employ a variety of different methods, be they questionnaire-based, observation-based, or test-based. It would also be advantageous to consider different elements of the lexicographical resource both individually and as a unit. The situation in EAP lexicography is further complicated by the need to consider a variety of disciplines, language backgrounds, proficiency-levels, and EAP contexts. In spite of these difficulties, a more complete EAP lexicographical project would have to involve research into the views and needs of users.

8.4 Concluding Remarks

The present study was born out of frustration with current methods of vocabulary selection for English for academic purposes. It was felt that general wordlist and phraselist solutions were inadequate, since they underestimate difference in word behaviour between academic disciplines. Furthermore, discrete-item wordlists were considered inadequate as they could not reflect the influence of collocates on verb meaning. The phraseological turn towards the extraction of multi-word units was seen as a positive development. However, the subordination of semantic concerns to distributional ones was regarded as problematic. A corpus-based procedure involving the semantic tagging of verb arguments was posited as a solution. This procedure indicated that frequently occurring words in a corpus of published academic writing do indeed demonstrate differences in meaning in different disciplines and that many of these differences in meaning would not have been brought to light by distributional approaches to phrase extraction. The CPA-based approach demonstrated in the present study is more resource intensive and complex than traditional approaches to the selection of vocabulary for EAP course and materials. However, it does provide greater insight into the language that EAP users actually need and use.

It is understood that lexicography, course design, and material design require a judicious equilibrium between what is possible and what is feasible with the resources available. In light of the findings set out in the present study it is clear that, as far as the selection and presentation of EAP vocabulary is concerned, the point of equilibrium has moved towards an approach that respects differences between disciplines and the phraseological nature of language. Given the current dominance of English in academic communication, the implications of this conclusion are extremely significant.

BIBLIOGRAPHY

Dictionaries and Pedagogical Grammars

Al-Baalbaki, R. (Ed.). (1990). *Al-Mawrid* (2nd ed.). Beirut: Dar el-Ilm Lil-Malayen.

Benson, M., Benson, E., & Ilson, R. (Eds.). (1986). *The BBI combinatory dictionary* (1st ed.). Amsterdam: John Benjamins.

Butterfield, J. (Ed.). (1990). *Harper Collins Spanish dictionary : Spanish-English; English-Spanish*. London: Harper Collins.

Carver, D., Wallace, M., & Cameron, J.(Eds.). (1974). *Collins English learner's dictionary*. London: Collins.

Cowie, A. P. (Ed.). (1989). *Oxford Advanced learner's dictionary of current English* (4th ed.). Oxford: Oxford University Press.

Cowie, A. P., Mackay, I. M., & McCaig, I. R. (Eds.). (1993). *Oxford dictionary of English idioms*. Oxford: Oxford University Press.

Cowie, A. P., Mackin, R., & McCaig, I. (Eds.). (1983). *Oxford dictionary of current idiomatic English. Vol II: Phrase, clause & sentence idioms*. Oxford: Oxford University Press.

Cowther, J. (Ed.). (1995). *Oxford advanced learner's dictionary of current English* (5th ed.). Oxford: Oxford University Press.

Deuter, M. (Ed.). (2015). *Oxford Advanced learner's dictionary of current English* (9th ed.). Oxford: Oxford University Press.

Dignen, S. (Ed.). (2000). *Longman business English dictionary*. Harlow: Longman.

Dubois, J. (Ed.). (1983). *Dictionnaire du français langue étrangère. Niveau 1*. Frankfurt am Main: Moritz Diesterweg.

Dutton, B., Harvey, L., & Walker, R. (Eds.). (1981). *Diccionario esencial inglés-español español-inglés Diáfora*. Barcelona: Cassell.

Francis, G., Hunston, S., & Manning, E. (1996). *Collins COBUILD grammar patterns 1: Verbs*. London: HarperCollins.

Francis, G., Hunston, S., & Manning, E. (1998). *Collins COBUILD grammar patterns 2: Nouns and Adjectives*. London: HarperCollins.

Granger, S., & Paquot, M. (2010). The Louvain EAP dictionary (LEAD). In A. Dykstra & T. Schoonheim (Eds.), *Proceedings of the 14th EURALEX International Congress* (pp. 321–326). Leeuwarden/Ljouwert, The Netherlands: Fryske Akademy.

Gray, A. (Ed.). (1983). *Longman dictionary of American English*. New York: Longman.

Gurulnik, D. (Ed.). (1976). *Webster's new world dictionary of the American language* (2nd ed.). Cleveland: The World Publishing Co.

Hanks, P. (2001). The pattern dictionary of English verbs. Masaryk, Czech Republic: Masaryk University, Czech Republic. Retrieved from http://www.pdev.org.uk

Hornby, A. S., Gatenby, E. V, & Cowie, A. P. (Eds.). (1974). *Oxford advanced learner's dictionary of current English* (3rd ed.). Oxford: Oxford University Press.

Hornby, A. S., Gatenby, E. V, & Wakefield, A. H. (Eds.). (1942). *Idiomatic and syntactic English Dictionary*. Tokyo: Language Institute of Education.

Hornby, A. S., Gatenby, E. V, & Wakefield, A. H. (Eds.). (1948). *Advanced learner's dictionary of current English* (1st ed.). Oxford: Oxford University Press.

Hornby, A. S., Gatenby, E. V, & Wakefield, A. H. (Eds.). (1963). *Advanced learner's dictionary of current English* (2nd ed.). Oxford: Oxford University Press.

Kirkpatrick, E. M. (Ed.). (1980). *Chambers universal learner's dictionary*. Edinburgh: Chambers.

Kozlowska, C., & Dzierzanowska, H. (1987). *Selected English collocations* (2nd ed.). Warsaw: Wydawnictwo Naukowe PWN.

Lea, D., Duncan, R., & Bull, V. (Eds.). (2014). *Oxford learner's dictionary of academic English*. Oxford: Oxford University Press.
Levenston, E., & Sivan, R. (Eds.). (1968). *The Megiddo modern dictionary, English-Hebrew*. Tel Aviv: Megiddo.
Long, T. (Ed.). (1979). *Longman dictionary of English idioms*. Harlow: Longman.
Mayor, M. (Ed.). (2009). *Longman dictionary of contemporary English*. Harlow: Longman.
OED online. (2017). *OED Online*. Oxford University Press. Retrieved from http://www.oed.com/
Palmer, H. (1924). *A grammar of spoken English*. Cambridge: W. Heffer and Sons.
Palmer, H. (1938). *A grammar of English words*. London: Longmans, Green.
Parkinson, D., & Noble, J. (Eds.). (2005). *Oxford business English dictionary for learners of English*. Oxford: Oxford University Press.
Pearsall, J. (Ed.). (1998). *The new Oxford dictionary of English*. Oxford: Clarendon Press.
Perrault, S. (Ed.). (2008). *Merriam-Webster's advanced learner's English dictionary*. Springfield: Merriam-Webster.
Proctor, P. (Ed.). (1978). *Longman dictionary of contemporary English* (1st ed.). Harlow: Longman.
Proctor, P. (Ed.). (1995). *Cambridge international dictionary of English*. Cambridge: Cambridge University Press.
Robert, P., Rey, A., & Rey-Debove, J. (Eds.). (1986). *Le petit Robert: Dictionnaire alphabétique et analogique de la langue français* (2nd ed.). Paris: Le Robert.
Runcie, M. (Ed.). (2002). *Oxford collocations dictionary for students of English*. Oxford: Oxford University Press.
Rundell, M. (Ed.). (2002). *Macmillan English dictionary for advanced learners*. London: Macmillan.
Rundell, M. (Ed.). (2007). *Macmillan English dictionary for advanced learners* (2nd ed.). London: Macmillan.
Sinclair, J. (Ed.). (1987). *Collins COBUILD English language dictionary*. Glasgow: Collins.
Sinclair, J. (Ed.). (1995). *Collins COBUILD English dictionary* (2nd ed.). London: Harper Collins.
Summers, D. (Ed.). (1983). *Longman active study dictionary of English*. Harlow: Longman.
Summers, D. (Ed.). (1987). *Longman dictionary of contemporary English* (2nd ed.). London: Longman.
Summers, D. (Ed.). (1988). *Longman active study dictionary* (3rd ed.). Harlow: Longman.
Summers, D. (Ed.). (1995). *Longman dictionary of contemporary English* (3rd ed.). Harlow: Longman.
Sykes, J. (Ed.). (1976). *The concise Oxford dictionary of current English*. Oxford: Clarendon Press.
Turnbull, J. (Ed.). (2010). *Oxford advanced learner's dictionary of current English* (8th ed.). Oxford: Oxford University Press.
Walter, E. (Ed.). (2005). *Cambridge advanced learner's Dictionary* (2nd ed.). Cambridge: Cambridge University Press.
Wehmeier, S. (Ed.). (2004). *Oxford advanced learner's English-Chinese dictionary* (6th ed.). Oxford: Oxford University Press.
Wehr, H., & Cowan, J. M. (Eds.). (1976). *A dictionary of modern written Arabic* (3rd ed.). Wlesbaden: Otto Harrassowitz.
West, M. P., & Endicott, G. (Eds.). (1935). *The new method English dictionary*. London: Longmans, Green & Co.

Other Resources

ABBYY. (2011). ABBYY FineReader 11. Milpitas, California: ABBYY Software.

Abraham, L. (2008). Computer-mediated glosses in second language reading comprehension and vocabulary learning: A meta-analysis. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588220802090246

Ackermann, K., & Chen, Y.-H. (2013). Developing the Academic Collocation List (ACL) – A corpus-driven and expert-judged approach. *Journal of English for Academic Purposes*, *12*(4), 235–247. https://doi.org/10.1016/j.jeap.2013.08.002

Ädel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes*, *31*(2), 81–92. https://doi.org/10.1016/j.esp.2011.08.004

Ahmed, M. (1989). Vocabulary learning strategies. In P. Meara (Ed.), *Beyond Words* (pp. 3–14). London: BAAL/CILT.

Al Ajmi, H. (1992). *The use of monolingual English and bilingual Arabic-English dictionaries in Kuwait: an experimental investigation into the dictionaries used and reference skills deployed by university students of arts and science* (unpublished doctoral dissertation). University of Leeds, Leeds.

Al Ajmi, H. (2008). The effectiveness of dictionary examples in decoding: The case of Kuwaiti learners of English. *Lexikos, 18,* 15-26.

Alsop, S., & Nesi, H. (2009). Issues in the development of the British Academic Written English (BAWE) corpus. *Corpora*, *4*(1), 71–83. https://doi.org/10.3366/E1749503209000227

Ames, W. (1966). The development of a classification scheme of contextual aids. *Reading Research Quarterly*, *2*(1), 57–82. https://doi.org/10.2307/747039

Anderson, J. (1980). The lexical difficulties of English medical discourse for Egyptian students. *English for Specific Purposes (Oregon State University)*, *4*(37).

Anderson, R., & Ortony, A. (1975). On putting apples into bottles - A problem of polysemy. *Cognitive Psychology*, *7*(2), 167–180. https://doi.org/10.1016/0010-0285(75)90008-0

Anderson, R., & Shifrin, Z. (1980). The meaning of words in context. In *Theoretical Issues in Reading Comprehension* (pp. 330–48). Mahwah, NJ: Lawrence Erlbaum Associates.

Anderson, R., Stevens, K., Shifrin, Z., & Osborn, J. (1978). Instantiation of word meanings in children. *Journal of Literacy Research*, *10*(2), 149–157. https://doi.org/10.1080/10862967809547264

Anthony, L. (2017). AntConc (version 3.4.4). Tokyo: Waseda University.

Apresjan, J. D. (2000). *Systematic lexicography*. Oxford: Oxford University Press.

Ard, J. (1982). The use of bilingual dictionaries by ESL students while writing. *ITL: Review of Applied Linguistics*, *58*, 1–27.

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, *62*(1), 67–82. https://doi.org/10.1016/j.jml.2009.09.005

Astika, G. (1993). Analytical assessments of foreign students' writing. *RELC Journal*, *24*(1), 61–70. https://doi.org/10.1177/003368829302400104

Atkins, B. T. S., & Knowles, F. (1990). Interim report on the EURALEX/AILA Research project into dictionary use. In T. Magay and J. Zigány (Eds.), *BudaLEX 88 Proceedings* (pp. 381–392). Budapest: Akadémiai Kiadó.

Atkins, B. T. S., & Varantola, K. (1998). Language learners using dictionaries: The final report on the EURALEX / AILA research project on dictionary use. In S. Atkins (Ed.), *Using Dictionaries: Studies of dictionary use by language learners and translators* (pp. 21–82). Berlin: de Gruyter.

ATLAS.ti. (2008). Berlin: Atlas.ti GmBH.
Baddeley, A. (1997). *Human memory. Theory and practice*. Hove: Psychology Press.
Baisa, V., El Maarouf, I., Rychlý, P., & Rambousek, A. (2015). Software and data for Corpus Pattern Analysis. In A. Horák, P. Rychlý, & A. Rambousek (Eds.) *Proceedings of the ninth workshop on recent advances in Slavonic Natural Language Processing* (pp. 75–86). Brno: Tribun EU.
Bank, C. (2010). *Die Usability von Online-Wörterbüchern und elektronischen Sprachportalen* (Unpublished MA thesis). Universität Hildesheim, Hildesheim.
Barber, C. (1962). Some measurable characteristics of modern scientific prose. In C. Barber (Ed.), *Contributions to English Syntax and Philology* (pp. 21–43). Göteburg: Acta Universitatis Gothoburgensis.
Barbieri, F. (2013). Involvement in university classroom discourse: Register variation and interactivity. *Applied Linguistics*, *36*(2), 151–173. https://doi.org/10.1093/applin/amt030
Barcroft, J. (2002). Semantic and structural elaboration in L2 lexical acquisition. *Language Learning*, *52*(2), 323–363. https://doi.org/10.1111/0023-8333.00186
Barcroft, J. (2009). Strategies and performance in intentional L2 vocabulary learning. *Language Awareness*, *18*(1), 74–89. https://doi.org/10.1080/09658410802557535
Bareggi, C. (1989). Gi studenti e il dizionario: un'inchiesta presso gli studenti di inglese del corso di laurea in lingue e letterature straniere moderne della facolta di lettere di Torino. In M. T. Prat Zagrebelsky (Ed.), *Dal dizionario ai dizionari. Orientamento e guida all`uso per studenti di lingua inglese* (pp. 155–179). Turin: Tirrentia Stampatori.
Barker, D. (2007). A personalized approach to analyzing "cost" and "benefit" in vocabulary selection. *System*, *35*, 523–533. https://doi.org/http://dx.doi.org/10.1016/j.system.2007.09.001
Barnhart, C. (1962). Problems in editing commercial monolingual dictionaries. In F. Householder & S. Saporta (Eds.), *Problems in Lexicography* (pp. 161–181). Bloomington: Indiana University.
Battenburg, J. (1989). *A study of English monolingual learners' dictionaries and their users* (unpublished doctoral dissertation). Purdue University, West Lafayette, IN.
Bauer, L., & Nation, P. (1993). Word Families. *International Journal of Lexicography*, *6*, 253–279.
Baxter, J. (1980). The dictionary and vocabulary behavior: A single word or a handful? *TESOL Quarterly*, *14*(3), 325–336. https://doi.org/10.2307/3586597
Bazerman, C. (1994). *Constructing experience*. Carbondale: Southern Illinois University Press.
Becher, T., & Trowler, P. (2001). *Academic tribes and territories*. Milton Keynes: SRHE and Open University Press.
Beck, I., McKeown, M., & Omanson, R. (1987). The effects and uses of diverse vocabulary instructional techniques. In M. McKeown & M. Curtis (Eds.), *The Nature of Vocabulary Acquisition* (pp. 147–63). Mahwah, NJ: Lawrence Erlbaum Associates.
Béjoint, H. (1981). The foreign student's use of monolingual English dictionaries: a study of language needs and reference skills. *Applied Linguistics*, *2*(3), 207–222.
Bensoussan, M., Sim, D., & Weiss, R. (1984). The effect of dictionary use on EFL test performance compared with student and teacher attitude and expectations. *Reading in a Foreign Language*, *2*(2), 262–276.
Bergenholtz, H., & Johnsen, M. (2005). Log files as a tool for improving internet dictionaries. *Hermes*, *34*, 117–141. Retrieved from http://download2.hermes.asb.dk/archive/download/Hermes-34-6-HB-Johnsen.pdf
Berman, R., & Cheng, L. (2010). English academic language skills: Perceived difficulties by undergraduate and graduate students, and their academic achievement. *Canadian Journal of Applied Linguistics*, *4*(1), 25–40. https://doi.org/10.1002/wdev.86

Biber, D. (1994). Representativeness in corpus design. In A. Zampolli, N. Calzolari, & M. Palmer (Eds.), *Current Issues in Comptutational Linguistics: in Honour of Don Walker* (pp. 377–408). Pisa: Giardini Editori e Stampatori.

Biber, D. (2006). *University language: A corpus based study of written and spoken registers*. Amsterdam: John Benjamins.

Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers, *26*(3), 263–286. https://doi.org/10.1016/j.esp.2006.08.003

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at …: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, *25*, 371–405.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow: Longman.

Björk, B.-C., & Solomon, D. (2012). Open access versus subscription journals: a comparison of scientific impact. *BMC Medicine*, *10*(1), 73. https://doi.org/10.1186/1741-7015-10-73

Black, A. (1986). *The effect on comprehension and memory of providing different types of defining information for new vocabulary. A report on two experiments conducted for Longman Dictionaries and Reference Division.* Cambridge.

Black, M. (1962). *Models and metaphors: Studies in language and philosophy*. Ithaca, NY: Cornell University Press.

Bloomfield, L. (1933). *Language*. London: Allen & Unwin.

Boers, F. (2001). Remembering figurative idioms by hypothesising about their origin. *Prospect*, *16*(3), 35–43.

Boers, F., & Demecheleer, M. (2001). Measuring the impact of cross-cultural differences on learners' comprehension of imageable idioms. *ELT Journal*, *55*(3), 255–262. https://doi.org/10.1093/elt/55.3.255

Boers, F., Demecheleer, M., & Eyckmans, J. (2004). Etymological elaboration as a strategy for learning idioms. In B. Laufer-Dvorkin (Ed.), *Vocabulary in a Second Language: Selection, Acquisition and Testing* (pp. 53–78). Amsterdam: John Benjamins.

Boers, F., Eyckmans, J., & Stengers, H. (2007). Presenting figurative idioms with a touch of etymology: More than mere mnemonics? *Language Teaching Research*, *11*(1), 43–62. https://doi.org/10.1177/1362168806072460

Boers, F., & Lindstromberg, S. (2009). *Optimizing a lexical approach to instructed Second Language Acquisition*. London: Palgrave Macmillan.

Boersma, P., & Weenink, D. (2017). Praat: doing phonetics by computer (Version 6.0.32) [Computer program]. Retrieved from http://www.praat.org/

Bogaards, P. (1991). Dictionnaires pédagogiques et apprentissage du vocabulaire. *Cahiers de Lexicologie*, *59*, 93–109.

Bogaards, P. (1992). French dictionary users and word frequency. In H. Tommola, K Varantola, T. Salmi-Tolonen & J. Shopp (Eds.), *EURALEX '92 Proceedings*.

Bogaards, P. (1999). Access structures of learners' dictionaries. In T. Herbst & K. Popp (Eds.), *The Perfect Learners' Dictionary (?)* (pp. 113–130). Tübingen: Max Niemeyer Verlag.

Bogaards, P. (2001). Lexical units and the learning of foreign Language Vocabulary. *Studies in Second Language Acquisition*, *23*(3), 321–343. https://doi.org/10.1017/S0272263101003011

Bogaards, P., & Van der Kloot, W. (2002). Verb constructions in learners' dictionaries. In A. Braasch & C. Povlsen (Eds.), *Proceedings of the 10th EURALEX International Congress* (pp. 747–757). København, Denmark: Center for Sprogteknologi.

Bogaards, P., & Van Der Kloot, W. A. (2001). The use of grammatical information in learners' dictionaries. *International Journal of Lexicography*. *14*(2), 97–121. https://doi.org/10.1093/ijl/14.2.97

Bolander, M. (1989). Prefabs, patterns and rules in interaction? Formulaic speech in adult learners' L2 Swedish. In K. Hyltenstam & L. K. Obler (Eds.), *Bilingualism across the lifespan*. (pp. 73–86). Cambridge: Cambridge University Press.
Bolinger, D. (1976). Meaning and memory. *Forum Linguisticum*, *1*(1), 1–14.
Bowles, M. (2004). L2 glossing: To CALL or not CALL. *Hispania*, *87*, 541–552.
Bradbury, J., & El Maarouf, I. (2013). An empirical classification of verbs based on Semantic Types: the case of the "poison" verbs. In *Proceedings of JSSP* (pp. 70–74).
Bramki, D., & Williams, R. (1984). Lexical familiarization in economics texts, and its pedegogic implications in reading comprehension. *Reading in a Foreign Language*, *2*(1), 169–81.
Brazil, D. (1995). *A grammar of speech*. Oxford: Oxford University Press.
Bresnan, J. (1978). A realistic transformational grammar. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic Theory and Psychological Reality* (pp. 1–59). Cambridge, MA: MIT Press.
Bresnan, J. (2007). Is syntactic knowledge probabilistic? Experiments with the English dative alternation. In S. Featherstone & W. Sternefeld (Eds.), *Roots: Linguistics in search of its evidential base* (pp. 77–96). Berlin: Mounton de Gruyter.
Brett, A., Rothlein, L., & Hurley, M. (1996). Vocabulary acquisition from listening to stories and explanations of target words. *The Elementary School Journal*, *96*(4), 415–422. Retrieved from http://www.jstor.org/stable/1001864
Brown, D. (1974). Advanced vocabulary teaching: the problem of collocations. *RELC Journal*, *5*(2), 1–11.
Brown, R., & McNeill, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior*, *5*(1934), 325–337. https://doi.org/10.1016/S0022-5371(66)80040-3
Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading , reading-while-listening , and listening to stories. *Reading in a Foreign Language*, *20*(2), 136–163.
Buehl, D. (2017). *Developing readers in the academic disciplines*. Portland, ME: Stenhouse.
Burns, A. (2010). *Doing action research in English Language Teaching: A Guide for Practitioners*. New York: Routledge.
Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.
Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, *82*(4), 711–733. https://doi.org/10.1353/lan.2006.0186
Cabré, M. T. (1999). *La terminología. Representación y comunicación. Elementos para una teoría de base comunicativa y otros artículos*. Barcelona: IULA.
Cameron, B. D. (2005). Trends in the usage of ISI bibliometric data: Uses, abuses, and implications. *Portal: Libraries and the Academy*, *5*(1), 105–125. https://doi.org/10.1353/pla.2005.0003
Campion, M., & Elley, W. (1971). *The Academic Vocabulary List*. Wellington: New Zealand Council for Educational Research.
Campoy Cubillo, M. C. (2002). Dictionary use and dictionary needs of ESP students: An experimental approach. *International Journal of Lexicography*, *15*(3), 206–228. https://doi.org/10.1093/ijl/15.3.206
Cantos Gómez, P. (2013). *Statistical methods in language and linguistic research*. Sheffield: Equinox.
Carter, R. (2004). *Language and creativity: The art of common talk*. London: Routledge.
Chang, Y., Chang, J., Chen, H., & Liou, H. (2008). An automatic collocation writing assistant for Taiwanese EFL learners: A case of corpus-based NLP technology. *Computer Assisted Language Learning*, *21*(3), 283–299. https://doi.org/10.1080/09588220802090337

Chen, M. (2013). Overuse or underuse: A corpus study of English phrasal verb use by Chinese, British and American university students. *International Journal of Corpus Linguistics*, *18*(3), 418–442.
Chen, Q., & Ge, G. (2007). A corpus-based lexical study on frequency and distribution of Coxhead's AWL word families in medical research articles (RAs). *English for Specific Purposes*, *26*, 502–514. https://doi.org/10.1016/j.esp.2007.04.003
Chen, Y. (2010). Dictionary use and EFL learning. A contrastive study of pocket electronic dictionaries and paper dictionaries. *International Journal of Lexicography*, *23*(3), 275–306. https://doi.org/10.1093/ijl/ecq013
Chen, Y. (2012). Dictionary use and vocabulary learning in the context of reading. *International Journal of Lexicography*, *25*(2), 216–247. https://doi.org/10.1093/ijl/ecr031
Cheng, W., Greaves, C., & Warren, M. (2006). From n-gram to skipgram to concgram. *International Journal of Corpus Linguistics*, *11*(4), 411–433. https://doi.org/10.1075/ijcl.11.4.04che
Chihara, T., Oller, J., Weaver, K., & Chavez-Oller, M. (1977). Are cloze items sensitive to discourse constraints? *Language Learning*, *27*, 63–73.
Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
Chomsky, N. (1970). Remarks on nominalization. In R. Jacobs & P. Rosenbaum (Eds.), *Readings in English transformational grammar* (pp. 184–221). Waltham, MA: Ginn.
Chomsky, N. (1981). *Lectures on government and binding: The Pisa lectures*. Berlin: De Gruyter.
Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
Chon, Y. V. (2009). The Electronic Dictionary for Writing: A Solution or a Problem? *International Journal of Lexicography*, *22*(1), 23–54. https://doi.org/10.1093/ijl/ecn034
Church, K., & Hanks, P. (1990). Word association norms, mutual information, and lexicography. *Computational Linguistics*, *16*, 22–29.
Cinková, S., & Hanks, P. (2010). Validation of Corpus Pattern Analysis - Assigning pattern numbers to random verb samples. Retrieved from http://nlp.fi.muni.cz/projects/cpa/CPA_valiman.pdf
Cinková, S., Holub, M., & Kríz, V. (2012). Managing uncertainty in semantic tagging. In W. Daelmans (Ed.) *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics* (pp. 840–850). Avignon: Association for Computational Linguistics
Cobb, T. (n.d.). The Compleat Lexical Tutor. Retrieved September 20, 2016, from http://www.lextutor.ca/
Cobb, T., & Horst, M. (2004). Is there room for an academic word list in French? In P. Bogaards & B. Laufer-Dvorkin (Eds.), *Vocabulary in a second language: Selection, acquisition and testing* (pp. 15–38). Amsterdam: John Benjamins.
Cock, S. De. (2006). Getting down to business: Monolingual learners' dictionaries and business English. In E. Corino, C. Marello & C. Onesti (Eds.), *Proceedings of the 12th EURALEX International Congress* (pp. 819–824). Torino, Italy: Edizioni dell'Orso.
Coffey, S. (2016). A Dictionary of Academic English: A Further Resource for Students in Higher Education. *International Journal of Lexicography*, *29*(1), 100–123. https://doi.org/10.1093/ijl/ecv020
Coffin, C. (2006). *Historical discourse*. London: Continuum.
Čolovic-Markovič, J. (2016 June). *To teach or not to teach explicitly: A qualitative examination of ESL writers' production of academic formulas in controlled and uncontrolled situations*. Paper presented at the 7th Formulaic Language Research Network conference, Vilnius.
Comte, A. (1896). *The positivist Philosophy of Auguste Comte*. London: George Bell & Sons.
Coniam, D. (1998). Partial parsing: Boundary marking. *International Journal of Corpus Linguistics*, *3*(2), 229–249. https://doi.org/10.1075/ijcl.3.2.04con

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, *29*(1), 72–89. https://doi.org/10.1093/applin/amm022
Cope, B., & Kalantzis, M. (1993). *The powers of literacy: A genre approach to teaching writing*. London: The Falmer Press.
Corson, D. (1985). *The lexical bar*. Oxford: Pergamon Press.
Corson, D. (1995). *Using English words*. Dordrecht: Kluwer Academic Publishers.
Corson, D. (1997). The learning and use of academic english words. *Language Learning*, *47*, 671–718. https://doi.org/10.1111/0023-8333.00025
Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, *23*, 397–423. https://doi.org/http://dx.doi.org/10.1016/j.esp.2003.12.001
Cowan, J. R. (1974). Lexical and syntactic research for the design of EFL reading materials. *TESOL Quarterly*, *8*(4), 389–399. https://doi.org/10.2307/3585470
Cowie, A. P. (1984). EFL dictionaries – past achievements and present need. In R. K. K. Hartmann (Ed.), *LEXeter '83 proceedings: Papers from the international conference on lexicography at Exeter. 9 - 12 September 1983* (pp. 155–164). Tübingen: Max Niemeyer Verlag.
Cowie, A. P. (1989). The language of examples in learner's dictionaries. In G. James (Ed.), *Lexicographers and their Works* (pp. 55–65). Exeter: University of Exeter.
Cowie, A. P. (1998). *Phraseology : Theory, analysis, and applications*. Oxford: Clarendon Press.
Cowie, A. P. (2002). *English Dictionaries for Foreign Learners: A History*. Oxford: Clarendon Press.
Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*, 213–238. Retrieved from http://dx.doi.org/10.2307/3587951
Coxhead, A., & Hirsh, D. (2007). A pilot science-specific word list. *Revue Française de Linguistique Appliquée, 2*(7), 65–78.
Creamer, T. (1987). Beyond the definition: Some problems with examples in recent Chinese-English and English-Chinese bilingual dictionaries. In A. P. Cowie (Ed.), *The Dictionary and the Language Learner: Papers from the EURALEX Seminar at the University of Leeds, 1-3 April 1985* (pp. 238–245). Tübingen: Max Niemeyer Verlag.
Crewe, W. J. (1990). The illogic of logical connectives. *ELT Journal*, *44*(4), 316–25. Retrieved from http://eltj.oxfordjournals.org/content/44/4/316.abstract
Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
Croft, W. (2005). The Darwinization of linguistics. *Selection*, *3*(1), 75–91. https://doi.org/10.1556/Select.3.2002.1.7
Crystal, D. (1986). The ideal dictionary, lexicographer and user. In R. Ilson (Ed.), *Lexicography an Emerging International Profession* (pp. 72–81). Manchester: University Press.
Cumming, G., Cropp, S., & Sussex, R. (1994). On-line lexical resources for language learners: Assessment of some approaches to word definition. *System*, *22*(3), 369–377. https://doi.org/10.1016/0346-251X(94)90022-1
Cummins, J. (1986). Language proficiency and language achievement. In J. Cummins & M. Swain (Eds.), *Bilingualism in Education* (pp. 138–61). London: Longman.
D'Anna, C., Zechmeister, E., & Hall, J. (1991). Toward a meaningful definition of vocabulary size. *Journal of Literacy Research*, *23*(1), 109–122. https://doi.org/10.1080/10862969109547729
Dagut, M., & Laufer, B. (1985). Avoidance of Phrasal Verbs—A Case for Contrastive Analysis. *Studies in Second Language Acquisition*, *7*(1), 73–79. https://doi.org/10.1017/S0272263100005167

Davis, P. M., & Walters, W. H. (2011). The impact of free access to the scientific literature: a review of recent research. *Journal of the Medical Library Association : JMLA*, *99*(3), 208–217. https://doi.org/10.3163/1536-5050.99.3.008

de Groot, A. M. B. (2006). Effects of stimulus characteristics and background music on foreign language vocabulary learning and forgetting. *Language Learning*, *56*(3), 463–506. https://doi.org/10.1111/j.1467-9922.2006.00374.x

de Jong, J., & Benigno, V. (2016, November). *The CEFR in higher education: developing descriptors of academic English*. Paper presented at Language Testing Forum 2016 University of Reading, Reading.

De Schryver, G.-M. (2013). The concept of simultaneous feedback. In R. Gouws, U. Heid, W. Schweickard & H. Wiegand (Eds.), *Dictionaries. An International Encyclopedia of Lexicography Supplementary Volume: Recent Developments with Focus on Electronic and Computational Lexicography* (pp. 548–556). Berlin: de Gruyter.

De Schryver, G.-M., & Joffe, D. (2004). On how electronic dictionaries are really used. In G. Williams & S. Vessier (Eds.), *Proceedings of the 11th EURALEX International Congress* (pp. 187–196). Lorient, France: Université de Bretagne-Sud, Faculté des lettres et des sciences humaines.

De Schryver, G.-M., Joffe, D., Joffe, P., & Hillewaert, S. (2006). Do dictionary users really look up frequent words? - On the overestimation of the value of corpus-based lexicography. *Lexikos*, *16*, 67–83. https://doi.org/10.4314/lex.v16i1.51504

DeCesaris, J. (2012). On the nature of signposts. In R. V. Fjeld & J. M. Torjusen (Eds.), *Proceedings of the 15th EURALEX International Congress* (pp. 532–540). Oslo,Norway: Department of Linguistics and Scandinavian Studies, University of Oslo.

Deignan, A. (2005). *Metaphor and Corpus linguistics. Converging evidence in language and communication research*. Amsterdam: John Benjamins.

Díez-Bedmar, M. B. (2012). The use of the Common European Framework of Reference for Languages to evaluate the compositions in the English exam in the University Entrance Examination. *Revista de Educación*, *357*, 55–80.

Dong, P., Loh, M., & Mondry, A. (2005). The impact factor revisited. *Biomedical Digital Libraries*, *2*, 7. https://doi.org/10.1186/1742-5581-2-7

Drysdale, P. (1987). The role of examples in a learner's dictionary. In A. P. Cowie (Ed.), *The Dictionary and the Language Learner: Papers from the EURALEX Seminar at the University of Leeds, 1-3 April 1985* (pp. 213–223). Tübingen: Max Niemeyer Verlag.

Dunning, T. (1993). Accurate Methods for the Statistics of Surprise and Coincidence. *Computational Linguistics*, *19*, 61–74. Retrieved from http://portal.acm.org/citation.cfm?id=972454

Durrant, P. (2009). Investigating the viability of a collocation list for students of English for academic purposes. *English for Specific Purposes*, *28*(3), 157–169. https://doi.org/10.1016/j.esp.2009.02.002

Durrant, P. (2013). To what extent is the Academic Vocabulary List relevant to university student writing? *English for Specific Purposes*, *35*(3), 328–356. https://doi.org/10.1016/j.esp.2016.01.004

Durrant, P. (2014). Discipline and level specificity in university students' written vocabulary. *Applied Linguistics*, *35*(3), 328–356.

Durrant, P. (2016). To what extent is the Academic Vocabulary List relevant to university student writing? *English for Specific Purposes*, *43*, 49–61. https://doi.org/10.1016/j.esp.2016.01.004

Durrant, P. (2017). Lexical bundles and disciplinary variation in university students' writing: Mapping the territories. *Applied Linguistics*, *38*(2), 165–193 https://doi.org/10.1093/applin/amv011

Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *IRAL - International Review of Applied Linguistics in Language Teaching*, *47*(2), 157–177. https://doi.org/10.1515/iral.2009.007
Durrant, P., & Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second Language Research*, *26*(2), 163–188. https://doi.org/10.1177/0267658309349431
Dziemianko, A. (2006). *User-friendliness of verb syntax in pedagogical dictionaries of English*. Tübingen: Max Niemeyer Verlag.
Dziemianko, A. (2015). Colours in online dictionaries: A case of functional labels. *International Journal of Lexicography*, *28*(1), 27–61. https://doi.org/10.1093/ijl/ecu028
Dziemianko, A. (2016). An insight into the visual presentation of signposts in English learners' dictionaries online. *International Journal of Lexicography*, *29*(4), 490–524. https://doi.org/10.1093/ijl/ecv040
Editorial. (2005). Not-so-deep impact. *Nature*, *435*, 1003–1004. https://doi.org/10.1038/4351003b
El Maarouf, I., & Baisa, V. (2013). Automatic classification of patterns from the Pattern Dictionary of English Verbs. In *Proceedings of JSSP* (pp. 95–99).
El Maarouf, I., Baisa, V., Bradbury, J., & Hanks, P. (2014). Disambiguating Verbs by Collocation: Corpus Lexicography meets Natural Language Processing. In *Proceedings of LREC* (pp. 1001–1006). Reykjavik.
Elgort, I. (2011). Deliberate learning and vocabulary acquisition in a second language. *Language Learning*, *61*(2), 367–413. https://doi.org/10.1111/j.1467-9922.2010.00613.x
Elley, W. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly*, *24*(2), 174–187. https://doi.org/10.2307/747863
Ellis, N. (1996). Sequencing in SLA phonological Memory, chunking, and points of order. *Studies in Second Language Acquisition*, *18*, 91–126.
Ellis, N., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning*, *43*(4), 559–617. https://doi.org/10.1111/j.1467-1770.1993.tb00627.x
Ellis, N., & Schmitt, N. (1997). Morphology and longer distance dependencies: Laboratory research illuminating the A in SLA. *Studies in Second Language Acquisition*, *19*, 145–71.
Ellis, N., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly*, *42*(3), 375–396. https://doi.org/10.1002/j.1545-7249.2008.tb00137.x
Ellis, N., Frey, E., & Jalkanen, I. (2009). The psycholinguistic reality of collocation and semantic prosody (1): Lexical access. In R. Ute & R. Schulze (Eds.), *Exploring the Lexis-grammar Interface* (pp. 89–114). Amsterdam: John Benjamins.
Ellis, R. (1991). The interaction hypothesis: A critical evaluation. In E. Sadtono (Ed.), *Language Acquisition and the Second/Foreign Language Classroom* (pp. 179–211). Singapore: SEAMEO-RELC.
Ellis, R., Tanaka, Y., & Yamazaki, A. (1994). Classroom interaction, comprehension, and the acquisition of L2 word meanings. *Language Learning*, *44*(3), 449–491. https://doi.org/10.1111/j.1467-1770.1994.tb01114.x
ETS. (2017). TOEFL: Home. Retrieved October 7, 2017, from https://www.ets.org/toefl
Evans, S., & Green, C. (2007). Why EAP is necessary: A survey of Hong Kong tertiary students. *Journal of English for Academic Purposes*, *6*(1), 3–17. https://doi.org/10.1016/j.jeap.2006.11.005
Evans, S., & Morrison, B. (2011). The first term at university: Implications for EAP. *ELT Journal*, *65*(4), 387–397. https://doi.org/10.1093/elt/ccq072
Eysenbach, G. (2006). Citation advantage of open access articles. *PLoS Biology*, *4*(5), 692–698. https://doi.org/10.1371/journal.pbio.0040157

Fan, M. (2003). Frequency of use, perceived usefulness, and actual usefulness of second language vocabulary strategies: A study of Hong Kong learners. *The Modern Language Journal*, *87*(2), 222–241. https://doi.org/10.2307/1193034
Faucet, L., West, M., Palmer, H., & Thorndike, E. L. (1936). *Interim report on vocabulary selection for the teaching of English as a foreign language*. London: P.S. King.
Fauconnier, G., & Turner, M. (1998). Conceptual integration networks. *Cognitive Science*, *22*(2), 133–187. https://doi.org/10.1016/S0364-0213(99)80038-X
Fellbaum, C. (1998). *WordNet: An electronic lexical database. Language, speech and communication*. Cambridge, MA: MIT Press.
File, K. A., & Adams, R. (2010). Should vocabulary instruction be integrated or isolated? *TESOL Quarterly*, *44*(2), 222–249. https://doi.org/10.5054/tq.2010.219943
Fillmore, C. J. (1976). Frame semantics and the nature of language. In S. R. Harnad, H. D. Steklis, & J. Lancaster (Eds.), *Origens and Evolution of Language and Speech* (pp. 20–32). New York: The New York Academy of Sciences.
Fillmore, C. J. (1979). On fluency. In C. Fillmore, D. Kempler & W. S-Y. Yang (Eds.) *Individual differences in language ability & language behavior* (pp. 85–101). New York: Academic Press.
Fillmore, C. J. (1985). Frames and the semantics of understanding. *Quaderni Di Semantica*, *6*, 222–254.
Fillmore, C. J. (1988). The Mechanisms of "Construction Grammar." In S. Axmaker, A. Jaisser & H. Singmaster (Eds.), *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society* (pp. 35–55). https://doi.org/10.3765/bls.v14i0.1794
Fillmore, C. J., & Atkins, B. T. S. (1992). The semantics of RISK and its neighbors. In A. Lehrer & E. F. Kittay (Eds.), *Frames, Fields and Concepts* (pp. 75–102). Hillsdale, NJ: Routledge.
Fillmore, C. J., Johnson, C. R., & Petruck, M. R. L. (2003). Background to Framenet. *International Journal of Lexicography*, *16*, 235–250. Retrieved from http://ijl.oxfordjournals.org/content/16/3/235.abstract
Fillmore, C. J., Kay, P., & O'Connor, M. C. (1988). Regularity and idiomaticity in Grammatical Constructions: The case of let alone. *Language, 64*(3), 501–508. https://doi.org/10.2307/414531
Firth, J. R. (1937). *The tongues of men and speech*. Oxford: Oxford University Press.
Flower, L. S., & Hayes, J. R. (1977). Problem-solving strategies and the writing process. *College English*, *39*(4), 449–461. https://doi.org/10.2307/375768
Flowerdew, J. (1992). Definitions in science lectures. *Applied Linguistics*, *13*(2), 202–221. https://doi.org/10.1093/applin/13.2.202
Flowerdew, J. (2001). Attitudes of journal editors to non-native speaker contributions. *TESOL Quarterly*, *35*(1), 121. https://doi.org/10.2307/3587862
Flowerdew, J. (2013). English for Research and Publication Purposes. In *The Handbook of English for Specific Purposes* (pp. 301–321). Hoboken, NJ: John Wiley & Sons.
Fontenelle, T. (2012). WordNet, FrameNet and other semantic networks in the International Journal of Lexicography - The Net Result? *International Journal of Lexicography*, *25*(4), 437–449. https://doi.org/10.1093/ijl/ecs027
Francis, W. N., & Kučera, H. (1979). Brown Corpus Manual. Providence, Rhode Island: Brown University. Retrieved from http://khnt.hit.uib.no/icame/manuals/brown/INDEX.HTM
Francis, W. N., & Kučera, H. (1982). *Frequency analysis of English usage: lexicon and grammar*. Boston: Houghton Mifflin.
Frankenberg-Garcia, A. (2012). Learners' Use of Corpus Examples. *International Journal of Lexicography*, *25*(3), 273–296. https://doi.org/10.1093/ijl/ecs011

Frankenberg-Garcia, A. (2014). The use of corpus examples for language comprehension and production. *ReCALL*, *26*(2), 128–146. https://doi.org/DOI: 10.1017/S0958344014000093

Frankenberg-Garcia, A. (2015). Dictionaries and encoding examples to support language production. *International Journal of Lexicography*, *28*(4), 490–512. https://doi.org/10.1093/ijl/ecv013

Frankenberg-Garcia, A. (2017, September). *Assessing the productive collocation repertoire of writers for the development of dedicated writing assistant tools*. Poster session presented at eLex 2017, Leiden.

Frankenberg-Garcia, A., Lew, R., Roberts, J., Rees, G., & Pereda, J. (2017). *Best practices in combining learner needs, lexicographic data and text editors to help learners write more idiomatically*. Poster session presented at eLex 2017, Leiden.

Fraser, C. A. (1999). Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition*, *21*(2), 225–241. https://doi.org/10.1017/S0272263199002041

Fries, P. (1995). A personal view of theme. In M. Ghadessy (Ed.), *Thematic development in English texts* (pp. 1–9). London: Pinter.

Fries, U. (1984). Theme and Rheme revisited. In R. J. Watts & U. Weidmann (Eds.), *Modes of Interpretation: Essays Presented to Ernst Leisi on the Occasion of his 65th Birthday* (pp. 177–192). Tübingen: Gunter Narr Verlag.

Fritz, E., & Ruegg, R. (2013). Rater sensitivity to lexical accuracy, sophistication and range when assessing writing. *Assessing Writing*, *18*(2), 173–181. https://doi.org/10.1016/j.asw.2013.02.001

Furetière, A. (1701). *Dictionnaire universel, contenant généralement tous les mots françois tant vieux que modernes, & les termes des sciences et des arts. Tome 1 / ,... par feu messire Antoine Furetière,... 2e édition revue, corrigée et augmentée par M. Basnage de Bauval*. La Hague and Rotterdam: Arnoud et Reinier Leers.

Gardner, D. (2004). Vocabulary input through extensive reading: A comparison of words found in children's narrative and expository reading materials. *Applied Linguistics*, *25*(1), 1–37. https://doi.org/10.1093/applin/25.1.1

Gardner, D. (2008). Validating the construct of word in applied corpus-based vocabulary research: A critical survey. *Applied Linguistics*, *28*(2), 241–265. https://doi.org/10.1093/applin/amm010

Gardner, D. (2013). *Exploring vocabulary language in action*. Oxford: Routledge.

Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics*, *35*(3), 305–327. https://doi.org/10.1093/applin/amt015

Gardner, S., & Nesi, H. (2012). A classification of genre families in university student writing. *Applied Linguistics*. Retrieved from http://applij.oxfordjournals.org/content/early/2012/06/26/applin.ams024.abstract

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, *28*(2), 200–213. https://doi.org/10.1016/0749-596X(89)90044-2

Geeraerts, D., & Cuyckens, H. (2010). Introducing cognitive linguistics. In D. Geeraerts & H. Cuyckens (Eds.), *Oxford Handbook of Cognitive Linguistics* (pp. 3–21). Oxford: Oxford University Press.

Ghadessy, P. (1979). Frequency counts, words lists, and materials preparation: A new approach. *English Teaching Forum*, *17*, 24–27.

Gheitasi, P. (2016, June). *Formulaicity in the oral language production of young foreign language learning*. Paper presented at the 7th Formulaic Language Research Network conference, Vilnius.

Gibbs, R. W., & Colston, H. L. (1995). The cognitive psychological reality of image schemas and their transformations. *Cognitive Linguistics*, *6*(4), 347–378. https://doi.org/10.1515/cogl.1995.6.4.347

Gildea, P., & Millller, G. (1987). Education as a new way to teach children words. Retrieved July 24, 2017, from https://www.washingtonpost.com/archive/opinions/1987/08/30/education-a-new-way-to-teach-children-words/f487a31f-427e-4638-84bc-d1d7f0ba3861/?utm_term=.bff9a3ad73ca

Gilner, L. (2011). A primer on the General Service List. *Reading in a Foreign Language*, *23*(1), 65–83.

Gilquin, G., Granger, S., & Paquot, M. (2007). Learner corpora: The missing link in EAP pedagogy. *Journal of English for Academic Purposes*, *6*(4), 319–335. https://doi.org/10.1016/j.jeap.2007.09.007

Glasser, B., & Strauss, A. (1967). *The discovery of grounded theory*. Chicago: Aldine.

Gledhill, C. J. (2000). *Collocations in science writing*. Tübingen: Gunter Narr Verlag.

Glucksberg, S., & Keysar, B. (1990). Understanding metaphorical comparisons: Beyond similarity. *Psychological Review*, *97*(1), 3–18. https://doi.org/10.1037/0033-295X.97.1.3

Glucksberg, S., Keysar, B., & McGlone, M. S. (1992). Metaphor understanding and accessing conceptual schema: Reply to Gibbs (1992). *Psychological Review*, *99*(3), 578–581. https://doi.org/10.1037/0033-295X.99.3.578

Goh, C. (2000). A cognitive perspective on language learners' listening comprehension problems. *System*, *28*(1), 55–75. https://doi.org/10.1016/S0346-251X(99)00060-3

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

González-Fernández, B. (2016). *The acquisition of collocations in relation to other components of vocabulary knowledge*. Paper presented at the 7th Formulaic Language Research Network conference, Vilnius.

Grabowski, Ł. (2015). Keywords and lexical bundles within English pharmaceutical discourse: A corpus-driven description. *English for Specific Purposes*, *38*, 23–33. https://doi.org/10.1016/j.esp.2014.10.004

Granger, S. (1998). Prefabricated patterns in advanced EFL writing: collocations and formulae. In A. P. Cowie (Ed.), *Phraseology: theory, analysis and applications*. Oxford: Oxford University Press.

Granger, S., Dagneaux, E., Meunier, F., & Paquot, M. (2009). *International Corpus of Learner English*. Louvain-la-Neuve: Presses universitaires de Louvain.

Granger, S., & Paquot, M. (2008). Disentangling the phraseological web. In S. Granger & F. Meunier (Eds.), *Phraseology: an interdisciplinary perspective* (pp. 27–47). Amsterdam: John Benjamins.

Granger, S., & Paquot, M. (2010a). Customising a general EAP dictionary to meet learner needs. In *eLexicography in the 21st century: New challenges, new applications. Proceedings of ELEX2009* (pp. 87–96). Louvain-la-Neuve: Presses universitaires de Louvain.

Granger, S., & Paquot, M. (2010b). The Louvain EAP dictionary (LEAD). In A. Dykstra & T. Schoonheim (Eds.), *Proceedings of the 14th EURALEX International Congress* (pp. 321–326). Leeuwarden/Ljouwert, the Netherlands: Fryske Akademy.

Granger, S., & Paquot, M. (2017). Putting the Louvain English for Academic Purposes Dictionary to the test: Implicit vs. explicit feedback. Paper presented at eLex 2017, Leiden.

Granger, S., & Tribble, C. (1998). Learner corpus data in the foreign language classroom: form-focused instruction and data-driven learning. In S. Granger & G. Leech (Eds.), *Learner English on Computer* (pp. 119–131). London: Longman.
Granger, S., & Tyson, S. (1996). Connector usage in the English essay writing of native and non-native EFL speakers of English. *World Englishes*, *15*(1), 17–27. https://doi.org/10.1111/j.1467-971X.1996.tb00089.x
Grant, L., & Bauer, L. (2004). Criteria for re-defining idioms: Are we barking up the wrong Tree? *Applied Linguistics*. https://doi.org/10.1093/applin/25.1.38
Grant, L., & Nation, P. (2006). How many idioms are there in English? *ITL - International Journal of Applied Linguistics*, (151), 1–14.
Gries, S. T. (2005). Null-hypothesis significance testing of word frequencies: a follow-up on Kilgarriff. *Corpus Linguistics and Linguistic Theory*, *1*(2), 277–294. https://doi.org/10.1515/cllt.2005.1.2.277
Gries, S. T. (2008). Phraseology and linguistic theory. In S. Granger & F. Meunier (Eds.), *Phraseology: an interdisciplinary perspective* (pp. 3–25). Amsterdam: John Benjamins.
Gu, P. Y. (2003a). Fine brush and freehand: The vocabulary-learning art of two successful Chinese EFL learners. *TESOL Quarterly*, *37*(1), 73–104. https://doi.org/10.2307/3588466
Gu, P. Y. (2003b). Vocabulary learning in a second language: Person, task, context and strategies. *Tesl-Ej*, *7*, 1–25. https://doi.org/ISSN 1072-4303
Gu, P. Y., & Johnson, R. K. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning*, *46*(4), 643–679. https://doi.org/10.1111/j.1467-1770.1996.tb01355.x
Gupta, P., & Tisdale, J. (2009). Does phonological short-term memory causally determine vocabulary learning? Toward a computational resolution of the debate. *Journal of Memory and Language*, *61*(4), 481–502. https://doi.org/10.1016/j.jml.2009.08.001
Ha, M.-J. (2016). Linking adverbials in first-year Korean university EFL learners' writing: a corpus-informed analysis. *Computer Assisted Language Learning*, *29*(6), 1090–1101. https://doi.org/10.1080/09588221.2015.1068814
Halliday, M. A. K. (1979). Modes and meaning and modes of expression: types of grammatical structure and their determination by different semantic functions. In D. J. Allerton, E. Carney, & D. Holdcroft (Eds.), *Function and context in linguistic analysis: essays offered to William Haas* (pp. 57–79). Cambridge: Cambridge University Press.
Halliday, M. A. K. (1989). *Spoken and written language*. Oxford: Oxford University Press.
Halliday, M. A. K. (1994). *An introduction to Functional Grammar*. London: Arnold.
Halliday, M. A. K., & Hasan, R. (1989). *Language, context and text: Aspects of language in a social-semiotic perspective*. Oxford: Oxford University Press.
Halliday, M. A. K., & Martin, J. R. (1993). *Writing science: literacy and discursive power*. London: The Falmer Press.
Hamada, M., & Koda, K. (2008). Influence of first language orthographic experience on second language decoding and word learning. *Language Learning*, *58*(1), 1–31. https://doi.org/10.1111/j.1467-9922.2007.00433.x
Hancioğlu, N., Neufeld, S., & Eldridge, J. (2008). Through the looking glass and into the land of lexico-grammar. *English for Specific Purposes*, *27*, 459–479. https://doi.org/http://dx.doi.org/10.1016/j.esp.2008.08.001
Hanks, P. (1987). Definitions and explanations. In J. Sinclair (Ed.), *Looking Up: An Account of the COBUILD Project in Lexical Computing and the Development of the Collins COBUILD English Language Dictionary* (pp. 123–136). London: Collins ELT.
Hanks, P. (1994). Linguistic norms and pragmatic exploitations or, why lexicographers need prototype theory, and vice versa. In F. Kiefer, G. Kiss & J. Pajzs (Eds.), *Papers in Computational Lexicography: Complex 94* (pp. 89–113).

Hanks, P. (2004). Corpus Pattern Analysis. In G. Williams & S. Vessier, (Eds.) *Proceedings of the Eleventh EURALEX International Congress*. Universite de Bretagne-Sud, Lorient: EURALEX.
Hanks, P. (2006). Metaphoricity is gradable. In A. Stefanowitsch & S. T. Gries (Eds.), *Corpus-Based Approaches to Metaphor and Metonymy* (pp. 17–35). Berlin: Mouton de Gruyter.
Hanks, P. (2008). *Mapping meaning onto use: a Pattern Dictionary of English Verbs*. Paper presented at 7th Conference of the American Association for Corpus Linguistics (AACL 2008), Brigham Young University, Provo, Utah.
Hanks, P. (2009). Stephen J. Perrault (Ed.). Merriam-Webster's Advanced Learner's English Dictionary. *International Journal of Lexicography*, *22*(3), 301–315. https://doi.org/10.1093/ijl/ecp019
Hanks, P. (2010). Nine issues in metaphor theory and analysis. *International Journal of Corpus Linguistics*, *15*(1), 133–150. https://doi.org/10.1075/ijcl.15.1.06han
Hanks, P. (2013). *Lexical analysis: Norms and exploitations*. Cambridge, MA: MIT Press.
Hanks, P., & Pustejovsky, J. (2005). A pattern dictionary for natural language processing. *Revue Française de Linguistique Appliquée*, *10*, 63–82.
Hargens, L. L. (1988). Scholarly consensus and journal rejection rates. *American Sociological Review*, *53*(1), 139–51. https://doi.org/10.2307/2095739
Hartmann, R. (1987). Four perspectives on dictionary use: a critical review of research methods. In A. P. Cowie (Ed.), *The Dictionary and the Language Learner: Papers from the EURALEX Seminar the University of Leeds, 1-3 April 1985. Lexicographica Series Maior 17* (pp. 11–28). Tübingen: Niemeyer.
Harvey, D. (2006). Editorial: The geographies of critical geography. *Transactions of the Institute of British Geographers*, *31*, 409–412.
Harvey, K., & Yuill, D. (1997). A Study of the Use of a Monolingual Pedagogical Dictionary by Learners of English Engaged in Writing. *Applied Linguistics*, *18*(3), 253–278. Retrieved from http://dx.doi.org/10.1093/applin/18.3.253
Harzing, A. W., & Wal, R. Van Der. (2009). A google scholar h-index for journals: An alternative metric to measure journal impact in economics and business. *Journal of the American Society for Information Science and Technology*, *60*(1), 41–46. https://doi.org/10.1002/asi.20953
Hasegawa, T., Sekine, S., & Grishman, R. (2004). Discovering relations among named entities from large corpora. In *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics* (Article No.: 415). Stroudsburg, PA: Association for Computational Linguistics. https://doi.org/10.3115/1218955.1219008
Hatherall, G. (1984). Studying dictionary use: some findings and proposals. In R. Hartmann (Ed.), *LEX'eter '83 Proceedings: Papers from the International Conference on Lexicography at Exeter, 9-12 September 1983. Lexicographica Series Maior I* (pp. 183–189). Tübingen: Niemeyer.
Hausman, F. (1985). Kollokationen im deutshen Wörterbuch: Ein Beitrag zur Theorie des lexikographisscen Beispiels. In B. Henning & J. Mugdan (Eds.), *Lexikographie und Grammatik* (pp. 118–129). Tübingen: Max Niemeyer Verlag.
Haywood, S. (n.d.). Using the Academic Word List. Retrieved July 24, 2017, from https://www.nottingham.ac.uk/alzsh3/acvocab/index.htm#background
Hazenburg, S., & Hulstijn, J. (1996). Defining a minimal receptive second-language vocabulary for non-native university students. *Applied Linguistics*, *17*, 145–163.
Heath, D. (1982). The treatment of grammar and syntax in monolingual English dictionaries of advanced learners. *Linguistik Und Didaktik*, *13*(49), 95–107.
Heid, U., & Zimmermann, J. T. (2012). Usability testing as a tool for e-dictionary design: collocations as a case in point. In R. V. Fjeld & J. M. Torjusen (Eds.), *Proceedings of the*

15th EURALEX International Congress (pp. 661–671). Oslo: Department of Linguistics and Scandinavian Studies, University of Oslo.
Herbst, T. (1986). Defining with a controlled defining vocabulary in foreign learners' dictionaries. *Lexicographica*, *2*, 101–119. https://doi.org/doi:10.1515/9783110244069.101
Herbst, T. (1996). On the way to the perfect learners' dictionary: A first comparison of OALD5, LDOCE3, COBUILD2 and CIDE. *International Journal of Lexicography, 9*(4), 321–357. https://doi.org/10.1093/ijl/9.4.321
HESA. (2017). Students in Higher Education 2015/16. Retrieved July 25, 2017, from https://www.hesa.ac.uk/data-and-analysis/publications/students-2015-16/introduction
Hicks, D. (2004). The four literatures of the social sciences. In H. F. Moed (Ed.), *Handbook of quantitative science and technology research* (pp. 473–496). Dordrecht: Kluwer Academic Publishers.
Higgins, J. (1966). Hard Facts: Notes on teaching English to science dtudents. *ELT Journal*, *21*(1), 55–60.
Hill, L. (1969). Delayed copying. *ELT Journal*, *23*(3), 238–9.
Hirsch, J. E. (2005). An index to quantify an individual's s scientific research output. *Proceedings of the National Academy of Sciences*, *102*(46), 16569–16572. https://doi.org/10.1073/pnas.0507655102
Hoey, M. (2005). *Lexical priming*. London: Routledge.
Hornby, A. S. (1954). *A guide to patterns and usage in English*. London: Oxford University Press.
Horst, M., Cobb, T., & Meara, P. (1998). Beyond a clockwork orange: acquiring second language vocabulary through reading. *Reading in a Foreign Language*, *11*(2), 207–23.
Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, *19*(1), 24–44. https://doi.org/10.1093/applin/19.1.24
Howatt, A. P. R. (1984). *A history of English language teaching*. Oxford: Oxford University Press.
Hsu, W. (2011). A business word list for prospective EFL business postgraduates. *The Asian ESP Journal*, *7*(4), 63–99.
Hsu, W. (2014). A formulaic sequences list for prospective EFL business postgraduates. *The Asian ESP Journal*, *10*(2), 114–162.
Hu, G., & Cao, F. (2011). Hedging and boosting in abstracts of applied linguistics articles: A comparative study of English- and Chinese-medium journals. *Journal of Pragmatics*, *43*(11), 2795–2809. https://doi.org/10.1016/j.pragma.2011.04.007
Hu, H., & Nation, P. (2000). Unknown vocabulary density and reading Comprehension. *Reading in a Foreign Language*, *13*(1), 403–30.
Huang, F., Finkelstein, M., & Rostan, M. (2013). *The internationalization of the academy: Changes, realities and prospects*. New York: Springer.
Huang, G. F. (1985). The productive use of EFL dictionaries. *RELC Journal*, *16*(2), 54–71. https://doi.org/10.1177/003368828501600209
Huang, M., & Chang, Y. (2008). Characteristics of research output in social sciences and Humanities: From a research evaluation perspective. *Journal of the American Society for Information Science and Technology*, *59*(11), 1819–1828. https://doi.org/10.1002/asi.20885
Hulstijn, J. (1988). Experiments with semi-artificial input in second language acquisition research. *Scandinavian Working Papers on Bilingualism*, *8*, 28–40.
Hulstijn, J. (1993). When do foreign-language readers look up the meaning of unfamiliar words? The influence of task and learner variables. *The Modern Language Journal*, *77*(2), 139–147. https://doi.org/10.2307/328937

Hulstijn, J., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: the influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal*, *80*(3), 327–339. https://doi.org/10.1177/1745691612459060.

Humphrey, S. (2016). EAP in school settings. In K. Hyland & P. Shaw (Eds.), *The Routledge Handbook of English for Academic Purposes* (pp. 447–460). Oxford: Routledge.

Hunston, S., & Francis, G. (2000). *Pattern Grammar*. Amsterdam: John Benjamins.

Hyland, K. (1999). Academic attribution: Citation and the construction of disciplinary knowledge. *Applied Linguistics*, *20*(3), 341–367. https://doi.org/10.1093/applin/20.3.341

Hyland, K. (2006). *English for Academic Purposes: An advanced resource book*. Oxford: Routledge.

Hyland, K. (2008a). Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics*, *18*(1), 41–62. https://doi.org/10.1111/j.1473-4192.2008.00178.x

Hyland, K. (2008b). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, *27*(1), 4–21. https://doi.org/http://dx.doi.org/10.1016/j.esp.2007.06.001

Hyland, K. (2016a). *Academic Publishing: Issues and challenges in the construction of knowledge*. Oxford: Oxford University Press.

Hyland, K. (2016b). General and specific EAP. In K. Hyland & P. Shaw (Eds.), *The Routledge Handbook of English for Academic Purposes* (pp. 17–29). Abingdon: Routledge.

Hyland, K., & Tse, P. (2007). Is There an "Academic Vocabulary"? *TESOL Quarterly*, *41*, 235–253. https://doi.org/10.1002/j.1545-7249.2007.tb00058.x

Hyland, K., & Tse, P. (2009). Academic lexis and disciplinary practice: corpus evidence for specificity. *International Journal of English Studies*, *9*, 111–130.

Hymes, D. H. (1962). The ethnography of speaking. In T. Gladwin & W. C. Strurtevant (Eds.), *Anthropology & Human Behavior* (pp. 13–53). Washington, DC.

IELTS. (2017). IELTS home of the English language test. Retrieved October 7, 2017, from https://www.ielts.org/

Ilson, R. (1983). Etymological information: can it help our students? *ELT Journal*, *37*(1), 76–82. https://doi.org/10.1093/elt/37.1.76

Jackendoff, R. (1997). Twistin' the Night Away. *Language*, *73*, 534–559. https://doi.org/10.2307/415883

Jackendoff, R. (2002). *Foundations of Language*. Oxford: Oxford University Press.

Jacobs, G., Dufon, P., & Hong, F. C. (1994). L1 and L2 vocabulary glosses in L2 reading passages: Their effectiveness for increasing comprehension and vocabulary knowledge. *Journal of Research in Reading*. https://doi.org/10.1111/j.1467-9817.1994.tb00049.x

Jacobs, G., Zingraf, S., Wormuth, D., Hartfiel, V., & Hughey, J. (1981). *Testing ESL composition: A practical approach*. Rowley, MA: Newbury House.

Jain, P. (1981). On meaning in the foreign learner's dictionary. *Applied Linguistics*, *2*(3), 274–286. Retrieved from http://dx.doi.org/10.1093/applin/II.3.274

Jansen, J., Margeai, J., & Vanandroye, J. (1987). Controlling LDOCE's controlled vocabulary. In A. P. Cowie (Ed.), *The Dictionary and the Language Learner: Papers from the EURALEX Seminar at the University of Leeds, 1-3 April 1985* (pp. 78–94). Tübingen: Max Niemeyer Verlag.

Jenkins, J. (2011). Accommodating (to) ELF in the international university. *Journal of Pragmatics*, *43*(4), 926–936. https://doi.org/10.1016/j.pragma.2010.05.011

Jenkins, J. (2014). *English as a lingua franca in the international university: The politics of academic english language policy*. Abingdon: Routledge.

Jenkins, J., Cogo, A., & Dewey, M. (2011). Review of developments in research into English as a lingua franca. *Language Teaching*, *44*(3), 281–315. https://doi.org/10.1017/S0261444811000115
Jespersen, O. (1924). Living grammar. In *The Philosophy of Grammar* (pp. 17–29). London: Allen & Unwin.
Jiang, N., & Nekrasova, T. (2007). The processing of formulaic sequences by second language speakers. *Modern Language Journal*, *91*(3), 433–445. https://doi.org/10.1111/j.1540-4781.2007.00589.x
Joe, A. (1995). Text-based tasks and incidental vocabulary learning. *Second Language Research*, *11*(2), 149–158. https://doi.org/10.1177/026765839501100206
Johnson, M. (1987). *The body in the mind*. Chicago: University of Chicago Press.
Jordan, R. (1997). *English for Academic Purposes: a guide and resource book for teachers*. Cambridge: Cambridge University Press.
Jordan, R. (2002). The growth of EAP in Britain. *Journal of English for Academic Purposes*, *1*(1), 69–78. https://doi.org/10.1016/S1475-1585(02)00004-8
Jung, J. (2016). Effects of glosses on learning of L2 grammar and vocabulary. *Language Teaching Research*. https://doi.org/10.1177/1362168815571151
Kaneta, T. (2011). Folded or unfolded: Eye-tracking analysis of L2 learners' reference behavior with different types of dictionary. In K. Akasu & S. Uchida (Eds.), *ASIALEX2011 Proceedings Lexicography: Theoretical and practical perspectives* (pp. 219–224). Kyoto: Asian Association for Lexicography.
Kaplan, R. M., & Bresnan, J. (1982). Lexical-Functional grammar: A formal system for grammatical representation. In J. Bresnan (Ed.), *The mental representation of grammatical representation* (pp. 173–281). Cambridge, MA: MIT Press.
Kashiha, H. (2015). Recurrent formulas and moves in writing research article conclusions among native and nonnative writers. *The Southeast Asian Journal of English Language Studies*, *21*(1), 47–59.
Kass, R., & Raftery, A. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572
Keating, G. D. (2008). Task effectiveness and word learning in a second language: The involvement load hypothesis on trial. *Language Teaching Research*, *12*(3), 365–386. https://doi.org/10.1177/1362168808089922
Kilgarriff, A. (2001). Comparing corpora. *International Journal of Corpus Linguistics*, *6*(1), 97–133. https://doi.org/10.1075/ijcl.6.1.05kil
Kim, Y. (2008). The contribution of collaborative and individual tasks to the acquisition of L2 vocabulary. *Modern Language Journal*, *92*(1), 114–130. https://doi.org/10.1111/j.1540-4781.2008.00690.x
Knight, S. (1994). Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities. *The Modern Language Journal*, *78*(3), 285–299. https://doi.org/10.2307/330108
Ko, M. H. (2012). Glossing and second language vocabulary learning. *TESOL Quarterly*. https://doi.org/10.1002/tesq.3
Kojic-Sabo, I., & Lightbown, P. (1999). Students' approaches to vocabulary learning and their relationship to success. *The Modern Language Journal*, *83*(2), 176–192. https://doi.org/10.1111/0026-7902.00014
Konstantakis, N. (2007). Creating a business word list for teaching Business English. *Estudios de Lingüística Inglesa Aplicada (ELIA)*, *7*, 79–102. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.121.5133
Koprowski, M. (2005). Investigating the usefulness of lexical phrases in contemporary coursebooks. *ELT Journal*, *59*(4), 322–32. https://doi.org/10.1093/elt/cci061

Kovalyuk, Y. (2016, June). *Finding the common ground: phraseological theory and terminology revisited*. Paper presented at the 7th Formulaic Language Research Network conference, Vilnius

Krashen, S. (2010). Academic language proficiency: acquired or learned? In Y. Leung (Ed.), *Proceedings from the 19th International Synposium on English Language Teaching* (pp. 34–43). Taipei.

Kremmel, B., Brunfaut, T., & Alderson, C. (2016, June). How important is knowledge of multi-word expressions for reading comprehension? Paper presented at the 7th Formulaic Language Research Network conference, Vilnius.

Kučera, H., & Francis, W. N. (1978). *Computational analysis of present-day American English*. Gloucestershire: Brown University Press.

Kuteeva, M., & McGrath, L. (2013). The theoretical research article as a reflection of disciplinary practices: The case of pure mathematics. *Applied Linguistics*, *36*(2), 215–235. https://doi.org/10.1093/applin/amt042

Laakso, M., Welling, P., Bukvova, H., Nyman, L., Björk, B. C., & Hedlund, T. (2011). The development of open access journal publishing from 1993 to 2009. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0020961

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. *Language* (Vol. 64). https://doi.org/10.2307/415440

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Landau, S. I. (2001). *Dictionaries: The art and craft of lexicography*. Cambridge: Cambridge University Press.

Langacker, R. W. (1986). Abstract motion. In V. Nikiforidou, M. VanClay, V. Nikiforidou, & D. Feder (Eds.), *Proceedings of the 12th Annual Meeting of the Berkeley Linguistics Society* (pp. 445–471).

Langacker, R. W. (1987). *Foundations of Cognitive Grammar. Vol. 1: Theoretical prerequisites*. Redwood, CA: Stanford University Press.

Langacker, R. W. (1990). Subjectification. *Cognitive Linguistics*, *1*(1), 5–38. https://doi.org/10.1515/cogl.1990.1.1.5

Langacker, R. W. (1991). *Foundations of cognitive grammar. Vol. 2: Descriptive application*. Redwood, CA: Stanford University Press.

Langacker, R. W. (1999). A view from cognitive linguistics. *Behavioural and Brain Sciences, 2*(4). https://doi.org/10.1017/S0140525X99392141

Langacker, R. W. (2005). Construction grammars: Cognitive, radical, and less so. In F. J. Ruiz de Mendoza Ibáñez & M. S. Peña Cervel (Eds.), C*ognitive Linguistics: Internal Dynamics and Interdisciplinary Interaction* (pp. 101–159). Berlin: Mounton de Gruyter.

Langacker, R. W. (2008). *Cognitive Grammar: A Basic Introduction*. New York: Oxford University Press

Laufer, B. (1989a). A factor of difficulty in vocabulary learning: deceptive transparency. *AILA Review*, *6*, 10–20.

Laufer, B. (1989b). What percentage of text-lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans to thinking machines* (pp. 126–32). Clevedon: Multilingual Matters.

Laufer, B. (1993). The effect of dictionary definitions and examples on the use and comprehension of new L2 words. *Cahiers de Lexicologie*, *63*(2), 131–142.

Laufer, B. (1998). The development of passive and active vocabulary in a second language: same or different? *Applied Linguistics*, *19*(2), 255–271. Retrieved from http://applij.oxfordjournals.org/content/19/2/255.abstract

Laufer, B. (2000). Electronic dictionaries and incidental vocabulary acquisition: does technology make a difference? In. U. Heid, S. Evert. E. Lehmann, & C. Rohrer

(Eds.), *Proceedings of the 9th EURALEX International Congress* (pp. 849–854). Stuttgart: Stuttgart University Press
Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, *22*(1), 1–26. https://doi.org/10.1093/applin/22.1.1
Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, *16*(3), 307–322. https://doi.org/10.1093/applin/16.3.307
Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, *16*(1), 36–55. https://doi.org/10.1191/026553299672614616
Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES--ELAN system. *Behavior Research Methods*, *41*(3), 841–9. https://doi.org/10.3758/BRM.41.3.841
Lawrence, S. (2001). Free online availability substantially increases a paper's impact. *Nature*, *411*, 521. Retrieved from http://www.nature.com/nature/journal/v411/n6837/full/411521a0.html?foxtrotcallback=true
Lawson, M., & Hogben, D. (1996). The Vocabulary-Learning Strategies of Foreign-Language Students. *Language Learning*. https://doi.org/10.1111/j.1467-1770.1996.tb00642.x
Le, T. N. P., & Harrington, M. (2015). Phraseology used to comment on results in the discussion section of applied linguistics quantitative research articles, *39*, 45–61. https://doi.org/10.1016/j.esp.2015.03.003
Lea, D. (2014a). Making a Learner's Dictionary of Academic English. In A. Abel, C. Vettori, & N. Ralli (Eds.), *Proceedings of the 16th EURALEX International Congress* (pp. 181–189). Bolzano, Italy: EURAC research.
Lea, D. (2014b). Oxford Corpus of Academic English Archives - Oxford University Press. Retrieved September 1, 2017, from https://oupeltglobalblog.com/tag/oxford-corpus-of-academic-english/
Lea, M. R., & Street, B. V. (1998). Student writing in higher education: An academic literacies approach. *Studies in Higher Education*, *23*(2), 157–172. https://doi.org/10.1080/03075079812331380364
Leech, G. (1991). The state of the art in corpus linguistics. In K. Aijmer & B. Altenberg (Eds.), *English corpus linguistics: studies in honour of Jan Svartvik* (pp. 8–29). London: Longman.
Leech, G. (1992). Corpora and the theories of linguistic performance. In J. Svartvik (Ed.), *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991* (pp. 105–122). Berlin: Mouton de Gruyter.
Leech, G., Rayson, P., & Wilson, A. (2001). *Word frequencies in written and spoken English*. Harlow: Longman.
Leeke, P., & Shaw, P. (2000). Learners' independent records of vocabulary. *System*, *28*(2), 271–289. https://doi.org/10.1016/S0346-251X(00)00012-9
Lei, L., & Liu, D. (2016). A new medical academic word list: A corpus-based study with enhanced methodology. *Journal of English for Academic Purposes*, *22*, 42–53. https://doi.org/10.1016/j.jeap.2016.01.008
Leki, I., & Carson, J. (1994). Students' perceptions of EAP writing instruction and writing needs across the disciplines. *TESOL Quarterly*, *28*(1), 81–101. https://doi.org/10.2307/3587199
Lemmens, M., & Wekker, H. (1986). *Grammar in English Learners' Dictionaries*. Tübingen: Max Niemeyer Verlag.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago: The University of Chicago Press.
Lew, R. (2002). Questionnaires in dictionary use research: A re-examination. In A. Braasch & C. Povlsen (Eds.), *Proceedings of the 10th EURALEX International Congress* (pp. 267–271). København, Denmark: Center for Sprogteknologi.
Lew, R. (2004). How do Polish learners of English rate bilingual and monolingual dictionaries? In G. Williams & S. Vessier (Eds.), *Proceedings of the 11th EURALEX International Congress* (pp. 697–706). Lorient, France: Université de Bretagne-Sud.
Lew, R. (2010). Users take shortcuts: Navigating dictionary entries. In A. Dykstra & T. Schoonheim (Eds.), *Proceedings of the 14th EURALEX International Congress* (pp. 1121–1132). Leeuwarden/Ljouwert, The Netherlands: Fryske Akademy.
Lew, R. (2011). User Studies: Opportunities and limitations. In K. Akasu & S. Uchida (Eds.), *ASIALEX2011 Proceedings Lexicography: Theoretical and practical perspectives* (pp. 7–16). Kyoto: Asian Association for Lexicography.
Lew, R. (2016). Can a dictionary help you write better? A user study of an active bilingual dictionary for polish learners of English. *International Journal of Lexicography*, *29*(3), 353–366. Retrieved from http://dx.doi.org/10.1093/ijl/ecw024
Lew, R., & De Schryver, G.-M. (2014). Dictionary users in the digital revolution. *International Journal of Lexicography*, *27*(4), 341–359. https://doi.org/10.1093/ijl/ecu011
Lew, R., & Doroszewska, J. (2009). Electronic dictionary entries with animated pictures: lookup preferences and word retention. *International Journal of Lexicography, 22*(3), 239–257. https://doi.org/10.1093/ijl/ecp022
Lew, R., & Dziemianko, A. (2006a). A new type of folk-inspired definition in English monolingual learners' dictionaries and its usefulness for conveying syntactic information. *International Journal of Lexicography*, *19*(3), 225–242. https://doi.org/10.1093/ijl/ecl011
Lew, R., & Dziemianko, A. (2006b). Non-standard dictionary definitions: What they cannot tell native speakers of Polish. *Cadernos de Traduçao*, *18*, 275–294. Retrieved from http://www.periodicos.ufsc.br/index.php/traducao/article/view/6953/6460
Lew, R., & Dziemianko, A. (2012). Single-clause when-definitions: Take three. In *Proceedings of the 15th EURALEX International Congress* (pp. 997–1002). https://doi.org/http://hdl.handle.net/10593/3697
Lew, R., Grzelak, M., & Leszkowicz, M. (2013). How dictionary users choose senses in bilingual dictionary entries: An eye-tracking study. *Lexikos*, *23*, 228–254.
Lewis, M. (1993). *The lexical approach: the state of ELT and a way forward*. Hove: Language Teaching Publications.
Li, S., & Pemberton, R. (1994). An investigation of students' knowledge of academic and subtechnical vocabulary. In *Proceedings of the Joint Seminar on Corpus Linguistics and Lexicography* (pp. 183–196). Hong Kong: Hong Kong University of Science and Technology.
Li, Y., & Qian, D. D. (2010). Profiling the Academic Word List (AWL) in a financial corpus. *System*, *38*, 402–411. https://doi.org/10.1016/j.system.2010.06.015
Liao, Y., & Fukuya, Y. (2002). Avoidance of phrasal verbs: The case of Chinese learners of English. *Language Learning*, *54*(2), 193–226. https://doi.org/10.1111/j.1467-9922.2004.00254.x
Lillis, T. M. (2003). Student writing as "academic literacies": Drawing on Bakhtin to move from critique to design. *Language and Education*, *17*(3), 192–207. https://doi.org/10.1080/09500780308666848
Liu, D. (2010). Going beyond patterns: Involving cognitive analysis in the learning of collocations. *TESOL Quarterly*, *44*(1), 4–30. https://doi.org/10.5054/tq.2010.214046

Liu, J., & Han, L. (2015). A corpus-based environmental academic word list building and its validity test. *English for Specific Purposes*, *39*, 1–11. https://doi.org/10.1016/j.esp.2015.03.001

Long, M., & Ross, S. (1993). Modifications That Preserve Language and Content. In M. Tickoo (Ed.), *Simplification: Theory and Application. RELC Anthology Series 31* (pp. 29–52). Singapore: SEAMEO-RELC

López-Jiménez, M. (2010). The treatment of lexical aspects in commercial textbooks for l2 teaching and learning. In R. Chacón-Beltrán, C. Abello-Contesse, & M. Torreblanca-López (Eds.), *Insights into non-native vocabulary teaching and learning* (pp. 156–74). Bristol: Multilingual Matters.

Louw, W. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and Technology in Honour of John Sinclair* (pp. 157–176). Amsterdam: John Benjamins.

Luppescu, S., & Day, R. (1993). Reading, dictionaries, and vocabulary learning. *Language Learning*, (42), 263–287.

Lynn, R. (1973). Preparing Word-Lists: A Suggested Method. *RELC Journal*, *4*, 25–28.

MacFarquhar, P., & Richards, J. (1983). On dictionaries and definitions. *RELC Journal*, *14*(1), 111–124. https://doi.org/10.1177/003368828301400112

Maci, S. M. (2015). "These data support the provocative view that..." Evaluation in medical academic posters. In M. Bondi, S. Cacchiani, & D. Mazzi (Eds.), *Discourse In and Through the Media: Recontextualizing and Reconceptualizing Expert Discourse* (pp. 119–145). Newcastle Upon Tyne: Cambridge Scholars Publishing.

Maingay, S., & Rundell, M. (1987). Anticipating learners' errors: implications for dictionary writers. In A. P. Cowie (Ed.), *The Dictionary and the Language Learner: Papers from the EURALEX Seminar at the University of Leeds, 1-3 April 1985* (pp. 128–135). Tübingen: Max Niemeyer Verlag.

Manning, C., & Schütze, H. (1999). *Foundations of Statistical Natural Language Processing*. Cambridge, Mass: MIT Press.

Manzo, A. V. (1970). CAT- A game for extending vocabulary and knowledge of allusions. *Journal of Reading*, *13*(5), 367–369.

Marco, M. J. L. (2000). Collocational frameworks in medical research papers: a genre-based study. *English for Specific Purposes*, *19*(2000), 63–86. https://doi.org/10.1016/S0889-4906(98)00013-1

Marcus, M., Santorini, B., & Marcinkiewicz, M. (1993). Building a large annotated corpus of English: the Penn Treebank. *Computational Linguistics*, *19*, 313–330. https://doi.org/citeulike-article-id:461221

Marello, C. (1987). Examples in contemporary italian bilingual dictionaries. In A. P. Cowie (Ed.), *The Dictionary and the Language Learner: Papers from the EURALEX Seminar the University of Leeds, 1-3 April 1985. Lexicographica Series Maior 17* (pp. 224–237). Berlin: Walter de Gruyter.

Marello, C. (1989). *Dizionari bilingui: con schede sui dizionari italiani per francese, inglese, spagnolo, tedesco*. Bologna: Zanichelli.

Martin, A. (1976). Teaching academic vocabulary to foreign graduate students. *TESOL Quarterly*, *10*(1), 91–97. Retrieved from http://www.jstor.org/stable/10.2307/3585942

Martínez, I., Beck, S., & Panza, C. (2009). Academic vocabulary in agriculture research articles: A corpus-based study. *English for Specific Purposes*, *28*(3), 183–198. https://doi.org/10.1016/j.esp.2009.04.003

Martinez, R., & Murphy, V. (2011). Effect of frequency and idiomaticity on second language reading comprehension. *TESOL Quarterly*, *40*(1), 235–257. https://doi.org/10.5054/tq.201

Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics*, *33*(3), 299–320. https://doi.org/10.1093/applin/ams010
Mazaherylaghab, H. (2013). *Iranian learner English: a corpus based study of phrasal verb usage* (Unpublished doctoral dissertation). University of Freiburg, Freiburg.
McCarten, J., & McCarthy, M. (2010). Bridging the gap between corpus and course book: the case of conversation strategies. In *Perspectives of language learning materials development* (pp. 11–32). Berne: Peter Lang.
McCarthy, M. (1993). *Discourse analysis for language teachers*. Cambridge: Cambridge University Press.
McCarthy, M., & Carter, R. (1994). *Language as discourse*. Cambridge: Cambridge University Press.
McDaniel, M. a., & Pressley, M. (1989). Keyword and context instruction of new vocabulary meanings: Effects on text comprehension and memory. *Journal of Educational Psychology*, *81*(2), 204–213. https://doi.org/10.1037/0022-0663.81.2.204
McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. London: Routledge.
McKeown, M., Beck, I., Omanson, R., & Pople, M. (1985). Some effects of the nature and frequency of vocabulary instruction on the knowledge and use of words. *Reading Research Quarterly*, *20*(5), 522–535. https://doi.org/10.2307/747940
McLaughlin, B. (1990). Restructuring. *Applied Linguistics*, *11*(2), 113–128. https://doi.org/10.1093/applin/11.2.113
McLoskey, D. (1994). *Knowledge and persuasion in economics*. Cambridge: Cambridge University Press.
McWilliam, N. (1998). *What's in a word? Vocabulary development in multilingual classrooms*. Stoke-on-Trent: Trentham Books.
Meara, P. (1990). A note on passive vocabulary. *Second Language Research*, *6*, 150–54.
Meara, P., & English, F. (1988). *Lexical errors and learners' dictionaries* (Unpublished report). Birkbeck College, London. Retrieved from http://files.eric.ed.gov/fulltext/ED290322.pdf
Mel'čuk, I. A. (1988). Semantic description of lexical units in an explanatory combinatorial dictionary. *International Journal of Lexicography*, *1*(3), 165–188.
Mel'čuk, I. A. (1995). Phrasemes in language and phraseology in linguistics. In M. Everaert & E.-J. van der Linden (Eds.), *Idioms: Structural and Psychological perspectives* (pp. 167–232). Hillsdale, NJ: Lawrence Erlbaum.
Mel'čuk, I. A. (2006). The explanatory combinatorial dictionary. In G. Sica (Ed.), *Open Problems in Linguistics and Lexicography* (pp. 225–355). Monza: Polimetrica.
Mel'čuk, I. A. (2013). *Semantics: From meaning to text*. (D. Beck & A. Polguère, Eds.). Amsterdam: John Benjamins.
Melka-Teichroew, F. J. (1982). Receptive vs. productive vocabulary: A survey. *Interlanguage Studies Bulletin (Utrecht)*, *6*, 5–33.
Metcalfe, T. S. (2006). The citation impact of digital preprint archives for solar physics papers. *Solar Physics*, *239*(1–2), 549–553. https://doi.org/10.1007/s11207-006-0262-7
Meyer, P. (1990). Non-technical vocabulary in technical language. In *AILA*. Thessalonica.
Michaelis, L. A. (1994). A case of constructional polysemy in latin. *Studies in Language*, *18*(1), 45–70.
Michaelis, L. A., & Lambrecht, K. (1996). toward a construction-based theory of language function: The case of nominal extraposition. *Language: Journal of the Linguistic Society of America*, *72*(2), 215–247. https://doi.org/10.2307/416650
Miller, G. (1984). How to misread a dictionary. In Y. Wilks (Ed.), *Proceedings of Coling '84* (p. 462). Morristown, NJ: Association for Computational Linguistics.
Miller, G., & Gildea, P. (1987). How children learn words. *Scientific American*, *257*(3), 94–99.

Min, H. (2008). EFL vocabulary acquisition and retention: Reading plus vocabulary enhancement activities and narrow reading. *Language Learning*, *58*(1), 73–115. https://doi.org/10.1111/j.1467-9922.2007.00435.x

Moon, R. (1998). *Fixed Expressions and Idioms in English: A Corpus-based Approach*. Oxford: Clarendon Press.

Morton, T., Walsh, S., & O'Keeffe, A. (2011). Analyzing university spoken interaction: a corpus linguistics/conversation analysis approach. *International Journal of Corpus Linguistics 16*(3), 326–354.

Mudraya, O. (2006). Engineering English: A lexical frequency instructional model. *English for Specific Purposes*, *25*(2), 235–256. https://doi.org/10.1016/j.esp.2005.05.002

Müller-Spitzer, C., Koplenig, A., & Töpel, A. (2011). What makes a good online dictionary? Empirical insights from an interdisciplinary research project. In I. Kosem & K. Kosem (Eds.), *Proceedings of eLEX2011* (pp. 203–208). Trojina: Institute for Applied Slovene Studies.

Müller-Spitzer, C., Koplenig, A., & Töpel, A. (2012). Online dictionary use: Key findings from an empirical research project. In S. Granger & M. Paquot (Eds.), *Electronic Lexicography* (pp. 425–457). Oxford.

Müller-Spitzer, C., Michaelis, F., & Koplenig, A. (2014). Evaluation of a new web design for the dictionary portal OWID. An attempt at using eye-tracking technology. In *Using Online Dictionaries* (pp. 207–228). Berlin: de Gruyter.

Müller-Spitzer, C., Wolfer, S., & Koplenig, A. (2015). Observing online dictionary users: studies using Wiktionary log files. *International Journal of Lexicography*, *28*(1), 1–26. https://doi.org/10.1093/ijl/ecu029

Müllich, H. (1990). *Die Definition Ist Blod!: Herubersetzen Mit Dem Einsprachigen Worterbuch; Das Franzosische Und Englische Lernerworterbuch in Der Hand Der Deutschen Schuler*. Tübingen: Max Niemeyer Verlag.

Munslow, A. (1997). *Deconstructing history*. London: Routledge.

Murali, N. S., Murali, H. R., Auethavekiat, P., Erwin, P. J., Mandrekar, J. N., Manek, N. J., & Ghosh, A. K. (2004). Impact of FUTON and NAA bias on visibility of research. *Mayo Clinic Proceedings. Mayo Clinic*, *79*(8), 1001–1006. https://doi.org/10.4065/79.8.1001

Nagy, W. (1985). Learning words from context. *Reading Research Quarterly*, *20*(2), 233–253. https://doi.org/10.2307/747758

Nagy, W. (1997). On the role of context in first- and second-language vocabulary learning. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary Description acquisition and Pedagogy* (pp. 64–83). Cambridge: Cambridge University Press

Nagy, W., Anderson, R., & Herman, P. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, *24*(2), 237–270. https://doi.org/10.3102/00028312024002237

Nagy, W., & Herman, P. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. McKeown & M. Curtis (Eds.) *The nature of vocabulary acquisition* (pp. 19–35). London: Routledge.

Nakata, T. (2017). Does repeated practice make perfect? The effects of within-session repeated retrieval on second language vocabulary learning. *Studies in Second Language Acquisition*, *39*(4), online preview. https://doi.org/10.1017/S0272263116000280

Nation, P. (1990). *Teaching and learning vocabulary*. Rowley, MA: Newbury House.

Nation, P. (1993). Using dictionaries to estimate vocabulary size: essential, but rarely followed, procedures. *Language Testing*, *10*(1), 27–40. https://doi.org/10.1177/026553229301000102

Nation, P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, *63*, 59–82.

Nation, P. (2013). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
Nation, P. (2016). *Making and using word lists for language learning and testing*. Amsterdam: John Benjamins.
Nation, P., & Crabbe, D. (1991). A survival language learning syllabus for foreign travel. *System*, *19*, 191–201.
Nation, P., & Kyongho, H. (1995). Where would general service vocabulary stop and special purposes vocabulary begin? *System*, *23*, 35–41. https://doi.org/10.1016/0346-251X(94)00050-G
Nation, P., & Moir, J. (2008). Vocabulary learning and the good language learner. In C. Griffiths (Ed.), *Lessons from Good Language Learners* (pp. 159–73). Cambridge: Cambridge University Press.
Nation, P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Boston: Heinle Cengage Learning.
Nation, P., & Yamamoto, A. (2012). Applying the four strands to language learning. *The International Journal of Innovation in English Language Teaching and Research*, *1*(2), 1–15.
Nattinger, J., & DeCarrico, J. (1992). *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.
Nesi, H. (1987). Do dictionaries help students write? In T. Bloor & J. Norrish (Eds.), *Written Language* (pp. 85–97). London: Centre for Information on Language Teaching Research for the British Association for Applied Linguistics.
Nesi, H. (1996). The role of illustrative examples in productive dictionary use. *Dictionaries: Journal of the Dictionary Society of North America*, *17*(1), 198–206. https://doi.org/10.1353/dic.1996.0008
Nesi, H. (2000a). A review of research into dictionary use and requirements. In *The Use and Abuse of EFL Dictionaries: how learners of English as a foreign language read and interpret dictionary entries* (pp. 3–54). Tübingen: Max Niemeyer Verlag.
Nesi, H. (2000b). The effect of dictionary use on performance in a multiple-choice reading comprehension test. In *The Use and Abuse of EFL Dictionaries: How learners of English as a foreign language read and interpret dictionary entries* (pp. 57–69). Tübingen: Max Niemeyer Verlag.
Nesi, H. (2000c). The effect of different dictionary defining styles of productive dictionary use. In *The Use and Abuse of EFL Dictionaries: How learners of English as a foreign language read and interpret dictionary entries* (pp. 71–93). Tübingen: Max Niemeyer Verlag.
Nesi, H. (2014). *Corpora for English for Academic Purposes*. Retrieved from http://www2.warwick.ac.uk/fac/soc/al/research/collect/bawe/papers/corpora_and_eap.pdf (cited in Timmis 2015)
Nesi, H., & Haill, R. (2002). A study of dictionary use by international students. *International Journal of Lexicography*, *14*(4), 277–305.
Nesi, H., & Meara, P. (1994). Patterns of misinterpretation in the productive use of EFL dictionary definitions. *System*, *22*(1), 1–15. https://doi.org/10.1016/0346-251X(94)90036-1
Nesi, H., & Tan, K. H. (2011). The effect of menus and signposting on the speed and accuracy of sense selection. *International Journal of Lexicography*, *24*(1), 79–96. https://doi.org/10.1093/ijl/ecq040
Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, *24*(2), 223–242. https://doi.org/10.1093/applin/24.2.223
Nesselhauf, N. (2005). *Collocations in a Learner Corpus. Studies in corpus linguistics* (Vol. 14). Amsterdam: John Benjamins.

Neubach, A., & Cohen, V. (1988). Processing strategies and problems encountered in the use of dictionaries. *Dictionaries*, *10*, 1–19.
Neubauer, K., & Clahsen, H. (2009). Decomposition of inflected words in a second language. *Studies in Second Language Acquisition*, *31*(3), 403–435.
Newton, J. (2001). Options for vocabulary learning through communication tasks. *ELT Journal*, *55*(1), 30–37. https://doi.org/10.1093/elt/55.1.30
Newton, J. (2013). Incidental vocabulary learning in classroom communication tasks. *Language Teaching Research*, *17*(2), 164–187. https://doi.org/10.1177/1362168812460814
Nist, S., & Olejnik, S. (1995). The role of context and dictionary definitions on varying levels of word knowledge. *Reading Research Quarterly*, *30*(2), 172–193. https://doi.org/10.2307/748031
O'Halloran, K. (2011). Investigating argumentation in reading groups: Combining manual qualitative coding and automated corpus analysis tools. *Applied Linguistics*, *32*(2), 172–196. Retrieved from http://dx.doi.org/10.1093/applin/amq041
O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From Corpus to Classroom*. Cambridge: Cambridge University Press.
Oakes, M. (1998). *Statistics for Corpus Linguistics*. Edinburgh: Edinburgh University Press.
Owen, C. (1993). Corpus-based grammar and the Heineken effect: Lexico-grammatical description for language learners. *Applied Linguistics*, *14*(2), 167–187. https://doi.org/10.1093/applin/14.2.167
Oxford, R. (1990). *Language learning strategies: what every teacher should know*. New York: Newbury House.
Oxford Text Archive. (2015). Creating the BNC - Designing and creating the BNC. Retrieved October 4, 2017, from http://www.natcorp.ox.ac.uk/corpus/creating.xml?ID=design
Palincsar, A., & Brown, A. (1986). Interactive promote learning teaching independent from text to. *The Reading Teacher*, *39*(8), 771–777. https://doi.org/10.2307/20199221
Palmer, D. (1982). Information transfer for listening and reading. *English Teaching Forum*, *20*(1), 29–33.
Palmer, H. (1989). Conversation. In *The Writing of Harold Palmer: an overview* (pp. 185–91). Tokyo: Hon-no-Tomosha.
Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language*, *30*(3), 331–347. https://doi.org/10.1016/0749-596X(91)90040-Q
Paquot, M. (2007). Towards a productively-oriented academic word list. In J. Walinski, K. Kredens, & S. Gozdz-Roszkowski (Eds.), *Corpora and ICT in Language Studies. PALC 2005* (Vol. 15, pp. 127–140). Frankfurt am Main: Peter Lang.
Paquot, M. (2010). *Academic Vocabulary in Learner Writing: From Extraction to Analysis*. London & New York: Continuum.
Paquot, M. (2012). The LEAD dictionary-cum-writing aid: an integrated dictionary and corpus tool. In S. Granger & M. Paquot (Eds.), *Electronic Lexicography* (pp. 163–185). Oxford: Oxford University Press.
Paquot, M. (2014). Cross-linguistic influence and formulaic language: Recurrent word sequences in French learner writing. *EUROSLA Yearbook*, *14*(1), 240–261. https://doi.org/http://dx.doi.org/10.1075/eurosla.14.10paq
Paquot, M., & de Cock, S. (2010). The monolingual learners' dictionary as a productive tool: the contribution of learner corpora. In M. C. Campoy, B. Belles-Fortuno, & M. L. Gea-Valor (Eds.), *Corpus-Based Approaches to English Language Teaching* (pp. 195–204). London: Bloomsbury.

Paquot, M., & Granger, S. (2015). Electronic lexicography goes local: Design and structures of a needs-driven online academic writing aid. *Lexicographica*, *31*(1), 118–141.
Paribakht, T., & Wesche, M. (1996). Enhancing vocabulary acquisition through reading: A hierarchy of text-related exercise types. *Canadian Modern Language Review*, *52*(2), 155–78.
Paulus, T. M., & Lester, J. N. (2015). ATLAS.ti for conversation and discourse analysis studies. *International Journal of Social Research Methodology*, *19*(4), 405–428. https://doi.org/10.1080/13645579.2015.1021949
Paulus, T. M., Lester, J. N., & Dempster, P. (2014). *Digital tools for qualitative research*. London: Sage.
Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike competence. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191–227). Harlow: Longman.
Peacock, M., & Flowerdew, J. (2001). *Research Perspectives on English for Academic Purposes*. Cambridge: Cambridge University Press.
Pearsall, J. (Ed.). (1998). *The New Oxford Dictionary of English*. Oxford: Clarendon Press.
Pecman, M. (2008). Compilation, formalisation and presentation of bilingual phraseology: Problems and possible solutions. In F. Meunier & S. Granger (Eds.) *Phraseology in foreign language learning and teaching* (pp. 203–222). Amsterdam: John Benjamins
Pérez-Llantada, C. (2014). Formulaic language in L1 and L2 expert academic writing: Convergent and divergent usage, *Journal of English for Academic Purposes 14*, 84–94. https://doi.org/10.1016/j.jeap.2014.01.002
Pérez-Llantada, C., Plo, R., & Ferguson, G. R. (2011). "You don't say what you know, only what you can": The perceptions and practices of senior Spanish academics regarding research dissemination in English. *English for Specific Purposes*, *30*(1), 18–30. https://doi.org/10.1016/j.esp.2010.05.001
Perfetti, C. (2010). Decoding, vocabulary, and comprehension. In M. McKeown & L. Kucan (Eds.), *Bringing reading research to life* (pp. 291–303). New York: Guilford Press.
Perfetti, C., & Hart, L. (2001). The lexical basis of comprehension skill. In D. Gorfien (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity* (pp. 67–86). Washington, DC: American Psychological Association.
Perfetti, C., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of Functional Literacy* (pp. 189–213). Amsterdam: John Benjamins.
Praninskas, J. (1972). *American University Word List*. Harlow: Longman. Retrieved from
Prince, P. (1996). Second language vocabulary learning: The role of context versus translations as a function of proficiency. *The Modern Language Journal*, *80*(4), 478–493. https://doi.org/doi: 10.1111/j.1540-4781.1996.tb05468.x
Prodromou, L. (2008). *English as a lingua franca: A corpus-based analysis*. London: Continuum.
Pustejovsky, J. (1995). *The generative lexicon*. Cambridge, MA: MIT Press.
Qian, D. (2004). Second language lexical inferencing: Preferences, prescriptions, and practices. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a Second Language: Selection, Acquisition and Testing* (pp. 155–69). Amsterdam: John Benjamins.
Qian, L. (2015). The processing on different types of formulaic sequences. *English Language Teaching*, *8*(8), 56–62.
Quirk, R. (1974). The image of the dictionary. In R. Quirk (Ed.), *The Linguist and the English Language* (pp. 148–163). London: Edward Arnold.
Rankin, E., & Overholser, B. (1969). Reaction of intermediate grade children to contextual clues. *Journal of Literacy Research*, *1*(3), 50–73. https://doi.org/10.1080/10862966909546857

Rayson, P. (2008). From key words to key semantic domains. *International Journal of Corpus Linguistics*, *13*(4), 519–549. https://doi.org/10.1075/ijcl.13.4.06ray
Rayson, P. (2016). *Log-likelihood and effect size calculator* (Excel spreadsheet). Retrieved from http://ucrel.lancs.ac.uk/llwizard.html
Rayson, P., & Garside, R. (2000). Comparing corpora using frequency profiling. In *Proceedings of the workshop on comparing corpora held in conjunction with the 38th annual meeting of the Association for Computational Linguistics (ACL 2000)* (Vol. 9, pp. 1–6). https://doi.org/10.3115/1117729.1117730
Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press. Retrieved from
Read, J., & Nation, P. (2004). Measurement of formulaic sequences. In *Formulaic sequences: Acquisition, processing, and use* (pp. 23–36).
Rees, G. P. (2013). *Adiscipline-specific phraseological approach to vocabulary selection for English for academic purposes* (Unpublished Masters Thesis). Universitat Pompeu Fabra, Barcelona.
Rees, G. P. (2016). Corpus evidence for a discipline-specific phraseological approach to academic vocabulary. *Research in Corpus Linguistics*, *4*, 61–74.
Renau, I., & Battaner, P. (2012). Using CPA to represent Spanish pronominal verbs in a learners' dictionary. In R. V. Fjeld & J. M. Torjusen (Eds.), *Proceedings of the 15th EURALEX International Congress* (pp. 350–361). Oslo, Norway: Department of Linguistics and Scandinavian Studies, University of Oslo.
Richards, B. J., Malvern, D. D., & Meara, P. (2009). *Vocabulary studies in first and second language acquisition: The interface between theory and application*. London: Palgrave Macmillan.
Ridder, I. De. (2002). Visible or invisible links : Does the highlighting of hyperlinks affect incidental vocabulary learning, text comprehension, and the reading process? *Language Learning & Technology 6*(1), 123–146.
Roberts, J., Frankenberg-Garcia, A., Lew, R., Rees, G., & Pereda, J. (2017, October). Visualisation and graphical techniques to help writers write more idiomatically. Poster session presented at IEEE VIS2017. Pheonix, AZ.
Roby, W. B. (1999). "What's in a gloss?" A Response to Lara L. Lomicka's "To gloss or not to gloss: An investigation of reading comprehension online. *Language Learning and Technology 2*(2), 94–101
Rozycki, W., & Johnson, N. H. (2013). Non-canonical grammar in Best Paper award winners in engineering. *English for Specific Purposes*, *32*(3), 157–169. https://doi.org/10.1016/j.esp.2013.04.002
Ruegg, R., Fritz, E., & Holland, J. (2011). Rater sensitivity to qualities of lexis in writing. *TESOL Quarterly*, *45*(1), 63–80. https://doi.org/10.5054/tq.201
Rundell, M. (2006). Learners' dictionaries. In K. Brown (Ed.), *Elsevier Encyclopedia of Language and Linguistics* (2nd ed., pp. 739–43). Amsterdam: Elsevier.
Ruppenhofer, J., Ellsworth, M., Petruck, M. R. L., Johnson, C. R., Baker, C. F., & Scheffczyk, J. (2016). *FrameNet II: Extended theory and practice*. Retrieved from https://framenet.icsi.berkeley.edu/fndrupal/the_book
Rye, J. (1985). Are cloze items sensitive to constraints across sentences? A review. *Journal of Research in Reading*, *8*(2), 94–105. https://doi.org/10.1111/j.1467-9817.1985.tb00310.x
Sanaoui, R. (1995). Adult learners' approaches to learning vocabulary in second languages. *The Modern Language Journal*, *79*(1), 15–28. https://doi.org/doi: 10.1111/j.1540-4781.1995.tb05410.x
Santos, T. (1988). Professors' reactions to the academic writing of nonnative-speaking students. *TESOL Quarterly*, *22*(1), 69–90. https://doi.org/10.2307/3587062
Saussure, F. de. (2013). *Course in general linguistics*. New York: Bloomsbury Publishing.

Savage, W. (2010). Monographs: The sounds of silence: The ransom. *Journal of Scholarly Publishing*, *41*, 483–488.
Sawcer, S. (2010). Bayes factors in complex genetics. *European Journal of Human Genetics*, *18*(7), 746–750. https://doi.org/10.1038/ejhg.2010.17
Schmid, H. (1994). Probabilistic Part-of-Speech Tagging Using Decision Trees. In D. Jones (Ed.), *Proceedings of the International Conference on New Methods in Language Processing*. Manchester: The University of Manchester Institute of Science and Technology.
Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*(2), 129–158. https://doi.org/10.1093/applin/11.2.129
Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition*. https://doi.org/10.1017/S0272263100011189
Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In *Talking to learn: Conversation in second language acquisition* (pp. 237–326).
Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary Description Acquisition and Pedagogy* (pp. 199–227). Cambridge: Cambridge University Press.
Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
Schmitt, N. (2010). *Researching vocabulary: a vocabulary research manual*. London: Palgrave Macmillan.
Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *Modern Language Journal*, *95*(1), 26–43. https://doi.org/10.1111/j.1540-4781.2011.01146.x
Schmitt, N., & Redwood, S. (2011). Learner knowledge of phrasal verbs: A corpus informed study. In F. Meunier, S. De Cock, G. Gilquin, & M. Paquot (Eds.), *A Taste for Corpora* (pp. 173–205). Amsterdam: John Benjamins.
Schmitt, N., & Schmitt, D. (2012). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, *47*(4), 484–503. https://doi.org/10.1017/S0261444812000018
Schmitt, N., & Underwood, G. (2004). Exploring the processing of formulaic sequences through a self-paced reading task. In N. Schmitt (Ed.), *Formulaic Sequences: acquisition, processing and use* (pp. 173–188). Amsterdam: John Benjamins.
Schmitt, N., & Zimmerman, C. (2002). Derivative word forms: What do learners know? *TESOL Quarterly*, *36*(2), 145–171.
SCImago. (2007). SJR — SCImago journal & country rank. Retrieved from http://www.scimagojr.com
Scott, J., & Nagy, W. (1997). Understanding the definitions of unfamiliar verbs. *Reading Research Quarterly*, *32*(2), 184–200. https://doi.org/10.1598/RRQ.32.2.4
Scott, J., & Nagy, W. (2004). Developing word consciousness. In J. Baumann & E. Kame'enui (Eds.), *Vocabulary Instruction: Research to Practice* (pp. 201–17). New York: Guilford Press.
Scott, M. (2017). WordSmith Tools version 7. Stroud: Lexical Analysis Software.
Seglen, P. O. (1992). Evaluation of scientists by journal impact. In *Representations of science and technology* (pp. 240–252). Leiden: DSWO Press.
Seglen, P. O. (1997). Why the impact factor should not be used for evaluating research. *British Medical Journal*, *314*(7079), 498–502.
Shin, D. (2009). *A collocation inventory for beginners: Spoken collocations of English*. Saarbrücken: Lap Lambert Academic Publishing.
Shin, D., & Nation, P. (2008). Beyond single words: The most frequent collocations in spoken English. *ELT Journal*, *62*(4), 339–348. https://doi.org/10.1093/elt/ccm091

Shonell, F., Meddleton, I., & Shaw, B. (1956). A study of the oral vocabulary of adults. Brisbane: University of Queensland Press.
Siepmann, D. (2008). Phraseology in learners' dictionaries: What, where and how? . In F. Meunier & S. Granger (Eds.) *Phraseology in foreign language learning and teaching* (pp. 185–202). Amsterdam: John Benjamins
Silver, C., & Lewins, A. (2014). *Using software in qualitative research: A step-by-step guide*. London: Sage.
Simonsen, H. (2009). *Vertical or horizontal? That is the question: an eye-track study of data presentation in internet dictionaries*. Frederiksberg: Copenhagen Business School.
Simonsen, H. (2011). User consultation behaviour in Internet dictionaries: An eye-tracking study. *Hermes*, *46*, 75–101.
Simpson-Vlach, R., & Ellis, N. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, *31*, 487–512.
Simpson, R., Briggs, L., Ovens, J., & Swales, J. (2002). The Michigan corpus of academic Spoken English. Ann Arbor, MI: The Regents of the University of Michigan.
Sinclair, J. (1982). Reflections on computer corpora in English language research. In S. Johansson (Ed.), *Computer corpora in English language research* (pp. 1–6). Bergen: Norwegian Computing Centre for the Humanities.
Sinclair, J. (1991). *Corpus, concordance, collocation. Describing English language*. Oxford: Oxford University Press.
Sinclair, J. (1996a). *EAGLES interim report corpus typology a framework for classification*. Retrieved from http://www.ilc.cnr.it/EAGLES/corpustyp/corpustyp.html
Sinclair, J. (1996b). The empty lexicon. *International Journal of Corpus Linguistics*, *1*(1), 99–119.
Sinclair, J. (1998). The lexical Item. In E. Weigand (Ed.), *Contrastive Lexical Semantics* (pp. 1–24). Amsterdam: John Benjamins.
Sinclair, J. (2001). Lexical grammar. In M. Gellerstram, K. Jóhanesson, B. Ralph, & L. Rogström (Eds.), *Nordiska Studier i Lexikografi 5* (pp. 323–343). Göteburg: Nordisk forening for leksikografi.
Sinclair, J. (2004a). *How to use corpora in language teaching*. Amsterdam: John Benjamins.
Sinclair, J. (2004b). The lexical item. In R. Carter (Ed.), *Trust the Text* (pp. 131–148). London: Routledge.
Sinclair, J. (2004c). The empty lexicon. In R. Carter (Ed.), *Trust the Text* (pp. 149–164). London: Routledge.
Sinclair, J. (2005). Corpus and text—Basic principles. In M. Wynne (Ed.), *Developing linguistic corpora: A guide to good practice* (pp. 1–16). Oxford: Oxbow Books.
Sinclair, J., & Mauranen, A. (2006). *Linear Unit Grammar: Integrating speech and writing*. Amsterdam: John Benjamins.
Siyanova-Chanturia, A. (2015). On the "holistic"nature of formulaic language. *Corpus Linguistics and Linguistic Theory*, *11*(2), 285–301.
Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: an eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research*, *27*(2), 251–272. https://doi.org/10.1177/0267658310382068
Smith, L. D., Best, L. A., Stubbs, D. A., Johnston, J., & Archibald, A. B. (2000). Scientific Graphs and the Hierarchy of the Sciences: *Social Studies of Science*, *30*(1), 73–94. https://doi.org/10.1177/030631200030001003
Smith, R. (2003). *Teaching english as a foreign language, 1912-1936: Pioneers of ELT*. London: Routledge.
Sokmen, A. J. (1992). Students as vocabulary generators. *TESOL Journal*, (Summer), 16–18.
Sonbul, S., & Schmitt, N. (2009). Direct teaching of vocabulary after reading: Is it worth the effort? *ELT Journal*, *64*(3), 253–260. https://doi.org/10.1093/elt/ccp059

Sparks, S. (2005). *JISC Disciplinary Differences Report*. Retrieved from: https://www.webarchive.org.uk/wayback/archive/20141125013123/http://www.jisc.ac.uk/media/documents/themes/infoenvironment/disciplinarydifferencesneeds.pdf

Stæhr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in english as a foreign language. *Studies in Second Language Acquisition*. https://doi.org/10.1017/S0272263109990039

Strevens, P. (1973). Technical, technological, and scientific English. *ELT Journal*, *27*(3), 223–34.

Stubbs, M. (2009). The search for units of meaning: Sinclair on empirical semantics. *Applied Linguistics*, *30*(1), 115–137. https://doi.org/10.1093/applin/amn052

Summers, D. (1988). The role of dictionaries in language learning. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 111–125). London: Longman.

Sutarsyah, C., Nation, P., & Kennedy, G. (1994). How Useful Is EAP Vocabulary for ESP? A corpus-based case study. *RELC Journal*, *25*(2), 34–50. https://doi.org/10.1177/003368829402500203

Swales, J. (1990). *Genre Analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Swales, J. (1998). *Other floors, other voices: A textography of a small university building*. Mahwah, NJ: Lawrence Erlbaum.

Swan, M., & Smith, B. (2001). *Learner English: a teacher's guide to interference and other problems*. *Cambridge handbooks for language teachers*. Cambridge University Press.

Swanborn, M. S. L., & de Glopper, K. (1999). Incidental word learning while reading: a meta-analysis. *Review of Educational Research*, *69*(3), 261–285. https://doi.org/10.3102/00346543069003261

Swenson, E., & West, M. P. (1934). *On the counting of new words in textbooks for the teaching of foreign languages*. *Bulletin of the Department of Educational Research*. Toronto: University of Toronto.

Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning & Technology*, *14*(2), 50–73. Retrieved from http://llt.msu.edu

Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science*, *12*(1), 49–100. https://doi.org/10.1016/0364-0213(88)90008-0

Tarone, E., Dwyer, S., Gillette, S., & Icke, V. (1998). On the use of the passive and active voice in astrophysics journal papers: With extensions to other languages and other fields. *English for Specific Purposes*, *17*(1), 113–132. https://doi.org/10.1016/S0889-4906(97)00032-X

Taylor, J. R. (2002). *Cognitive Grammar*. Oxford: Oxford University Press.

Teubert, W. (2001). Corpus Linguistics and Lexicography. *International Journal of Corpus Linguistics*, *6*(1), 125–153. https://doi.org/10.1075/ijcl.6.3.11teu

Teubert, W. (2005). My version of corpus linguistics. *International Journal of Corpus Linguistics*, *10*(1), 1–13. https://doi.org/10.1075/ijcl.10.1.01teu

Thompson, G. (1996). *Introducing functional grammar*. London: Arnold.

Thomson Reuters. (2017). InCitesTM Journal Citation Reports. Retrieved July 28, 2017, from https://jcr.incites.thomsonreuters.com

Thorndike, E. L., & Lorge, I. (1944). *The Teacher's Word Book of 30,000 Words*. New York: Teachers College Columbia University.

Timmis, I. (2015). *Corpus Linguistics for ELT*. Abingdon: Routledge.

Tognini-Bonelli, E. (2001). *Coprus Linguistics at Work*. Amsterdam: John Benjamins.

Tomaszczyk, J. (1979). Dictionaries: users and uses. *Glottodidactica*, *12*, 103–199.

Tono, Y. (2011). Application of eye-tracking in EFL learners' dictionary look-up process Research. *International Journal of Lexicography*, *24*(1), 124–153. https://doi.org/10.1093/ijl/ecq043
Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics*, *17*(1), 84–119. https://doi.org/10.1093/applin/17.1.84
treat, v. (2017). Retrieved October 9, 2017, from http://www.oed.com/view/Entry/205381?rskey=lpUuxb&result=3&isAdvanced=false
Tribble, C. (2002). Corpora and corpus analysis. In J. Flowerdew (Ed.), *Academic Discourse* (pp. 131–49). London: Longman, Pearson Education.
Tribble, C. (2017). ELFA vs. Genre: A new paradigm war in EAP writing instruction? *Journal of English for Academic Purposes*, *25*, 30–44.
Trimble, L. (1985). *English for Science and Technology: A Discourse Approach*. Cambridge: Cambridge University Press.
Turner, J. (2012). Academic literacies: Providing a space for the socio-political dynamics of EAP. *Journal of English for Academic Purposes*, *11*(1), 17–25. https://doi.org/10.1016/j.jeap.2011.11.007
Underwood, G., Schmitt, N., & Galpin, A. (2004). The eyes have it: An eye-movement study into the processing of formulaic sequences. In N. Schmitt (Ed.), *Formulaic Sequences: acquisition, processing and use* (pp. 153–172). Amsterdam: John Benjamins.
Van Noorden, R. (2012). Record number of journals banned for boosting impact factor with self-citations : News blog. Retrieved July 28, 2017, from http://blogs.nature.com/news/2012/06/record-number-of-journals-banned-for-boosting-impact-factor-with-self-citations.html
Van Zeeland, H., & Schmitt, N. (2013). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics*, *34*(4), 457–479. https://doi.org/10.1093/applin/ams074
Vilkaitè, L. (2016, June). *Reading collocations in a second language: an eye-tracking study*. Paper presented at the 7th Formulaic Language Research Network conference, Vilnius.
Vongpumivitch, V., Huang, J., & Chang, Y.-C. (2009). Frequency analysis of the words in the Academic Word List (AWL) and non-AWL content words in applied linguistics research papers. *English for Specific Purposes*, *28*, 33–41. https://doi.org/http://dx.doi.org/10.1016/j.esp.2008.08.003
Waibel, B. (2007). *Phrasal verbs in learner English: A corpus based study of German and Italian students* (Unpublished dissertation). University of Freiburg, Freiburg.
Walker, C. (2009). The treatment of collocation by learners' dictionaries, collocational dictionaries and dictionaries of business English. *International Journal of Lexicography*, *22*(3), 281–299. https://doi.org/10.1093/ijl/ecp016
Walker, C. (2011). A corpus-based study of the linguistic features and processes which influence the way collocations are formed: Some implications for the learning of collocations. *TESOL Quarterly*, *45*(2), 291–312. https://doi.org/10.5054/tq.2011.247710
Walsh, S., & Knight, D. (2016). Analysing spoken discourse in university small group teaching. In *Creating and Digitizing Language Corpora* (pp. 291–319). London: Palgrave Macmillan UK. https://doi.org/10.1057/978-1-137-38645-8_11
Wang, J., Liang, S., & Ge, G. (2008). Establishment of a medical academic word list. *English for specific purposes*, *27*(4), 442–458. https://doi.org/10.1016/j.esp.2008.05.003
Wang, M., & Nation, P. (2004). Word meaning in academic English: Homography in the Academic Word List. *Applied Linguistics*, *25*(3), 291–314. https://doi.org/10.1093/applin/25.3.291

Ward, J. (1999). How large a vocabulary do EAP engineering students need? *Reading in a Foreign Language*, *12*(2), 309–324.
Ward, J. (2007). Collocation and technicality in EAP engineering. *Journal of English for Academic Purposes*, *6*(1), 18–35. https://doi.org/10.1016/j.jeap.2006.10.001
Ward, J. (2009). A basic engineering English word list for less proficient foundation engineering undergraduates. *English for Specific Purposes*, *28*, 170–182. https://doi.org/http://dx.doi.org/10.1016/j.esp.2009.04.001
Ward, J., & Chuenjundaeng, J. (2009). Suffix knowledge: Acquisition and applications. *System*, *37*(3), 461–469. https://doi.org/10.1016/j.system.2009.01.004
Ward, K., Johnston, R., Richards, K., Gandy, M., Taylor, Z., Paasi, McDowell, L. (2009). The future of research monographs: an international set of perspectives. *Progress in Human Geography*, *33*(1), 101–126. https://doi.org/10.1177/0309132508100966
Waring, R. (2002). Scales of vocabulary knowledge in second language vocabulary assessment. *Kiyo the Occasional Papers of Notre Dame Seishin University*, *26*(1), 40–54.
Watanabe, Y. (1997). Input, Intake, and Retention: Effects of Increased Processing on Incidental Learning of Foreign Language Vocabulary. *Studies in Second Language Acquisition*, *19*(3), 287–307.
Webb, S., & Chang, A. (2012). Vocabulary Learning through Assisted and Unassisted Repeated Reading. *Canadian Modern Language Review*, *68*(3), 267–290. https://doi.org/10.3138/cmlr.1204.1
Webb, S., & Hang, N. (2016, June). *Examining second language receptive knowledge of collocation and the factors that affect learning.* Paper presented at the 7th Formulaic Language Research Network conference, Vilnius.
Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocations. *Language Learning*, *63*(1), 91–120. https://doi.org/10.1111/j.1467-9922.2012.00729.x
Weir, C., Hawkey, R., Green, A., Unaldi, A., & Devi, S. (2009). The relationship between the academic reading construct as measured by IELTS and the reading experiences of students in their first year of study at a British university. In P. Thompson, J. Holliday & L. Taylor (Eds.), *IELTS Research Reports* (Vol. 9, pp. 97–156). Canberra: IELTS Australia.
West, M. (1960). *Teaching English in difficult circumstances*... London: Longmans Green.
West, M. (1953). *A general service list of English words: with semantic frequencies and a supplementary word-list for the writing of popular science and technology*. London: Longmans, Green & Co.
Williams, G. (1998). Collocational networks: Interlocking patterns of lexis in a corpus of plant biology research articles. *International Journal of Corpus Linguistics*, *3*(1), 151–171. https://doi.org/10.1075/ijcl.3.1.07wil
Williams, G. (2002). Corpus-driven lexicography and the specialised dictionary: Headword extraction for the parasitic plant research dictionary. In A. Braasch & C. Povlsen (Eds.), *Proceedings of the 10th EURALEX International Congress* (pp. 859–864). København, Denmark: Center for Sprogteknologi.
Williams, G. (2006). Advanced ESP and the learner's dictionary: Tools for the non-language specialist. In E. Corino, C. Marello & C. Onesti (Eds.), *Proceedings of the 12th EURALEX International Congress* (pp. 795–801). Torino, Italy: Edizioni dell'Orso.
Williams, G. (2008). The good lord and his works: A corpus-driven study of collocational resonance. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 159–173). Amsterdam: John Benjamins Pub. Company.
Williams, G. (2011). The learner's dictionary and the sciences. In N. Kuebler (Ed.), *Corpora, Language, Teaching, and Resources: From theory to practice* (pp. 323–340). Bern: Peter Lang.
Williams, G. (2017). Mapping a dictionary: Using Atlas ti and XML to analyse a late 17th century dictionary. *RiCOGNIZIONI. Rivista Di Lingue E Letterature Straniere E Culture*

Moderne, *4*(7), 161–180. Retrieved from http://www.ojs.unito.it/index.php/ricognizioni/article/view/2104/2024
Williams, G., & Millon, C. (2010). Going organic: Building an experimental bottom-up dictionary of verbs in science. In A. Dykstra & S. Tanneke (Eds.), *Proceedings of the 14th EURALEX International Congress* (pp. 1251–1257). Leeuwarden: Euralex.
Williams, G., & Millon, C. (2014). Les verbes et la science : construction d'un dictionnaire organique. In A. Tutin & F. Grossmann (Eds.), *L'écrit scientifique : du lexique au discours* (pp. 45–66). Rennes: Presses universitaires de Rennes.
Williams, G., Millon, C., & Alonso, A. (2012). Growing naturally: The DicSci Organic E-Advanced Learner's Dictionary of Verbs in Science. In A. Abel, C. Vettori & N. Ralli (Eds.), *Proceedings of the 15th EURALEX International Congress* (pp. 1008–1013). Oslo: University of Oslo.
Williams, P., Stevenson, I., Nicholas, D., Watkinson, A., & Rowlands, I. (2009). The role and future of the monograph in arts and humanities research. *ASLIB Proceedings*, *61*(1), 67–82. https://doi.org/10.1108/00012530910932294
Wilson, A. (2013). Embracing Bayes factors for key item analysis in corpus linguistics. In M. Bieswanger & A. Koll-Stobbe, *New approaches to the study of linguistic variability : Language Competence and Language Awareness in Europe* (pp. 3–11). Frankfurt: Peter Lang.
Wilson, N. (1999). *History in crisis? Recent directions in historiography*. New Jersey: Prentice Hall.
Wolter, B., & Gyllstad, H. (2011). Collocational links in the L2 mental lexicon and the influence of l1 intralexical knowledge. *Applied Linguistics*, *32*(4), 430–449. https://doi.org/10.1093/applin/amr011
Wray, A. (2000). Formulaic sequences in second language teaching: principle and practice. *Applied Linguistics*, *21*(4), 463–89. https://doi.org/10.1093/applin/21.4.463
Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.
Wray, A. (2004). "Here's one I prepared earlier': Formulaic language learning on television. In N. Schmitt (Ed.), *Formulaic Sequences: acquisition, processing and use* (pp. 249–68). Amsterdam: John Benjamins.
Wu, W., & Hammond, M. (2011). Challenges of university adjustment in the UK: a study of East Asian Master's degree students. *Journal of Further and Higher Education*, *35*(3), 423–438. https://doi.org/10.1080/0309877X.2011.569016
Wüster, E. (1979). *Einführung in die allgemeine Terminologielehre und terminologische Lexikographie*. Berlin: Springer.
Xue, G., & Nation, P. (1984). A University Word List. *Language Learning and Communication*, *3*, 215–229.
Yamada, S. (2013). Monolingual learners' dictionaries – Where now? In H. Jackson (Ed.), *The Bloomsbury companion to lexicography* (pp. 188–212). London: Bloomsbury Publishing.
Yamashita, J., & Jiang, N. (2010). L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly*, *44*(4), 647–668. https://doi.org/10.5054/tq.2010.235998
Yang, H. (1986). A new technique for identifying scientific/technical terms. *Literary and Linguistic Computing*, *1*, 93–103.
Yang, M. (2015). A nursing academic word list. *English for Specific Purposes*, *37*, 27–38. https://doi.org/10.1016/j.esp.2014.05.003
Yano, Y., Long, M., & Ross, S. (1994). The effects of simplified and elaborated texts on foreign language reading comprehension. *Language Learning*, *44*(2), 189–219. https://doi.org/10.1111/j.1467-1770.1994.tb01100.x

Yoiro, C. (1989). Idiomaticity as an indicator of second language proficiency. In K. Hyltenstam & L. Obler (Eds.), *Bilingualism across the Lifespan* (pp. 55–72). Cambridge: Cambridge University Press.

Zhang, B., & Li, C. (2011). Classification of L2 vocabulary learning strategies: evidence from exploratory and confirmatory factor analyses. *RELC Journal*, *42*(2), 141–154. https://doi.org/10.1177/0033688211405180

Zhou, A. (2009). What adult ESL learners say about improving grammar and vocabulary in their writing for academic purposes. *Language Awareness*, *18*(1), 31–46. https://doi.org/10.1080/09658410802307923

Zoubir-Shaw, S., & Oxford, R. (1995). Gender differences in language strategy use in university-level introductory French classes. In C. Clee (Ed.), *Faces in a Crowd: The Individual Learner in Multisection Courses* (pp. 181–213). Boston: Heinle & Heinle.

APPENDICES

Appendix A: Summary of studies included in meta-analysis of corpus composition

Table A1

Summary of studies included in meta-analysis

Study	Aim(s)	Study Type	Corpus Composition
Campion and Elley (1971)	To assist assessment designers in selecting vocabulary for a university entry exam for overseas students	WL	Exams, Lectures, TBs
Praninskas (1972)	To discover what words foreign students of English need to know for study in the US.	WL	TBs
Lynn (1973)	To discover which words caused difficulties for students of Accounting, Business Administration and Economics at Nanyang University	WL	TBs
Ghadessey (1979)	To create a list of vocabulary items suitable for first-year university students in Iran	WL	TBs
Xu and Nation (1984)	To complement to the GSL for EAP learners	WL	Amalgamation of previous four lists
H. Yang (1986)	To demonstrate a new technique for identifying Scientific/Technical terms and describing science texts	PE	TBs
G. Williams (1998)	To investigate collocational networks in a corpus of parasitic plant biology	PE	RAs
Coxhead (2000)	To develop and evaluate an academic word list	WL	RAs, Academic articles from the WWW, TBs, LA
Biber et al. (2004)	To investigate lexical bundles in classroom teaching and textbooks	DA	TBs
Cortes (2004)	To compare the use of lexical bundles by published authors in history and biology and by students at three different levels in those disciplines	DA	RA, UG student writing
Marco (2000)	To analyze the use of collocational frameworks in medical RAs	DA	RAs
Kashiha (2005)	To investigate the use of lexical bundles in RA conclusions written by native and Iranian non-native writers	DA	RA conclusions
Mudraya (2006)	To establish a frequency-based corpus of student engineering lexis	WL	TBs
Hyland and Tse (2007)	To examine the behavior of AWL words across disciplines	WL	RAs, TBs, Book reviews, Scientific letters, Theses (Phd, MA, UG)
Q. Chen and Ge (2007)	To examine the frequency and coverage of the AWL in medical RAs	WL	RAs
Konstantakis (2007)	To create complement to the AWL and GSL for business students to reach the 95% goal for reading comprehension	WL	TBs
Paquot (2007)	To examine why the AWL is not suited to productive purposes while demonstrating a new extraction methodology for a productively orientated list	WL	Published academic prose, NSW
Coxhead and Hirsh (2007)	To discover if there is a list of scientific words outside the GSL and the AWL	WL	Manuals, TBs, Lecture notes, Study guides, Handouts

Note. DA = Discourse Analysis; PE = Phrase Extraction; WL = Worldist; LA = Learned or Academic section of existing corpus; NSW = Native student writing; NNSW = Non-native student writing; RA = Research Article; TB = Textbook; UG = Undergraduate.

Table A1 (cont.)

Summary of studies included in meta-analysis

Study	Aim(s)	Study Type	Corpus Composition
Hyland (2008a. 2008b)	To examine variation in use of lexical bundles across academic genres	DA	RAs, Theses (Phd, MA/MSc)
J. Wang, Liang, and Ge (2008)	To develop a medical AWL	WL	RAs
Durrant (2009)	To create and evaluate a of a list of "positionally-variable academic collocations"	WL	RAs
Martinez et. al (2009)	To identify to what extent the AWL is representative of agricultural science RAs	WL	RAs
Vongpumivitch, Huang, and Chang (2009)	To explore the the use of AWL words in the field of applied linguistics.	WL	RAs
J. Ward (2009)	To create a receptively orientated word list for foundation engineers	WL	TBs
Li & Qian (2010)	The examine AWL items in the Hong Kong Financial Services Corpus	WL	Annual Reports, Brochures, Fund descriptions, Ordinances, Speeches
Simpson-Vlach & Ellis (2010)	To create a list of academic formulae	WL	LA
Hsu (2011)	To create a business wordlist to aid postgraduate business students in reading comprehension	WL	RAs
Ädel & Erman (2012)	To investigate the use of lexical bundles in writing by L1 speakers of Swedish university linguistics students	DA	NSW, NNSW
Gardner and Nesi (2012)	To understand the type of texts SS need to produce	DA	NSW
Rees (2013; 2016)	To examine the bahaviour of AWL words in different disciplines	WL	RAs
Gardner & Davies (2014)	To develop an academic wordlist	WL	LA
Durrant (2014)	To investigate student writing and vocabulary needs across disciplines	DA	LA
Hsu (2014)	To create an business wordlist to aid postgraduates in reading comprehension	WL	TBs
Pérez-Llantada (2014)	To compare use of lexical bundles in Spanish and English between L1 and L2 writers	DA	RAs
Durrant (2017)	To understand disciplinary variation specifically in student academic writing.	DA	LA
Grabowski (2015)	To analyse key vocabulary and lexical bundles in pharmacology literature	WL	Patient information leaflets, Product summaries, Clinical protocols, TBs
Liu & Han (2015)	To investigate compilation criteria field-specific academic word lists in environmental science	WL	RAs
Le and Harrington (2015)	To investigate word clusters used to comment on results in applied linguistics Ras	DA	RAs
Yang (2015)	To examine the most frequently-used vocabulary across different sub-disciplines in nursing	WL	RAs
Lei & Liu (2016)	To create a wordlist to serve the needs of medical students	WL	RAs, TBs

Note. DA = Discourse Analysis; PE = Phrase Extraction; WL = Worldist; LA = Learned or Academic section of existing corpus; NSW = Native student writing; NNSW = Non-native student writing; RA = Research Article; TB = Textbook; UG = Undergraduate.

Appendix B: Bibliographic references for articles included in the AJACX2

History

Abalahin, A. J. (2011). "Sino-Pacifica": Conceptualizing Greater Southeast Asia as a Sub-Arena of World History. *Journal of World History*, *22*, 659–691. https://doi.org/10.2307/41508014

Alfred J. Rieber. (2004). The Crack in the Plaster: Crisis in Romania and the Origins of the Cold War. *The Journal of Modern History*, *76*(1), 62–106 https://doi.org/10.1086/421185

Amelang, J. S. (2005). Mourning Becomes Eclectic: Ritual Lament and the Problem of Continuity. *Past & Present*, *187*, 3–31. Retrieved from http://past.oxfordjournals.org/content/187/1/3.short

Amrith, S., & Sluga, G. (2008). New Histories of the United Nations. *Journal of World History*, *19*, 251–274. https://doi.org/10.2307/40542615

Anderson, J. (2003). `Turned into Taxpayers': Paraplegia, Rehabilitation and Sport at Stoke Mandeville, 1944-56. *Journal of Contemporary History, 38*(3), 461–475. https://doi.org/10.1177/0022009403038003007

Andrade, T. (2010). A Chinese Farmer, Two African Boys, and a Warlord: Toward a Global Microhistory. *Journal of World History*, *21*(4), 573–591. https://doi.org/citeulike-article-id:12446560

Arnold, K. J., & Lübbers, G. C. (2007). The Meeting of the Staatssekretäre on 2 May 1941 and the Wehrmacht: A Document up for Discussion. *Journal of Contemporary History, 42*(4), 613–626. https://doi.org/10.1177/0022009407081485

Aslanian, S. (2008). "The Salt in a Merchant's Letter": The Culture of Julfan Correspondence in the Indian Ocean and the Mediterranean. *Journal of World History*, *19*, 127–188. https://doi.org/10.2307/20079471

Auerbach, S. (2012). "The Law Has No Feeling for Poor Folks Like Us!": Everyday Responses to Legal Compulsion in England's Working-Class Communities, 1871–1904. *Journal of Social History, 45*(3), 686–708. https://doi.org/10.1093/jsh/shr060

Bagchi, K. (2003). An Orientalist in the Orient: Richard Garbe's Indian Journey, 1885-1886. *Journal of World History*, *14*, 281–325. https://doi.org/10.2307/20079220

Bailkin, B. J. (2009). The Postcolonial Family? West African Children, Private Fostering, and the British State. *The Journal of Modern History*, *81*(1), 87–121 https://doi.org/10.1086/593156

Balabanlilar, L. (2007). Lords of the Auspicious Conjunction: Turco-Mongol Imperial Identity on the Subcontinent. *Journal of World History*, *18*, 1–39. https://doi.org/10.2307/20079408

Bang, P. F. (2007). Trade and Empire — In Search of Organizing Concepts for the Roman Economy. *Past & Present*, *195*, 3–54. Retrieved from http://past.oxfordjournals.org/content/195/1/3.short

Barkawi, T. (2006). Culture and Combat in the Colonies: The Indian Army in the Second World War. *Journal of Contemporary History, 41*(2), 325–355. https://doi.org/10.1177/0022009406062071

Beaujard, P. (2010). From Three Possible Iron-Age World-Systems to a Single Afro-Eurasian World-System. *Journal of World History*, *21*(1), 1–43. https://doi.org/10.2307/20752924

Bell, P. (2002). The Foreign Office and the 1939 Royal Visit to America: Courting the USA in an Era of Isolationism. *Journal of Contemporary History, 37*(4), 599–616. https://doi.org/10.1177/0022009402037004060l

Bennett, J. M. (2010). Compulsory Service in Late Medieval England. *Past & Present*, *209*, 7–51. Retrieved from http://past.oxfordjournals.org/content/209/1/7.short
Biess, F. (2012). Moral Panic in Postwar Germany: The Abduction of Young Germans into the Foreign Legion and French Colonialism in the 1950s. *The Journal of Modern History*, *84*(4), 789–832 https://doi.org/10.1086/667681
Biller, P. (2006). Goodbye to Waldensianism? *Past & Present*, *192*, 3–33. Retrieved from http://past.oxfordjournals.org/content/192/1/3.short
Bishop, J. (2011). Utopia and Civic Politics in Mid-Sixteenth-Century London. *The Historical Journal*, *54*, 933–953.
Blauvelt, M. T. (2002). The Work of the Heart: Emotion in the 1805–35 Diary of Sarah Connell Ayer. *Journal of Social History, 35*(3), 577–592. https://doi.org/10.1353/jsh.2002.0003
Boeck, B. B. J. (2008). When Peter I Was Forced to Settle for Less: Coerced Labor and Resistance in a Failed Russian Colony (1695–1711). *The Journal of Modern History*, *80*(3), 485–514 https://doi.org/10.1086/589589
Brown, T. S. (2004). Subcultures, Pop Music and Politics: Skinheads and "Nazi Rock" in England and Germany. *Journal of Social History, 38*(1), 157–178. https://doi.org/10.1353/jsh.2004.0079
Burke, E. (2012). Toward a Comparative History of the Modern Mediterranean, 1750–1919. *Journal of World History*, *23*(4), 907–939. https://doi.org/10.1353/jwh.2012.0133
Butler, S. (2006). Runaway Wives: Husband Desertion in Medieval England. *Journal of Social History, 40*(2), 337–359. https://doi.org/10.1353/jsh.2007.0007
Byars, J. (2008). The Long and Varied Relationship of Andrea Mora and Anzola Davide: Concubinage, Marriage and the Authorities in the Early Modern Veneto. *Journal of Social History, 41*(3), 667–690. https://doi.org/10.1353/jsh.2008.0054
Callaghan, J. (2005). The Plan to Capture the British Labour Party and its Paradoxical Results, 1947-91. *Journal of Contemporary History, 40*(4), 707–725. https://doi.org/10.1177/0022009405056126
Cambers, A. (2009). Demonic Possession, Literacy and "Superstition" in Early Modern England. *Past & Present*, *202*, 3–35. Retrieved from http://past.oxfordjournals.org/content/202/1/3.short
Campbell, B. M. S. (2005). The Agrarian Problem in the Early Fourteenth Century. *Past & Present*, *188*, 3–70. Retrieved from http://past.oxfordjournals.org/content/188/1/3.short
Capp, B. (2009). Bigamous Marriage in Early Modern England. *The Historical Journal*, *52*, 537–556.
Carol Helstosky, B. (2009). Giovanni Bastianini, Art Forgery, and the Market in Nineteenth-Century Italy. *The Journal of Modern History*, *81*(4), 793–823 https://doi.org/10.1086/605486
Caron, V. (2009). Catholic Political Mobilization and Antisemitic Violence in Fin de Siècle France: The Case of the Union Nationale. *The Journal of Modern History*, *81*(2), 294–346 https://doi.org/10.1086/598923
Castillo, G. (2005). Domesticating the Cold War: Household Consumption as Propaganda in Marshall Plan Germany. *Journal of Contemporary History, 40*(2), 261–288. https://doi.org/10.1177/0022009405051553
Caunce, S. A. (2012). The Hiring Fairs of Northern England, 1890–1930: A Regional Analysis of Commercial and Social Networking in Agriculture. *Past & Present*, *217*(1), 213–246. https://doi.org/10.1093/pastj/gts024
Chambers, V. A. (2009). `Informed By, but Not Guided By, the Concerns of the Present': Contemporary History in UK Higher Education — Its Teaching and Assessment.

Journal of Contemporary History, 44(1), 89–105. https://doi.org/10.1177/0022009408098648
Charipova, L. (2003). Peter Mohyla's Translation of the Imitation of Christ. *The Historical Journal*, *46*, 237–261.
Chivvis, C. S. (2006). Charles de Gaulle, Jacques Rueff and French International Monetary Policy under Bretton Woods. *Journal of Contemporary History, 41*(4), 701–720. https://doi.org/10.1177/0022009406067753
Christian, D. (2003). World History in Context. *Journal of World History*, *14*, 437–458. https://doi.org/10.2307/20079239
Clark, H. R. (2009). Frontier Discourse and China's Maritime Frontier: China's Frontiers and the Encounter with the Sea through Early Imperial History. *Journal of World History*, *20*, 1–33. https://doi.org/10.2307/40542719
Clark, M. (2011). The Gentry, the Commons, and the Politics of Common Right in Enfield, c. 1558 – c. 1603. *The Historical Journal*, *54*, 609–629.
Clarke, F. (2007). So Lonesome I Could Die: Nostalgia and Debates Over Emotional Control in the Civil War North. *Journal of Social History, 41*(2), 253–282. https://doi.org/10.1353/jsh.2008.0001
Clarke, J. (2012). Work, Consumption and Subjectivity in Postwar France: Moulinex and the Meanings of Domestic Appliances 1950s–70s. *Journal of Contemporary History, 47*(4), 838–859. https://doi.org/10.1177/0022009412451292
Claydon, T. (2008). Latitudinarianism and Apocalyptic History in the Worldview of Gilbert Burnet, 1643–1715*. *The Historical Journal*, *51*, 577–597.
Cockayne, E. (2003). Experiences of the Deaf In Early Modern England. *The Historical Journal*, *46*, 493–510.
Cohn, S. K. (2007). The Black Death and the Burning of Jews. *Past & Present*, *196*, 3–36. Retrieved from http://past.oxfordjournals.org/content/196/1/3.short
Connelly, J. (2007). Catholic Racism and Its Opponents. *The Journal of Modern History*, *79*(4), 813–847 https://doi.org/10.1086/521066
Cooper, J. E. (2007). Thomas Hobbes on the Political Theorist's Vocation. *The Historical Journal*, *50*, 519–547.
Cooper, K. (2007). Closely Watched Households: Visibility, Exposure and Private Power in the Roman Domus. *Past & Present*, *197*, 3–33. Retrieved from http://past.oxfordjournals.org/content/197/1/3.short
Cortada, J. W. (2009). Public Policies and the Development of National Computer Industries in Britain, France, and the Soviet Union, 1940—80. *Journal of Contemporary History, 44*(3), 493–512. https://doi.org/10.1177/0022009409104120
Cressy, D. (2002). The Protestation Protested, 1641 AND 1642. *The Historical Journal*, *45*, 251–279.
Crooks, P. (2011). State of the Union: Perspectives on English Imperialism in the Late Middle Ages. *Past & Present*, *212*, 3–42. Retrieved from http://past.oxfordjournals.org/content/212/1/3.short
Dan Stone. (2003). The English Mistery, the BUF, and the Dilemmas of British Fascism. *The Journal of Modern History*, *75*(2), 336–358 https://doi.org/10.1086/380138
Daryaee, T. (2003). The Persian Gulf Trade in Late Antiquity. *Journal of World History*, *14*, 1–16. https://doi.org/10.2307/20079006
David Kammerling Smith. (2002). Structuring Politics in Early Eighteenth-Century France: The Political Innovations of the French Council of Commerce. *The Journal of Modern History*, *74*(3), 490–537 https://doi.org/10.1086/345110
de Mieroop, M. van. (2003). Revenge, Assyrian Style. *Past & Present*, *179*, 3–23. Retrieved from http://past.oxfordjournals.org/content/179/1/3.short

Dejung, C. (2010). "Switzerland must be a special democracy": Sociopolitical Compromise, Military Comradeship, and the Gender Order in 1930s and 1940s Switzerland. *The Journal of Modern History*, *82*(1), 101–126 https://doi.org/10.1086/650508
Dháibhéid, C. N. I. C. (2012). The Irish National Aid Association and the Radicalization of Public Opinion in Ireland, 1916–1918. *The Historical Journal*, *55*(3), 705–729.
Dialeti, A. (2011). Defending Women, Negotiating Masculinity in Early Modern Italy. *The Historical Journal*, *54*, 1–23.
Di-Capua, Y. (2008). Common Skies Divided Horizons: Aviation, Class and Modernity in Early Twentieth Century Egypt. *Journal of Social History, 41*(4), 917–942. https://doi.org/10.1353/jsh.0.0024
Dickinson, E. R. (2003). The Men's Christian Morality Movement in Germany, 1880–1914: Some Reflections on Politics, Sex, and Sexual Politics. *The Journal of Modern History*, *75*(1), 59–110 https://doi.org/10.1086/377749
Dirlik, A. (2005). Performing the World: Reality and Representation in the Making of World Histor(ies). *Journal of World History*, *16*, 391–410. https://doi.org/10.2307/20079345
Dossey, L. (2008). Wife Beating and Manliness in Late Antiquity. *Past & Present*, *199*, 3–40. Retrieved from http://past.oxfordjournals.org/content/199/1/3.short
Durrill, W. K. (2006). Ritual, Community and War: Local Flag Presentation Ceremonies and Disunity in the Early Confederacy. *Journal of Social History, 39*(4), 1105–1122. https://doi.org/10.1353/jsh.2006.0034
Eagle Glassheim. (2006). Ethnic Cleansing, Communism, and Environmental Devastation in Czechoslovakia's Borderlands, 1945–1989. *The Journal of Modern History*, *78*(1), 65–92 CR–Copyright © 2006 The University o. https://doi.org/10.1086/499795
Eidinow, E. (2010). Patterns of Persecution: "Witchcraft" Trials in Classical Athens. *Past & Present*, *208*, 9–35. Retrieved from http://past.oxfordjournals.org/content/208/1/9.short
Elder, S. (2006). Murder, Denunciation and Criminal Policing in Weimar Berlin. *Journal of Contemporary History, 41*(3), 401–419. https://doi.org/10.1177/0022009406064647
Evangelisti, S. (2004). Monastic Poverty and Material Culture in Early Modern Italian Convents. *The Historical Journal*, *47*, 1–20.
Fass, P. S. (2005). Children in Global Migrations. *Journal of Social History, 38*(4), 937–953. https://doi.org/10.1353/jsh.2005.0057
Fentress, E. (2006). Romanizing the Berbers. *Past & Present*, *190*, 3–33. Retrieved from http://past.oxfordjournals.org/content/190/1/3.short
Finlay, R. (2007). Weaving the Rainbow: Visions of Color in World History. *Journal of World History*, *18*, 383–431. https://doi.org/10.2307/20079447
Fletcher, C. (2005). Manhood and Politics in the Reign of Richard II. *Past & Present*, *189*, 3–39. Retrieved from http://past.oxfordjournals.org/content/189/1/3.short
Forsdyke, S. (2008). Street Theatre and Popular Justice in Ancient Greece: Shaming, Stoning and Starving Offenders inside and outside the Courts. *Past & Present*, *201*, 3–50. Retrieved from http://past.oxfordjournals.org/content/201/1/3.short
Frank, A. G., & Thompson, W. R. (2005). Afro-Eurasian Bronze Age Economic Expansion and Contraction Revisited. *Journal of World History*, *16*, 115–172. https://doi.org/10.2307/20079314
Frost, G. (2003). "The Black Lamb of the Black Sheep": Illegitimacy in the English Working Class, 1850–1939. *Journal of Social History, 37*(2), 293–322. https://doi.org/10.1353/jsh.2003.0178
Fynn-Paul, J. (2009). Empire, Monotheism and Slavery in the Greater Mediterranean Region from Antiquity to the Early Modern Era. *Past & Present*, *205*, 3–40. Retrieved from http://past.oxfordjournals.org/content/205/1/3.short

Gaggio, D. (2011). Before the Exodus: The Landscape of Social Struggle in Rural Tuscany, 1944–1960. *The Journal of Modern History*, *83*(2), 319–345 https://doi.org/10.1086/659062
Gajda, A. (2009). Debating War and Peace in Late Elizabethan England. *The Historical Journal*, *52*, 851–878.
Gardner, N. (2007). The Harold Wilson Government, Airwork Services Limited, and the Saudi Arabian Air Defence Scheme, 1965—73. *Journal of Contemporary History, 42*(2), 345–363. https://doi.org/10.1177/0022009407075547
Gerwarth, R., & Horne, J. (2011). Vectors of Violence: Paramilitarism in Europe after the Great War, 1917–1923. *The Journal of Modern History*, *83*(3), 489–512 https://doi.org/10.1086/660364
Goddard, R. (2011). Small Boroughs and the Manorial Economy: Enterprise Zones or Urban Failures? *Past & Present*, *210*, 3–31. Retrieved from http://past.oxfordjournals.org/content/210/1/3.short
Goeschel, C., & Wachsmann, N. (2010). Before Auschwitz: The Formation of the Nazi Concentration Camps, 1933-9. *Journal of Contemporary History, 45*(3), 515–534. https://doi.org/10.1177/0022009410366554
Goodman, M. (2004). Trajan and the Origins of Roman Hostility to the Jews. *Past & Present*, *182*, 3–29. Retrieved from http://past.oxfordjournals.org/content/182/1/3.short
Goose, N., & Looijesteijn, H. (2012). Almshouses in England and the Dutch Republic circa 1350–1800: A Comparative Perspective. *Journal of Social History, 45*(4), 1049–1073. https://doi.org/10.1093/jsh/shr146
Gorman, D. (2010). Ecumenical Internationalism: Willoughby Dickinson, the League of Nations and the World Alliance for Promoting International Friendship through the Churches. *Journal of Contemporary History, 45*(1), 51–73. https://doi.org/10.1177/0022009409348020
Graff, H. J. (2010). The Literacy Myth at Thirty. *Journal of Social History, 43*(3), 635–661. https://doi.org/10.1353/jsh.0.0316
Greenstone, D. (2005). Frightened George: How the Pediatric-Educational Complex Ruined the Curious George Series. *Journal of Social History, 39*(1), 221–228. https://doi.org/10.1353/jsh.2005.0103
Groves, T. (2012). Everyday Struggles against Franco's Authoritarian Legacy: Pedagogical Social Movements and Democracy in Spain. *Journal of Social History, 46*(2), 305–334. https://doi.org/10.1093/jsh/shs094
Gullace, N. F. (2005). Friends, Aliens, and Enemies: Fictive Communities and the Lusitania Riots of 1915. *Journal of Social History, 39*(2), 345–367. https://doi.org/10.1353/jsh.2005.0137
Gurney, P. (2005). The Battle of the Consumer in Postwar Britain. *The Journal of Modern History*, *77*(4), 956–987 https://doi.org/10.1086/499831
Halevi, L. (2004). Wailing for the Dead: The Role of Women in Early Islamic Funerals. *Past & Present*, *183*, 3–39. Retrieved from http://past.oxfordjournals.org/content/183/1/3.short
Hanebrink, P. (2008). Transnational Culture War: Christianity, Nation, and the Judeo-Bolshevik Myth in Hungary, 1890–1920. *The Journal of Modern History*, *80*(1), 55–80 https://doi.org/10.1086/529077
Harari, Y. N. (2007). The Concept of "Decisive Battles" in World History. *Journal of World History*, *18*, 251–266. https://doi.org/10.2307/20079432
Harrison, C. E. (2007). Zouave Stories: Gender, Catholic Spirituality, and French Responses to the Roman Question. *The Journal of Modern History*, *79*(2), 274–305 https://doi.org/10.1086/519315

Henry Clark. (2004). Grain Trade Information: Economic Conflict and Political Culture under Terray, 1770–1774. *The Journal of Modern History*, *76*(4), 793–834 https://doi.org/10.1086/427569
Hense, A. (2011). Limitation of Economic Damages as a "Humanitarian Gesture": The German Foundation "Remembrance, Responsibility and the Future." *Journal of Contemporary History, 46*(2), 407–424. https://doi.org/10.1177/0022009410392358
Herman, E. (2002). The Paradoxical Rationalization of Modern Adoption. *Journal of Social History, 36*(2), 339–385. https://doi.org/10.1353/jsh.2003.0017
Hexham, I. (2007). Inventing "Paganists": a Close Reading of Richard Steigmann-Gall's the Holy Reich. *Journal of Contemporary History, 42*(1), 59–78. https://doi.org/10.1177/0022009407071632
Holt, M. P. (2003). The Social History of the Reformation: Recent Trends and Future Agendas. *Journal of Social History, 37*(1), 133–144. https://doi.org/10.1353/jsh.2003.0142
Hooper, J. (2011). Pirates and Kings: Power on the Shores of Early Modern Madagascar and the Indian Ocean. *Journal of World History*, *22*, 215–242. https://doi.org/citeulike-article-id:12446561
Horodowich, E. (2003). Civic Identity and the Control of Blasphemy in Sixteenth-Century Venice. *Past & Present*, *181*, 3–33. Retrieved from http://past.oxfordjournals.org/content/181/1/3.short
Horowitz, R. S. (2004). International Law and State Transformation in China, Siam, and the Ottoman Empire during the Nineteenth Century. *Journal of World History*, *15*(4), 445–486. https://doi.org/10.1353/jwh.2005.0126
Horster, M. (2004). The Trade in Political Prisoners between the Two German States, 1962-89. *Journal of Contemporary History, 39*(3), 403–424. https://doi.org/10.1177/0022009404044447
Huang, E. C. (2012). From the Imperial Court to the International Art Market: Jingdezhen Porcelain Production as Global Visual Culture. *Journal of World History*, *23*(1), 115–145. https://doi.org/10.1353/jwh.2012.0028
Humphrey, C., & Hürelbaatar, A. (2005). Regret as a Political Intervention: An Essay in the Historical Anthropology of the Early Mongols. *Past & Present*, *186*, 3–45. Retrieved from http://past.oxfordjournals.org/content/186/1/3.short
Hung, C. (2012). The Cultural Palace of Nationalities: Ethnicities Under One Roof? *Journal of Contemporary History, 47*(3), 572–593. https://doi.org/10.1177/0022009412441649
Hurl-Eamon, J. (2004). Policing Male Heterosexuality: The Reformation of Manners Societies' Campaign Against the Brothels in Westminster, 1690–1720. *Journal of Social History, 37*(4), 1017–1035. https://doi.org/10.1353/jsh.2004.0052
Iii, E. B. (2009). Islam at the Center: Technological Complexes and the Roots of Modernity. *Journal of World History*, *20*(2), 165–186. https://doi.org/10.2307/40542756
Imlay, T. (2003). From Villain to Partner: British Labour Party Leaders, France and International Policy During the Phoney War, 1939-40. *Journal of Contemporary History, 38*(4), 579–596. https://doi.org/10.1177/00220094030384004
James Van Horn Melton. (2004). School, Stage, Salon: Musical Cultures in Haydn's Vienna. *The Journal of Modern History*, *76*(2), 251–279 https://doi.org/10.1086/422930
Johnson, C. R. (2006). Renaissance German Cosmographers and the Naming of America. *Past & Present*, *191*, 3–43. Retrieved from http://past.oxfordjournals.org/content/191/1/3.short
Jonas, M. (2012). "Can One Go Along With This?" German Diplomats and the Changes of 1918–19 and 1933–4. *Journal of Contemporary History, 47*(2), 240–269. https://doi.org/10.1177/0022009411431712

Jones, P. M. (2012). The Challenge of Land Reform in Eighteenth- and Nineteenth-Century France*. *Past & Present* , *216*(1), 107–142. https://doi.org/10.1093/pastj/gts003
Joshi, V. (2002). The `Private' became `Public': Wives as Denouncers in the Third Reich. *Journal of Contemporary History, 37*(3), 419–435. https://doi.org/10.1177/0022009402037003050 1
Jr, J. E. W. (2009). Putnam, Dennett, and Others: Philosophical Resources for the World Historian. *Journal of World History*, *20*, 491–522. https://doi.org/10.2307/40542848
Kaiser, W. (2005). Cultural Transfer of Free Trade at the World Exhibitions, 1851–1862. *The Journal of Modern History*, *77*(3), 563–590 https://doi.org/10.1086/497717
Kallis, A. (2012). The "Third Rome" of Fascism: Demolitions and the Search for a New Urban Syntax. *The Journal of Modern History*, *84*(1), 40–79 CR–Copyright © 2012 The University o. https://doi.org/10.1086/663287
Keita, M. (2005). Africans and Asians: Historiography and the Long View of Global Interaction. *Journal of World History*, *16*, 1–30. https://doi.org/10.2307/20079302
Kershaw, I. (2004). Hitler and the Uniqueness of Nazism. *Journal of Contemporary History, 39*(2), 239–254. https://doi.org/10.1177/0022009404042130
Kitson, P. M. (2009). Religious Change and the Timing of Baptism in England, 1538–1750. *The Historical Journal*, *52*, 269–294.
Knepper, P. (2009). The `White Slave Trade' and the Music Hall Affair in 1930s Malta. *Journal of Contemporary History, 44*(2), 205–220. https://doi.org/10.1177/0022009408101248
Knight, R. (2007). Denazification and Integration in the Austrian Province of Carinthia. *The Journal of Modern History*, *79*(3), 572–612 https://doi.org/10.1086/517982
Koehlinger, A. (2004). "Let us Live for Those Who Love Us": Faith, Family, and the Contours of Manhood among the Knights of Columbus in Late Nineteenth-Century Connecticut. *Journal of Social History, 38*(2), 455–469. https://doi.org/10.1353/jsh.2004.0126
Kostroun, D. (2003). A Formula for Disobedience: Jansenism, Gender, and the Feminist Paradox. *The Journal of Modern History*, *75*(3), 483–522 https://doi.org/10.1086/380236
Lake, M. (2010). Chinese Colonists Assert Their "Common Human Rights": Cosmopolitanism as Subject and Method of History. *Journal of World History*, *21*(3), 375–392. https://doi.org/10.1353/jwh.2010.0011
Leff, L. M. (2002). Jewish Solidarity in Nineteenth-Century France: The Evolution of a Concept. *The Journal of Modern History*, *74*(1), 33–61 CR–Copyright © 2002 The University o. https://doi.org/10.1086/343367
Levillain, C.-E. (2005). William III's Military and Political Career in Neo-Roman Context, 1672-1702. *The Historical Journal*, *48*, 321–350.
Levine, E. J. (2011). PanDora, or Erwin and Dora Panofsky and the Private History of Ideas. *The Journal of Modern History*, *83*(4), 753–787 https://doi.org/10.1086/662217
Lewis, B. M. D. (2008). Geographies of Power: The Tunisian Civic Order, Jurisdictional Politics, and Imperial Rivalry in the Mediterranean, 1881–1935. *The Journal of Modern History*, *80*(4), 791–830 https://doi.org/10.1086/591111
Linstrum, E. (2012). The Politics of Psychology in the British Empire, 1898–1960. *Past & Present* , *215*(1), 195–233. https://doi.org/10.1093/pastj/gtr044
Liu, X. (2001). Migration and Settlement of the Yuezhi-Kushan: Interaction and Interdependence of Nomadic and Sedentary Societies. *Journal of World History*, *12*, 261–292. https://doi.org/10.2307/20078910

Liu, Y. (2011). The Spiritual Journey of an Independent Thinker: The Conversion of Li Zhizao to Catholicism. *Journal of World History*, *22*, 433–453. https://doi.org/citeulike-article-id:12446579

Love, G. (2007). `What's the Big Idea?': Oswald Mosley, the British Union of Fascists and Generic Fascism. *Journal of Contemporary History, 42*(3), 447–468. https://doi.org/10.1177/0022009407078334

Lovell, S. (2002). The Making of the Stalin-Era Dacha. *The Journal of Modern History*, *74*(2), 253–288 https://doi.org/10.1086/343408

Lowenthal, D. (2008). "The Marriage of Choice and the Marriage of Covenance": A New England Puritan Views Risorgimento Italy. *Journal of Social History, 42*(1), 157–174. https://doi.org/10.1353/jsh.0.0062

Ma, J. (2003). Peer Polity Interaction in the Hellenistic Age. *Past & Present*, *180*, 9–39. Retrieved from http://past.oxfordjournals.org/content/180/1/9.short

Machaqueiro, M. (2012). The Islamic Policy of Portuguese Colonial Mozambique, 1960–1973. *The Historical Journal*, *55*(4), 1097–1116.

MacLean, S. (2003). Queenship, Nunneries and Royal Widowhood in Carolingian Europe. *Past & Present*, *178*, 3–38. https://doi.org/10.1093/past/178.1.3

Maginn, C. (2004). The Baltinglass Rebellion, 1580: English Dissent or A Gaelic Uprising? *The Historical Journal*, *47*, 205–232.

Magnússon, S. G. (2006). Social History as "Sites of Memory"? The Insitutionalization of History: Microhistory and the Grand Narrative. *Journal of Social History, 39*(3), 891–913. https://doi.org/10.1353/jsh.2006.0012

Maguire, M. J., & Cinnéide, S. Ó. (2005). "A Good Beating Never Hurt Anyone": The Punishment and Abuse of Children in Twentieth Century Ireland. *Journal of Social History, 38*(3), 635–652. https://doi.org/10.1353/jsh.2005.0023

Malkiel, D. (2007). Jews and Apostates in Medieval Europe — Boundaries Real and Imagined. *Past & Present*, *194*, 3–34. Retrieved from http://past.oxfordjournals.org/content/194/1/3.short

Manchester, L. (2011). Gender and Social Estate as National Identity: The Wives and Daughters of Orthodox Clergymen as Civilizing Agents in Imperial Russia. *The Journal of Modern History*, *83*(1), 48–77 CR–Copyright © 2011 The University o. https://doi.org/10.1086/658051

Manning, P. (2006). Homo sapiens Populates the Earth: A Provisional Synthesis, Privileging Linguistic Evidence. *Journal of World History*, *17*, 115–158. https://doi.org/10.2307/20079372

Marshall, P. (2012). The Naming of Protestant England*. *Past & Present* , *214*(1), 87–128. https://doi.org/10.1093/pastj/gtr030

Maskiell, M. (2002). Consuming Kashmir: Shawls and Empires, 1500-2000. *Journal of World History*, *13*, 27–65. https://doi.org/10.2307/20078943

Maxwell, M. J. (2006). Afanasii Nikitin: An Orthodox Russian's Spiritual Voyage in the Dar al-Islam, 1468-1475. *Journal of World History*, *17*, 243–266. https://doi.org/10.2307/20079384

McCahill, M., & Wasson, E. A. (2003). The New Peerage: Recruitment to the House of Lords, 1704-1847. *The Historical Journal*, *46*, 1–38.

McCormack, J. (2011). Social Memories in (Post) Colonial France: Remembering the Franco-Algerian War. *Journal of Social History, 44*(4), 1129–1138. https://doi.org/10.1353/jsh.2011.0048

McCormack, M. (2006). Citizenship, Nationhood, and Masculinity in the Affair of The Hanoverian Soldier, 1756. *The Historical Journal*, *49*, 971–993.

McCormick, M. (2002). New Light on the "Dark Ages": How the Slave Trade Fuelled the Carolingian Economy. *Past & Present*, *177*, 17–54. Retrieved from http://past.oxfordjournals.org/content/177/1/17.short
McCullough, P. (2008). Print, Publication, and Religious Politics in Caroline England. *The Historical Journal*, *51*, 285–313.
McLellan, J. (2007). State Socialist Bodies: East German Nudism from Ban to Boom. *The Journal of Modern History*, *79*(1), 48–79 https://doi.org/10.1086/517544
Mengel, D. C. (2011). A Plague on Bohemia? Mapping the Black Death. *Past & Present*, *211*, 3–34. Retrieved from http://past.oxfordjournals.org/content/211/1/3.short
Miller, P. N. (2002). Nazis and Neo-Stoics: Otto Brunner and Gerhard Oestreich Before and after the Second World War. *Past & Present*, *176*, 144–186. Retrieved from http://past.oxfordjournals.org/content/176/1/144.short
Moore, S. T. (2002). "Justifiable Provocation": Violence against Women in Essex County, New York, 1799–1860. *Journal of Social History, 35*(4), 889–918. https://doi.org/10.1353/jsh.2002.0057
Mouton, M. (2010). From Adventure and Advancement to Derailment and Demotion: Effects of Nazi Gender Policy on Women's Careers and Lives. *Journal of Social History, 43*(4), 945–971. https://doi.org/10.1353/jsh.0.0361
Müller, J.-W. (2011). European Intellectual History as Contemporary History. *Journal of Contemporary History, 46*(3), 574–590. https://doi.org/10.1177/0022009411403339
Nancy L. Green. (2005). The Politics of Exit: Reversing the Immigration Paradigm. *The Journal of Modern History*, *77*(2), 263–289 https://doi.org/10.1086/431815
Naor, M. (2008). Israel's 1948 War of Independence as a Total War. *Journal of Contemporary History, 43*(2), 241–257. https://doi.org/10.1177/0022009408089031
Nash, D. (2007). Analyzing the History of Religious Crime: Models of "Passive" and "Active" Blasphemy since the Medieval Period. *Journal of Social History, 41*(1), 5–29. https://doi.org/10.1353/jsh.2007.0145
Nelson, R. K. (2004). "The Forgetfulness of Sex": Devotion and Desire in the Courtship Letters of Angelina Grimke and Theodore Dwight Weld. *Journal of Social History, 37*(3), 663–679. https://doi.org/10.1353/jsh.2004.0018
Neville, P. (2005). A Prophet Scorned? Ralph Wigram, the Foreign Office and the German Threat, 1933-36. *Journal of Contemporary History, 40*(1), 41–54. https://doi.org/10.1177/0022009405049265
Nicoletta, J. (2010). Art Out of Place: International Art Exhibits at the New York World's Fair of 1964–1965. *Journal of Social History, 44*(2), 499–519. https://doi.org/10.1353/jsh.2010.0066
Nightingale, P. (2010). The Rise and Decline of Medieval York: A Reassessment. *Past & Present*, *206*, 3–42. Retrieved from http://past.oxfordjournals.org/content/206/1/3.short
Nirenberg, D. (2002). Mass Conversion and Genealogical Mentalities: Jews and Christians in Fifteenth-Century Spain. *Past & Present*, *174*, 3–41. Retrieved from http://past.oxfordjournals.org/content/174/1/3.short
Northrup, D. (2005). Globalization and the Great Convergence: Rethinking World History in the Long Term. *Journal of World History*, *16*, 249–267. https://doi.org/10.2307/20079329
Nowakowska, N. (2011). From Strassburg to Trent: Bishops, Printing and Liturgical Reform in the Fifteenth Century. *Past & Present*, *213*, 3–39. Retrieved from http://past.oxfordjournals.org/content/213/1/3.short
O'Hara, G. (2009). "This is What Growth Does": British Views of the European Economies in the Prosperous "Golden Age" of 1951—73. *Journal of Contemporary History, 44*(4), 697–718. https://doi.org/10.1177/0022009409340647

Opler, D. (2002). Monkey Business in Union Square: A Cultural Analysis of the Klein's-Ohrbach's Strikes of 1934–5. *Journal of Social History, 36*(1), 149–164. https://doi.org/10.1353/jsh.2002.0101
Owen, N. (2012). "Facts Are Sacred": The Manchester Guardian and Colonial Violence, 1930–1932. *The Journal of Modern History*, *84*(3), 643–678 https://doi.org/10.1086/666052
Pack, S. D. (2010). Revival of the Pilgrimage to Santiago de Compostela: The Politics of Religious, National, and European Patrimony, 1879–1988. *The Journal of Modern History*, *82*(2), 335–367 https://doi.org/10.1086/651613
Parry, G. (2006). John Dee and the Elizabethan British Empire in Its European Context. *The Historical Journal*, *49*, 643–675.
Partington, J. S. (2004). The Pen as Sword: George Orwell, H.G. Wells and Journalistic Parricide. *Journal of Contemporary History, 39*(1), 45–56. https://doi.org/10.1177/0022009404039883
Partridge, M. (2007). Thomas Hoby's English Translation of Castiglione's Book of the Courtier. *The Historical Journal*, *50*, 769–786.
Pearson, M. N. (2006). Littoral Society: The Concept and the Problems. *Journal of World History*, *17*, 353–373. https://doi.org/10.2307/20079396
Pedaliu, E. G. H. (2004). Britain and the "Hand-Over" of Italian War Criminals to Yugoslavia, 1945–48. *Journal of Contemporary History, 39*(4), 503–529. https://doi.org/10.1177/0022009404046752
Peter Kenez. (2003). The Hungarian Communist Party and the Catholic Church, 1945–1948. *The Journal of Modern History*, *75*(4), 864–889 https://doi.org/10.1086/383356
Peterson, J. L. (2009). Holy Heretics in Later Medieval Italy. *Past & Present*, *204*, 3–31. Retrieved from http://past.oxfordjournals.org/content/204/1/3.short
Pettegree, A., & Hall, M. (2004). The Reformation and the Book: A Reconsideration. *The Historical Journal*, *47*, 785–808.
Pieris, A. (2011). The "Other" Side of Labor Reform: Accounts of Incarceration and Resistance in the Straits Settlements Penal System, 1825–1873. *Journal of Social History, 45*(2), 453–479. https://doi.org/10.1093/jsh/shr082
Piirim, Auml, E, P., & Rtel. (2002). Just War in Theory and Practice: The Legitimation of Swedish Intervention in the Thirty Years War. *The Historical Journal*, *45*, 499–523.
Pittaway, M. (2002). The Reproduction of Hierarchy: Skill, Working-Class Culture, and the State in Early Socialist Hungary. *The Journal of Modern History*, *74*(4), 737–769 https://doi.org/10.1086/376211
Pollard, E. A. (2009). Pliny's Natural History and the Flavian Templum Pacis: Botanical Imperialism in First-Century C. E. Rome. *Journal of World History*, *20*, 309–338. https://doi.org/10.2307/40542802
Popkin, J. D. (2007). Worlds Turned Upside Down: Bourgeois Experience in the 19th-Century Revolutions. *Journal of Social History, 40*(4), 821–839. https://doi.org/10.1353/jsh.2007.0109
Prazniak, R. (2010). Siena on the Silk Roads: Ambrogio Lorenzetti and the Mongol Global Century, 1250-1350. *Journal of World History*, *21*(2), 177–217. https://doi.org/10.2307/20752947
Prestholdt, J. (2007). Similitude and Empire: On Comorian Strategies of Englishness. *Journal of World History*, *18*, 113–138. https://doi.org/10.2307/20079420
Prezioso, S. (2008). Antifascism and Anti-totalitarianism: The Italian Debate. *Journal of Contemporary History, 43*(4), 555–572. https://doi.org/10.1177/0022009408095415
Prior, C. W. A. (2005). Ecclesiology and Political Thought in England, 1580 - 1630. *The Historical Journal*, *48*, 855–884.
Pursell, B. C. (2002). The End of the Spanish Match. *The Historical Journal*, *45*, 699–726.

Ra, J. (2005). Governing North Korea. Some Afterthoughts on the Autumn of 1950. *Journal of Contemporary History, 40*(3), 521–546. https://doi.org/10.1177/0022009405054570
Raath, A., & De Freitas, S. (2005). Rebellion, Resistance, and a Swiss Brutus? *The Historical Journal*, *48*, 1–26.
Randall, D. (2008). Epistolary Rhetoric, the Newspaper, and the Public Sphere. *Past & Present*, *198*, 3–32. Retrieved from http://past.oxfordjournals.org/content/198/1/3.short
Rein, R. (2008). Echoes of the Spanish Civil War in Palestine: Zionists, Communists and the Contemporary Press. *Journal of Contemporary History, 43*(1), 9–23. https://doi.org/10.1177/0022009407084555
Reinisch, J. (2008). `We Shall Rebuild Anew a Powerful Nation': UNRRA, Internationalism and National Reconstruction in Poland. *Journal of Contemporary History, 43*(3), 451–476. https://doi.org/10.1177/0022009408091835
Reiss, M. (2012). The Importance of Being Men: The Afrika-Korps in American Captivity. *Journal of Social History, 46*(1), 23–47. https://doi.org/10.1093/jsh/shs030
Rio, A. (2006). Freedom and Unfreedom in Early Medieval Francia: The Evidence of the Legal Formulae. *Past & Present*, *193*, 7–40. Retrieved from http://past.oxfordjournals.org/content/193/1/7.short
Rivers, I. (2007). The First Evangelical Tract Society. *The Historical Journal*, *50*, 1–22.
Roberts, D. D., De Grand, A., Antliff, M., & Linehan, T. (2002). Comments on Roger Griffin, `The Primacy of Culture: The Current Growth (or Manufacture) of Consensus within Fascist Studies'. *Journal of Contemporary History, 37*(2), 259–274. https://doi.org/10.1177/00220094020370020601
Roberts, G. (2011). Moscow's Cold War on the Periphery: Soviet Policy in Greece, Iran, and Turkey, 1943—8. *Journal of Contemporary History, 46*(1), 58–81. https://doi.org/10.1177/0022009410383292
Roberts, P. (2004). Royal Authority and Justice during the French Religious Wars. *Past & Present*, *184*, 3–32. Retrieved from http://past.oxfordjournals.org/content/184/1/3.short
Romani, R. (2012). Reluctant Revolutionaries: Moderate Liberalism in The Kingdom Of Sardinia, 1849–1859. *The Historical Journal*, *55*(1), 45–73.
Rose, J. (2005). John Locke and the Restoration of the Church of England. *The Historical Journal*, *48*, 601–621.
Rowe, S. E. (2006). Writing Modern Selves: Literacy and the French Working Class in the Early Nineteenth Century. *Journal of Social History, 40*(1), 55–83. https://doi.org/10.1353/jsh.2006.0094
Ruiz Ibáñez, J. J., & Sabatini, G. (2009). Monarchy as Conquest: Violence, Social Opportunity, and Political Stability in the Establishment of the Hispanic Monarchy. *The Journal of Modern History*, *81*(3), 501–536 https://doi.org/10.1086/599270
Ryrie, A. (2009). Paths Not Taken in the British Reformations. *The Historical Journal*, *52*, 1–22.
Sandberg, B. (2006). Beyond Encounters: Religion, Ethnicity, and Violence in the Early Modern Atlantic World, 1492-1700. *Journal of World History*, *17*, 1–25. https://doi.org/10.2307/20079358
Sanford, G. (2006). The Katyn Massacre and Polish-Soviet Relations, 1941-43. *Journal of Contemporary History, 41*(1), 95–111. https://doi.org/10.1177/0022009406058676
Sarah Hanley. (2006). The Family, the State, and the Law in Seventeenth- and Eighteenth-Century France: The Political Ideology of Male Right versus an Early Theory of Natural Rights. *The Journal of Modern History*, *78*(2), 289–332 https://doi.org/10.1086/505799

Scardaville, M. C. (2003). Justice by Paperwork: A Day in the Life of a Court Scribe in Bourbon Mexico City. *Journal of Social History, 36*(4), 979–1007. https://doi.org/10.1353/jsh.2003.0114
Schmidt, A. (2010). Irenic Patriotism in Sixteenth- and Seventeenth-Century German Political Discourse. *The Historical Journal*, *53*, 243–269.
Schwarcz, V. (2008). Truthfulness at Dawn, Truthfulness at Night: Reflections on a Common Striving in Chinese and Jewish Traditions. *Journal of World History*, *19*, 403–430. https://doi.org/10.2307/40542677
Sen, T. (2001). In Search of Longevity and Good Karma: Chinese Diplomatic Missions to Middle India in the Seventh Century. *Journal of World History*, *12*, 1–28. https://doi.org/10.2307/20078877
Shaw, D. (2006). Thomas Wentworth and Monarchical Ritual in Early Modern Ireland. *The Historical Journal*, *49*, 331–355.
Shimazu, N. (2003). Popular Representations of the Past: The Case of Postwar Japan. *Journal of Contemporary History, 38*(1), 101–116. https://doi.org/10.1177/0022009403038001966
Shpayer-Makov, H. (2009). Journalists and Police Detectives in Victorian and Edwardian England: An Uneasy Reciprocal Relationship. *Journal of Social History, 42*(4), 963–987. https://doi.org/10.1353/jsh/42.4.963
Shubert, A. (2012). Women Warriors and National Heroes: Agustina de Aragón and Her Indian Sisters. *Journal of World History*, *23*(2), 279–313. https://doi.org/10.1353/jwh.2012.0039
Siebenhüner, K. (2008). Conversion, Mobility and the Roman Inquisition in Italy Around 1600. *Past & Present*, *200*, 5–35. Retrieved from http://past.oxfordjournals.org/content/200/1/5.short
Siegelbaum, L. (2012). Sputnik Goes to Brussels: The Exhibition of a Soviet Technological Wonder. *Journal of Contemporary History, 47*(1), 120–136. https://doi.org/10.1177/0022009411422372
Sigel, L. Z. (2011). Censorship in Inter-War Britain: Obscenity, Spectacle, and the Workings of the Liberal State. *Journal of Social History, 45*(1), 61–83. https://doi.org/10.1093/jsh/shr004
Singy, P. (2010). The Popularization of Medicine in the Eighteenth Century: Writing, Reading, and Rewriting Samuel Auguste Tissot's Avis au peuple sur sa santé. *The Journal of Modern History*, *82*(4), 769–800 https://doi.org/10.1086/656073
Sizgorich, T. (2004). Narrative and Community in Islamic Late Antiquity. *Past & Present*, *185*, 9–42. Retrieved from http://past.oxfordjournals.org/content/185/1/9.short
Skaff, J. K. (2004). Survival in the Frontier Zone: Comparative Perspectives on Identity and Political Allegiance in China's Inner Asian Borderlands during the Sui-Tang Dynastic Transition (617-630). *Journal of World History*, *15*(2), 117–153. https://doi.org/10.2307/20068610
Smith, A. K. (2009). Authority in a Serf Village: Peasants, Managers, and the Role of Writing in Early Nineteenth Century Russia. *Journal of Social History, 43*(1), 157–170. https://doi.org/10.1353/jsh.0.0226
Sowerby, T. A. (2011). Richard Pate, the Royal Supremacy, and Reformation Diplomacy. *The Historical Journal*, *54*, 265–285.
Speidel, M. P. (2002). Berserks: A History of Indo-European "Mad Warriors." *Journal of World History*, *13*, 253–290. https://doi.org/10.2307/20078974
Stephens, I. (2008). The Courtship and Singlehood of Elizabeth Isham, 1630–1634. *The Historical Journal*, *51*, 1–25.

Strathern, A. (2009). The Vijaya Origin Myth of Sri Lanka and the Strangeness of Kingship. *Past & Present*, *203*, 3–28. Retrieved from http://past.oxfordjournals.org/content/203/1/3.short
Strub, W. (2007). Black and White and Banned All Over: Race, Censorship and Obscenity in Postwar Memphis. *Journal of Social History, 40*(3), 685–715. https://doi.org/10.1353/jsh.2007.0072
Stuurman, S. (2008). Herodotus and Sima Qian: History and the Anthropological Turn in Ancient Greece and Han China. *Journal of World History*, *19*, 1–40. https://doi.org/10.2307/20079459
Susan R. Grayzel. (2006). "The Souls of Soldiers": Civilians under Fire in First World War France. *The Journal of Modern History*, *78*(3), 588–622 https://doi.org/10.1086/509209
Tague, I. H. (2008). Eighteenth-Century English Debates on A Dog Tax. *The Historical Journal*, *51*, 901–920.
Tomlinson, B. R. (2003). What was the Third World? *Journal of Contemporary History, 38*(2), 307–321. https://doi.org/10.1177/0022009403038002135
Tremml, B. M. (2012). The Global and the Local: Problematic Dynamics of the Triangular Trade in Early Modern Manila. *Journal of World History*, *23*(3), 555–586. https://doi.org/10.1353/jwh.2012.0095
Trivellato, F. (2012). Credit, Honor, and the Early Modern French Legend of the Jewish Invention of Bills of Exchange. *The Journal of Modern History*, *84*(2), 289–334 https://doi.org/10.1086/664732
Tyacke, N. (2010). The Puritan Paradigm of English Politics, 1558–1642. *The Historical Journal*, *53*, 527–550.
Underwood, L. (2012). Youth, Religious Identity, and Autobiography at the English Colleges in Rome and Valladolid, 1592–1685. *The Historical Journal*, *55*(2), 349–374.
Underwood, W. (2004). Thomas Cromwell and William Marshall's Protestant Books. *The Historical Journal*, *47*, 517–539.
Valge, J. (2011). Foreign Involvement and Loss of Democracy, Estonia 1934. *Journal of Contemporary History, 46*(4), 788–808. https://doi.org/10.1177/0022009411413408
van Daalen, R. (2008). Children and Their Educators: Changing Relations in a Meritocratizing World. *Journal of Social History, 42*(2), 403–423. https://doi.org/10.1353/jsh.0.0108
Vester, K. (2010). Regime Change: Gender, Class, and the Invention of Dieting in Post-Bellum America. *Journal of Social History, 44*(1), 39–70. https://doi.org/10.1353/jsh.2010.0032
Vick, B. (2010). Liberalism, Nationalism, and Gender Dichotomy in Mid-Nineteenth-Century Germany: The Contested Case of German Civil Law. *The Journal of Modern History*, *82*(3), 546–584 https://doi.org/10.1086/653041
Vikström, L. (2011). Before and after Crime: Life-Course Analyses of Young Offenders Arrested in Nineteenth-Century Northern Sweden. *Journal of Social History, 44*(3), 861–888. https://doi.org/10.1353/jsh.2011.0001
Vink, M. (2003). "The World's Oldest Trade": Dutch Slavery and Slave Trade in the Indian Ocean in the Seventeenth Century. *Journal of World History*, *14*, 131–177. https://doi.org/10.2307/20079204
Viswanathan, G. (2004). Ireland, India, and the Poetics of Internationalism. *Journal of World History*, *15*(1), 7–30. https://doi.org/10.1353/jwh.2004.0015
Von Friedeburg, R. (2010). The Juridification of Natural Law: Christoph Besold's Claim for a Natural Right to Believe what one Wants. *The Historical Journal*, *53*, 1–19.
Vopa, A. J. L. (2005). Thinking about Marriage: Kant's Liberalism and the Peculiar Morality of Conjugal Union. *The Journal of Modern History*, *77*(1), 1–34 CR–Copyright © 2005 The University of. https://doi.org/10.1086/429427

Waite, R. G. (2010). "Ish bin ein Bearleener" - JFK's 26 June 1963 Visit to Berlin: The Views from East Germany. *Journal of Contemporary History, 45*(4), 844–865. https://doi.org/10.1177/0022009410375259
Walker, G. (2002). Rethinking the Fall of Anne Boleyn. *The Historical Journal*, *45*, 1–29.
Walsham, A. (2003). Miracles and the Counter-Reformation Mission to England. *The Historical Journal*, *46*, 779–815.
Warner, J., & Griller, R. (2003). "My Pappa is Out, and My Mamma is Asleep." Minors, Their Routine Activities, and Interpersonal Violence in an Early Modern Town, 1653–1781. *Journal of Social History, 36*(3), 561–584. https://doi.org/10.1353/jsh.2003.0073
Watkins, C. S. (2002). Sin, Penance and Purgatory in the Anglo-Norman Realm: The Evidence of Visions and Ghost Stories. *Past & Present*, *175*, 3–33. Retrieved from http://past.oxfordjournals.org/content/175/1/3.short
Whitelock, A., & Macculloch, D. (2007). Princess Mary's Household and the Succession Crisis, July 1553. *The Historical Journal*, *50*, 265–287.
Whitley, C. B., & Kramer, K. (2010). A New Explanation for the Reproductive Woes and Midlife Decline of Henry VIII. *The Historical Journal*, *53*, 827–848.
Whittle, J. (2010). Lords and Tenants in Kett's Rebellion 1549. *Past & Present*, *207*, 3–52. Retrieved from http://past.oxfordjournals.org/content/207/1/3.short
Wilford, H. (2002). American Labour Diplomacy and Cold War Britain. *Journal of Contemporary History, 37*(1), 45–65. https://doi.org/10.1177/0022009402037001080 1
William Caferro. (2008). Continuity, Long-Term Service, and Permanent Forces: A Reassessment of the Florentine Army in the Fourteenth Century. *The Journal of Modern History*, *80*(2), 219–251 https://doi.org/10.1086/588856
William Wagner. (2006). The Transformation of Female Orthodox Monasticism in Nizhnii Novgorod Diocese, 1764–1929, in Comparative Perspective. *The Journal of Modern History*, *78*(4), 793–845 https://doi.org/10.1086/511203
Wood, J. C. (2009). "Those Who Have Had Trouble Can Sympathise with You": Press Writing, Reader Responses and a Murder Trial in Interwar Britain. *Journal of Social History, 43*(2), 439–462. https://doi.org/10.1353/jsh.0.0277
Wray, S. K. (2009). Instruments of Concord: Making Peace and Settling Disputes through a Notary in the City and Contado of Late Medieval Bologna. *Journal of Social History, 42*(3), 733–760. https://doi.org/10.1353/jsh/42.3.733
Yang, B. (2011). The Rise and Fall of Cowrie Shells: The Asian Story. *Journal of World History*, *22*, 1–25. https://doi.org/10.2307/23011676
Yang, B. (2004). Horses, Silver, and Cowries: Yunnan in Global Perspective. *Journal of World History*, *15*(3), 281–322. https://doi.org/10.2307/20079276
Yanni Kotsonis. (2004). "No Place to Go": Taxation and State Transformation in Late Imperial and Early Soviet Russia. *The Journal of Modern History*, *76*(3), 531–577 https://doi.org/10.1086/425440
Ziv, G. (2010). Shimon Peres and the French-Israeli Alliance, 1954—9. *Journal of Contemporary History*, *45*(2), 406–429. https://doi.org/10.1177/0022009409356915
Zmora, H. (2006). The Princely State and the Noble Family: Conflict and Co-Operation in the Margraviates Kulmbach in the Early Sixteenth Century. *The Historical Journal*, *49*, 1–21.

Management

Adams, R. B., Licht, A. N., & Sagiv, L. (2011). Shareholders and stakeholders: How do directors decide? *Strategic Management Journal*, *32*(12), 1331–1355. https://doi.org/10.1002/smj.940

Ahuja, R. K., Orlin, J. B., Pallottino, S., Scaparra, M. P., & Scutellà, M. G. (2004). A Multi-Exchange Heuristic for the Single-Source Capacitated Facility Location Problem. *Management Science*, *50*(6), 749–760. https://doi.org/10.1287/mnsc.1030.0193

Aken, J. E. van. (2004). Management Research Based on the Paradigm of the Design Sciences: The Quest for Field-Tested and Grounded Technological Rules. *Journal of Management Studies*, *41*, 219–246. https://doi.org/10.1111/j.1467-6486.2004.00430.x

Albuquerque, P., Bronnenberg, B. J., & Corbett, C. J. (2007). A Spatiotemporal Analysis of the Global Diffusion of ISO 9000 and ISO 14000 Certification. *Management Science*, *53*(3), 451–468. https://doi.org/10.1287/mnsc.1060.0633

Alvesson, M., & Kärreman, D. (2007). Unraveling HRM: Identity, Ceremony, and Control in a Management Consulting Firm. *Organization Science*, *18*(4), 711–723. https://doi.org/10.1287/orsc.1070.0267

Alvesson, M., & Sveningsson, S. (2003). Good Visions, Bad Micro-management and Ugly Ambiguity: Contradictions of (Non-)Leadership in a Knowledge-Intensive Organization. *Organization Studies, 24*(6), 961–988. https://doi.org/10.1177/0170840603024006007

Anderson, S. W., & Dekker, H. C. (2005). Management Control for Market Transactions: The Relation between Transaction Characteristics, Incomplete Contract Design, and Subsequent Performance. *Management Science*, *51*(12), 1734–1752. https://doi.org/10.1287/mnsc.1050.0456

Armbrüster, T. (2004). Rationality and Its Symbols: Signalling Effects and Subjectification in Management Consulting*. *Journal of Management Studies*, *41*, 1247–1269. https://doi.org/10.1111/j.1467-6486.2004.00474.x

Arnaud, G. (2007). Poweract and Organizational Work: Gérard Mendel's Socio-psychoanalysis. *Organization Studies, 28*(3), 409–428. https://doi.org/10.1177/0170840607076010

Bakshi, G., & Wu, L. (2010). The Behavior of Risk and Market Prices of Risk Over the Nasdaq Bubble Period. *Management Science*, *56*(12), 2251–2264. https://doi.org/10.1287/mnsc.1100.1256

Bakshi, G., & Madan, D. (2006). A Theory of Volatility Spreads. *Management Science*, *52*(12), 1945–1956. https://doi.org/10.1287/mnsc.1060.0579

Ballinger, G. A., & Marcel, J. J. (2010). The use of an interim CEO during succession episodes and firm performance. *Strategic Management Journal*, *31*(3), 262–283. https://doi.org/10.1002/smj.808

Barnett, M. L. (2006). Finding a Working Balance Between Competitive and Communal Strategies*. *Journal of Management Studies*, *43*, 1753–1773. https://doi.org/10.1111/j.1467-6486.2006.00661.x

Baron, R. A., & Jintong Tang. (2009). Entrepreneurs' Social Skills and New Venture Performance: Mediating Mechanisms and Cultural Generality. *Journal of Management, 35*(2), 282–306. https://doi.org/10.1177/0149206307312513

Barry, D., & Rerup, C. (2006). Going Mobile: Aesthetic Design Considerations from Calder and the Constructivists. *Organization Science*, *17*(2), 262–276. https://doi.org/10.1287/orsc.1050.0165

Benjamin, O., & Goclaw, R. (2005). Narrating the Power of Non-Standard Employment: The Case of the Israeli Public Sector*. *Journal of Management Studies*, *42*, 737–759. https://doi.org/10.1111/j.1467-6486.2005.00517.x

Birkinshaw, J., & Lingblad, M. (2005). Intrafirm Competition and Charter Evolution in the Multibusiness Firm. *Organization Science*, *16*(6), 674–686. https://doi.org/10.1287/orsc.1050.0142

Birkinshaw, J., Nobel, R., & Ridderstråle, J. (2002). Knowledge as a Contingency Variable: Do the Characteristics of Knowledge Predict Organization Structure? *Organization Science*, *13*(3), 274–289. https://doi.org/10.1287/orsc.13.3.274.2778

Blair-Loy, M., Wharton, A. S., & Goodstein, J. (2011). Exploring the Relationship between Mission Statements and Work-Life Practices in Organizations. *Organization Studies, 32*(3), 427–450. https://doi.org/10.1177/0170840610397480

Bogaert, S., Boone, C., & van Witteloostuijn, A. (2012). Social Value Orientation and Climate Strength as Moderators of the Impact of Work Group Cooperative Climate on Affective Commitment. *Journal of Management Studies*, *49*(5), 918–944. https://doi.org/10.1111/j.1467-6486.2011.01029.x

Bollapragada, R., Rao, U. S., & Zhang, J. (2004). Managing Inventory and Supply Performance in Assembly Systems with Random Supply Capacity and Demand. *Management Science*, *50*(12), 1729–1743. https://doi.org/10.1287/mnsc.1040.0314

Bonazzi, G., & Antonelli, C. (2003). To Make or to Sell? The Case of In-House Outsourcing at Fiat Auto. *Organization Studies, 24*(4), 575–594. https://doi.org/10.1177/0170840603024004004

Boswell, W. R., & Olson-Buchanan, J. B. (2007). The Use of Communication Technologies After Hours: The Role of Work Attitudes and Work-Life Conflict. *Journal of Management, 33*(4), 592–610. https://doi.org/10.1177/0149206307302552

Boulding, W., & Christen, M. (2009). Pioneering Plus a Broad Product Line Strategy: Higher Profits or Deeper Losses? *Management Science*, *55*(6), 958–967. https://doi.org/10.1287/mnsc.1090.0997

Bowen, H. P. (2012). Testing Moderating Hypotheses in Limited Dependent Variable and Other Nonlinear Models: Secondary Versus Total Interactions. *Journal of Management, 38*(3), 860–889. https://doi.org/10.1177/0149206309356324

Bowman, C., & Swart, J. (2007). Whose Human Capital? The Challenge of Value Capture When Capital is Embedded. *Journal of Management Studies*, *44*, 488–505. https://doi.org/10.1111/j.1467-6486.2006.00635.x

Boyce, A. S., Ryan, A. M., Imus, A. L., & Morgeson, F. P. (2007). "Temporary Worker, Permanent Loser?"A Model of the Stigmatization of Temporary Workers. *Journal of Management, 33*(1), 5–29. https://doi.org/10.1177/0149206306296575

Boyd, B. K., Bergh, D. D., & Ketchen, D. J. (2010). Reconsidering the Reputation—Performance Relationship: A Resource-Based View. *Journal of Management, 36*(3), 588–609. https://doi.org/10.1177/0149206308328507

Boyd, B. K., Finkelstein, S., & Gove, S. (2005). How advanced is the strategy paradigm? The role of particularism and universalism in shaping research outcomes. *Strategic Management Journal*, *26*(9), 841–854. https://doi.org/10.1002/smj.477

Brady, T., & Davies, A. (2004). Building Project Capabilities: From Exploratory to Exploitative Learning. *Organization Studies, 25*(9), 1601–1621. https://doi.org/10.1177/0170840604048002

Brammer, S. J., Pavelin, S., & Porter, L. A. (2009). Corporate Charitable Giving, Multinational Companies and Countries of Concern. *Journal of Management Studies*, *46*, 575–596. https://doi.org/10.1111/j.1467-6486.2008.00827.x

Broadie, M., Du, Y., & Moallemi, C. C. (2011). Efficient Risk Estimation via Nested Sequential Simulation. *Management Science*, *57*(6), 1172–1194. https://doi.org/10.1287/mnsc.1110.1330

Brouthers, K. D., & Hennart, J.-F. (2007). Boundaries of the Firm: Insights From International Entry Mode Research. *Journal of Management, 33*(3), 395–425. https://doi.org/10.1177/0149206307300817
Brown, A. D., & Humphreys, M. (2006). Organizational Identity and Place: A Discursive Exploration of Hegemony and Resistance. *Journal of Management Studies*, *43*, 231–257. https://doi.org/10.1111/j.1467-6486.2006.00589.x
Buell, R. W., & Norton, M. I. (2011). The Labor Illusion: How Operational Transparency Increases Perceived Value. *Management Science*, *57*(9), 1564–1579. https://doi.org/10.1287/mnsc.1110.1376
Cachon, G. P., & Zhang, F. (2006). Procuring Fast Delivery: Sole Sourcing with Information Asymmetry. *Management Science*, *52*(6), 881–896. https://doi.org/10.1287/mnsc.1060.0510
Cao, Q., Gedajlovic, E., & Zhang, H. (2009). Unpacking Organizational Ambidexterity: Dimensions, Contingencies, and Synergistic Effects. *Organization Science*, *20*(4), 781–796. https://doi.org/10.1287/orsc.1090.0426
Carare, O., & Rothkopf, M. (2005). Slow Dutch Auctions. *Management Science*, *51*(3), 365–373. https://doi.org/10.1287/mnsc.1040.0328
Carney, M., & Gedajlovic, E. (2002). The Co-evolution of Institutional Environments and Organizational Strategies: The Rise of Family Business Groups in the ASEAN Region. *Organization Studies, 23*(1), 1–29. https://doi.org/10.1177/0170840602231001
Caruana, R., & Crane, A. (2008). Constructing Consumer Responsibility: Exploring the Role of Corporate Communications. *Organization Studies, 29*(12), 1495–1519. https://doi.org/10.1177/0170840607096387
Cattani, G., Pennings, J. M., & Wezel, F. C. (2003). Spatial and Temporal Heterogeneity in Founding Patterns. *Organization Science*, *14*(6), 670–685. https://doi.org/10.1287/orsc.14.6.640.24874
Chen, C. C., Peng, M. W., & Saparito, P. A. (2002). Individualism, Collectivism, and Opportunism: A Cultural Perspective on Transaction Cost Economics. *Journal of Management, 28*(4), 567–583. https://doi.org/10.1177/014920630202800405
Chen, G., Liu, C., & Tjosvold, D. (2005). Conflict Management for Effective Top Management Teams and Innovation in China*. *Journal of Management Studies*, *42*, 277–300. https://doi.org/10.1111/j.1467-6486.2005.00497.x
Child, J., & Rodrigues, S. B. (2011). How Organizations Engage with External Complexity: A Political Action Perspective. *Organization Studies, 32*(6), 803–824. https://doi.org/10.1177/0170840611410825
Chod, J., Rudi, N., & Van Mieghem, J. A. (2010). Operational Flexibility and Financial Hedging: Complements or Substitutes? *Management Science*, *56*(6), 1030–1045. https://doi.org/10.1287/mnsc.1090.1137
Choi, Y. R., & Shepherd, D. A. (2005). Stakeholder Perceptions of Age and Other Dimensions of Newness. *Journal of Management, 31*(4), 573–596. https://doi.org/10.1177/0149206304272294
Christ, D., & Avi-Itzhak, B. (2002). Strategic Equilibrium for a Pair of Competing Servers with Convex Cost and Balking. *Management Science*, *48*(6), 813–820. https://doi.org/10.1287/mnsc.48.6.813.191
Clegg, S., e Cunha, M. P., & Rego, A. (2012). The Theory and Practice of Utopia in a Total Institution: The Pineapple Panopticon. *Organization Studies, 33*(12), 1735–1757. https://doi.org/10.1177/0170840612464611
Cohen, M. A., Ho, T. H., Ren, Z. J., & Terwiesch, C. (2003). Measuring Imputed Cost in the Semiconductor Equipment Supply Chain. *Management Science*, *49*(12), 1653–1670. https://doi.org/10.1287/mnsc.49.12.1653.25115

Colman, H. L., & Lunnan, R. (2011). Organizational Identification and Serendipitous Value Creation in Post-Acquisition Integration. *Journal of Management, 37*(3), 839–860. https://doi.org/10.1177/0149206309354645
Cottrell, T., & Nault, B. R. (2004). Product variety and firm survival in the microcomputer software industry. *Strategic Management Journal*, *25*(10), 1005–1025. https://doi.org/10.1002/smj.408
Crilly, D., & Sloan, P. (2012). Enterprise logic: explaining corporate attention to stakeholders from the "inside-out." *Strategic Management Journal*, *33*(10), 1174–1193. https://doi.org/10.1002/smj.1964
Crouch, C. (2006). Modelling the Firm in its Market and Organizational Environment: Methodologies for Studying Corporate Social Responsibility. *Organization Studies, 27*(10), 1533–1551. https://doi.org/10.1177/0170840606068255
Daskalaki, M. (2010). Building "Bonds" and "Bridges": Linking Tie Evolution and Network Identity in the Creative Industries. *Organization Studies, 31*(12), 1649–1666. https://doi.org/10.1177/0170840610380805
de Bettignies, J.-E., & Chemla, G. (2008). Corporate Venturing, Allocation of Talent, and Competition for Star Managers. *Management Science*, *54*(3), 505–521. https://doi.org/10.1287/mnsc.1070.0758
Desai, V. M. (2008). Constrained Growth: How Experience, Legitimacy, and Age Influence Risk Taking in Organizations. *Organization Science*, *19*(4), 594–608. https://doi.org/10.1287/orsc.1070.0335
Devaraj, S., & Kohli, R. (2003). Performance Impacts of Information Technology: Is Actual Usage the Missing Link? *Management Science*, *49*(3), 273–289. https://doi.org/10.1287/mnsc.49.3.273.12736
Ding, M., & Eliashberg, J. (2002). Structuring the New Product Development Pipeline. *Management Science*, *48*(3), 343–363. https://doi.org/10.1287/mnsc.48.3.343.7727
Dobrev, S. D., & Carroll, G. R. (2003). Size (and competition) among organizations: modeling scale-based selection among automobile producers in four major countries, 1885–1981. *Strategic Management Journal*, *24*(6), 541–558. https://doi.org/10.1002/smj.317
Dreyfus, H. L., & Dreyfus, S. E. (2005). Peripheral Vision: Expertise in Real World Contexts. *Organization Studies, 26*(5), 779–792. https://doi.org/10.1177/0170840605053102
Dulebohn, J. H. (2002). An Investigation of the Determinants of Investment Risk Behavior in Employer-Sponsored Retirement Plans. *Journal of Management, 28*(1), 3–26. https://doi.org/10.1177/014920630202800102
Dunn, P. (2004). The Impact of Insider Power on Fraudulent Financial Reporting. *Journal of Management, 30*(3), 397–412. https://doi.org/10.1016/j.jm.2003.02.004
Durand, R., Bruyaka, O., & Mangematin, V. (2008). Do science and money go together? The case of the French biotech industry. *Strategic Management Journal*, *29*(12), 1281–1299. https://doi.org/10.1002/smj.707
Easterby-Smith, M., Lyles, M. A., & Tsang, E. W. K. (2008). Inter-Organizational Knowledge Transfer: Current Themes and Future Prospects. *Journal of Management Studies*, *45*, 677–690. https://doi.org/10.1111/j.1467-6486.2008.00773.x
Eckhardt, J. T., & Shane, S. A. (2003). Opportunities and Entrepreneurship. *Journal of Management, 29*(3), 333–349. https://doi.org/10.1177/014920630302900304
Eggers, J. P. (2012). All experience is not created equal: learning, adapting, and focusing in product portfolio management. *Strategic Management Journal*, *33*(3), 315–335. https://doi.org/10.1002/smj.956

Egri, C. P., & Ralston, D. A. (2004). Generation Cohorts and Personal Values: A Comparison of China and the United States. *Organization Science*, *15*(2), 210–220. https://doi.org/10.1287/orsc.1030.0048
Eisenmann, T. R. (2006). Internet companies' growth strategies: determinants of investment intensity and long-term performance. *Strategic Management Journal*, *27*(12), 1183–1204. https://doi.org/10.1002/smj.567
Elfring, T., & Hulsink, W. (2007). Networking by Entrepreneurs: Patterns of Tie—Formation in Emerging Organizations. *Organization Studies, 28*(12), 1849–1872. https://doi.org/10.1177/0170840607078719
Elicker, J. D., Levy, P. E., & Hall, R. J. (2006). The Role of Leader-Member Exchange in the Performance Appraisal Process. *Journal of Management, 32*(4), 531–551. https://doi.org/10.1177/0149206306286622
Elsbach, K. D., Barr, P. S., & Hargadon, A. B. (2005). Identifying Situated Cognition in Organizations. *Organization Science*, *16*(4), 422–433. https://doi.org/10.1287/orsc.1050.0138
Eraker, B. (2008). Affine General Equilibrium Models. *Management Science*, *54*(12), 2068–2080. https://doi.org/10.1287/mnsc.1070.0796
Fagart, M.-C., & Sinclair-Desgagné, B. (2007). Ranking Contingent Monitoring Systems. *Management Science*, *53*(9), 1501–1509. https://doi.org/10.1287/mnsc.1060.0693
Felici, G., & Gentile, C. (2004). A Polyhedral Approach for the Staff Rostering Problem. *Management Science*, *50*(3), 381–393. https://doi.org/10.1287/mnsc.1030.0142
Firth, M., Fung, P. M. Y., & Rui, O. M. (2006). Firm Performance, Governance Structure, and Top Management Turnover in a Transitional Economy*. *Journal of Management Studies*, *43*, 1289–1330. https://doi.org/10.1111/j.1467-6486.2006.00621.x
Fiss, P. C., Kennedy, M. T., & Davis, G. F. (2012). How Golden Parachutes Unfolded: Diffusion and Variation of a Controversial Practice. *Organization Science*, *23*(4), 1077–1099. https://doi.org/10.1287/orsc.1110.0685
Flanagan, D. J., & O'Shaughnessy, K. C. (2005). The Effect of Layoffs on Firm Reputation. *Journal of Management, 31*(3), 445–463. https://doi.org/10.1177/0149206304272186
Fong, E. A., Misangyi, V. F., & Tosi, H. L. (2010). The effect of CEO pay deviations on CEO withdrawal, firm size, and firm profits. *Strategic Management Journal*, *31*(6), 629–651. https://doi.org/10.1002/smj.827
Fortune, A., & Mitchell, W. (2012). Unpacking Firm Exit at the Firm and Industry Levels: The Adaptation and Selection of Firm Capabilities. *Strategic Management Journal*, *33*(7), 794–819. https://doi.org/10.1002/smj.972
Fox, C. R., & Clemen, R. T. (2005). Subjective Probability Assessment in Decision Analysis: Partition Dependence and Bias toward the Ignorance Prior. *Management Science*, *51*(9), 1417–1432. https://doi.org/10.1287/mnsc.1050.0409
Fox, S. (2008). `That Miracle of Familiar Organizational Things': Social and Moral Order in the MBA Classroom. *Organization Studies, 29*(5), 733–761. https://doi.org/10.1177/0170840608088765
Fritz, C., & Sonnentag, S. (2009). Antecedents of Day-Level Proactive Behavior: A Look at Job Stressors and Positive Affect During the Workday†. *Journal of Management, 35*(1), 94–111. https://doi.org/10.1177/0149206307308911
Fuentelsaz, L., Gomez, J., & Polo, Y. (2002). Followers' entry timing: evidence from the Spanish banking sector after deregulation. *Strategic Management Journal*, *23*(3), 245–264. https://doi.org/10.1002/smj.222
Gainey, T. W., & Klaas, B. S. (2003). The Outsourcing of Training and Development: Factors Impacting Client Satisfaction. *Journal of Management, 29*(2), 207–229. https://doi.org/10.1177/014920630302900205

García-Canal, E., Valdés-Llaneza, A., & Ariño, A. (2003). Effectiveness of Dyadic and Multi-Party Joint Ventures. *Organization Studies, 24*(5), 743–770. https://doi.org/10.1177/0170840603024005004
George, E., Chattopadhyay, P., & Zhang, L. L. (2012). Helping Hand or Competition? The Moderating Influence of Perceived Upward Mobility on the Relationship between Blended Workgroups and Employee Attitudes and Behaviors. *Organization Science*, *23*(2), 355–372. https://doi.org/10.1287/orsc.1100.0606
Geppert, M., & Matten, D. (2006). Institutional Influences on Manufacturing Organization in Multinational Corporations: The "Cherrypicking" Approach. *Organization Studies, 27*(4), 491–515. https://doi.org/10.1177/0170840605059452
Ghemawat, P., & Levinthal, D. (2008). Choice Interactions and Business Strategy. *Management Science*, *54*(9), 1638–1651. https://doi.org/10.1287/mnsc.1080.0883
Gherardi, S., & Nicolini, D. (2002). Learning in A Constellation of Interconnected Practices: Canon or Dissonance? *Journal of Management Studies*, *39*, 419–436. https://doi.org/10.1111/1467-6486.t01-1-00298
Gilson, L. L., & Shalley, C. E. (2004). A Little Creativity Goes a Long Way: An Examination of Teams' Engagement in Creative Processes. *Journal of Management, 30*(4), 453–470. https://doi.org/10.1016/j.jm.2003.07.001
Goldman, B. M. (2003). The Application of Referent Cognitions Theory to Legal-Claiming by Terminated Workers: The Role of Organizational Justice and Anger. *Journal of Management, 29*(5), 705–728. https://doi.org/10.1016/S0149-2063_03_00032-1
Goodrick, E. (2002). From Management as a Vocation to Management as a Scientific Activity: An Institutional Account of a Paradigm Shift. *Journal of Management, 28*(5), 649–668. https://doi.org/10.1177/014920630202800505
Graffin, S. D., Wade, J. B., Porac, J. F., & McNamee, R. C. (2008). The Impact of CEO Status Diffusion on the Economic Outcomes of Other Senior Managers. *Organization Science*, *19*(3), 457–474. https://doi.org/10.1287/orsc.1080.0354
Greenberg, J., & Tomlinson, E. C. (2004). Situated Experiments in Organizations: Transplanting the Lab to the Field. *Journal of Management, 30*(5), 703–724. https://doi.org/10.1016/j.jm.2003.11.001
Greve, H. R. (2008). Multimarket contact and sales growth: evidence from insurance. *Strategic Management Journal*, *29*(3), 229–249. https://doi.org/10.1002/smj.662
Gulati, R., & Puranam, P. (2009). Renewal through Reorganization: The Value of Inconsistencies between Formal and Informal Organization. *Organization Science*, *20*(2), 422–440. https://doi.org/10.1287/orsc.1090.0421
Haas, M. R., & Park, S. (2010). To Share or Not to Share? Professional Norms, Reference Groups, and Information Withholding Among Life Scientists. *Organization Science*, *21*(4), 873–891. https://doi.org/10.1287/orsc.1090.0500
Hankins, K. W. (2011). How Do Financial Firms Manage Risk? Unraveling the Interaction of Financial and Operational Hedging. *Management Science*, *57*(12), 2197–2212. https://doi.org/10.1287/mnsc.1090.1068
Hargadon, A. B., & Bechky, B. A. (2006). When Collections of Creatives Become Creative Collectives: A Field Study of Problem Solving at Work. *Organization Science*, *17*(4), 484–500. https://doi.org/10.1287/orsc.1060.0200
Harrison, S. H., & Corley, K. G. (2011). Clean Climbing, Carabiners, and Cultural Cultivation: Developing an Open-Systems Perspective of Culture. *Organization Science*, *22*(2), 391–412. https://doi.org/10.1287/orsc.1100.0538
Harvey, J., Pettigrew, A., & Ferlie, E. (2002). The determinants of research group performance: Towards mode 2?*. *Journal of Management Studies*, *39*, 747–774. https://doi.org/10.1111/1467-6486.00310

Hassard, J., Morris, J., & McCann, L. (2012). "My Brilliant Career"? New Organizational Forms and Changing Managerial Careers in Japan, the UK, and USA. *Journal of Management Studies*, *49*(3), 571–599. https://doi.org/10.1111/j.1467-6486.2011.01032.x

He, Z.-L., & Wong, P.-K. (2004). Exploration vs. Exploitation: An Empirical Test of the Ambidexterity Hypothesis. *Organization Science*, *15*(4), 481–494. https://doi.org/10.1287/orsc.1040.0078

Heavey, C., Simsek, Z., Roche, F., & Kelly, A. (2009). Decision Comprehensiveness and Corporate Entrepreneurship: The Moderating Role of Managerial Uncertainty Preferences and Environmental Dynamism. *Journal of Management Studies*, *46*(8), 1289–1314. https://doi.org/10.1111/j.1467-6486.2009.00858.x

Henderson, B. J., & Tookes, H. (2012). Do Investment Banks' Relationships with Investors Impact Pricing? The Case of Convertible Bond Issues. *Management Science*, *58*(12), 2272–2291. https://doi.org/10.1287/mnsc.1120.1553

Higgins, M. C., & Gulati, R. (2003). Getting Off to a Good Start: The Effects of Upper Echelon Affiliations on Underwriter Prestige. *Organization Science*, *14*(3), 244–263. https://doi.org/10.1287/orsc.14.2.244.15160

Hinds, P. J., & Mortensen, M. (2005). Understanding Conflict in Geographically Distributed Teams: The Moderating Effects of Shared Identity, Shared Context, and Spontaneous Communication. *Organization Science*, *16*(3), 290–307. https://doi.org/10.1287/orsc.1050.0122

Ho, J. L. Y., Wu, A., & Xu, S. X. (2011). Corporate Governance and returns on information technology investment: evidence from an emerging market. *Strategic Management Journal*, *32*(6), 595–623. https://doi.org/10.1002/smj.886

Holburn, G. L. F., & Zelner, B. A. (2010). Political capabilities, policy risk, and international investment strategy: evidence from the global electric power generation industry. *Strategic Management Journal*, *31*(12), 1290–1315. https://doi.org/10.1002/smj.860

Holmqvist, M. (2004). Experiential Learning Processes of Exploitation and Exploration Within and Between Organizations: An Empirical Study of Product Development. *Organization Science*, *15*(1), 70–81. https://doi.org/10.1287/orsc.1030.0056

Hong Ren, Gray, B., & Kim, K. (2009). Performance of International Joint Ventures: What Factors Really Make a Difference and How? *Journal of Management, 35*(3), 805–832. https://doi.org/10.1177/0149206308331165

Hsu, V. N., Lee, C. Y., & So, K. C. (2006). Optimal Component Stocking Policy for Assemble-to-Order Systems with Lead-Time-Dependent Component and Product Pricing. *Management Science*, *52*(3), 337–351. https://doi.org/10.1287/mnsc.1050.0438

Huault, I., & Rainelli-Le Montagner, H. (2009). Market Shaping as an Answer to Ambiguities: The Case of Credit Derivatives. *Organization Studies, 30*(5), 549–575. https://doi.org/10.1177/0170840609106113

Hutzschenreuter, T., & Kleindienst, I. (2006). Strategy-Process Research: What Have We Learned and What Is Still to Be Explored. *Journal of Management, 32*(5), 673–720. https://doi.org/10.1177/0149206306291485

Iedema, R. (2007). On the Multi-modality, Materially and Contingency of Organization Discourse. *Organization Studies, 28*(6), 931–946. https://doi.org/10.1177/0170840607075264

Ireland, R. D., Hitt, M. A., & Sirmon, D. G. (2003). A Model of Strategic Entrepreneurship: The Construct and its Dimensions. *Journal of Management, 29*(6), 963–989. https://doi.org/10.1016/S0149-2063_03_00086-2

Ireland, R. D., Hitt, M. A., & Vaidyanath, D. (2002). Alliance Management as a Source of Competitive Advantage. *Journal of Management, 28*(3), 413–446. https://doi.org/10.1177/014920630202800308

Jagannathan, R., Marakani, S., Takehara, H., & Wang, Y. (2012). Calendar Cycles, Infrequent Decisions, and the Cross Section of Stock Returns. *Management Science*, *58*(3), 507–522. https://doi.org/10.1287/mnsc.1110.1427

Jane Zhao, Z., & Anand, J. (2009). A multilevel perspective on knowledge transfer: evidence from the Chinese automotive industry. *Strategic Management Journal*, *30*(9), 959–983. https://doi.org/10.1002/smj.780

Jansen, K. J. (2004). From Persistence to Pursuit: A Longitudinal Examination of Momentum during the Early Stages of Strategic Change. *Organization Science*, *15*(3), 276–294. https://doi.org/10.1287/orsc.1040.0064

Jelinek, M., & Griffith, T. L. (2005). Organizational Science and the NSF: Funding for Mutual Benefit. *Organization Science*, *16*(5), 550–559. https://doi.org/10.1287/orsc.1050.0155

Jensen, M., & Zajac, E. J. (2004). Corporate elites and corporate strategy: how demographic preferences and structural position shape the scope of the firm. *Strategic Management Journal*, *25*(6), 507–524. https://doi.org/10.1002/smj.393

Jiang, Z., Sarkar, S., De, P., & Dey, D. (2007). A Framework for Reconciling Attribute Values from Multiple Data Sources. *Management Science*, *53*(12), 1946–1963. https://doi.org/10.1287/mnsc.1070.0745

Johnson, J. P., Korsgaard, M. A., & Sapienza, H. J. (2002). Perceived fairness, decision control, and commitment in international joint venture management teams. *Strategic Management Journal*, *23*(12), 1141–1160. https://doi.org/10.1002/smj.277

Julian, S. D., Ofori-Dankwa, J. C., & Justis, R. T. (2008). Understanding strategic responses to interest group pressures. *Strategic Management Journal*, *29*(9), 963–984. https://doi.org/10.1002/smj.698

Kamoche, K., Pang, M., & Wong, A. L. Y. (2011). Career Development and Knowledge Appropriation: A Genealogical Critique. *Organization Studies, 32*(12), 1665–1679. https://doi.org/10.1177/0170840611421249

Kane, A. A. (2010). Unlocking Knowledge Transfer Potential: Knowledge Demonstrability and Superordinate Social Identity. *Organization Science*, *21*(3), 643–660. https://doi.org/10.1287/orsc.1090.0469

Katok, E., & Wu, D. Y. (2009). Contracting in Supply Chains: A Laboratory Investigation. *Management Science*, *55*(12), 1953–1968. https://doi.org/10.1287/mnsc.1090.1089

Kay, N. (2008). Reappraising the Nature of the Firm: The Role and Influence of Lexical and Structural Ambiguity. *Organization Studies, 29*(8–9), 1209–1226. https://doi.org/10.1177/0170840608094777

Kenny, K. (2012). "Someone Big and Important": Identification and Affect in an International Development Organization. *Organization Studies, 33*(9), 1175–1193. https://doi.org/10.1177/0170840612448156

Kim, J.-Y., & Finkelstein, S. (2009). The effects of strategic and market complementarity on acquisition performance: evidence from the U.S. commercial banking industry, 1989–2001. *Strategic Management Journal*, *30*(6), 617–646. https://doi.org/10.1002/smj.754

Kishimoto, N. (2004). Pricing Path-Dependent Securities by the Extended Tree Method. *Management Science*, *50*(9), 1235–1248. https://doi.org/10.1287/mnsc.1030.0198

Klaas, B. S., Gainey, T. W., McClendon, J. A., & Yang, H. (2005). Professional Employer Organizations and Their Impact on Client Satisfaction with Human Resource Outcomes: A Field Study of Human Resource Outsourcing in Small and Medium Enterprises. *Journal of Management, 31*(2), 234–254. https://doi.org/10.1177/0149206304271761

Knott, A. M., Posen, H. E., & Wu, B. (2009). Spillover Asymmetry and Why It Matters. *Management Science*, *55*(3), 373–388. https://doi.org/10.1287/mnsc.1080.0950

Koenigsberg, O., Kohli, R., & Montoya, R. (2010). Package Size Decisions. *Management Science*, *56*(3), 485–494. https://doi.org/10.1287/mnsc.1090.1119
Korczynski, M., & Ott, U. (2004). When Production and Consumption Meet: Cultural Contradictions and the Enchanting Myth of Customer Sovereignty. *Journal of Management Studies*, *41*, 575–599. https://doi.org/10.1111/j.1467-6486.2004.00445.x
Lamertz, K., & Heugens, P. P. M. A. R. (2009). Institutional Translation through Spectatorship: Collective Consumption and Editing of Symbolic Organizational Texts by Firms and their Audiences. *Organization Studies, 30*(11), 1249–1279. https://doi.org/10.1177/0170840609337935
Laursen, K., Masciarelli, F., & Prencipe, A. (2012). Regions Matter: How Localized Social Capital Affects Innovation and External Knowledge Acquisition. *Organization Science*, *23*(1), 177–193. https://doi.org/10.1287/orsc.1110.0650
Lavie, D. (2007). Alliance portfolios and firm performance: A study of value creation and appropriation in the U.S. software industry. *Strategic Management Journal*, *28*(12), 1187–1212. https://doi.org/10.1002/smj.637
Lavie, D., Haunschild, P. R., & Khanna, P. (2012). Organizational differences, relational mechanisms, and alliance performance. *Strategic Management Journal*, *33*(13), 1453–1479. https://doi.org/10.1002/smj.1987
Lee, P. M., & James, E. H. (2007). She'-e-os: gender effects and investor reactions to the announcements of top executive appointments. *Strategic Management Journal*, *28*(3), 227–241. https://doi.org/10.1002/smj.575
Leiblein, M. J., & Miller, D. J. (2003). An empirical examination of transaction- and firm-level influences on the vertical boundaries of the firm. *Strategic Management Journal*, *24*(9), 839–859. https://doi.org/10.1002/smj.340
Levin, D. Z., Walter, J., & Murnighan, J. K. (2010). Dormant Ties: The Value Of Reconnecting. *Organization Science*, *22*(4), 923–939. https://doi.org/10.1287/orsc.1100.0576
Levy, P. E., & Williams, J. R. (2004). The Social Context of Performance Appraisal: A Review and Framework for the Future. *Journal of Management, 30*(6), 881–905. https://doi.org/10.1016/j.jm.2004.06.005
Lichtenthaler, U., Ernst, H., & Hoegl, M. (2010). Not-Sold-Here: How Attitudes Influence External Knowledge Exploitation. *Organization Science*, *21*(5), 1054–1071. https://doi.org/10.1287/orsc.1090.0499
Linnehan, F., & De Carolis, D. (2005). Strategic frameworks for understanding employer participation in school-to-work programs. *Strategic Management Journal*, *26*(6), 523–539. https://doi.org/10.1002/smj.467
Liu, Q., & van Ryzin, G. J. (2008). Strategic Capacity Rationing to Induce Early Purchases. *Management Science*, *54*(6), 1115–1131. https://doi.org/10.1287/mnsc.1070.0832
Loughry, M. L., & Tosi, H. L. (2008). Performance Implications of Peer Monitoring. *Organization Science*, *19*(6), 876–890. https://doi.org/10.1287/orsc.1080.0356
Luo, Y. (2008). Structuring interorganizational cooperation: the role of economic integration in strategic alliances. *Strategic Management Journal*, *29*(6), 617–637. https://doi.org/10.1002/smj.677
Lutgen-Sandvik, P., Tracy, S. J., & Alberts, J. K. (2007). Burned by Bullying in the American Workplace: Prevalence, Perception, Degree and Impact*. *Journal of Management Studies*, *44*, 837–862. https://doi.org/10.1111/j.1467-6486.2007.00715.x
Lux, S., Crook, T. R., & Woehr, D. J. (2011). Mixing Business With Politics: A Meta-Analysis of the Antecedents and Outcomes of Corporate Political Activity. *Journal of Management, 37*(1), 223–247. https://doi.org/10.1177/0149206310392233
Lyon, F. (2006). Managing Co-operation: Trust and Power in Ghanaian Associations. *Organization Studies, 27*(1), 31–52. https://doi.org/10.1177/0170840605056392

Macher, J. T., & Boerner, C. S. (2006). Experience and scale and scope economies: trade-offs and performance in development. *Strategic Management Journal*, *27*(9), 845–865. https://doi.org/10.1002/smj.540

Madhok, A. (2002). Reassessing the fundamentals and beyond: Ronald Coase, the transaction cost and resource-based theories of the firm and the institutional structure of production. *Strategic Management Journal*, *23*(6), 535–550. https://doi.org/10.1002/smj.247

Madjar, N., & Shalley, C. E. (2008). Multiple Tasks' and Multiple Goals' Effect on Creativity: Forced Incubation or Just a Distraction? *Journal of Management, 34*(4), 786–805. https://doi.org/10.1177/0149206308318611

Makhija, M. V. (2004). The value of restructuring in emerging economies: the case of the Czech Republic. *Strategic Management Journal*, *25*(3), 243–267. https://doi.org/10.1002/smj.378

Mantere, S. (2008). Role Expectations and Middle Manager Strategic Agency. *Journal of Management Studies*, *45*, 294–316. https://doi.org/10.1111/j.1467-6486.2007.00744.x

March, J. G. (2006). Rationality, foolishness, and adaptive intelligence. *Strategic Management Journal*, *27*(3), 201–214. https://doi.org/10.1002/smj.515

Martínez-Sánchez, A., Vela-Jiménez, M.-J., Pérez-Pérez, M., & de-Luis-Carnicer, P. (2011). The Dynamics of Labour Flexibility: Relationships between Employment Type and Innovativeness. *Journal of Management Studies*, *48*, 715–736. https://doi.org/10.1111/j.1467-6486.2010.00935.x

Massey, C., & Wu, G. (2005). Detecting Regime Shifts: The Causes of Under- and Overreaction. *Management Science*, *51*(6), 932–947. https://doi.org/10.1287/mnsc.1050.0386

McFadyen, M. A., Semadeni, M., & Cannella, A. A. (2009). Value of Strong Ties to Disconnected Others: Examining Knowledge Creation in Biomedicine. *Organization Science*, *20*(3), 552–564. https://doi.org/10.1287/orsc.1080.0388

McInerney, P.-B. (2008). Showdown at Kykuit: Field-Configuring Events as Loci for Conventionalizing Accounts. *Journal of Management Studies*, *45*, 1089–1116. https://doi.org/10.1111/j.1467-6486.2008.00784.x

McMullen, J. S., Shepherd, D. A., & Patzelt, H. (2009). Managerial (In)attention to Competitive Threats. *Journal of Management Studies*, *46*, 157–181. https://doi.org/10.1111/j.1467-6486.2008.00799.x

Mehra, A., Dixon, A. L., Brass, D. J., & Robertson, B. (2006). The Social Network Ties of Group Leaders: Implications for Group Performance and Leader Reputation. *Organization Science*, *17*(1), 64–79. https://doi.org/10.1287/orsc.1050.0158

Mero, N. P., Guidice, R. M., & Brownlee, A. L. (2007). Accountability in a Performance Appraisal Context: The Effect of Audience and Form of Accounting on Rater Response and Behavior. *Journal of Management, 33*(2), 223–252. https://doi.org/10.1177/0149206306297633

Meyer, K. E., Mudambi, R., & Narula, R. (2011). Multinational Enterprises and Local Contexts: The Opportunities and Challenges of Multiple Embeddedness. *Journal of Management Studies*, *48*, 235–252. https://doi.org/10.1111/j.1467-6486.2010.00968.x

Misangyi, V. F., Elms, H., Greckhamer, T., & Lepine, J. A. (2006). A new perspective on a fundamental debate: a multilevel approach to industry, corporate, and business unit effects. *Strategic Management Journal*, *27*(6), 571–590. https://doi.org/10.1002/smj.530

Molloy, J. C., Ployhart, R. E., & Wright, P. M. (2011). The Myth of "the" Micro-Macro Divide: Bridging System-Level and Disciplinary Divides. *Journal of Management, 37*(2), 581–609. https://doi.org/10.1177/0149206310365000

Moran, P. (2005). Structural vs. relational embeddedness: social capital and managerial performance. *Strategic Management Journal*, *26*(12), 1129–1151. https://doi.org/10.1002/smj.486
Morgan, R. E., & Berthon, P. (2008). Market Orientation, Generative Learning, Innovation Strategy and Business Performance Inter-Relationships in Bioscience Firms. *Journal of Management Studies*, *45*, 1329–1353. https://doi.org/10.1111/j.1467-6486.2008.00778.x
Mudambi, S. M., & Tallman, S. (2010). Make, Buy or Ally? Theoretical Perspectives on Knowledge Process Outsourcing through Alliances. *Journal of Management Studies*, *47*, 1434–1456. https://doi.org/10.1111/j.1467-6486.2010.00944.x
Mueller, B. A., Titus, V. K., Covin, J. G., & Slevin, D. P. (2012). Pioneering Orientation and Firm Growth: Knowing When and to What Degree Pioneering Makes Sense . *Journal of Management, 38*(5), 1517–1549. https://doi.org/10.1177/0149206310380249
Munir, K. A., & Phillips, N. (2005). The Birth of the "Kodak Moment": Institutional Entrepreneurship and the Adoption of New Technologies. *Organization Studies, 26*(11), 1665–1687. https://doi.org/10.1177/0170840605056395
Munyon, T. P., Hochwarter, W. A., Perrewé, P. L., & Ferris, G. R. (2010). Optimism and the Nonlinear Citizenship Behavior—Job Satisfaction Relationship in Three Studies. *Journal of Management, 36*(6), 1505–1528. https://doi.org/10.1177/0149206309350085
Nachum, L. (2003). Liability of foreignness in global competition? Financial service affiliates in the city of London. *Strategic Management Journal*, *24*(12), 1187–1208. https://doi.org/10.1002/smj.347
Nasrallah, W., Levitt, R., & Glynn, P. (2003). Interaction Value Analysis: When Structured Communication Benefits Organizations. *Organization Science*, *14*(5), 541–557. https://doi.org/10.1287/orsc.14.5.541.16764
Nemeth, C., O'Connor, M., Klock, P. A., & Cook, R. (2006). Discovering Healthcare Cognition: The Use of Cognitive Artifacts to Reveal Cognitive Work. *Organization Studies, 27*(7), 1011–1035. https://doi.org/10.1177/0170840606065708
Nicolini, D. (2010). Practice as the Site of Knowing: Insights from the Field of Telemedicine. *Organization Science*, *22*(3), 602–620. https://doi.org/10.1287/orsc.1100.0556
Okhuysen, G. A., & Eisenhardt, K. M. (2002). Integrating Knowledge in Groups: How Formal Interventions Enable Flexibility. *Organization Science*, *13*(4), 370–386. https://doi.org/10.1287/orsc.13.4.370.2947
O'Mahoney, J. (2007). The Diffusion of Management Innovations: The Possibilities and Limitations of Memetics. *Journal of Management Studies*, *44*, 1324–1348. https://doi.org/10.1111/j.1467-6486.2007.00734.x
Orsato, R. J., den Hond, F., & Clegg, S. R. (2002). The Political Ecology of Automobile Recycling in Europe. *Organization Studies, 23*(4), 639–665. https://doi.org/10.1177/0170840602234006
Özen, Ş., & Akkemik, K. A. (2012). Does Illegitimate Corporate Behaviour Follow the Forms of Polity? The Turkish Experience. *Journal of Management Studies*, *49*(3), 515–537. https://doi.org/10.1111/j.1467-6486.2011.01007.x
Ozman, M. (2010). The Knowledge Base of Products: Implications for Organizational Structures. *Organization Studies, 31*(8), 1129–1154. https://doi.org/10.1177/0170840610376148
Pablo, Z., & Hardy, C. (2009). Merging, Masquerading and Morphing: Metaphors and the World Wide Web. *Organization Studies, 30*(8), 821–843. https://doi.org/10.1177/0170840609334951
Pacheco, D. F., York, J. G., Dean, T. J., & Sarasvathy, S. D. (2010). The Coevolution of Institutional Entrepreneurship: A Tale of Two Theories. *Journal of Management, 36*(4), 974–1010. https://doi.org/10.1177/0149206309360280

Padilla, M., & Van Roy, B. (2012). Intermediated Blind Portfolio Auctions. *Management Science*, *58*(9), 1747–1760. https://doi.org/10.1287/mnsc.1120.1521
Palmer, I., & Dunford, R. (2002). Managing Discursive Tension: the Co–Existence of Individualist and Collaborative Discourses in Flight Centre. *Journal of Management Studies*, *39*, 1045–1069. https://doi.org/10.1111/1467-6486.00323
Pangarkar, N. (2009). Do Firms Learn from Alliance Terminations? An Empirical Examination. *Journal of Management Studies*, *46*, 982–1004. https://doi.org/10.1111/j.1467-6486.2009.00844.x
Parmigiani, A., & Rivera-Santos, M. (2011). Clearing a Path through the Forest: A Meta-Review of Interorganizational Relationships. *Journal of Management, 37*(4), 1108–1136. https://doi.org/10.1177/0149206311407507
Pelzer, P. (2005). Essai: Contempt and Organization: Present in Practice — Ignored by Research?. *Organization Studies, 26*(8), 1217–1227. https://doi.org/10.1177/0170840605053537
Pennings, J. M. E., & Smidts, A. (2003). The Shape of Utility Functions and Organizational Behavior. *Management Science*, *49*(9), 1251–1263. https://doi.org/10.1287/mnsc.49.9.1251.16566
Pérez-Nordtvedt, L., Payne, G. T., Short, J. C., & Kedia, B. L. (2008). An Entrainment-Based Model of Temporal Organizational Fit, Misfit, and Performance. *Organization Science*, *19*(5), 785–801. https://doi.org/10.1287/orsc.1070.0330
Perrone, V., Zaheer, A., & McEvily, B. (2003). Free to Be Trusted? Organizational Constraints on Trust in Boundary Spanners. *Organization Science*, *14*(4), 422–439. https://doi.org/10.1287/orsc.14.4.422.17487
Philippe, D., & Durand, R. (2011). The impact of norm-conforming behaviors on firm reputation. *Strategic Management Journal*, *32*(9), 969–993. https://doi.org/10.1002/smj.919
Piderit, S. K., & Ashford, S. J. (2003). Breaking Silence: Tactical Choices Women Managers Make in Speaking Up About Gender-Equity Issues*. *Journal of Management Studies*, *40*, 1477–1502. https://doi.org/10.1111/1467-6486.00388
Pierce, L. (2012). Organizational Structure and the Limits of Knowledge Sharing: Incentive Conflict and Agency in Car Leasing. *Management Science*, *58*(6), 1106–1121. https://doi.org/10.1287/mnsc.1110.1472
Powell, T. C. (2003). Strategy without ontology. *Strategic Management Journal*, *24*(3), 285–291. https://doi.org/10.1002/smj.284
Quick, J. C., Macik-Frey, M., & Cooper, C. L. (2007). Managerial Dimensions of Organizational Health: The Healthy Leader at Work*. *Journal of Management Studies*, *44*, 189–205. https://doi.org/10.1111/j.1467-6486.2007.00684.x
Quigley, N. R., Tesluk, P. E., Locke, E. A., & Bartol, K. M. (2007). A Multilevel Investigation of the Motivational Mechanisms Underlying Knowledge Sharing and Performance. *Organization Science*, *18*(1), 71–88. https://doi.org/10.1287/orsc.1060.0223
Rafferty, A. E., & Restubog, S. L. D. (2010). The Impact of Change Process and Context on Change Reactions and Turnover during a Merger. *Journal of Management, 36*(5), 1309–1338. https://doi.org/10.1177/0149206309341480
Raisch, S., & Birkinshaw, J. (2008). Organizational Ambidexterity: Antecedents, Outcomes, and Moderators. *Journal of Management, 34*(3), 375–409. https://doi.org/10.1177/0149206308316058
Rawley, E., & Simcoe, T. S. (2010). Diversification, Diseconomies of Scope, and Vertical Contracting: Evidence from the Taxicab Industry. *Management Science*, *56*(9), 1534–1550. https://doi.org/10.1287/mnsc.1100.1207

Reed, M. I. (2012). Masters of the Universe: Power and Elites in Organization Studies. *Organization Studies, 33*(2), 203–221. https://doi.org/10.1177/0170840611430590
Remington Abramson, N., & Senyshyn, Y. (2010). Effective Punishment through Forgiveness: Rediscovering Kierkegaard's Knight of Faith in the Abraham Story. *Organization Studies, 31*(5), 555–581. https://doi.org/10.1177/0170840610372202
Robinson, W. T., & Chiang, J. (2002). Product development strategies for established market pioneers, early followers, and late entrants. *Strategic Management Journal*, *23*(9), 855–866. https://doi.org/10.1002/smj.257
Rosenkopf, L., & Almeida, P. (2003). Overcoming Local Search through Alliances and Mobility. *Management Science*, *49*(6), 751–766. https://doi.org/10.1287/mnsc.49.6.751.16026
Rupp, D. E., Silke McCance, A., Spencer, S., & Sonntag, K. (2008). Customer (In)Justice and Emotional Labor: The Role of Perspective Taking, Anger, and Emotional Regulation†. *Journal of Management, 34*(5), 903–924. https://doi.org/10.1177/0149206307309261
Salaman, G., & Storey, J. (2002). Managers' Theories about the Process of Innovation. *Journal of Management Studies*, *39*, 147–165. https://doi.org/10.1111/1467-6486.00286
Sanchez-Burks, J., & Huy, Q. N. (2009). Emotional Aperture and Strategic Change: The Accurate Recognition of Collective Emotions. *Organization Science*, *20*(1), 22–34. https://doi.org/10.1287/orsc.1070.0347
Sandholtz, K. W. (2012). Making Standards Stick: A Theory of Coupled vs. Decoupled Compliance. *Organization Studies, 33*(5–6), 655–679. https://doi.org/10.1177/0170840612443623
Santaló, J., & Kock, C. J. (2009). Division Director versus CEO Compensation: New Insights into the Determinants of Executive Pay. *Journal of Management, 35*(4), 1047–1077. https://doi.org/10.1177/0149206308329965
Sarasvathy, S. D., & Dew, N. (2008). Is Effectuation Lachmannian? A Response to Chiles, Bluedorn, and Gupta (2007). *Organization Studies, 29*(2), 239–245. https://doi.org/10.1177/0170840607088153
Sarathy, R., Muralidhar, K., & Parsa, R. (2002). Perturbing Nonnormal Confidential Attributes: The Copula Approach. *Management Science*, *48*(12), 1613–1627. https://doi.org/10.1287/mnsc.48.12.1613.439
Scarbrough, H., Swan, J., Laurent, S., Bresnen, M., Edelman, L., & Newell, S. (2004). Project-Based Learning and the Role of Learning Boundaries. *Organization Studies, 25*(9), 1579–1600. https://doi.org/10.1177/0170840604048001
Schilling, M. A. (2009). Understanding the alliance data. *Strategic Management Journal*, *30*(3), 233–260. https://doi.org/10.1002/smj.731
Schreyögg, G., & Sydow, J. (2010). CROSSROADS—Organizing for Fluidity? Dilemmas of New Organizational Forms. *Organization Science*, *21*(6), 1251–1262. https://doi.org/10.1287/orsc.1100.0561
Schulze, W. S., & Gedajlovic, E. (2010). Whither Family Business? *Journal of Management Studies*, *47*, 191–204. https://doi.org/10.1111/j.1467-6486.2009.00887.x
Shamir, B., & Lapidot, Y. (2003). Trust in Organizational Superiors: Systemic and Collective Considerations. *Organization Studies, 24*(3), 463–491. https://doi.org/10.1177/0170840603024003912
Shi, W. (Stone), Sun, S. L., & Peng, M. W. (2012). Sub-National Institutional Contingencies, Network Positions, and IJV Partner Selection. *Journal of Management Studies*, *49*(7), 1221–1245. https://doi.org/10.1111/j.1467-6486.2012.01058.x
Sidhu, J. S., Volberda, H. W., & Commandeur, H. R. (2004). Exploring Exploration Orientation and its Determinants: Some Empirical Evidence*. *Journal of Management Studies*, *41*, 913–932. https://doi.org/10.1111/j.1467-6486.2004.00460.x

Siegel, J. I., & Larson, B. Z. (2009). Labor Market Institutions and Global Strategic Adaptation: Evidence from Lincoln Electric. *Management Science*, *55*(9), 1527–1546. https://doi.org/10.1287/mnsc.1090.1028

Slangen, A. H. L. (2011). A Communication-Based Theory of the Choice between Greenfield and Acquisition Entry. *Journal of Management Studies*, *48*, 1699–1726. https://doi.org/10.1111/j.1467-6486.2011.01013.x

Snell, R., & Tseng, C. S. (2002). Moral Atmosphere and Moral Influence under China's Network Capitalism. *Organization Studies, 23*(3), 449–478. https://doi.org/10.1177/0170840602233006

Somaya, D. (2012). Patent Strategy and Management: An Integrative Review and Research Agenda. *Journal of Management, 38*(4), 1084–1114. https://doi.org/10.1177/0149206312444447

Song, M., Droge, C., Hanvanich, S., & Calantone, R. (2005). Marketing and technology resource complementarity: an analysis of their interaction effect in two environmental contexts. *Strategic Management Journal*, *26*(3), 259–276. https://doi.org/10.1002/smj.450

Spencer, J., & Gomez, C. (2011). MNEs and corruption: the impact of national institutions and subsidiary strategy. *Strategic Management Journal*, *32*(3), 280–300. https://doi.org/10.1002/smj.874

Spicer, A., & Sewell, G. (2010). From National Service to Global Player: Transforming the Organizational Logic of a Public Broadcaster. *Journal of Management Studies*, *47*, 913–943. https://doi.org/10.1111/j.1467-6486.2009.00915.x

Starbuck, W. H. (2004). Vita Contemplativa: Why I Stopped Trying to Understand the Real World. *Organization Studies, 25*(7), 1233–1254. https://doi.org/10.1177/0170840604046361

Strang, D., & Jung, D.-I. (2009). Participatory Improvement at a Global Bank: The Diffusion of Quality Teams and the Demise of a Six Sigma Initiative. *Organization Studies, 30*(1), 31–53. https://doi.org/10.1177/0170840608100517

Sullivan, D. M., & Marvel, M. R. (2011). Knowledge Acquisition, Network Reliance, and Early-Stage Technology Venture Outcomes. *Journal of Management Studies*, *48*, 1169–1193. https://doi.org/10.1111/j.1467-6486.2010.00998.x

Tallman, S., & Phene, A. (2007). Leveraging Knowledge across Geographic Boundaries. *Organization Science*, *18*(2), 252–260. https://doi.org/10.1287/orsc.1060.0227

Taylor, J. R. (2011). Organization as an (Imbricated) Configuring of Transactions. *Organization Studies, 32*(9), 1273–1294. https://doi.org/10.1177/0170840611411396

Thomas, P. (2003). The Recontextualization of Management: A Discourse-based Approach to Analysing the Development of Management Thinking*. *Journal of Management Studies*, *40*, 775–801. https://doi.org/10.1111/1467-6486.00360

Thompson, P. (2007). Adler's Theory of the Capitalist Labour Process: A Pale(o) Imitation. *Organization Studies, 28*(9), 1359–1368. https://doi.org/10.1177/0170840607080745

Thonemann, U. W., Brown, A. O., & Hausman, W. H. (2002). Easy Quantification of Improved Spare Parts Inventory Policies. *Management Science*, *48*(9), 1213–1225. https://doi.org/10.1287/mnsc.48.9.1213.173

Tiwana, A., & Keil, M. (2007). Does peripheral knowledge complement control? An empirical test in technology outsourcing alliances. *Strategic Management Journal*, *28*(6), 623–634. https://doi.org/10.1002/smj.623

Tosi, H. L., Brownlee, A. L., Silva, P., & Katz, J. P. (2003). An Empirical Exploration of Decision-making under Agency Controls and Stewardship Structure. *Journal of Management Studies*, *40*(8), 2053–2071. https://doi.org/10.1046/j.1467-6486.2003.00411.x

Tourish, D., & Robson, P. (2006). Sensemaking and the Distortion of Critical Upward Communication in Organizations. *Journal of Management Studies*, *43*, 711–730. https://doi.org/10.1111/j.1467-6486.2006.00608.x

Treviño, L. K., Weaver, G. R., & Reynolds, S. J. (2006). Behavioral Ethics in Organizations: A Review. *Journal of Management, 32*(6), 951–990. https://doi.org/10.1177/0149206306294258

Tsai, W.-C., Chen, C.-C., & Chiu, S.-F. (2005). Exploring Boundaries of the Effects of Applicant Impression Management Tactics in Job Interviews. *Journal of Management, 31*(1), 108–125. https://doi.org/10.1177/0149206304271384

Tsai, W. (2002). Social Structure of "Coopetition" Within a Multiunit Organization: Coordination, Competition, and Intraorganizational Knowledge Sharing. *Organization Science*, *13*(2), 179–190. https://doi.org/10.1287/orsc.13.2.179.536

Tsui-Auch, L. S. (2005). Unpacking Regional Ethnicity and the Strength of Ties in Shaping Ethnic Entrepreneurship. *Organization Studies, 26*(8), 1189–1216. https://doi.org/10.1177/0170840605055343

Tuggle, C. S., Sirmon, D. G., Reutzel, C. R., & Bierman, L. (2010). Commanding board of director attention: investigating how organizational performance and CEO duality affect board members' attention to monitoring. *Strategic Management Journal*, *31*(9), 946–968. https://doi.org/10.1002/smj.847

Tyler, M., & Cohen, L. (2010). Spaces that Matter: Gender Performativity and Organizational Space. *Organization Studies, 31*(2), 175–198. https://doi.org/10.1177/0170840609357381

Uhlenbruck, K., Meyer, K. E., & Hitt, M. A. (2003). Organizational Transformation in Transition Economies: Resource-based and Organizational Learning Perspectives. *Journal of Management Studies*, *40*, 257–282. https://doi.org/10.1111/1467-6486.00340

Uhlenbruck, K., Rodriguez, P., Doh, J., & Eden, L. (2006). The Impact of Corruption on Entry Strategy: Evidence from Telecommunication Projects in Emerging Economies. *Organization Science*, *17*(3), 402–414. https://doi.org/10.1287/orsc.1060.0186

Ullrich, J., Wieseke, J., & Dick, R. Van. (2005). Continuity and Change in Mergers and Acquisitions: A Social Identity Case Study of a German Industrial Merger*. *Journal of Management Studies*, *42*, 1549–1569. https://doi.org/10.1111/j.1467-6486.2005.00556.x

Vaara, E. (2002). On the Discursive Construction of Success/Failure in Narratives of Post-Merger Integration. *Organization Studies, 23*(2), 211–248. https://doi.org/10.1177/0170840602232003

van de Kuilen, G., & Wakker, P. P. (2011). The Midweight Method to Measure Attitudes toward Risk and Ambiguity. *Management Science*, *57*(3), 582–598. https://doi.org/10.1287/mnsc.1100.1282

von Nordenflycht, A. (2011). Firm Size and Industry Structure under Human Capital Intensity: Insights from the Evolution of the Global Advertising Industry. *Organization Science*, *22*(1), 141–157. https://doi.org/10.1287/orsc.1090.0518

Wijnberg, N. M. (2004). Innovation and Organization: Value and Competition in Selection Systems. *Organization Studies, 25*(8), 1413–1433. https://doi.org/10.1177/0170840604046350

Williams, C. (2007). Transfer in context: replication and adaptation in knowledge transfer relationships. *Strategic Management Journal*, *28*(9), 867–889. https://doi.org/10.1002/smj.614

Winter, S. G., Cattani, G., & Dorsch, A. (2007). The Value of Moderate Obsession: Insights from a New Model of Organizational Search. *Organization Science*, *18*(3), 403–419. https://doi.org/10.1287/orsc.1070.0273

Withers, M. C., Corley, K. G., & Hillman, A. J. (2012). Stay or Leave: Director Identities and Voluntary Exit from the Board during Organizational Crisis. *Organization Science*, *23*(3), 835–850. https://doi.org/10.1287/orsc.1110.0660

Wolfe, R. A., & Putler, D. S. (2002). How Tight Are the Ties that Bind Stakeholder Groups? *Organization Science*, *13*(1), 64–80. https://doi.org/10.1287/orsc.13.1.64.544

Wood, M. (2005). The Fallacy of Misplaced Leadership*. *Journal of Management Studies*, *42*, 1101–1121. https://doi.org/10.1111/j.1467-6486.2005.00535.x

Wooldridge, B., Schmid, T., & Floyd, S. W. (2008). The Middle Management Perspective on Strategy Process: Contributions, Synthesis, and Future Research. *Journal of Management, 34*(6), 1190–1221. https://doi.org/10.1177/0149206308324326

Xia, J., Boal, K., & Delios, A. (2009). When experience meets national institutional environmental change: foreign entry attempts of U.S. firms in the Central and Eastern European region. *Strategic Management Journal*, *30*(12), 1286–1309. https://doi.org/10.1002/smj.794

Yang, M., & Hyland, M. (2006). Who Do Firms Imitate? A Multilevel Approach to Examining Sources of Imitation in the Choice of Mergers and Acquisitions. *Journal of Management, 32*(3), 381–399. https://doi.org/10.1177/0149206305280790

Yoo, B., Choudhary, V., & Mukhopadhyay, T. (2007). Electronic B2B Marketplaces with Different Ownership Structures. *Management Science*, *53*(6), 952–961. https://doi.org/10.1287/mnsc.1060.0685

Zhang, G. (2006). Market Valuation and Employee Stock Options. *Management Science*, *52*(9), 1377–1393. https://doi.org/10.1287/mnsc.1060.0539

Zhang, J., & Baden-Fuller, C. (2010). The Influence of Technological Knowledge Base and Organizational Structure on Technology Collaboration. *Journal of Management Studies*, *47*, 679–704. https://doi.org/10.1111/j.1467-6486.2009.00885.x

Zollo, M., & Singh, H. (2004). Deliberate learning in corporate acquisitions: post-acquisition strategies and integration capability in U.S. bank mergers. *Strategic Management Journal*, *25*(13), 1233–1256. https://doi.org/10.1002/smj.426

Zubac, A., Hubbard, G., & Johnson, L. (2012). Extending Resource-Based Logic: Applying the Resource-Investment Concept to the Firm from a Payments Perspective. *Journal of Management, 38*(6), 1867–1891. https://doi.org/10.1177/0149206310383907

Microbiology

Abe, T., Sakaki, K., Fujihara, A., Ujiie, H., Ushida, C., Himeno, H., Muto, A. (2008). tmRNA-dependent trans-translation is required for sporulation in Bacillus subtilis. *Molecular Microbiology*, *69*(6), 1491–1498. https://doi.org/10.1111/j.1365-2958.2008.06381.x

Aguirre, J., Ríos-Momberg, M., Hewitt, D., & Hansberg, W. (2005). Reactive oxygen species and development in microbial eukaryotes. *Trends in Microbiology*, *13*(3), 111–8. https://doi.org/10.1016/j.tim.2005.01.007

Ahuja, U., Kjelgaard, P., Schulz, B. L., Thöny-Meyer, L., & Hederstedt, L. (2009). Haem-delivery proteins in cytochrome c maturation System II. *Molecular Microbiology*, *73*(6), 1058–1071. https://doi.org/10.1111/j.1365-2958.2009.06833.x

Alix, E., & Blanc-Potard, A.-B. (2007). MgtC: a key player in intramacrophage survival. *Trends in Microbiology*, *15*(6), 252–256. https://doi.org/10.1016/j.tim.2007.03.007

Al-Khodor, S., Price, C. T., Kalia, A., & Abu Kwaik, Y. (2010). Functional diversity of ankyrin repeats in microbial proteins. *Trends in Microbiology*, *18*(3), 132–139. https://doi.org/10.1016/j.tim.2009.11.004

Allen, M. J., & Wilson, W. H. (2008). Aquatic virus diversity accessed through omic techniques: A route map to function. *Current Opinion in Microbiology*, *11*(3), 226–232. https://doi.org/10.1016/j.mib.2008.05.004

An, Q., Hückelhoven, R., Kogel, K.-H., & Van Bel, A. J. E. (2006). Multivesicular bodies participate in a cell wall-associated defence response in barley leaves attacked by the pathogenic powdery mildew fungus. *Cellular Microbiology*, *8*(6), 1009–1019. https://doi.org/10.1111/j.1462-5822.2006.00683.x

Anderson, I. C., & Cairney, J. W. G. (2007). Ectomycorrhizal fungi: exploring the mycelial frontier. *FEMS Microbiology Reviews*, *31*(4), 388–406. https://doi.org/10.1111/j.1574-6976.2007.00073.x

Anderson, J. P., Daifuku, R., & Loeb, L. A. (2004). Viral Error Catastrophe by Mutagenic Nucleosides. *Annual Review of Microbiology*, *58*(1), 183–205. https://doi.org/10.1146/annurev.micro.58.030603.123649

Anderson, K. L., & Lebepe-Mazur, S. (2003). Comparison of rapid methods for the extraction of bacterial DNA from colonic and caecal lumen contents of the pig. *Journal of Applied Microbiology*, *94*(6), 988–993. https://doi.org/10.1046/j.1365-2672.2003.01917.x

Andersson, J. O. (2009). Gene Transfer and Diversification of Microbial Eukaryotes. *Annual Review of Microbiology*, *63*(1), 177–193. https://doi.org/10.1146/annurev.micro.091208.073203

Angert, E. R. (2012). DNA Replication and Genomic Architecture of Very Large Bacteria. *Annual Review of Microbiology*, *66*(1), 197–212. https://doi.org/10.1146/annurev-micro-090110-102827

Aravind, L., Anantharaman, V., Balaji, S., Babu, M. M., & Iyer, L. M. (2005). The many faces of the helix-turn-helix domain: Transcription regulation and beyond⋆. *FEMS Microbiology Reviews*, *29*(2), 231–262. https://doi.org/10.1016/j.fmrre.2004.12.008

Aro, N., Pakula, T., & Penttilä, M. (2005). Transcriptional regulation of plant cell wall degradation by filamentous fungi. *FEMS Microbiology Reviews*, *29*(4), 719–739. https://doi.org/10.1016/j.femsre.2004.11.006

Atlas, R. M. (2002). Bioterrorism: From Threat to Reality. *Annual Review of Microbiology*, *56*(1), 167–185. https://doi.org/10.1146/annurev.micro.56.012302.160616

Atmakuri, K., Cascales, E., & Christie, P. J. (2004). Energetic components VirD4, VirB11 and VirB4 mediate early DNA transfer reactions required for bacterial type IV secretion. *Molecular Microbiology*, *54*(5), 1199–1211. https://doi.org/10.1111/j.1365-2958.2004.04345.x

Avet-Rochex, A., Bergeret, E., Attree, I., Meister, M., & Fauvarque, M.-O. (2005). Suppression of Drosophila cellular immunity by directed expression of the ExoS toxin GAP domain of Pseudomonas aeruginosa. *Cellular Microbiology*, *7*(6), 799–810. https://doi.org/10.1111/j.1462-5822.2005.00512.x

Ayres, J. S., & Schneider, D. S. (2006). Genomic dissection of microbial pathogenesis in cultured Drosophila cells. *Trends in Microbiology*, *14*(3), 101–4. https://doi.org/10.1016/j.tim.2006.01.008

Baines, J. D. (2011). Herpes simplex virus capsid assembly and DNA packaging: a present and future antiviral drug target. *Trends in Microbiology*, *19*(12), 606–13. https://doi.org/10.1016/j.tim.2011.09.001

Bakthavatsalam, D., Meijer, H. J. G., Noegel, A. A., & Govers, F. (2006). Novel phosphatidylinositol phosphate kinases with a G-protein coupled receptor signature are shared by Dictyostelium and Phytophthora. *Trends in Microbiology*, *14*(9), 378–82. https://doi.org/10.1016/j.tim.2006.07.006

Bapteste, E., Boucher, Y., Leigh, J., & Doolittle, W. (2004). Phylogenetic reconstruction and lateral gene transfer. *Trends in Microbiology*, *12*(9), 406–411. https://doi.org/10.1016/j.tim.2004.07.002

Belanger, P. H., Johnston, D. A., Fratti, R. A., Zhang, M., & G. Filler, S. (2002). Endocytosis of Candida albicans by vascular endothelial cells is associated with tyrosine phosphorylation of specific host cell proteins. *Cellular Microbiology*, *4*(12), 805–812. https://doi.org/10.1046/j.1462-5822.2002.00232.x

Benga, L., Goethe, R., Rohde, M., & Valentin-Weigand, P. (2004). Non-encapsulated strains reveal novel insights in invasion and survival of Streptococcus suis in epithelial cells. *Cellular Microbiology*, *6*(9), 867–881. https://doi.org/10.1111/j.1462-5822.2004.00409.x

Berkey, C. D., Blow, N., & Watnick, P. I. (2009). *Genetic analysis of Drosophila melanogaster susceptibility to intestinal Vibrio cholerae infection*. *Cellular Microbiology* (Vol. 11). Blackwell Publishing Ltd. https://doi.org/10.1111/j.1462-5822.2008.01267.x

Berman, J. (2006). Morphogenesis and cell cycle progression in Candida albicans. *Current Opinion in Microbiology*, *9*(6), 595–601. https://doi.org/10.1016/j.mib.2006.10.007

Binker, M. G., Cosen-Binker, L. I., Terebiznik, M. R., Mallo, G. V, McCaw, S. E., Eskelinen, E.-L., … Gray-Owen, S. D. (2007). *Arrested maturation of Neisseria-containing phagosomes in the absence of the lysosome-associated membrane proteins, LAMP-1 and LAMP-2*. *Cellular microbiology* (Vol. 9). https://doi.org/10.1111/j.1462-5822.2007.00946.x

Block, A., Guo, M., Li, G., Elowsky, C., Clemente, T. E., & Alfano, J. R. (2010). *The Pseudomonas syringae type III effector HopG1 targets mitochondria, alters plant development and suppresses plant innate immunity*. *Cellular Microbiology* (Vol. 12). Blackwell Publishing Ltd. https://doi.org/10.1111/j.1462-5822.2009.01396.x

Bodrossy, L., & Sessitsch, A. (2004). Oligonucleotide microarrays in microbial diagnostics. *Current Opinion in Microbiology*, *7*(3), 245–254. https://doi.org/10.1016/j.mib.2004.04.005

Boss, I. W., Plaisance, K. B., & Renne, R. (2009). Role of virus-encoded microRNAs in herpesvirus biology. *Trends in Microbiology*, *17*(12), 544–553. https://doi.org/10.1016/j.tim.2009.09.002

Botts, M. R., & Hull, C. M. (2010). Dueling in the lung: how Cryptococcus spores race the host for survival. *Current Opinion in Microbiology*, *13*(4), 437–42. https://doi.org/10.1016/j.mib.2010.05.003

Boven, M. (2004). Unveiling the burden of pertussis. *Trends in Microbiology*, *12*(3), 116–119. https://doi.org/10.1016/j.tim.2004.01.002

Brand, A., & Gow, N. A. R. (2009). Mechanisms of hypha orientation of fungi. *Current Opinion in Microbiology*, *12*(4), 350–7. https://doi.org/10.1016/j.mib.2009.05.007

Brantl, S. (2007). Regulatory mechanisms employed by cis-encoded antisense RNAs. *Current Opinion in Microbiology*, *10*(2), 102–9. https://doi.org/10.1016/j.mib.2007.03.012

Breitbach, K., Rottner, K., Klocke, S., Rohde, M., Jenzora, A., Wehland, J., & Steinmetz, I. (2003). Actin-based motility of Burkholderia pseudomallei involves the Arp 2/3 complex, but not N-WASP and Ena/VASP proteins. *Cellular Microbiology*, *5*(6), 385–393. https://doi.org/10.1046/j.1462-5822.2003.00277.x

Bruhn, D. F., Mozeleski, B., Falkin, L., & Klingbeil, M. M. (2010). Mitochondrial DNA polymerase POLIB is essential for minicircle DNA replication in African trypanosomes. *Molecular Microbiology*, *75*(6), 1414–1425. https://doi.org/10.1111/j.1365-2958.2010.07061.x

Buchanan, G., Leeuw, E. de, Stanley, N. R., Wexler, M., Berks, B. C., Sargent, F., & Palmer, T. (2002). Functional complexity of the twin-arginine translocase TatC component revealed by site-directed mutagenesis. *Molecular Microbiology*, *43*(6), 1457–1470. https://doi.org/10.1046/j.1365-2958.2002.02853.x

Bush, K. (2010). Alarming β-lactamase-mediated resistance in multidrug-resistant Enterobacteriaceae. *Current Opinion in Microbiology*, *13*(5), 558–64. https://doi.org/10.1016/j.mib.2010.09.006

Cannon, R. D., Lamping, E., Holmes, A. R., Niimi, K., Baret, P. V, Keniya, M. V, … Monk, B. C. (2009). Efflux-mediated antifungal drug resistance. *Clinical Microbiology Reviews*, *22*(2), 291–321. https://doi.org/10.1128/CMR.00051-08

Carpousis, A. J. (2007). The RNA Degradosome of Escherichia coli: An mRNA-Degrading Machine Assembled on RNase E. *Annual Review of Microbiology*, *61*(1), 71–87. https://doi.org/10.1146/annurev.micro.61.080706.093440

Castellano, P., Aristoy, M. C., Sentandreu, M. A., Vignolo, G., & Toldrá, F. (2012). Lactobacillus sakei CRL1862 improves safety and protein hydrolysis in meat systems. *Journal of Applied Microbiology*, *113*(6), 1407–1416. https://doi.org/10.1111/jam.12005

Castilla-Llorente, V., Salas, M., & Meijer, W. J. J. (2008). kinC/D-mediated heterogeneous expression of spo0A during logarithmical growth in Bacillus subtilis is responsible for partial suppression of φ29 development. *Molecular Microbiology*, *68*(6), 1406–1417. https://doi.org/10.1111/j.1365-2958.2008.06234.x

Chaconas, G., & Kobryn, K. (2010). Structure, Function, and Evolution of Linear Replicons in Borrelia. *Annual Review of Microbiology*, *64*(1), 185–202. https://doi.org/10.1146/annurev.micro.112408.134037

Chang, S.-S., Zhang, Z., & Liu, Y. (2012). RNA Interference Pathways in Fungi: Mechanisms and Functions. *Annual Review of Microbiology*, *66*(1), 305–323. https://doi.org/10.1146/annurev-micro-092611-150138

Chasteen, T. G., Fuentes, D. E., Tantaleán, J. C., & Vásquez, C. C. (2009). Tellurite: history, oxidative stress, and molecular mechanisms of resistance. *FEMS Microbiology Reviews*, *33*(4), 820–832. Retrieved from http://femsre.oxfordjournals.org/content/33/4/820.abstract

Cheeran, M. C.-J., Lokensgard, J. R., & Schleiss, M. R. (2009). Neuropathogenesis of congenital cytomegalovirus infection: disease mechanisms and prospects for intervention. *Clinical Microbiology Reviews*, *22*(1), 99–126. https://doi.org/10.1128/CMR.00023-08

Chen, H.-M., & Neiman, A. M. (2011). A conserved regulatory role for antisense RNA in meiotic gene expression in yeast. *Current Opinion in Microbiology*, *14*(6), 655–9. https://doi.org/10.1016/j.mib.2011.09.010

Chew, A. G. M., & Bryant, D. A. (2007). Chlorophyll Biosynthesis in Bacteria: The Origins of Structural and Functional Diversity. *Annual Review of Microbiology*, *61*(1), 113–129. https://doi.org/10.1146/annurev.micro.61.080706.093242

Claudianos, C., Dessens, J. T., Trueman, H. E., Arai, M., Mendoza, J., Butcher, G. A., … Sinden, R. E. (2002). A malaria scavenger receptor-like protein essential for parasite development. *Molecular Microbiology*, *45*(6), 1473–1484. https://doi.org/10.1046/j.1365-2958.2002.03118.x

Cohan, F. M. (2002). What are Bacterial Species? *Annual Review of Microbiology*, *56*(1), 457–487. https://doi.org/10.1146/annurev.micro.56.012302.160634

Cooper, T. G. (2002). Transmitting the signal of excess nitrogen in Saccharomyces cerevisiae from the Tor proteins to the GATA factors: connecting the dots. *FEMS Microbiology Reviews*, *26*(3), 223–238. https://doi.org/10.1111/j.1574-6976.2002.tb00612.x

Cortezzo, D. E., & Setlow, P. (2005). Analysis of factors that influence the sensitivity of spores of Bacillus subtilis to DNA damaging chemicals. *Journal of Applied Microbiology*, *98*(3), 606–617. https://doi.org/10.1111/j.1365-2672.2004.02495.x

Courtois, J. (2009). Oligosaccharides from land plants and algae: production and applications in therapeutics and biotechnology. *Current Opinion in Microbiology*, *12*(3), 261–273. https://doi.org/10.1016/j.mib.2009.04.007

Coustou, V., Plazolles, N., Guegan, F., & Baltz, T. (2012). *Sialidases play a key role in infection and anaemia in Trypanosoma congolense animal trypanosomiasis*. *Cellular Microbiology* (Vol. 14). https://doi.org/10.1111/j.1462-5822.2011.01730.x

Craun, G. F., Brunkard, J. M., Yoder, J. S., Roberts, V. A., Carpenter, J., Wade, T., … Roy, S. L. (2010). Causes of Outbreaks Associated with Drinking Water in the United States from 1971 to 2006. *Clinical Microbiology Reviews*, *23*(3), 507–528. https://doi.org/10.1128/CMR.00077-09

Criss, A. K., & Seifert, H. S. (2006). Gonococci exit apically and basally from polarized epithelial cells and exhibit dynamic changes in type IV pili. *Cellular Microbiology*, *8*(9), 1430–1443. https://doi.org/10.1111/j.1462-5822.2006.00722.x

De Buck, E., Lammertyn, E., & Anné, J. (2008). The importance of the twin-arginine translocation pathway for bacterial virulence. *Trends in Microbiology*, *16*(9), 442–453. https://doi.org/10.1016/j.tim.2008.06.004

de Gouw, D., Diavatopoulos, D. A., Bootsma, H. J., Hermans, P. W. M., & Mooi, F. R. (2011). Pertussis: a matter of immune modulation. *FEMS Microbiology Reviews*, *35*(3), 441–474. https://doi.org/10.1111/j.1574-6976.2010.00257.x

De Guzman, B. B., Hisatsune, J., Nakayama, M., Yahiro, K., Wada, A., Yamasaki, E., … Hirayama, T. (2005). Cytotoxicity and recognition of receptor-like protein tyrosine phosphatases, RPTPα and RPTPβ, by Helicobacter pylori m2VacA. *Cellular Microbiology*, *7*(9), 1285–1293. https://doi.org/10.1111/j.1462-5822.2005.00556.x

De Jong, M. F., Sun, Y.-H., Den Hartigh, A. B., Van Dijl, J. M., & Tsolis, R. M. (2008). Identification of VceA and VceC, two members of the VjbR regulon that are translocated into macrophages by the Brucella type IV secretion system. *Molecular Microbiology*, *70*(6), 1378–1396. https://doi.org/10.1111/j.1365-2958.2008.06487.x

de la Cruz, F., Frost, L. S., Meyer, R. J., & Zechner, E. L. (2010). Conjugative DNA metabolism in Gram-negative bacteria. *FEMS Microbiology Reviews*, *34*(1), 18–40. https://doi.org/10.1111/j.1574-6976.2009.00195.x

De la Iglesia, R., Castro, D., Ginocchio, R., van der Lelie, D., & González, B. (2006). Factors influencing the composition of bacterial communities found at abandoned copper-tailings dumps. *Journal of Applied Microbiology*, *100*(3), 537–544. https://doi.org/10.1111/j.1365-2672.2005.02793.x

de Oliveira, D. E., Ballon, G., & Cesarman, E. (2010). NF-κB signaling modulation by EBV and KSHV. *Trends in Microbiology*, *18*(6), 248–257. https://doi.org/10.1016/j.tim.2010.04.001

Dean, P., Mühlen, S., Quitard, S., & Kenny, B. (2010). *The bacterial effectors EspG and EspG2 induce a destructive calpain activity that is kept in check by the co-delivered Tir effector. Cellular microbiology* (Vol. 12). https://doi.org/10.1111/j.1462-5822.2010.01469.x

Denner, J., & Tönjes, R. R. (2012). Infection barriers to successful xenotransplantation focusing on porcine endogenous retroviruses. *Clinical Microbiology Reviews*, *25*(2), 318–43. https://doi.org/10.1128/CMR.05011-11

Diemert, D. J. (2006). Prevention and self-treatment of traveler's diarrhea. *Clinical Microbiology Reviews*, *19*(3), 583–94. https://doi.org/10.1128/CMR.00052-05

Dmitriev, B., Toukach, F., & Ehlers, S. (2005). Towards a comprehensive view of the bacterial cell wall. *Trends in Microbiology*, *13*(12), 569–574. https://doi.org/10.1016/j.tim.2005.10.001

Dobrindt, U., & Hacker, J. (2008). Targeting virulence traits: potential strategies to combat extraintestinal pathogenic E. coli infections. *Current Opinion in Microbiology*, *11*(5), 409–413. https://doi.org/10.1016/j.mib.2008.09.005

Dodsworth, J. A., Cady, N. C., & Leigh, J. A. (2005). 2-Oxoglutarate and the PII homologues NifI1 and NifI2 regulate nitrogenase activity in cell extracts of Methanococcus maripaludis. *Molecular Microbiology*, *56*(6), 1527–1538. https://doi.org/10.1111/j.1365-2958.2005.04621.x

Dorer, M. S., Sessler, T. H., & Salama, N. R. (2011). Recombination and DNA Repair in Helicobacter pylori. *Annual Review of Microbiology*, *65*(1), 329–348. https://doi.org/10.1146/annurev-micro-090110-102931

Dover, L. G., Cerdeño-Tárraga, A. M., Pallen, M. J., Parkhill, J., & Besra, G. S. (2004). Comparative cell wall core biosynthesis in the mycolated pathogens, Mycobacterium tuberculosis and Corynebacterium diphtheriae. *FEMS Microbiology Reviews*, *28*(2), 225–250. Retrieved from http://femsre.oxfordjournals.org/content/28/2/225.abstract

Dramsi, S., Caliot, E., Bonne, I., Guadagnini, S., Prévost, M.-C., Kojadinovic, M., … Trieu-Cuot, P. (2006). Assembly and role of pili in group B streptococci. *Molecular Microbiology*, *60*(6), 1401–1413. https://doi.org/10.1111/j.1365-2958.2006.05190.x

Dramsi, S., Magnet, S., Davison, S., & Arthur, M. (2008). Covalent attachment of proteins to peptidoglycan. *FEMS Microbiology Reviews*, *32*(2), 307–320. https://doi.org/10.1111/j.1574-6976.2008.00102.x

Drummond, R. A., & Brown, G. D. (2011). The role of Dectin-1 in the host defence against fungal infections. *Current Opinion in Microbiology*, *14*(4), 392–9. https://doi.org/10.1016/j.mib.2011.07.001

Dubey, J. P., Schares, G., & Ortega-Mora, L. M. (2007). Epidemiology and control of neosporosis and Neospora caninum. *Clinical Microbiology Reviews*, *20*(2), 323–67. https://doi.org/10.1128/CMR.00031-06

Edwards, J. L., & Apicella, M. A. (2002). The role of lipooligosaccharide in Neisseria gonorrhoeae pathogenesis of cervical epithelia: lipid A serves as a C3 acceptor molecule. *Cellular Microbiology*, *4*(9), 585–598. https://doi.org/10.1046/j.1462-5822.2002.00212.x

Ehira, S., & Ohmori, M. (2006). NrrA, a nitrogen-responsive response regulator facilitates heterocyst development in the cyanobacterium Anabaena sp. strain PCC 7120. *Molecular Microbiology*, *59*(6), 1692–1703. https://doi.org/10.1111/j.1365-2958.2006.05049.x

Ehrenkaufer, G. M., Hackney, J. A., & Singh, U. (2009). *A developmentally regulated Myb domain protein regulates expression of a subset of stage-specific genes in Entamoeba histolytica. Cellular microbiology* (Vol. 11). https://doi.org/10.1111/j.1462-5822.2009.01300.x

Eksi, S., Stump, A., Fanning, S. L., Shenouda, M. I., Fujioka, H., & Williamson, K. C. (2002). Targeting and sequestration of truncated Pfs230 in an intraerythrocytic

compartment during Plasmodium falciparum gametocytogenesis. *Molecular Microbiology*, *44*(6), 1507–1516. https://doi.org/10.1046/j.1365-2958.2002.02986.x

Esimone, C. O., Grunwald, T., Wildner, O., Nchinda, G., Tippler, B., Proksch, P., & Überla, K. (2005). In vitro pharmacodynamic evaluation of antiviral medicinal plants using a vector-based assay technique. *Journal of Applied Microbiology*, *99*(6), 1346–1355. https://doi.org/10.1111/j.1365-2672.2005.02732.x

Eto, D. S., Gordon, H. B., Dhakal, B. K., Jones, T. A., & Mulvey, M. A. (2008). *Clathrin, AP-2, and the NPXY-binding subset of alternate endocytic adaptors facilitate FimH-mediated bacterial invasion of host cells. Cellular microbiology* (Vol. 10). https://doi.org/10.1111/j.1462-5822.2008.01229.x

Etxebeste, O., Garzia, A., Espeso, E. A., & Ugalde, U. (2010). Aspergillus nidulans asexual development: making the most of cellular modules. *Trends in Microbiology*, *18*(12), 569–76. https://doi.org/10.1016/j.tim.2010.09.007

Ewing, S. A., & Panciera, R. J. (2003). American Canine Hepatozoonosis. *Clinical Microbiology Reviews*, *16*(4), 688–697. https://doi.org/10.1128/CMR.16.4.688-697.2003

Facklam, R. (2002). What Happened to the Streptococci: Overview of Taxonomic and Nomenclature Changes. *Clinical Microbiology Reviews*, *15*(4), 613–630. https://doi.org/10.1128/CMR.15.4.613-630.2002

Ferris, H. U., & Minamino, T. (2006). Flipping the switch: bringing order to flagellar assembly. *Trends in Microbiology*, *14*(12), 519–26. https://doi.org/10.1016/j.tim.2006.10.006

Feurer, C., Irlinger, F., Spinnler, H. E., Glaser, P., & Vallaeys, T. (2004). Assessment of the rind microbial diversity in a farmhouse-produced vs a pasteurized industrially produced soft red-smear cheese using both cultivation and rDNA-based methods. *Journal of Applied Microbiology*, *97*(3), 546–556. https://doi.org/10.1111/j.1365-2672.2004.02333.x

Fijalkowska, I. J., Schaaper, R. M., & Jonczyk, P. (2012). DNA replication fidelity in Escherichia coli: a multi-DNA polymerase affair. *FEMS Microbiology Reviews*, *36*(6), 1105–1121. https://doi.org/10.1111/j.1574-6976.2012.00338.x

Fischbach, M. A. (2009). Antibiotics from microbes: converging to kill. *Current Opinion in Microbiology*, *12*(5), 520–7. https://doi.org/10.1016/j.mib.2009.07.002

Forchhammer, K. (2004). Global carbon/nitrogen control by PII signal transduction in cyanobacteria: from signals to targets. *FEMS Microbiology Reviews*, *28*(3), 319–333. Retrieved from http://femsre.oxfordjournals.org/content/28/3/319.abstract

Fournier, P.-E., & Raoult, D. (2011). Prospects for the Future Using Genomics and Proteomics in Clinical Microbiology. *Annual Review of Microbiology*, *65*(1), 169–188. https://doi.org/10.1146/annurev-micro-090110-102922

Frank, D. N., Zhu, W., Sartor, R. B., & Li, E. (2011). Investigating the biological and clinical significance of human dysbioses. *Trends in Microbiology*, *19*(9), 427–34. https://doi.org/10.1016/j.tim.2011.06.005

Fujimoto, D. F., Higginbotham, R. H., Sterba, K. M., Maleki, S. J., Segall, A. M., Smeltzer, M. S., & Hurlburt, B. K. (2009). Staphylococcus aureus SarA is a regulatory protein responsive to redox and pH that can support bacteriophage lambda integrase-mediated excision/recombination. *Molecular Microbiology*, *74*(6), 1445–1458. https://doi.org/10.1111/j.1365-2958.2009.06942.x

Fujinami, R. S., von Herrath, M. G., Christen, U., & Whitton, J. L. (2006). Molecular mimicry, bystander activation, or viral persistence: infections and autoimmune disease. *Clinical Microbiology Reviews*, *19*(1), 80–94. https://doi.org/10.1128/CMR.19.1.80-94.2006

García Véscovi, E., Sciara, M. I., & Castelli, M. E. (2010). Two component systems in the spatial program of bacteria. *Current Opinion in Microbiology*, *13*(2), 210–8. https://doi.org/10.1016/j.mib.2009.12.012

Garner, M. J., Hayward, R. D., & Koronakis, V. (2002). The Salmonella pathogenicity island 1 secretion system directs cellular cholesterol redistribution during mammalian cell entry and intracellular trafficking. *Cellular Microbiology*, *4*(3), 153–165. https://doi.org/10.1046/j.1462-5822.2002.00181.x

Gatti, M., Lazzi, C., Rossetti, L., Mucchetti, G., & Neviani, E. (2003). Biodiversity in Lactobacillus helveticus strains present in natural whey starter used for Parmigiano Reggiano cheese. *Journal of Applied Microbiology*, *95*(3), 463–470. https://doi.org/10.1046/j.1365-2672.2003.01997.x

Genisset, C., Puhar, A., Calore, F., de Bernard, M., Dell'Antone, P., & Montecucco, C. (2007). The concerted action of the Helicobacter pylori cytotoxin VacA and of the v-ATPase proton pump induces swelling of isolated endosomes. *Cellular Microbiology*, *9*(6), 1481–1490. https://doi.org/10.1111/j.1462-5822.2006.00886.x

Georlette, D., Blaise, V., Collins, T., D'Amico, S., Gratia, E., Hoyoux, A., … Gerday, C. (2004). Some like it cold: biocatalysis at low temperatures. *FEMS Microbiology Reviews*, *28*(1), 25–42. Retrieved from http://femsre.oxfordjournals.org/content/28/1/25.abstract

Ghosh, S., Rao, K. H., Sengupta, M., Bhattacharya, S. K., & Datta, A. (2011). Two gene clusters co-ordinate for a functional N-acetylglucosamine catabolic pathway in Vibrio cholerae. *Molecular Microbiology*, *80*(6), 1549–1560. https://doi.org/10.1111/j.1365-2958.2011.07664.x

Godish, D. R., & Godish, T. J. (2007). Relationship between sampling duration and concentration of culturable airborne mould and bacteria on selected culture media. *Journal of Applied Microbiology*, *102*(6), 1479–1484. https://doi.org/10.1111/j.1365-2672.2006.03200.x

Goldmann, W., Baylis, M., Chihota, C., Stevenson, E., & Hunter, N. (2005). Frequencies of PrP gene haplotypes in British sheep flocks and the implications for breeding programmes. *Journal of Applied Microbiology*, *98*(6), 1294–1302. https://doi.org/10.1111/j.1365-2672.2005.02568.x

González-Novo, A., Labrador, L., Pablo-Hernando, M. E., Correa-Bordes, J., Sánchez, M., Jiménez, J., & De Aldana, C. R. V. (2009). Dbf2 is essential for cytokinesis and correct mitotic spindle formation in Candida albicans. *Molecular Microbiology*, *72*(6), 1364–1378. https://doi.org/10.1111/j.1365-2958.2009.06729.x

Goodrich-Blair, H. (2007). They've got a ticket to ride: Xenorhabdus nematophila-Steinernema carpocapsae symbiosis. *Current Opinion in Microbiology*, *10*(3), 225–30. https://doi.org/10.1016/j.mib.2007.05.006

Göpel, Y., & Görke, B. (2012). Rewiring two-component signal transduction with small RNAs. *Current Opinion in Microbiology*, *15*(2), 132–9. https://doi.org/10.1016/j.mib.2011.12.001

Gornik, S. G., Albalat, A., Macpherson, H., Birkbeck, H., & Neil, D. M. (2011). The effect of temperature on the bacterial load and microbial composition in Norway lobster (Nephrops norvegicus) tail meat during storage. *Journal of Applied Microbiology*, *111*(3), 582–592. https://doi.org/10.1111/j.1365-2672.2011.05081.x

Gottesman, S. (2004). The Small RNA Regulators of Escherichia coli: Roles and Mechanisms. *Annual Review of Microbiology*, *58*(1), 303–328. https://doi.org/10.1146/annurev.micro.58.030603.123841

Gouin, E., Welch, M. D., & Cossart, P. (2005). Actin-based motility of intracellular pathogens. *Current Opinion in Microbiology*, *8*(1), 35–45. https://doi.org/10.1016/j.mib.2004.12.013

Graczyk, T. K., Knight, R., & Tamang, L. (2005). Mechanical transmission of human protozoan parasites by insects. *Clinical Microbiology Reviews*, *18*(1), 128–32. https://doi.org/10.1128/CMR.18.1.128-132.2005

Gradman, R. J., Ptacin, J. L., Bhasin, A., Reznikoff, W. S., & Goryshin, I. Y. (2008). A bifunctional DNA binding region in Tn5 transposase. *Molecular Microbiology*, *67*(3), 528–540. https://doi.org/10.1111/j.1365-2958.2007.06056.x

Grahn, A. M., Daugelavičius, R., & Bamford, D. H. (2002). Sequential model of phage PRD1 DNA delivery: active involvement of the viral membrane. *Molecular Microbiology*, *46*(5), 1199–1209. https://doi.org/10.1046/j.1365-2958.2002.03250.x

Gray, M. J., & Escalante-Semerena, J. C. (2010). A new pathway for the synthesis of α-ribazole-phosphate in Listeria innocua. *Molecular Microbiology*, *77*(6), 1429–1438. https://doi.org/10.1111/j.1365-2958.2010.07294.x

Greub, G., & Raoult, D. (2004). Microorganisms Resistant to Free-Living Amoebae. *Clinical Microbiology Reviews*, *17*(2), 413–433. https://doi.org/10.1128/CMR.17.2.413-433.2004

Griffin, D. W., Donaldson, K. A., Paul, J. H., & Rose, J. B. (2003). Pathogenic Human Viruses in Coastal Waters. *Clinical Microbiology Reviews*, *16*(1), 129–143. https://doi.org/10.1128/CMR.16.1.129-143.2003

Gulley, M. L., & Tang, W. (2010). Using Epstein-Barr Viral Load Assays To Diagnose, Monitor, and Prevent Posttransplant Lymphoproliferative Disorder. *Clinical Microbiology Reviews*, *23*(2), 350–366. https://doi.org/10.1128/CMR.00006-09

Gutiérrez-Fernández, J., Lara, A., Bautista, M. F., de Dios Luna, J., Polo, P., Miranda, C., & Navarro, J. M. (2012). Performance of the Sysmex UF1000i system in screening for significant bacteriuria before quantitative culture of aerobic/facultative fast-growth bacteria in a reference hospital. *Journal of Applied Microbiology*, *113*(3), 609–614. https://doi.org/10.1111/j.1365-2672.2012.05369.x

Hagemann, M. (2011). Molecular biology of cyanobacterial salt acclimation. *FEMS Microbiology Reviews*, *35*(1), 87–123. https://doi.org/10.1111/j.1574-6976.2010.00234.x

Hammer, N. D., & Skaar, E. P. (2012). The impact of metal sequestration on Staphylococcus aureus metabolism. *Current Opinion in Microbiology*, *15*(1), 10–4. https://doi.org/10.1016/j.mib.2011.11.004

Haramoto, E., Katayama, H., Oguma, K., & Ohgaki, S. (2007). Quantitative analysis of human enteric adenoviruses in aquatic environments. *Journal of Applied Microbiology*, *103*(6), 2153–2159. https://doi.org/10.1111/j.1365-2672.2007.03453.x

Hariri, S., & McKenna, M. T. (2007). Epidemiology of Human Immunodeficiency Virus in the United States. *Clinical Microbiology Reviews*, *20*(3), 478–488. https://doi.org/10.1128/CMR.00006-07

Harrison, E., & Brockhurst, M. A. (2012). Plasmid-mediated horizontal gene transfer is a coevolutionary process. *Trends in Microbiology*, *20*(6), 262–267. https://doi.org/10.1016/j.tim.2012.04.003

Hatfull, G. F. (2010). Mycobacteriophages: Genes and Genomes. *Annual Review of Microbiology*, *64*(1), 331–356. https://doi.org/10.1146/annurev.micro.112408.134233

Hausmann, C. D., & Ibba, M. (2008). Aminoacyl-tRNA synthetase complexes: molecular multitasking revealed. *FEMS Microbiology Reviews*, *32*(4), 705–721. https://doi.org/10.1111/j.1574-6976.2008.00119.x

He, Y.-W., Boon, C., Zhou, L., & Zhang, L.-H. (2009). Co-regulation of Xanthomonas campestris virulence by quorum sensing and a novel two-component regulatory system RavS/RavR. *Molecular Microbiology*, *71*(6), 1464–1476. https://doi.org/10.1111/j.1365-2958.2009.06617.x

Hecker, M., Pané-Farré, J., & Uwe, V. (2007). SigB-Dependent General Stress Response in Bacillus subtilis and Related Gram-Positive Bacteria. *Annual Review of Microbiology*, *61*(1), 215–236. https://doi.org/10.1146/annurev.micro.61.080706.093445

Henderson, B., Nair, S. P., Ward, J. M., & Wilson, M. (2003). Molecular Pathogenicity of the Oral Opportunistic Pathogen Actinobacillus actinomycetemcomitans. *Annual Review of Microbiology*, *57*(1), 29–55. https://doi.org/10.1146/annurev.micro.57.030502.090908

Hennekinne, J.-A., De Buyser, M.-L., & Dragacci, S. (2012). Staphylococcus aureus and its food poisoning toxins: characterization and outbreak investigation. *FEMS Microbiology Reviews*, *36*(4), 815–836. Retrieved from http://femsre.oxfordjournals.org/content/36/4/815.abstract

Henrickson, K. J. (2003). Parainfluenza Viruses. *Clinical Microbiology Reviews*, *16*(2), 242–264. https://doi.org/10.1128/CMR.16.2.242-264.2003

Herricks, T., Antia, M., & Rathod, P. K. (2009). *Deformability limits of Plasmodium falciparum - infected red blood cells*. *Cellular Microbiology* (Vol. 11). https://doi.org/10.1111/j.1462-5822.2009.01334.x

Higashide, W., Dai, S., Hombs, V. P., & Zhou, D. (2002). Involvement of SipA in modulating actin dynamics during Salmonella invasion into cultured epithelial cells. *Cellular Microbiology*, *4*(6), 357–365. https://doi.org/10.1046/j.1462-5822.2002.00196.x

Hoeijmakers, W. A. M., Flueck, C., Françoijs, K.-J., Smits, A. H., Wetzel, J., Volz, J. C., … Bártfai, R. (2012). *Plasmodium falciparum centromeres display a unique epigenetic makeup and cluster prior to and during schizogony*. *Cellular Microbiology* (Vol. 14). https://doi.org/10.1111/j.1462-5822.2012.01803.x

Hoekstra, K. A., & Paulton, R. J. L. (2002). Clinical prevalence and antimicrobial susceptibility of Staphylococcus aureus and Staph. intermedius in dogs. *Journal of Applied Microbiology*, *93*(3), 406–413. https://doi.org/10.1046/j.1365-2672.2002.01708.x

Holbrook, E. D., & Rappleye, C. A. (2008). Histoplasma capsulatum pathogenesis: making a lifestyle switch. *Current Opinion in Microbiology*, *11*(4), 318–24. https://doi.org/10.1016/j.mib.2008.05.010

Holland, C. A., & Kiechle, F. L. (2005). Point-of-care molecular diagnostic systems--past, present and future. *Current Opinion in Microbiology*, *8*(5), 504–9. https://doi.org/10.1016/j.mib.2005.08.001

Holmes, E. C. (2003). Error thresholds and the constraints to RNA virus evolution. *Trends in Microbiology*, *11*(12), 543–546. https://doi.org/10.1016/j.tim.2003.10.006

Horn, M. (2008). Chlamydiae as Symbionts in Eukaryotes. *Annual Review of Microbiology*, *62*(1), 113–131. https://doi.org/10.1146/annurev.micro.62.081307.162818

Hrincius, E. R., Dierkes, R., Anhlan, D., Wixler, V., Ludwig, S., & Ehrhardt, C. (2011). *Phosphatidylinositol-3-kinase (PI3K) is activated by influenza virus vRNA via the pathogen pattern receptor Rig-I to promote efficient type I interferon production*. *Cellular microbiology* (Vol. 13). https://doi.org/10.1111/j.1462-5822.2011.01680.x

Hrincius, E. R., Wixler, V., Wolff, T., Wagner, R., Ludwig, S., & Ehrhardt, C. (2010). *CRK adaptor protein expression is required for efficient replication of avian influenza A viruses and controls JNK-mediated apoptotic responses*. *Cellular microbiology* (Vol. 12). https://doi.org/10.1111/j.1462-5822.2010.01436.x

Huang, C.-H., Tsai, H.-H., Tsay, Y.-G., Chien, Y.-N., Wang, S.-L., Cheng, M.-Y., … Chen, C. W. (2007). The telomere system of the Streptomyces linear plasmid SCP1 represents a novel class. *Molecular Microbiology*, *63*(6), 1710–1718. https://doi.org/10.1111/j.1365-2958.2007.05616.x

Huang, J., Shi, J., Molle, V., Sohlberg, B., Weaver, D., Bibb, M. J., … Cohen, S. N. (2005). Cross-regulation among disparate antibiotic biosynthetic pathways of Streptomyces

coelicolor. *Molecular Microbiology*, *58*(5), 1276–1287. https://doi.org/10.1111/j.1365-2958.2005.04879.x
Huthoff, H., & Towers, G. J. (2008). Restriction of retroviral replication by APOBEC3G/F and TRIM5α. *Trends in Microbiology*, *16*(12), 612–619. https://doi.org/10.1016/j.tim.2008.08.013
Ito, K., & Akiyama, Y. (2005). CELLULAR FUNCTIONS, MECHANISM OF ACTION, AND REGULATION OF FTSH PROTEASE. *Annual Review of Microbiology*, *59*(1), 211–231. https://doi.org/10.1146/annurev.micro.59.030804.121316
Iwasaki, A. (2012). A Virological View of Innate Immune Recognition. *Annual Review of Microbiology*, *66*(1), 177–196. https://doi.org/10.1146/annurev-micro-092611-150203
Jacobsen, S. M., Stickler, D. J., Mobley, H. L. T., & Shirtliff, M. E. (2008). Complicated catheter-associated urinary tract infections due to Escherichia coli and Proteus mirabilis. *Clinical Microbiology Reviews*, *21*(1), 26–59. https://doi.org/10.1128/CMR.00019-07
Jantsch, J., Cheminay, C., Chakravortty, D., Lindig, T., Hein, J., & Hensel, M. (2003). Intracellular activities of Salmonella enterica in murine dendritic cells. *Cellular Microbiology*, *5*(12), 933–945. https://doi.org/10.1046/j.1462-5822.2003.00334.x
Jean, J.-S., Guo, H.-R., Chen, S.-H., Liu, C.-C., Chang, W.-T., Yang, Y.-J., & Huang, M.-C. (2006). The association between rainfall rate and occurrence of an enterovirus epidemic due to a contaminated well. *Journal of Applied Microbiology*, *101*(6), 1224–1231. https://doi.org/10.1111/j.1365-2672.2006.03025.x
Jeong, S.-Y., Choi, C.-H., Kim, J.-S., Park, S.-J., & Kang, S.-O. (2006). Thioredoxin reductase is required for growth and regulates entry into culmination of Dictyostelium discoideum. *Molecular Microbiology*, *61*(6), 1443–1456. https://doi.org/10.1111/j.1365-2958.2006.05329.x
Jin, S., Song, Y. C., Emili, A., Sherman, P. M., & Chan, V. L. (2003). JlpA of Campylobacter jejuni interacts with surface-exposed heat shock protein 90alpha and triggers signalling pathways leading to the activation of NF-kappaB and p38 MAP kinase in epithelial cells. *Cellular Microbiology*, *5*(3), 165–174. https://doi.org/10.1046/j.1462-5822.2003.00265.x
Jordan, S., Hutchings, M. I., & Mascher, T. (2008). Cell envelope stress response in Gram-positive bacteria. *FEMS Microbiology Reviews*, *32*(1), 107–146. https://doi.org/10.1111/j.1574-6976.2007.00091.x
Jousselin, A., Metzinger, L., & Felden, B. (2009). On the facultative requirement of the bacterial RNA chaperone, Hfq. *Trends in Microbiology*, *17*(9), 399–405. https://doi.org/10.1016/j.tim.2009.06.003
Jung, K.-H., Trivedi, V. D., & Spudich, J. L. (2003). Demonstration of a sensory rhodopsin in eubacteria. *Molecular Microbiology*, *47*(6), 1513–1522. https://doi.org/10.1046/j.1365-2958.2003.03395.x
Kabuki, T., Kawai, Y., Uenishi, H., Seto, Y., Kok, J., Nakajima, H., & Saito, T. (2011). Gene cluster for biosynthesis of thermophilin 1277 – a lantibiotic produced by Streptococcus thermophilus SBT1277, and heterologous expression of TepI, a novel immunity peptide. *Journal of Applied Microbiology*, *110*(3), 641–649. https://doi.org/10.1111/j.1365-2672.2010.04914.x
Kaguni, J. M. (2006). DnaA: Controlling the Initiation of Bacterial DNA Replication and More. *Annual Review of Microbiology*, *60*(1), 351–371. https://doi.org/10.1146/annurev.micro.60.080805.142111
Kaiser, D. (2004). Signaling in myxobacteria. *Annual Review of Microbiology*, *58*(1), 75–98. https://doi.org/10.1146/annurev.micro.58.030603.123620
Kalamorz, F., Reichenbach, B., März, W., Rak, B., & Görke, B. (2007). Feedback control of glucosamine-6-phosphate synthase GlmS expression depends on the small RNA

GlmZ and involves the novel protein YhbJ in Escherichia coli. *Molecular Microbiology*, *65*(6), 1518–1533. https://doi.org/10.1111/j.1365-2958.2007.05888.x
Karl, D. M. (2002). Nutrient dynamics in the deep blue sea. *Trends in Microbiology*, *10*(9), 410–418. https://doi.org/10.1016/S0966-842X(02)02430-7
Keiser, J., & Utzinger, J. (2009). Food-borne trematodiases. *Clinical Microbiology Reviews*, *22*(3), 466–83. https://doi.org/10.1128/CMR.00012-09
Kenney, L. (2002). Structure/function relationships in OmpR and other winged-helix transcription factors. *Current Opinion in Microbiology*, *5*(2), 135–141. https://doi.org/10.1016/S1369-5274(02)00310-7
Kent, A. D., & Triplett, E. W. (2002). Microbial Communities and Their Interactions in Soil and Rhizosphere Ecosystems. *Annual Review of Microbiology*, *56*(1), 211–236. https://doi.org/10.1146/annurev.micro.56.012302.161120
Kim, C. H., Lee, J. H., Heo, J. H., Kwon, O. S., Kang, H. A., & Rhee, S. K. (2004). Cloning and expression of a novel esterase gene cpoA from Burkholderia cepacia. *Journal of Applied Microbiology*, *96*(6), 1306–1316. https://doi.org/10.1111/j.1365-2672.2004.02262.x
Kim, J., Hetzel, M., Boiangiu, C. D., & Buckel, W. (2004). Dehydration of (R)-2-hydroxyacyl-CoA to enoyl-CoA in the fermentation of α-amino acids by anaerobic bacteria. *FEMS Microbiology Reviews*, *28*(4), 455–468. https://doi.org/10.1016/j.femsre.2004.03.001
Kitajima, M., Haramoto, E., Phanuwan, C., Katayama, H., & Furumai, H. (2012). Molecular detection and genotyping of human noroviruses in influent and effluent water at a wastewater treatment plant in Japan. *Journal of Applied Microbiology*, *112*(3), 605–613. https://doi.org/10.1111/j.1365-2672.2012.05231.x
Kiulia, N. M., Netshikweta, R., Page, N. A., Van Zyl, W. B., Kiraithe, M. M., Nyachieo, A., … Taylor, M. B. (2010). The detection of enteric viruses in selected urban and rural river water and sewage in Kenya, with special reference to rotaviruses. *Journal of Applied Microbiology*, *109*(3), 818–828. https://doi.org/10.1111/j.1365-2672.2010.04710.x
Kleerebezem, M., & Vaughan, E. E. (2009). Probiotic and Gut Lactobacilli and Bifidobacteria: Molecular Approaches to Study Diversity and Activity. *Annual Review of Microbiology*, *63*(1), 269–290. https://doi.org/10.1146/annurev.micro.091208.073341
Kok, J., Buist, G., Zomer, A. L., Van Hijum, S. A. F. T., & Kuipers, O. P. (2005). Comparative and functional genomics of lactococci. *FEMS Microbiology Reviews*, *29*(3), 411–433. https://doi.org/10.1016/j.fmrre.2005.04.004
Koo, I. C., Wang, C., Raghavan, S., Morisaki, J. H., Cox, J. S., & Brown, E. J. (2008). *ESX-1-dependent cytolysis in lysosome secretion and inflammasome activation during mycobacterial infection. Cellular microbiology* (Vol. 10). https://doi.org/10.1111/j.1462-5822.2008.01177.x
Körner, H., Sofia, H. J., & Zumft, W. G. (2003). Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: exploiting the metabolic spectrum by controlling alternative gene programs. *FEMS Microbiology Reviews*, *27*(5), 559–592. https://doi.org/10.1016/S0168-6445(03)00066-4
Kretzschmar, M., Gabriela M. Gomes, M., Coutinho, R. A., & Koopman, J. S. (2010). Unlocking pathogen genotyping information for public health by mathematical modeling. *Trends in Microbiology*, *18*(9), 406–412. https://doi.org/10.1016/j.tim.2010.06.008
Kreuzer, K. N. (2005). Interplay between DNA replication and recombination in prokaryotes. *Annual Review of Microbiology*, *59*(1), 43–67. https://doi.org/10.1146/annurev.micro.59.030804.121255

Kulp, A., & Kuehn, M. J. (2010). Biological Functions and Biogenesis of Secreted Bacterial Outer Membrane Vesicles. *Annual Review of Microbiology*, *64*(1), 163–184. https://doi.org/10.1146/annurev.micro.091208.073413

Kunicka-Styczyńska, A., Sikora, M., & Kalemba, D. (2009). Antimicrobial activity of lavender, tea tree and lemon oils in cosmetic preservative systems. *Journal of Applied Microbiology*, *107*(6), 1903–1911. https://doi.org/10.1111/j.1365-2672.2009.04372.x

Kunttu, H. M. T., Jokinen, E. I., Valtonen, E. T., & Sundberg, L.-R. (2011). Virulent and nonvirulent Flavobacterium columnare colony morphologies: characterization of chondroitin AC lyase activity and adhesion to polystyrene. *Journal of Applied Microbiology*, *111*(6), 1319–1326. https://doi.org/10.1111/j.1365-2672.2011.05149.x

Kuo, M. M.-C., Haynes, W. J., Loukin, S. H., Kung, C., & Saimi, Y. (2005). Prokaryotic K+ channels: From crystal structures to diversity. *FEMS Microbiology Reviews*, *29*(5), 961–985. https://doi.org/10.1016/j.femsre.2005.03.003

La, M.-V., Raoult, D., & Renesto, P. (2008). Regulation of whole bacterial pathogen transcription within infected hosts. *FEMS Microbiology Reviews*, *32*(3), 440–460. https://doi.org/10.1111/j.1574-6976.2008.00103.x

Lafont, F., Abrami, L., & van der Goot, F. G. (2004). Bacterial subversion of lipid rafts. *Current Opinion in Microbiology*, *7*(1), 4–10. https://doi.org/10.1016/j.mib.2003.12.007

Lang, A. S., Rise, M. L., Culley, A. I., & Steward, G. F. (2009). RNA viruses in the sea. *FEMS Microbiology Reviews*, *33*(2), 295–323. https://doi.org/10.1111/j.1574-6976.2008.00132.x

Lang, T., Goyard, S., Lebastard, M., & Milon, G. (2005). Bioluminescent Leishmania expressing luciferase for rapid and high throughput screening of drugs acting on amastigote-harbouring macrophages and for quantitative real-time monitoring of parasitism features in living mice. *Cellular Microbiology*, *7*(3), 383–92. https://doi.org/10.1111/j.1462-5822.2004.00468.x

Laranjo, M., Branco, C., Soares, R., Alho, L., Carvalho, M. D. E., & Oliveira, S. (2002). Comparison of chickpea rhizobia isolates from diverse Portuguese natural populations based on symbiotic effectiveness and DNA fingerprint. *Journal of Applied Microbiology*, *92*(6), 1043–1050. https://doi.org/10.1046/j.1365-2672.2002.01615.x

Lazar Adler, N. R., Govan, B., Cullinane, M., Harper, M., Adler, B., & Boyce, J. D. (2009). The molecular and cellular basis of pathogenesis in melioidosis: how does Burkholderia pseudomallei cause disease? *FEMS Microbiology Reviews*, *33*(6), 1079–1099. https://doi.org/10.1111/j.1574-6976.2009.00189.x

Leal, N. C., Sobreira, M., Leal-Balbino, T. C., de Almeida, A. M. P., da Silva, M. J. B., Mello, D. M., … Hofer, E. (2004). Evaluation of a RAPD-based typing scheme in a molecular epidemiology study of Vibrio cholerae O1, Brazil. *Journal of Applied Microbiology*, *96*(3), 447–454. https://doi.org/10.1111/j.1365-2672.2004.02090.x

Lee, K., Iwata, T., Shimizu, M., Taniguchi, T., Nakadai, A., Hirota, Y., & Hayashidani, H. (2009). A novel multiplex PCR assay for Salmonella subspecies identification. *Journal of Applied Microbiology*, *107*(3), 805–811. https://doi.org/10.1111/j.1365-2672.2009.04263.x

Lee, S.-J., Böhm, A., Krug, M., & Boos, W. (2007). The ABC of binding-protein-dependent transport in Archaea. *Trends in Microbiology*, *15*(9), 389–97. https://doi.org/10.1016/j.tim.2007.08.002

Lee, S.-J., Surma, M., Seitz, S., Hausner, W., Thomm, M., & Boos, W. (2007). Differential signal transduction via TrmB, a sugar sensing transcriptional repressor of Pyrococcus furiosus. *Molecular Microbiology*, *64*(6), 1499–1505. https://doi.org/10.1111/j.1365-2958.2007.05737.x

Lemonnier, M., Landraud, L., & Lemichez, E. (2007). Rho GTPase-activating bacterial toxins: from bacterial virulence regulation to eukaryotic cell biology. *FEMS*

Microbiology Reviews, *31*(5), 515–534. https://doi.org/10.1111/j.1574-6976.2007.00078.x

Leong, J. M., Parveen, N., & Goguen, J. D. (2002). New approaches and old problems in the shadow of bioterrorism. *Trends in Microbiology*, *10*(3), 112–114. https://doi.org/10.1016/S0966-842X(02)02310-7

Leung, D. W., Basler, C. F., & Amarasinghe, G. K. (2012). Molecular mechanisms of viral inhibitors of RIG-I-like receptors. *Trends in Microbiology*, *20*(3), 139–46. https://doi.org/10.1016/j.tim.2011.12.005

Lewis, K., Epstein, S., Godoy, V. G., & Hong, S.-H. (2008). Intact DNA in ancient permafrost. *Trends in Microbiology*, *16*(3), 92–4. https://doi.org/10.1016/j.tim.2008.01.002

Lilley, C. E., Schwartz, R. A., & Weitzman, M. D. (2007). Using or abusing: viruses and the cellular DNA damage response. *Trends in Microbiology*, *15*(3), 119–26. https://doi.org/10.1016/j.tim.2007.01.003

Ling, J., Reynolds, N., & Ibba, M. (2009). Aminoacyl-tRNA Synthesis and Translational Quality Control. *Annual Review of Microbiology*, *63*(1), 61–78. https://doi.org/10.1146/annurev.micro.091208.073210

Linke, D., Riess, T., Autenrieth, I. B., Lupas, A., & Kempf, V. A. J. (2006). Trimeric autotransporter adhesins: variable structure, common function. *Trends in Microbiology*, *14*(6), 264–270. https://doi.org/10.1016/j.tim.2006.04.005

Lobry, J. R., & Louarn, J.-M. (2003). Polarisation of prokaryotic chromosomes. *Current Opinion in Microbiology*, *6*(2), 101–108. https://doi.org/10.1016/S1369-5274(03)00024-9

Lommel, S., Benesch, S., Rohde, M., Wehland, J., & Rottner, K. (2004). Enterohaemorrhagic and enteropathogenic Escherichia coli use different mechanisms for actin pedestal formation that converge on N-WASP. *Cellular Microbiology*, *6*(3), 243–254. https://doi.org/10.1111/j.1462-5822.2004.00364.x

Long, S. W., & Faguy, D. M. (2004). Anucleate and titan cell phenotypes caused by insertional inactivation of the structural maintenance of chromosomes (smc) gene in the archaeon Methanococcus voltae. *Molecular Microbiology*, *52*(6), 1567–1577. https://doi.org/10.1111/j.1365-2958.2004.04084.x

Lorenz, M. C. (2002). Genomic approaches to fungal pathogenicity. *Current Opinion in Microbiology*, *5*(4), 372–378. https://doi.org/10.1016/S1369-5274(02)00336-3

Macheboeuf, P., Contreras-Martel, C., Job, V., Dideberg, O., & Dessen, A. (2006). Penicillin Binding Proteins: key players in bacterial cell cycle and drug resistance processes. *FEMS Microbiology Reviews*, *30*(5), 673–691. https://doi.org/10.1111/j.1574-6976.2006.00024.x

Mahlen, S. D. (2011). Serratia Infections: from Military Experiments to Current Practice. *Clinical Microbiology Reviews*, *24*(4), 755–791. https://doi.org/10.1128/CMR.00017-11

Majdalani, N., & Gottesman, S. (2005). THE RCS PHOSPHORELAY: A Complex Signal Transduction System. *Annual Review of Microbiology*, *59*(1), 379–405. https://doi.org/10.1146/annurev.micro.59.050405.101230

Marlovits, T. C., & Stebbins, C. E. (2010). Type III secretion systems shape up as they ship out. *Current Opinion in Microbiology*, *13*(1), 47–52. https://doi.org/10.1016/j.mib.2009.11.001

Martínez, J. M. G., Gómez, E. V., Pemán, J., Cantón, E., García, M. G., & Del Castillo Agudo, L. (2010). Identification of pathogenic yeast species by polymerase chain reaction amplification of the RPS0 gene intron fragment. *Journal of Applied Microbiology*, *108*(6), 1917–1927. https://doi.org/10.1111/j.1365-2672.2009.04595.x

Martínez-Rocha, A. L., Roncero, M. I. G., López-Ramirez, A., Mariné, M., Guarro, J., Martínez-Cadena, G., & Di Pietro, A. (2008). *Rho1 has distinct functions in morphogenesis,*

cell wall biosynthesis and virulence of Fusarium oxysporum. *Cellular Microbiology* (Vol. 10). https://doi.org/10.1111/j.1462-5822.2008.01130.x
Massey, R. C., & Buckling, A. (2002). Environmental regulation of mutation rates at specific sites. *Trends in Microbiology*, *10*(12), 580–584. https://doi.org/10.1016/S0966-842X(02)02475-7
Mathema, B., Kurepina, N. E., Bifani, P. J., & Kreiswirth, B. N. (2006). Molecular Epidemiology of Tuberculosis: Current Insights. *Clinical Microbiology Reviews*, *19*(4), 658–685. https://doi.org/10.1128/CMR.00061-05
McBride, A. A., McPhillips, M. G., & Oliveira, J. G. (2004). Brd4: tethering, segregation and beyond. *Trends in Microbiology*, *12*(12), 527–9. https://doi.org/10.1016/j.tim.2004.10.002
McManus, C. J., & Kelley, S. T. (2005). Molecular survey of aeroplane bacterial contamination. *Journal of Applied Microbiology*, *99*(3), 502–508. https://doi.org/10.1111/j.1365-2672.2005.02651.x
McPherson, S. A., Li, M., Kearney, J. F., & Turnbough, C. L. (2010). ExsB, an unusually highly phosphorylated protein required for the stable attachment of the exosporium of Bacillus anthracis. *Molecular Microbiology*, *76*(6), 1527–1538. https://doi.org/10.1111/j.1365-2958.2010.07182.x
Meckes, M. C., & MacDonald, J. A. (2003). Evaluation of a DNA probe test kit for detection of Salmonellae in biosolids. *Journal of Applied Microbiology*, *94*(3), 382–387. https://doi.org/10.1046/j.1365-2672.2003.01841.x
Messer, W. (2002). The bacterial replication initiator DnaA. DnaA and oriC, the bacterial mode to initiate DNA replication. *FEMS Microbiology Reviews*, *26*(4), 355–374. https://doi.org/10.1111/j.1574-6976.2002.tb00620.x
Misseri, Y., Labesse, G., Bucheton, A., & Terzian, C. (2003). Comparative sequence analysis and predictions for the envelope glycoproteins of insect endogenous retroviruses. *Trends in Microbiology*, *11*(6), 253–256. https://doi.org/10.1016/S0966-842X(03)00119-7
Moeller, R., Wassmann, M., Reitz, G., & Setlow, P. (2011). Effect of radioprotective agents in sporulation medium on Bacillus subtilis spore resistance to hydrogen peroxide, wet heat and germicidal and environmentally relevant UV radiation. *Journal of Applied Microbiology*, *110*(6), 1485–1494. https://doi.org/10.1111/j.1365-2672.2011.05004.x
Molina-Henares, A. J., Krell, T., Eugenia Guazzaroni, M., Segura, A., & Ramos, J. L. (2006). Members of the IclR family of bacterial transcriptional regulators function as activators and/or repressors. *FEMS Microbiology Reviews*, *30*(2), 157–186. https://doi.org/10.1111/j.1574-6976.2005.00008.x
Moliner, C., Fournier, P.-E., & Raoult, D. (2010). Genome analysis of microorganisms living in amoebae reveals a melting pot of evolution. *FEMS Microbiology Reviews*, *34*(3), 281–294. https://doi.org/10.1111/j.1574-6976.2009.00209.x
Moon, K., & Gottesman, S. (2011). Competition among Hfq-binding small RNAs in Escherichia coli. *Molecular Microbiology*, *82*(6), 1545–1562. https://doi.org/10.1111/j.1365-2958.2011.07907.x
Moon, S. L., Barnhart, M. D., & Wilusz, J. (2012). Inhibition and avoidance of mRNA degradation by RNA viruses. *Current Opinion in Microbiology*, *15*(4), 500–505. https://doi.org/10.1016/j.mib.2012.04.009
Mora, L., Zavialov, A., Ehrenberg, M., & Buckingham, R. H. (2003). Stop codon recognition and interactions with peptide release factor RF3 of truncated and chimeric RF1 and RF2 from Escherichia coli. *Molecular Microbiology*, *50*(5), 1467–1476. https://doi.org/10.1046/j.1365-2958.2003.03799.x

Moreno, F., & Herrero, P. (2002). The hexokinase 2-dependent glucose signal transduction pathway of Saccharomyces cerevisiae. *FEMS Microbiology Reviews*, *26*(1), 83–90. https://doi.org/10.1111/j.1574-6976.2002.tb00600.x

Morgan, W., & Kamoun, S. (2007). RXLR effectors of plant pathogenic oomycetes. *Current Opinion in Microbiology*, *10*(4), 332–8. https://doi.org/10.1016/j.mib.2007.04.005

Motley, S. T., Morrow, B. J., Liu, X., Dodge, I. L., Vitiello, A., Ward, C. K., & Shaw, K. J. (2004). Simultaneous analysis of host and pathogen interactions during an in vivo infection reveals local induction of host acute phase response proteins, a novel bacterial stress response, and evidence of a host-imposed metal ion limited environment. *Cellular Microbiology*, *6*(9), 849–865. https://doi.org/10.1111/j.1462-5822.2004.00407.x

Mottram, J. C., Coombs, G. H., & Alexander, J. (2004). Cysteine peptidases as virulence factors of Leishmania. *Current Opinion in Microbiology*, *7*(4), 375–81. https://doi.org/10.1016/j.mib.2004.06.010

Müller, S., & Nebe-von-Caron, G. (2010). Functional single-cell analyses: flow cytometry and cell sorting of microbial populations and communities. *FEMS Microbiology Reviews*, *34*(4), 554–587. https://doi.org/10.1111/j.1574-6976.2010.00214.x

Munksgaard, P. S., Vorup-Jensen, T., Reinholdt, J., Söderström, C. M., Poulsen, K., Leipziger, J., … Skals, M. (2012). Leukotoxin from Aggregatibacter actinomycetemcomitans causes shrinkage and P2X receptor-dependent lysis of human erythrocytes. *Cellular Microbiology*, *14*(12), 1904–20. https://doi.org/10.1111/cmi.12021

Nam, I. S., & Garnsworthy, P. C. (2007). Biohydrogenation of linoleic acid by rumen fungi compared with rumen bacteria. *Journal of Applied Microbiology*, *103*(3), 551–556. https://doi.org/10.1111/j.1365-2672.2007.03317.x

Nau, R., & Eiffert, H. (2002). Modulation of Release of Proinflammatory Bacterial Compounds by Antibacterials: Potential Impact on Course of Inflammation and Outcome in Sepsis and Meningitis. *Clinical Microbiology Reviews*, *15*(1), 95–110. https://doi.org/10.1128/CMR.15.1.95-110.2002

Nau, R., Sörgel, F., & Eiffert, H. (2010). Penetration of drugs through the blood-cerebrospinal fluid/blood-brain barrier for treatment of central nervous system infections. *Clinical Microbiology Reviews*, *23*(4), 858–83. https://doi.org/10.1128/CMR.00007-10

Nawani, N. N., Kapadnis, B. P., Das, A. D., Rao, A. S., & Mahajan, S. K. (2002). Purification and characterization of a thermophilic and acidophilic chitinase from Microbispora sp. V2. *Journal of Applied Microbiology*, *93*(6), 965–975. https://doi.org/10.1046/j.1365-2672.2002.01766.x

Nechaev, S., & Severinov, K. (2003). Bacteriophage-Induced Modifications of Host RNA Polymerase. *Annual Review of Microbiology*, *57*(1), 301–322. https://doi.org/10.1146/annurev.micro.57.030502.090942

Neri, F., Mari, M., Menniti, A. M., & Brigati, S. (2006). Activity of trans-2-hexenal against Penicillium expansum in "Conference" pears. *Journal of Applied Microbiology*, *100*(6), 1186–1193. https://doi.org/10.1111/j.1365-2672.2006.02873.x

Nettmann, E., Bergmann, I., Mundt, K., Linke, B., & Klocke, M. (2008). Archaea diversity within a commercial biogas plant utilizing herbal biomass determined by 16S rDNA and mcrA analysis. *Journal of Applied Microbiology*, *105*(6), 1835–1850. https://doi.org/10.1111/j.1365-2672.2008.03949.x

Nigro, G., Fazio, L. L., Martino, M. C., Rossi, G., Tattoli, I., Liparoti, V., … Bernardini, M. L. (2008). *Muramylpeptide shedding modulates cell sensing of Shigella flexneri*. *Cellular Microbiology* (Vol. 10). Blackwell Publishing Ltd. https://doi.org/10.1111/j.1462-5822.2007.01075.x

Nikaido, H., & Pagès, J.-M. (2012). Broad Specificity Efflux pumps and Their Role in Multidrug Resistance of Gram Negative Bacteria. *FEMS Microbiology Reviews*, *36*(2), 340–363. https://doi.org/10.1111/j.1574-6976.2011.00290.x

Nogales, B., Lanfranconi, M. P., Piña-Villalonga, J. M., & Bosch, R. (2011). Anthropogenic perturbations in marine microbial communities. *FEMS Microbiology Reviews*, *35*(2), 275–298. https://doi.org/10.1111/j.1574-6976.2010.00248.x

O'Driscoll, J., Fitzgerald, G. F., & van Sinderen, D. (2005). A dichotomous epigenetic mechanism governs expression of the LlaJI restriction/modification system. *Molecular Microbiology*, *57*(6), 1532–1544. https://doi.org/10.1111/j.1365-2958.2005.04769.x

Oleszak, E. L., Chang, J. R., Friedman, H., Katsetos, C. D., & Platsoucas, C. D. (2004). Theiler's Virus Infection: a Model for Multiple Sclerosis. *Clinical Microbiology Reviews*, *17*(1), 174–207. https://doi.org/10.1128/CMR.17.1.174-207.2004

Omsland, A., & Heinzen, R. A. (2011). Life on the Outside: The Rescue of Coxiella burnetii from Its Host Cell. *Annual Review of Microbiology*, *65*(1), 111–128. https://doi.org/10.1146/annurev-micro-090110-102927

Openshaw, P. J. M., & Tregoning, J. S. (2005). Immune responses and disease enhancement during respiratory syncytial virus infection. *Clinical Microbiology Reviews*, *18*(3), 541–55. https://doi.org/10.1128/CMR.18.3.541-555.2005

Ortega, Y. R., & Sanchez, R. (2010). Update on Cyclospora cayetanensis, a Food-Borne and Waterborne Parasite. *Clinical Microbiology Reviews*, *23*(1), 218–234. https://doi.org/10.1128/CMR.00026-09

Ortín, J., & Parra, F. (2006). Structure and Function of RNA Replication. *Annual Review of Microbiology*, *60*(1), 305–326. https://doi.org/10.1146/annurev.micro.60.080805.142248

Ottenhoff, T. H. M. (2012). New pathways of protective and pathological host defense to mycobacteria. *Trends in Microbiology*, *20*(9), 419–28. https://doi.org/10.1016/j.tim.2012.06.002

Pande, S. G., Pagliai, F. A., Gardner, C. L., Wrench, A., Narvel, R., Gonzalez, C. F., & Lorca, G. L. (2011). Lactobacillus brevis responds to flavonoids through KaeR, a LysR-type of transcriptional regulator. *Molecular Microbiology*, *81*(6), 1623–1639. https://doi.org/10.1111/j.1365-2958.2011.07796.x

Park, Y., Herbert, E. E., Cowles, C. E., Cowles, K. N., Menard, M. L., Orchard, S. S., & Goodrich-Blair, H. (2007). Clonal variation in Xenorhabdus nematophila virulence and suppression of Manduca sexta immunity. *Cellular Microbiology*, *9*(3), 645–56. https://doi.org/10.1111/j.1462-5822.2006.00815.x

Paterson, D. L., & Bonomo, R. A. (2005). Extended-Spectrum -Lactamases: a Clinical Update. *Clinical Microbiology Reviews*, *18*(4), 657–686. https://doi.org/10.1128/CMR.18.4.657-686.2005

Pavlova, S. I., Kilic, A. O., Kilic, S. S., So, J.-S., Nader-Macias, M. E., Simoes, J. A., & Tao, L. (2002). Genetic diversity of vaginal lactobacilli from women in different countries based on 16S rRNA gene sequences. *Journal of Applied Microbiology*, *92*(3), 451–459. https://doi.org/10.1046/j.1365-2672.2002.01547.x

Pelosof, L. C., Davis, P. H., Zhang, Z., Zhang, X., & Stanley, S. L. (2006). Co-ordinate but disproportionate activation of apoptotic, regenerative and inflammatory pathways characterizes the liver response to acute amebic infection. *Cellular Microbiology*, *8*(3), 508–22. https://doi.org/10.1111/j.1462-5822.2005.00642.x

Peng, F., Wang, Y., Sun, F., Liu, Z., Lai, Q., & Shao, Z. (2008). A novel lipopeptide produced by a Pacific Ocean deep-sea bacterium, Rhodococcus sp. TW53. *Journal of Applied Microbiology*, *105*(3), 698–705. https://doi.org/10.1111/j.1365-2672.2008.03816.x

Pesavento, C., & Hengge, R. (2009). Bacterial nucleotide-based second messengers. *Current Opinion in Microbiology*, *12*(2), 170–6. https://doi.org/10.1016/j.mib.2009.01.007
Pfaller, M. A., & Diekema, D. J. (2007). Epidemiology of invasive candidiasis: a persistent public health problem. *Clinical Microbiology Reviews*, *20*(1), 133–63. https://doi.org/10.1128/CMR.00029-06
Pfeiler, E. A., & Klaenhammer, T. R. (2007). The genomics of lactic acid bacteria. *Trends in Microbiology*, *15*(12), 546–553. https://doi.org/10.1016/j.tim.2007.09.010
Pichoff, S., & Lutkenhaus, J. (2005). Tethering the Z ring to the membrane through a conserved membrane targeting sequence in FtsA. *Molecular Microbiology*, *55*(6), 1722–1734. https://doi.org/10.1111/j.1365-2958.2005.04522.x
Pillich, H., Loose, M., Zimmer, K.-P., & Chakraborty, T. (2012). Activation of the unfolded protein response by Listeria monocytogenes. *Cellular Microbiology*, *14*(6), 949–64. https://doi.org/10.1111/j.1462-5822.2012.01769.x
Potrykus, K., & Cashel, M. (2008). (p)ppGpp: Still Magical? *Annual Review of Microbiology*, *62*(1), 35–51. https://doi.org/10.1146/annurev.micro.62.081307.162903
Prato, M., Gallo, V., Giribaldi, G., Aldieri, E., & Arese, P. (2010). *Role of the NF-κB transcription pathway in the haemozoin- and 15-HETE-mediated activation of matrix metalloproteinase-9 in human adherent monocytes*. *Cellular Microbiology* (Vol. 12). https://doi.org/10.1111/j.1462-5822.2010.01508.x
Price, C. E., Zeyniyev, A., Kuipers, O. P., & Kok, J. (2012). From meadows to milk to mucosa – adaptation of Streptococcus and Lactococcus species to their nutritional environments. *FEMS Microbiology Reviews*, *36*(5), 949–971. Retrieved from http://femsre.oxfordjournals.org/content/36/5/949.abstract
Price, P., Murdoch, D. M., Agarwal, U., Lewin, S. R., Elliott, J. H., & French, M. A. (2009). Immune Restoration Diseases Reflect Diverse Immunopathological Mechanisms. *Clinical Microbiology Reviews*, *22*(4), 651–663. https://doi.org/10.1128/CMR.00015-09
Puerta-Fernández, E., Romero-López, C., Barroso-delJesus, A., & Berzal-Herranz, A. (2003). Ribozymes: recent advances in the development of RNA tools. *FEMS Microbiology Reviews*, *27*(1), 75–97. Retrieved from http://femsre.oxfordjournals.org/content/27/1/75.abstract
Quardokus, E. M., & Brun, Y. V. (2003). Cell cycle timing and developmental checkpoints in Caulobacter crescentus. *Current Opinion in Microbiology*, *6*(6), 541–549. https://doi.org/10.1016/j.mib.2003.10.013
Rathbone, D. A., & Bruce, N. C. (2002). Microbial transformation of alkaloids. *Current Opinion in Microbiology*, *5*(3), 274–281. https://doi.org/10.1016/S1369-5274(02)00317-X
Rawlings, D. E. (2002). Heavy Metal Mining Using Microbes. *Annual Review of Microbiology*, *56*(1), 65–91. https://doi.org/10.1146/annurev.micro.56.012302.161052
Reams, A. B., & Neidle, E. L. (2004). Selection for Gene Clustering by Tandem Duplication. *Annual Review of Microbiology*, *58*(1), 119–142. https://doi.org/10.1146/annurev.micro.58.030603.123806
Reinhart, K., Bauer, M., Riedemann, N. C., & Hartog, C. S. (2012). New approaches to sepsis: molecular diagnostics and biomarkers. *Clinical Microbiology Reviews*, *25*(4), 609–34. https://doi.org/10.1128/CMR.00016-12
Rensing, C., & Grass, G. (2003). Escherichia coli mechanisms of copper homeostasis in a changing environment. *FEMS Microbiology Reviews*, *27*(2–3), 197–213. Retrieved from http://femsre.oxfordjournals.org/content/27/2-3/197.abstract
Richards, J., Mehta, P., & Karzai, A. W. (2006). RNase R degrades non-stop mRNAs selectively in an SmpB-tmRNA-dependent manner. *Molecular Microbiology*, *62*(6), 1700–1712. https://doi.org/10.1111/j.1365-2958.2006.05472.x

Rinttilä, T., Kassinen, A., Malinen, E., Krogius, L., & Palva, A. (2004). Development of an extensive set of 16S rDNA-targeted primers for quantification of pathogenic and indigenous bacteria in faecal samples by real-time PCR. *Journal of Applied Microbiology*, *97*(6), 1166–1177. https://doi.org/10.1111/j.1365-2672.2004.02409.x

Ritt, J.-F., Remize, F., Grandvalet, C., Guzzo, J., Atlan, D., & Alexandre, H. (2009). Peptidases specific for proline-containing peptides and their unusual peptide-dependent regulation in Oenococcus oeni. *Journal of Applied Microbiology*, *106*(3), 801–813. https://doi.org/10.1111/j.1365-2672.2008.04032.x

Roberts, A. P., & Mullany, P. (2009). A modular master on the move: the Tn916 family of mobile genetic elements. *Trends in Microbiology*, *17*(6), 251–258. https://doi.org/10.1016/j.tim.2009.03.002

Roberts, J. W., Shankar, S., & Filter, J. J. (2008). RNA Polymerase Elongation Factors. *Annual Review of Microbiology*, *62*(1), 211–233. https://doi.org/10.1146/annurev.micro.61.080706.093422

Rodriguez, M., & Fishman, J. A. (2004). Prevention of infection due to Pneumocystis spp. in human immunodeficiency virus-negative immunocompromised patients. *Clinical Microbiology Reviews*, *17*(4), 770–82, https://doi.org/10.1128/CMR.17.4.770-782.2004

Romberg, L., & Levin, P. A. (2003). Assembly Dynamics of the Bacterial Cell Division Protein FtsZ: Poised at the Edge of Stability. *Annual Review of Microbiology*, *57*(1), 125–154. https://doi.org/10.1146/annurev.micro.57.012903.074300

Rossi, M.-S., Paquelin, A., Ghigo, J. M., & Wandersman, C. (2003). Haemophore-mediated signal transduction across the bacterial cell envelope in Serratia marcescens: the inducer and the transported substrate are different molecules. *Molecular Microbiology*, *48*(6), 1467–1480. https://doi.org/10.1046/j.1365-2958.2003.03516.x

Rubin-de-Celis, S. S. C., Uemura, H., Yoshida, N., & Schenkman, S. (2006). Expression of trypomastigote trans-sialidase in metacyclic forms of Trypanosoma cruzi increases parasite escape from its parasitophorous vacuole. *Cellular Microbiology*, *8*(12), 1888–98. https://doi.org/10.1111/j.1462-5822.2006.00755.x

Ryan, R. P., & Dow, J. M. (2011). Communication with a growing family: diffusible signal factor (DSF) signaling in bacteria. *Trends in Microbiology*, *19*(3), 145–52. https://doi.org/10.1016/j.tim.2010.12.003

Saguir, F. M., Loto Campos, I. E., & Manca de Nadra, M. C. (2008). Utilization of amino acids and dipeptides by Lactobacillus plantarum from orange in nutritionally stressed conditions. *Journal of Applied Microbiology*, *104*(6), 1597–1604. https://doi.org/10.1111/j.1365-2672.2007.03708.x

Saliba, A. M., Nascimento, D. O., Silva, M. C. A., Assis, M. C., Gayer, C. R. M., Raymond, B., Plotkowski, M. C. (2005). Eicosanoid-mediated proinflammatory activity of Pseudomonas aeruginosa ExoU. *Cellular Microbiology*, *7*(12), 1811–1822. https://doi.org/10.1111/j.1462-5822.2005.00635.x

Samal, J., Kandpal, M., & Vivekanandan, P. (2012). Molecular mechanisms underlying occult hepatitis B virus infection. *Clinical Microbiology Reviews*, *25*(1), 142–63. https://doi.org/10.1128/CMR.00018-11

Sano, D., Wada, K., Imai, T., Masago, Y., & Omura, T. (2010). Norovirus-binding proteins recovered from activated sludge micro-organisms with an affinity to a noroviral capsid peptide. *Journal of Applied Microbiology*, *109*(6), 1923–1928. https://doi.org/10.1111/j.1365-2672.2010.04821.x

Satchell, K. J. F. (2011). Structure and Function of MARTX Toxins and Other Large Repetitive RTX Proteins. *Annual Review of Microbiology*, *65*(1), 71–90. https://doi.org/10.1146/annurev-micro-090110-102943

Scheidegger, F., Quebatte, M., Mistl, C., & Dehio, C. (2011). *The Bartonella henselae VirB/Bep system interferes with vascular endothelial growth factor (VEGF) signalling in human vascular*

endothelial cells. *Cellular Microbiology* (Vol. 13). https://doi.org/10.1111/j.1462-5822.2010.01545.x
Schildgen, O., Muller, A., Allander, T., Mackay, I. M., Volz, S., Kupfer, B., & Simon, A. (2008). Human Bocavirus: Passenger or Pathogen in Acute Respiratory Tract Infections? *Clinical Microbiology Reviews*, *21*(2), 291–304. https://doi.org/10.1128/CMR.00030-07
Schmidt, I., Sliekers, O., Schmid, M., Bock, E., Fuerst, J., Kuenen, J. G., … Strous, M. (2003). New concepts of microbial treatment processes for the nitrogen removal in wastewater. *FEMS Microbiology Reviews*, *27*(4), 481–492. https://doi.org/10.1016/S0168-6445(03)00039-1
Schneemann, A. (2006). The Structural and Functional Role of RNA in Icosahedral Virus Assembly. *Annual Review of Microbiology*, *60*(1), 51–67. https://doi.org/10.1146/annurev.micro.60.080805.142304
Schoolnik, G. (2002). Functional and comparative genomics of pathogenic bacteria. *Current Opinion in Microbiology*, *5*(1), 20–26. https://doi.org/10.1016/S1369-5274(02)00280-1
Schujman, G. E., & Mendoza, D. de. (2005). Transcriptional control of membrane lipid synthesis in bacteria. *Current Opinion in Microbiology*, *8*(2), 149–153. https://doi.org/10.1016/j.mib.2005.02.003
Schuster, F. L., & Ramirez-Avila, L. (2008). Current World Status of Balantidium coli. *Clinical Microbiology Reviews*, *21*(4), 626–638. https://doi.org/10.1128/CMR.00021-08
Scollard, D. M., Adams, L. B., Gillis, T. P., Krahenbuhl, J. L., Truman, R. W., & Williams, D. L. (2006). The continuing challenges of leprosy. *Clinical Microbiology Reviews*, *19*(2), 338–81. https://doi.org/10.1128/CMR.19.2.338-381.2006
Scott, J. R., & Barnett, T. C. (2006). Surface Proteins of Gram-Positive Bacteria and How They Get There. *Annual Review of Microbiology*, *60*(1), 397–423. https://doi.org/10.1146/annurev.micro.60.080805.142256
Sebaihia, M., Bentley, S., Thomson, N., Holden, M., & Parkhill, J. (2002). Tales of the unexpected. *Trends in Microbiology*, *10*(6), 261–262. https://doi.org/10.1016/S0966-842X(02)02379-X
Segal, S., & Hill, A. V. . (2003). Genetic susceptibility to infectious disease. *Trends in Microbiology*, *11*(9), 445–448. https://doi.org/10.1016/S0966-842X(03)00207-5
Seiboth, B., Karimi, R. A., Phatale, P. A., Linke, R., Hartl, L., Sauer, D. G., … Kubicek, C. P. (2012). The putative protein methyltransferase LAE1 controls cellulase gene expression in Trichoderma reesei. *Molecular Microbiology*, *84*(6), 1150–1164. https://doi.org/10.1111/j.1365-2958.2012.08083.x
Seligman, S. J., & Bucher, D. J. (2003). The importance of being outer: consequences of the distinction between the outer and inner surfaces of flavivirus glycoprotein E. *Trends in Microbiology*, *11*(3), 108–110. https://doi.org/10.1016/S0966-842X(03)00005-2
Shabala, L., Ross, T., McMeekin, T., & Shabala, S. (2006). Non-invasive microelectrode ion flux measurements to study adaptive responses of microorganisms to the environment. *FEMS Microbiology Reviews*, *30*(3), 472–486. https://doi.org/10.1111/j.1574-6976.2006.00019.x
Shankar, J., Restrepo, A., Clemons, K. V, & Stevens, D. A. (2011). Hormones and the resistance of women to paracoccidioidomycosis. *Clinical Microbiology Reviews*, *24*(2), 296–313. https://doi.org/10.1128/CMR.00062-10
Sharon, A., Finkelstein, A., Shlezinger, N., & Hatam, I. (2009). Fungal apoptosis: function, genes and gene function. *FEMS Microbiology Reviews*, *33*(5), 833–854. https://doi.org/10.1111/j.1574-6976.2009.00180.x
Sharp, S. E., & Elder, B. L. (2004). Competency assessment in the clinical microbiology laboratory. *Clinical Microbiology Reviews*, *17*(3), 681–94. https://doi.org/10.1128/CMR.17.3.681-694.2004

Shima, S., & Thauer, R. K. (2005). Methyl-coenzyme M reductase and the anaerobic oxidation of methane in methanotrophic Archaea. *Current Opinion in Microbiology*, *8*(6), 643–648. https://doi.org/10.1016/j.mib.2005.10.002

Shiomi, D., & Margolin, W. (2007). Dimerization or oligomerization of the actin-like FtsA protein enhances the integrity of the cytokinetic Z ring. *Molecular Microbiology*, *66*(6), 1396–1415. https://doi.org/10.1111/j.1365-2958.2007.05998.x

Shotland, Y., Krämer, H., & Groisman, E. A. (2003). The Salmonella SpiC protein targets the mammalian Hook3 protein function to alter cellular trafficking. *Molecular Microbiology*, *49*(6), 1565–1576. https://doi.org/10.1046/j.1365-2958.2003.03668.x

Smith, I. (2003). Mycobacterium tuberculosis Pathogenesis and Molecular Determinants of Virulence. *Clinical Microbiology Reviews*, *16*(3), 463–496. https://doi.org/10.1128/CMR.16.3.463-496.2003

Sokolova, O., Heppel, N., Jägerhuber, R., Kim, K. S., Frosch, M., Eigenthaler, M., & Schubert-Unkmeir, A. (2004). Interaction of Neisseria meningitidis with human brain microvascular endothelial cells: role of MAP- and tyrosine kinases in invasion and inflammatory cytokine release. *Cellular Microbiology*, *6*(12), 1153–66. https://doi.org/10.1111/j.1462-5822.2004.00422.x

Sologub, L., Kuehn, A., Kern, S., Przyborski, J., Schillig, R., & Pradel, G. (2011). Malaria proteases mediate inside-out egress of gametocytes from red blood cells following parasite transmission to the mosquito. *Cellular Microbiology*, *13*(6), 897–912. https://doi.org/10.1111/j.1462-5822.2011.01588.x

Son, S. W., Kim, H. Y., Choi, G. J., Lim, H. K., Jang, K. S., Lee, S. O., … Kim, J.-C. (2008). Bikaverin and fusaric acid from Fusarium oxysporum show antioomycete activity against Phytophthora infestans. *Journal of Applied Microbiology*, *104*(3), 692–698. https://doi.org/10.1111/j.1365-2672.2007.03581.x

Souza, T. C., Zacarías, M. F., Silva, A. M., Binetti, A., Reinheimer, J., Nicoli, J. R., & Vinderola, G. (2012). Cell viability and immunostimulating and protective capacities of Bifidobacterium longum 51A are differentially affected by technological variables in fermented milks. *Journal of Applied Microbiology*, *112*(6), 1184–1192. https://doi.org/10.1111/j.1365-2672.2012.05280.x

Spoering, A. L., & Gilmore, M. S. (2006). Quorum sensing and DNA release in bacterial biofilms. *Current Opinion in Microbiology*, *9*(2), 133–7. https://doi.org/10.1016/j.mib.2006.02.004

Springer, B., Sander, P., Sedlacek, L., Hardt, W.-D., Mizrahi, V., Schär, P., & Böttger, E. C. (2004). Lack of mismatch correction facilitates genome evolution in mycobacteria. *Molecular Microbiology*, *53*(6), 1601–1609. https://doi.org/10.1111/j.1365-2958.2004.04231.x

Srivastava, A., Talaue, M., Liu, S., Degen, D., Ebright, R. Y., Sineva, E., … Ebright, R. H. (2011). New target for inhibition of bacterial RNA polymerase: "switch region". *Current Opinion in Microbiology*, *14*(5), 532–43. https://doi.org/10.1016/j.mib.2011.07.030

Srivatsan, A., & Wang, J. D. (2008). Control of bacterial transcription, translation and replication by (p)ppGpp. *Current Opinion in Microbiology*, *11*(2), 100–5. https://doi.org/10.1016/j.mib.2008.02.001

Stahl, D. A., & de la Torre, J. R. (2012). Physiology and Diversity of Ammonia-Oxidizing Archaea. *Annual Review of Microbiology*, *66*(1), 83–101. https://doi.org/10.1146/annurev-micro-092611-150128

Starai, V., Takahashi, H., Boeke, J., & Escalante-Semerena, J. (2004). A link between transcription and intermediary metabolism: a role for Sir2 in the control of acetyl-coenzyme A synthetase. *Current Opinion in Microbiology*, *7*(2), 115–119. https://doi.org/10.1016/j.mib.2004.02.005

Stepanauskas, R. (2012). Single cell genomics: an individual look at microbes. *Current Opinion in Microbiology*, *15*(5), 613–20. https://doi.org/10.1016/j.mib.2012.09.001

Stevens, A., & Søgaard-Andersen, L. (2005). Making waves: pattern formation by a cell-surface-associated signal. *Trends in Microbiology*, *13*(6), 249–52. https://doi.org/10.1016/j.tim.2005.04.002

Subhadra, B., Hurwitz, I., Fieck, A., Rao, D. V. S., Subba Rao, G., & Durvasula, R. (2010). Development of paratransgenic Artemia as a platform for control of infectious diseases in shrimp mariculture. *Journal of Applied Microbiology*, *108*(3), 831–840. https://doi.org/10.1111/j.1365-2672.2009.04479.x

Takaya, A., Erhardt, M., Karata, K., Winterberg, K., Yamamoto, T., & Hughes, K. T. (2012). YdiV: a dual function protein that targets FlhDC for ClpXP-dependent degradation by promoting release of DNA-bound FlhDC complex. *Molecular Microbiology*, *83*(6), 1268–1284. https://doi.org/10.1111/j.1365-2958.2012.08007.x

Tamagnini, P., Leitão, E., Oliveira, P., Ferreira, D., Pinto, F., Harris, D. J., … Lindblad, P. (2007). Cyanobacterial hydrogenases: diversity, regulation and applications. *FEMS Microbiology Reviews*, *31*(6), 692–720. Retrieved from http://femsre.oxfordjournals.org/content/31/6/692.abstract

Tamma, P. D., Cosgrove, S. E., & Maragakis, L. L. (2012). Combination therapy for treatment of infections with gram-negative bacteria. *Clinical Microbiology Reviews*, *25*(3), 450–70. https://doi.org/10.1128/CMR.05041-11

Taylor, J. W., & Fisher, M. C. (2003). Fungal multilocus sequence typing — it's not just for bacteria. *Current Opinion in Microbiology*, *6*(4), 351–356. https://doi.org/10.1016/S1369-5274(03)00088-2

Taylor-Robinson, D., & Jensen, J. S. (2011). Mycoplasma genitalium: from Chrysalis to multicolored butterfly. *Clinical Microbiology Reviews*, *24*(3), 498–514. https://doi.org/10.1128/CMR.00006-11

Tebruegge, M., & Curtis, N. (2008). Epidemiology, etiology, pathogenesis, and diagnosis of recurrent bacterial meningitis. *Clinical Microbiology Reviews*, *21*(3), 519–37. https://doi.org/10.1128/CMR.00009-08

Teichmann, B., Labbé, C., Lefebvre, F., Bölker, M., Linne, U., & Bélanger, R. R. (2011). Identification of a biosynthesis gene cluster for flocculosin a cellobiose lipid produced by the biocontrol agent Pseudozyma flocculosa. *Molecular Microbiology*, *79*(6), 1483–1495. https://doi.org/10.1111/j.1365-2958.2010.07533.x

Tenaillon, O., Denamur, E., & Matic, I. (2004). Evolutionary significance of stress-induced mutagenesis in bacteria. *Trends in Microbiology*, *12*(6), 264–70. https://doi.org/10.1016/j.tim.2004.04.002

Teske, A. P. (2005). The deep subsurface biosphere is alive and well. *Trends in Microbiology*, *13*(9), 402–4. https://doi.org/10.1016/j.tim.2005.07.004

Thorsen, M., Jacobson, T., Vooijs, R., Navarrete, C., Bliek, T., Schat, H., & Tamás, M. J. (2012). Glutathione serves an extracellular defence function to decrease arsenite accumulation and toxicity in yeast. *Molecular Microbiology*, *84*(6), 1177–1188. https://doi.org/10.1111/j.1365-2958.2012.08085.x

Thoulouze, M.-I., & Alcover, A. (2011). Can viruses form biofilms? *Trends in Microbiology*, *19*(6), 257–262. https://doi.org/10.1016/j.tim.2011.03.002

Triantafilou, M., Uddin, A., Maher, S., Charalambous, N., Hamm, T. S. C., Alsumaiti, A., & Triantafilou, K. (2007). *Anthrax toxin evades Toll-like receptor recognition, whereas its cell wall components trigger activation via TLR2/6 heterodimers. Cellular microbiology* (Vol. 9). https://doi.org/10.1111/j.1462-5822.2007.01003.x

Trifonova, R., Postma, J., & Van Elsas, J. D. (2009). Interactions of plant-beneficial bacteria with the ascomycete Coniochaeta ligniaria. *Journal of Applied Microbiology*, *106*(6), 1859–1866. https://doi.org/10.1111/j.1365-2672.2009.04163.x

Tumpey, T. M., & Belser, J. A. (2009). Resurrected Pandemic Influenza Viruses. *Annual Review of Microbiology*, *63*(1), 79–98. https://doi.org/10.1146/annurev.micro.091208.073359

Turnbough, C. L. (2011). Regulation of gene expression by reiterative transcription. *Current Opinion in Microbiology*, *14*(2), 142–7. https://doi.org/10.1016/j.mib.2011.01.012

Ueno, N., Bratt, C. L., Rodriguez, N. E., & Wilson, M. E. (2009). *Differences in human macrophage receptor usage, lysosomal fusion kinetics and survival between logarithmic and metacyclic Leishmania infantum chagasi promastigotes*. *Cellular Microbiology* (Vol. 11). Blackwell Publishing Ltd. https://doi.org/10.1111/j.1462-5822.2009.01374.x

Van Baarlen, P., Van Belkum, A., Summerbell, R. C., Crous, P. W., & Thomma, B. P. H. J. (2007). Molecular mechanisms of pathogenicity: how do pathogenic microorganisms develop cross-kingdom host jumps? *FEMS Microbiology Reviews*, *31*(3), 239–277. https://doi.org/10.1111/j.1574-6976.2007.00065.x

van Crevel, R., Ottenhoff, T. H. M., & van der Meer, J. W. M. (2002). Innate Immunity to Mycobacterium tuberculosis. *Clinical Microbiology Reviews*, *15*(2), 294–309. https://doi.org/10.1128/CMR.15.2.294-309.2002

Van De Veerdonk, F. L., Netea, M. G., Joosten, L. A., Van Der Meer, J. W. M., & Kullberg, B. J. (2010). Novel strategies for the prevention and treatment of Candida infections: the potential of immunotherapy. *FEMS Microbiology Reviews*, *34*(6), 1063–1075. https://doi.org/10.1111/j.1574-6976.2010.00232.x

Van den Abbeele, P., Van de Wiele, T., Verstraete, W., & Possemiers, S. (2011). The host selects mucosal and luminal associations of coevolved gut microorganisms: a novel concept. *FEMS Microbiology Reviews*, *35*(4), 681–704. https://doi.org/10.1111/j.1574-6976.2011.00270.x

van den Burg, B. (2003). Extremophiles as a source for novel enzymes. *Current Opinion in Microbiology*, *6*(3), 213–218. https://doi.org/10.1016/S1369-5274(03)00060-2

Van Der Sar, A. M., Musters, R. J. P., Van Eeden, F. J. M., Appelmelk, B. J., Vandenbroucke-Grauls, C. M. J. E., & Bitter, W. (2003). Zebrafish embryos as a model host for the real time analysis of Salmonella typhimurium infections. *Cellular Microbiology*, *5*(9), 601–611. https://doi.org/10.1046/j.1462-5822.2003.00303.x

Van Dooren, G. G., Stimmler, L. M., & McFadden, G. I. (2006). Metabolic maps and functions of the Plasmodium mitochondrion. *FEMS Microbiology Reviews*, *30*(4), 596–630. https://doi.org/10.1111/j.1574-6976.2006.00027.x

Van Etten, J. L., Lane, L. C., & Dunigan, D. D. (2010). DNA Viruses: The Really Big Ones (Giruses). *Annual Review of Microbiology*, *64*(1), 83–99. https://doi.org/10.1146/annurev.micro.112408.134338

Van Lanen, S. G., & Shen, B. (2006). Microbial genomics for the improvement of natural product discovery. *Current Opinion in Microbiology*, *9*(3), 252–60. https://doi.org/10.1016/j.mib.2006.04.002

van Vliet, A. H. M., Ketley, J. M., Park, S. F., & Penn, C. W. (2002). The role of iron in Campylobacter gene regulation, metabolism and oxidative stress defense. *FEMS Microbiology Reviews*, *26*(2), 173–186. https://doi.org/10.1111/j.1574-6976.2002.tb00609.x

Vareille, M., Kieninger, E., Edwards, M. R., & Regamey, N. (2011). The Airway Epithelium: Soldier in the Fight against Respiratory Viruses. *Clinical Microbiology Reviews*, *24*(1), 210–229. https://doi.org/10.1128/CMR.00014-10

Vautrin, E., & Vavre, F. (2009). Interactions between vertically transmitted symbionts: cooperation or conflict? *Trends in Microbiology*, *17*(3), 95–99. https://doi.org/10.1016/j.tim.2008.12.002

Veening, J.-W., Smits, W. K., Hamoen, L. W., & Kuipers, O. P. (2006). Single cell analysis of gene expression patterns of competence development and initiation of sporulation

in Bacillus subtilis grown on chemically defined media. *Journal of Applied Microbiology*, *101*(3), 531–541. https://doi.org/10.1111/j.1365-2672.2006.02911.x
Veening, J.-W., Smits, W. K., & Kuipers, O. P. (2008). Bistability, Epigenetics, and Bet-Hedging in Bacteria. *Annual Review of Microbiology*, *62*(1), 193–210. https://doi.org/10.1146/annurev.micro.62.081307.163002
Vega, L. A., & Caparon, M. G. (2012). Cationic antimicrobial peptides disrupt the Streptococcus pyogenes ExPortal. *Molecular Microbiology*, *85*(6), 1119–1132. https://doi.org/10.1111/j.1365-2958.2012.08163.x
Viboud, G. I., & Bliska, J. B. (2005). YERSINIA OUTER PROTEINS: Role in Modulation of Host Cell Signaling Responses and Pathogenesis. *Annual Review of Microbiology*, *59*(1), 69–89. https://doi.org/10.1146/annurev.micro.59.030804.121320
Visvesvara, G. S. (2002). In Vitro Cultivation of Microsporidia of Clinical Importance. *Clinical Microbiology Reviews*, *15*(3), 401–413. https://doi.org/10.1128/CMR.15.3.401-413.2002
Voth, D. E., & Ballard, J. D. (2005). Clostridium difficile toxins: mechanism of action and role in disease. *Clinical Microbiology Reviews*, *18*(2), 247–63. https://doi.org/10.1128/CMR.18.2.247-263.2005
Walsh, T. R. (2006). Combinatorial genetic evolution of multiresistance. *Current Opinion in Microbiology*, *9*(5), 476–82. https://doi.org/10.1016/j.mib.2006.08.009
Weaver, D., Karoonuthaisiri, N., Tsai, H.-H., Huang, C.-H., Ho, M.-L., Gai, S., … Kao, C. M. (2004). Genome plasticity in Streptomyces: identification of 1 Mb TIRs in the S. coelicolor A3(2) chromosome. *Molecular Microbiology*, *51*(6), 1535–1550. https://doi.org/10.1111/j.1365-2958.2003.03920.x
Webb, C. T., Heinz, E., & Lithgow, T. (2012). Evolution of the β-barrel assembly machinery. *Trends in Microbiology*, *20*(12), 612–20. https://doi.org/10.1016/j.tim.2012.08.006
Węgrzyn, G., & Czyż, A. (2003). Detection of mutagenic pollution of natural environment using microbiological assays. *Journal of Applied Microbiology*, *95*(6), 1175–1181. https://doi.org/10.1046/j.1365-2672.2003.02121.x
Werner, F. (2008). Structural evolution of multisubunit RNA polymerases. *Trends in Microbiology*, *16*(6), 247–50. https://doi.org/10.1016/j.tim.2008.03.008
Wexler, H. M. (2007). Bacteroides: the good, the bad, and the nitty-gritty. *Clinical Microbiology Reviews*, *20*(4), 593–621. https://doi.org/10.1128/CMR.00008-07
Wileman, T. (2007). Aggresomes and Pericentriolar Sites of Virus Assembly: Cellular Defense or Viral Design? *Annual Review of Microbiology*, *61*(1), 149–167. https://doi.org/10.1146/annurev.micro.57.030502.090836
Withey, J. H., & Friedman, D. I. (2003). A Salvage Pathway for Protein Synthesis: tmRNA and Trans-Translation. *Annual Review of Microbiology*, *57*(1), 101–123. https://doi.org/10.1146/annurev.micro.57.030502.090945
Wong, K.-W., & Jacobs, W. R. (2011). Critical role for NLRP3 in necrotic death triggered by Mycobacterium tuberculosis. *Cellular Microbiology*, *13*(9), 1371–84. https://doi.org/10.1111/j.1462-5822.2011.01625.x
Wright, H. T., & Reynolds, K. A. (2007). Antibacterial targets in fatty acid biosynthesis. *Current Opinion in Microbiology*, *10*(5), 447–53. https://doi.org/10.1016/j.mib.2007.07.001
Xie, Z., Okinaga, T., Niu, G., Qi, F., & Merritt, J. (2010). Identification of a novel bacteriocin regulatory system in Streptococcus mutans. *Molecular Microbiology*, *78*(6), 1431–1447. https://doi.org/10.1111/j.1365-2958.2010.07417.x
Yakandawala, N., Gawande, P. V, LoVetri, K., & Madhyastha, S. (2007). Effect of ovotransferrin, protamine sulfate and EDTA combination on biofilm formation by

catheter-associated bacteria. *Journal of Applied Microbiology*, *102*(3), 722–727. https://doi.org/10.1111/j.1365-2672.2006.03129.

Appendix C: Results of interdisciplinary comparison

Table A2

Pattern frequencies for three subcorpora and result of cross-corpora comparison for absorb

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
1	[[Physical_Object 1]] absorbs [[Stuff]] or [[Physical_Object 2]] <Stuff or Physical_Object 2 becomes part of Physical Object 1 by means of a natural process>	2	3.23	2.80	-2.02	0	0.00	9.84**	5.36†	4	15.38	3.52	-0.96
2	[[Human]] or [[Animal]] or [[Plant]] or [[Body_Part]] or [[MB_Entity]] absorbs [[Stuff]] <Human or Animal or Plant or Body_Part takes Stuff = Nutrient or Liquid into the body so that Stuff or Liquid becomes an integral part of Human or Animal or Plant or Cell or Cell part>	1	1.61	1.40	-3.43	0	0.00	24.61****	20.12†	10	38.46	18.38****	13.91†
3	[[Physical_Object]] absorbs [[Radiation]] <Physical_Object receives Radiation in large quantities so that Radiation becomes part of Physical Object>	0	0.00	0.00	-4.83	0	0.00	7.38**	2.89†	3	11.54	7.31**	2.84†
4	[[Physical_Object]] or [[Stuff]] absorbs [[Energy]] <Physical_Object receives and utilizes or stores Energy>	2	3.23	2.80	-2.02	0	0.00	12.30***	7.82†	5	19.23	5.21*	0.74
5	[[Human_Group 1]] or [[Institution 1]] absorbs [[Human_Group 2]] or [[Institution 2]] <Human_Group 2 or Institution 2 becomes a part of Human Group 1 or Institution 1>	18	29.03	8.01**	3.18†	5	7.94	3.46	-1.03	0	0.00	12.60***	8.13†
6	[[Eventuality 1]] or [[Institution]] absorbs [[Eventuality 2]] or [[Institution]] <Eventuality 2 becomes a part of Eventuality 1 or Institution>	5	8.06	2.21	-2.62	11	17.46	7.60**	3.11†	0	0.00	3.50	-0.98

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A2 (cont.)

Pattern frequencies for three subcorpora and result of cross-corpora comparison for absorb

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
7	[[Human]] or [[Abstract_Entity]] absorbs [[Asset]] <Human or Abstract_Entity uses Asset - This is not typically considered a good use of [[Asset]]>	5	8.06	7.01**	2.18	0	0.00	0.00	-4.49	0	0.00	3.50	-0.98
8	[[Institution]] or [[Human]] absorbs [[Deficit]] <Business | Economics. Jargon. Institution or Human = Business Person is robust or wealthy enough to bear the adverse effects of Deficit>	1	1.61	4.97*	0.14	7	11.11	4.83*	0.35	0	0.00	0.70	-3.78
9	[[Human]] or [[Institution]] absorbs [[Anything]] or [[Attitude]] <Human learns about, understands, and accepts Anything = Concept or Attitude>	25	40.32	2.50	-2.33	38	60.32	26.25***	21.77†	0	0.00	17.51****	13.03†
10	[[Human]] is or become absorbed in or by [[Activity]] or in or by [[Anything]] <Human is greatly interested in Activity or Anything - This is typically to the exclusion of any other [[Activity | Anything]]>	2	3.23	0.00	-4.83	2	3.17	1.38	-3.11	0	0.00	1.40	-3.08
11	[[Abstract_Entity]] absorbs [[property]] <Abstract_Entity is influenced by property of another abstract entity>	1	1.61	1.40	-3.43	0	0.00	0.00	-4.49	0	0.00	0.70	-3.78
12	[[Stuff]] absorbs {strongly} <stuff = liquid or other molecule enters MB entity via a microbiological process>	0	0.00	0.00	-4.83	0	0.00	4.92*	0.43	2	7.69	4.87*	0.40
13	[[Liquid 1]] is absorbed with [[Liquid 2]] <Human adds Liquid 2 to Liquid 2 = chemical reagent, to test if reaction occurs>	0	0.00	0.00	-4.83	0	0.00	4.92*	0.43	2	7.69	4.87*	0.40
	TOTAL	62				63				26			

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A3

Pattern frequencies for three subcorpora and result of cross-corpora comparison for accelerate

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
1	[[Anything]] accelerates [[Process]] or [[Pace]] <Anything causes Pace of Process to proceed more quickly>	16	36.63	4.11*	-0.35	29	67.44	1.56	-2.92	41	91.11	3.06	-1.43
2	[[Process]] or [[Pace]] accelerates <Pace of Process proceeds more and more quickly>	27	61.36	3.90*	-0.57	14	32.56	6.35*	1.87	4	8.89	19.65****	15.17†
3	[[Human]] accelerates [[Proposition]] <Human expresses Speech Act with greater urgency>	1	2.27	1.36	-3.10	0	0	1.34	-3.14	0	0	1.36	-3.12
	TOTAL	44				43					45		

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A4

Pattern frequencies for three subcorpora and result of cross-corpora comparison for accept

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
1	[[Human]] or [[Institution]] accepts [[Proposition]] or [[Concept]] or [[Eventuality]] <Human or Institution agrees that Proposition or Concept or Eventuality is correct and does not need to be contested>	213	71.00	0.00	-6.12	112	71.34	7.91*	2.51†	26	40	8.85**	2.95†
2	[[Human]] or [[Institution]] accepts <Human or Institution agrees that that [clause] is true or correct>	21	7.00	0.22	-5.9	13	8.28	19.99****	14.59†	24	36.92	28.88****	22.98†
3	[[Human]] or [[Institution]] accepts {responsibility} (for Eventuality or for Entity) <Human or Institution agrees that he, she, or it is or will be responsible for Eventuality or Entity>	7	2.33	0.02	-6.11	4	2.55	2.77	-2.63	0	0.00	2.75	-3.15
4	[[Human 1]] accepts [[Entity]] or [[Money]] <Human 1 consents to receive Entity = Valuable or Money as a gift (or bribe) or as part of buisiness transaction from Human 2>	13	4.33	0.00	-6.12	7	4.46	4.85*	-0.55	0	0.00	5.1*	-0.8
5	[[Human]] accepts [[Offer]] <Human 1 agrees to act on the Offer of Human 2>	5	1.67	9.25**	3.13†	12	7.64	8.31**	2.91†	0	0.00	1.96	-3.94
6	[[Human 1]] accepts {resignation} <Human 1 = Authority receives and consents to resignation of Human 2 = Employee>	0	0.00	2.14	-3.99	1	0.64	0.69	-4.71	0	0.00	0	-5.9

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A4 (cont.)

Pattern frequencies for three subcorpora and result of cross-corpora comparison for accept

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
7	[[Human]] accepts [[Human Role]] <Human agrees to undertake the work specified as required to fulfil appointment or job>	11	3.67	4.51*	-1.61	1	0.64	0.69	-4.71	0	0.00	4.31*	-1.59
8	[[Human 1]] or [[Institution 1]] accepts [[Human 2]] or [[Institution 2]] as [[Human_Role]] or as [[Institution_Role]] <Human 1 or Institution 1 acknowledges that Human 2 or Institution 2 is suitable for Human_Role or Institution_Role>	17	5.67	8.72**	2.6†	1	0.64	0.69	-4.71	0	0.00	6.67**	0.77
9	[[Human_Group 1]] or [[Human 1]] accepts [[Human 2]] or [[Human_Group 2]] <Human_Group 1 or Human 1 agrees to allow Human 2 or Human Group 2 to become a member of or be respected by Human Group 1>	12	4.00	5.19*	-0.94	1	0.64	0.38	-5.03	1	1.54	1.11	-4.79
10	[[Location]] accepts [[Human Group]] <Location is suitable for Human Group>	1	0.33	0.84	-5.28	0	0.00	0	-5.9	0	0.00	0.39	-5.51
11	[[MB Entity]] accepts {task} <MB Entity= part of cell undertakes task in cell>	0	0.00	0.00	-6.12	0	0.00	2.46	-2.95	1	1.54	3.45	-2.45
12	[[MB Entity1]] accepts [[MB Entity 2]] <MB Entity 1 takes MB Entity 2 which becomes an integral part of MB Entity 1>	0	0.00	0.00	-6.12	0	0.00	22.11****	16.71†	9	13.85	31.06****	25.16†

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A4 (cont.)

Pattern frequencies for three subcorpora and result of cross-corpora comparison for accept

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
13	[[MB_Entity 1]] accepts [[MB_Entity 2] as [[MB_Role]] <MB_Entity 2 plays a role in the normal functioning of MB_Entity 1>	0	0.00	0.00	-6.12	0	0.00	4.91*	-0.49	2	3.08	6.9**	1
14	[[Device]] accepts [[Stuff]] <Device is capable of processing Stuff = biological sample>	0	0.00	0.00	-6.12	0	0.00	2.46	-2.95	1	1.54	3.45	-2.45
15	[[MB Entity]] accepts {topology} <MB Entity=gene occurs in a given spacial configuration>	0	0.00	0.00	-6.12	0	0.00	2.46	-2.95	1	1.54	3.45	-2.45
16	[[Concept 1]] accepts [[Concept 2]] or [[Eventuality]] <Concept 1 is compatible with Concept 2 or Eventuality>	0	0.00	4.27*	-1.85	2	1.27	1.39	-4.02	0	0.00	0	-5.9
17	[[Human]] or [[Institution]] accepts [[Document]] <Human or Institution agrees that Document meets the standards required for publication>	0	0.00	6.41*	0.29	3	1.91	2.08	-3.32	0	0.00	0	-5.9
	TOTAL	300				157				65			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A5

Pattern frequencies for three subcorpora and result of cross-corpora comparison for accommodate

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
1	[[Human 1]] or [[Institution]] or [[Location]] accommodates [[Human 2]] or [[Human_Group]] or [[Animal_Group]] <Human 1 or Institution provides suitable space in Location for Human 2 or Human_Group or Animal_Group to live or work in>	10	10.53	6.05*	0.95	1	1.43	0.85	-3.82	0	0.00	6.58*	1.7
2	[[Location]] or [[Building]] or [[Machine]] accommodates [[Physical_Object]] or [[MB Entity]] or [[Activity]] <Location, in case of MB Entity part of cell, or Building or Machine or provides enough space for Physical_Object or MB_Entity to be located or for Activity to take place>	12	12.63	7.91**	2.81†	1	1.43	37.31****	32.64†	21	56.76	18.05****	13.17†
3	[[Human]] or [[Institution]] accommodates [[Eventuality]] <Human or Institution deliberately adapts the way they think or behave to take account of Eventuality>	47	49.47	0.45	-4.66	40	57.14	26.67****	22†	1	2.70	23.74****	18.86†
4	[[Human 1]] accommodates [[Human 2]] <Human 1 makes a special effort meet the needs or wishes of Human 2>	4	4.21	0.00	-5.11	3	4.29	2.55	-2.13	0	0.00	2.63	-2.25
5	[[Human]] or [[Institution]] accommodates [[Self]]to [[Abstract_Entity]] or to [[Eventuality]] <Human or Institution deliberately adapts the way they think or behave to take account of Abstract_Entity or Eventuality>	7	7.37	1.62	-3.48	2	2.86	1.7	-2.98	0	0.00	4.6*	-0.28

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A5 (cont.)

Pattern frequencies for three subcorpora and result of cross-corpora comparison for accommodate

PT #	Pattern <implicature>	History Freq.	History %	Hist/ Man LL	BIC	Management Freq.	Management %	Man/ Micro LL	BIC	Microbiology Freq.	Microbiology %	Micro/ Hist LL	BIC
6	[[Animate]] or [[Plant]] or [[MB Entity]] accommodates (Eventuality) <Animate or Plant or MB Entity makes adjustments in behaviour or structure in order to cope with Eventuality>	0	0.00	0.00	-5.11	0	0.00	12.74***	8.07†	6	16.22	15.26****	10.38†
7	[[Human]] or [[Activity]] accommodates [[Eventuality 1]] to [[Eventuality 2]] <Human or Activity adjusts Eventuality 1 in order to cope with Eventuality 2>	3	3.16	3.31	-1.79	0	0.00	2.55	-2.13	0	0.00	1.97	-2.91
8	[[Concept]] or [[State of affairs]] accommodates [[Eventuality]] <Concept or state of affairs provides the conditions necessary for Eventuality to be realised>	12	12.63	7.69**	2.58†	23	32.86	14.03**	9.36†	9	24.32	2.11	-2.78
	TOTAL	95				70				37			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A6

Pattern frequencies for three subcorpora and result of cross-corpora comparison for accompany

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
1	[[Human 1]] or [[Vehicle 1]] or [[Human_Group 1]] or [[MB Entity1]] accompanies [[Human 2]] or [[Animal]] or [[Vehicle 2]] or [[Human_Group 2]] or [[MB Entity 2]] (Direction) <Human 1 or Human_Group 1 or Vehicle 1 or MB_Entity 2 goes with Human 2 or Human Group 2 or Animal or Vehicle 2 or MB Entity 2, either when travelling Direction or (if Human 1, 2) also at a social function>	51	34.00	27.04****	21.64†	2	2.82	1.68	-3.44	7	7.37	20.58****	15.08†
2	[[Document 1]] accompanies [[Document 2]] or [[Artifact]] <Document 1 is transmitted or published at the same time as Document 2 or Artifact = Valuable, in order to explain, verify, or supplement it>	26	17.33	4.11*	-1.28	5	7.04	1.26	-3.86	3	3.16	11.91***	6.4†
3	[[Eventuality 1]] accompanies [[Eventuality 2]] <Eventuality 1 occurs at the same time as Eventuality 2 and may be associated with it in some way>	69	46.00	12.28***	6.88†	61	85.92	0.06	-5.05	85	89.47	16.94****	11.44†
4	[[Human 1]] or [[Musical_Performance 1]] accompanies [[Human 2]] or [[Musical_Performance 2]] <Human 1 = Musician plays Musical_Performance 1 at the same time as Musical Performance 2 is played by Human 2 = Musician>	4	2.67	3.10	-2.3	0	0.00	0	-5.11	0	0.00	3.92*	-1.58
5	[[Abstract Entity 1]] accompanies [[Abstract Entity 2]] <Abstract Enitity 1 forms an important part of Abstract Entity 2>	0	0.00	6.81*	1.41	3	4.23	5.1*	-0.02	0	0.00	0	-5.5
	TOTAL	150				71				95			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A7

Pattern frequencies for three subcorpora and result of cross-corpora comparison for accomplish

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
1	[[Human]] or [[Institution]] accomplishes [[Activity]]or [[Eventuality]] <Human successfully completes Activity = Task or brings about Eventuaity>	73	100.00	0.00	-5.35	138	100.00	0.00	-5.32	67	100.00	0.00	-4.95
	TOTAL	73				138				67			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A8

Pattern frequencies for three subcorpora and result of cross-corpora comparison for account

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
1	[[Entity 1]] or [[Eventuality 1]] accounts for [[Numerical_Value]] of [[Entity 2]] or of [[Eventuality 2]] <Entity 1 or Eventuality 1 makes up a certain Numerical_Value = Percentage of the total number of Entity 2 or of Eventuality 2>	27	24.32	5.53*	-0.49	40	13.33	9.53**	3.53†	29	29.00	0.43	-4.92
2	[[Eventuality 1]] accounts for [[Eventuality 2]] <Eventuality 1 was a factor in causing Eventuality 2 to happen>	45	40.54	0.02	-5.99	125	41.67	1.16	-4.84	50	50.00	1.04	-4.31
3	[[Human]]or [[Concept]] accounts for [[Eventuality]] <Human or Concept explains or justifies Eventuality>	38	34.23	1.70	-4.33	130	43.33	74.8****	68.81†	0	0.00	48.82****	43.47 †
4	[[Human]] accounts for [[Entity]] <Human is able to explain where Entity is or how it was used>	1	0.90	0.36	-5.66	5	1.67	2.88	-3.11	0	0.00	1.28	-4.07
5	[[MB Entity]] accounts for [[Eventuality]] <the presence of MB_Entity explains Eventuality>	0	0.00	0.00	-6.02	0	0.00	58.22****	52.23†	21	21.00	31.36****	26.01 †
	TOTAL	111				300				100			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A9

Pattern frequencies for three subcorpora and result of cross-corpora comparison for accumulate

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Frequency	%			Frequency	%			Frequency	%		
1	[[Human]] or [[Institution]] or [[MB Entity]] accumulates [[Entity]] <Human or Institution acquires or experiences an amount of Entity over a period of time>	24	57.14	0.01	-5.02	62	55.86	17.48****	11.93†	34	23.45	10.31**	5.08†
2	[[Body_Part]] accumulates [[Stuff]] <Stuff collects and builds up in Body_Part - Typically, this is undesirable>	0	0.00	0.00	-5.03	0	0.00	1.14	-4.41	1	0.69	0.51	-4.72
3	[[Stuff]] accumulates in [[Location]] or on [[Location]] <A quantity of Stuff collects and builds up in Location or on Location>	0	0.00	0.00	-5.03	0	0.00	81.86****	76.31†	72	49.66	36.63****	31.4†
4	[[Abstract Entity]] or [[Document]] or [[MB Entity]] accumulates <The amount of Abstract_Entity available or experienced increases>	18	42.86	0.01	-5.02	49	44.14	5.88*	0.34	38	26.21	2.77	-2.46
	TOTAL	42				111				145			

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A10

Pattern frequencies for three subcorpora and result of cross-corpora comparison for act

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
1	[[Human]] or [[Institution]] or [[Animal]] or [[Machine]] or [[MB unit]] acts <Human or Institution or Animal or Machine = Agent performs a motivated Action>	50	16.67	0.01	-6.39	51	17.00	0.52	-5.88	44	14.67	0.38	-6.01
2	[[Human]] or [[Institution]] or [[MB unit]] acts [[Manner]] <Human or Institution behaves in the manner specified>	115	38.33	1.45	-4.95	134	44.67	16.88****	10.49†	75	25.00	8.48**	2.09†
3	[[Entity 1]] acts as or like [[Anything]] <In a particular situation, Entity 1 performs the role or function specified>	99	33.00	1.41	-4.99	83	27.67	12.95****	6.56†	136	45.33	5.85*	-0.55
4	[[Human 1]] or [[Institution 1]] acts for [[Human 2]] or for [[Institution 2]] <Human 1 = Lawyer or Agent or Institution 1 is employed to perform certain tasks for Human 2 or Institution 2>	2	0.67	2.77	-3.62	0	0.00	0	-6.4	0	0.00	2.77	-3.62
5	[[Human 1]] or [[Institution 1]] acts on behalf of [[Human 2]] or on behalf of [[Institution 2]] <Human 1 or Institution 1 performs a motivated Action for the benefit and/or at the request of Human 2 or Institution 2>	6	2.00	2.09	-4.3	2	0.67	2.77	-3.62	0	0.00	8.32**	1.92

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A10 (cont.)
Pattern frequencies for three subcorpora and result of cross-corpora comparison for act

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
6	[[Human]] or [[Institution]] acts on [[Eventuality]] <Human or Institution behaves in accordance with Eventuality=Motivation>	15	5.00	0.00	-6.4	15	5.00	20.79****	14.4†	0	0.00	20.79****	14.4†
7	[[Human]] or [[Institution]] acts under [[Rule]] or under [[Command]] <Human or Institution behaves in accordance with Rule or Command>	1	0.33	1.39	-5.01	0	0.00	0	-6.4	0	0.00	1.39	-5.02
8	[[Entity 1]] acts on or upon [[Entity 2]] <Entity1 has a particular effect on Entity 2>	8	2.67	0.22	-6.17	10	3.33	23.11****	16.71†	44	14.67	27.44****	21.04†
9	[[Human]] acts (Role) (in Performance) <Human plays Role = Theatrical (in Performance)>	1	0.33	1.39	-5.01	0	0.00	0	-6.4	0	0.00	1.39	-5.01
10	Phrasal verb. [[Human]] acts [[Event]] or [[Human_Role]] or [[Emotion]] out <Human performs Role, not necessarily sincerely, or behaves as if feeling Emotion>	2	0.67	0.34	-6.06	1	0.33	1.39	-5.01	0	0.00	2.77	-3.62
11	[[Human]] acts [[ADJ]] <Human behaves in the manner specified by ADJ>	1	0.33	0.00	-6.4	1	0.33	1.39	-5.01	0	0.00	1.39	-5.01
12	[[Drug]] acts <Drug has an effect>	0	0.00	1.39	-5.01	1	0.33	1.39	-5.01	0	0.00	0	-6.4
13	[[Process]] acts <Process exerts an influence>	0	0.00	2.77	-3.62	2	0.67	0.34	-6.06	1	0.33	1.39	-5.01
	TOTAL	300				300				300			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A11

Pattern frequencies for three subcorpora and result of cross-corpora comparison for answer

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
1	[[Human]] answers quote or that [clause] <Human says that [clause] or [quote] in response to a question or statement by someone else>	24	16.55	5.98*	0.52	5	5.62	3.79	-1.08	0	0.00	11.95***	6.73†
2	[[Human 1]] or [[Institution]] or [[Document]] answers [[Question]] or [[Human 2]] or [[Speech_Act]] < Human 1 or Institution or Document says or writes Proposition that is intended to provide relevant information in response to Question or Speech_Act that has been asked by Human 2>	81	55.86	8.49**	3.03†	79	88.76	0.38	-4.49	41	100.00	8.57**	3.35†
3	[[Human]] or [[Institution]] or [[Document]] answers [[Speech_Act]] <Human or Institution says something or writes something in Document intended to refute Speech_Act = Accusation>	17	11.72	16.27****	10.82†	0	0.00	0	-4.87	0	0.00	8.47**	3.24†
4	[[Human]] answers {telephone} or {call} <Human accepts incoming call and speaks into telephone after it rings>	0	0.00	7.73**	2.28†	4	4.49	3.03	-1.84	0	0.00	0	-5.23
5	[[Human]] answers [[Document]] <Human 1 writes an email or letter in response to Document = Mail from human 2>	6	4.14	5.74*	0.29	0	0.00	0	-4.87	0	0.00	2.99	-2.24
6	[[Human 1]] or [[Institution 1]] answers to [[Deity]] or to [[Human 2]] or to [[Institution 2]] <Human 1 has an obligation to account responsibly for his/her actions to Deity or Human 2 or Institution 2>	4	2.76	3.83	-1.63	0	0.00	0	-4.87	0	0.00	1.99	-3.23

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A11 (cont.)

Pattern frequencies for three subcorpora and result of cross-corpora comparison for answer

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
7	[[Human 1]] answers ([[Human 2]]) back Human 1 says something rude, impertinent, or forthright and unexpected in response to something that (Human 2) has said	1	0.69	0.12	-5.34	1	1.12	0.76	-4.11	0	0.00	0.5	-4.73
8	[[Artifact]] or [[Proposition]] answers need or purpose <Artifact or Proposition provides what is necessary for some purpose>	3	2.07	2.87	-2.58	0	0.00	0	-4.87	0	0.00	1.49	-3.73
9	[[Deity]] or [[Eventuality]] answers {prayer} <Eventuality desired by Human happens>	3	2.07	2.87	-2.58	0	0.00	0	-4.87	0	0.00	1.49	-3.73
10	[[Human]] answers for [[Eventuality]] <Human takes responsibility for Eventuality>	3	2.07	2.87	-2.58	0	0.00	0	-4.87	0	0.00	1.49	-3.73
11	[[Human 1]] answers {the description of} [[Human 2]] <Human 1 has the features listed in the description of Human 2>	2	1.38	1.91	-3.54	0	0.00	0	-4.87	0	0.00	1	-4.23
12	[[Human]] answers {calling} <Human undertakes activity in response to impulse or inclination>	1	0.69	0.96	-4.5	0	0.00	0	-4.87	0	0.00	0.5	-4.73
	TOTAL	145				89				41			

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = ≥ positive evidence against H_0

Table A12

Pattern frequencies for three subcorpora and result of cross-corpora comparison for appear

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
1	[[Physical Object]] appears	8	2.67	3.85	-2.54	2	0.67	0.2	-6.2	3	1.00	2.36	-4.04
2	<Physical_Object becomes visible> [[Anything]] appears [ADJ] <Anything seems to be ADJ>	46	15.33	4.63*	-1.77	69	23.00	13.88***	7.48†	32	10.67	2.53	-3.87
3	[[Abstract Entity]] or [[State of Affairs]] appears <Abstract_Entity or State_of_Affairs becomes noticeable or comes into existence>	16	5.33	0.26	-6.14	19	6.33	8.71**	2.31†	5	1.67	6.06**	-0.34
4	[[Stuff]] OR [[Physical Object]] appears [[Time Period]] <Stuff or Physical_Object comes into use or becomes available Time_Period>	10	3.33	13.86***	7.47†	0	0.00	6.93**	0.53	5	1.67	1.7	-4.7
5	[[Human]] OR [[Animal]] or [[MB Entity]] appears ([[Location]]) <Human or Animal arrives at Location>	3	1.00	0.20	-6.2	2	0.67	2.09	-4.3	6	2.00	1.02	-5.38
6	[[Human]] appears in [[Performance]] on [[TV Programme]] on [[Radio Programme]] or on [[Stage]] <Human takes part in Performance>	3	1.00	4.16*	-2.24	0	0.00	0	-6.4	0	0.00	4.16*	-2.24
7	[[Document 1]] OR [[Image]] appears (in [[Document 2]] or [[Document Part]]) <Document 1 or Image is published (in Document 2 or Document Part)>	54	18.00	21.78****	15.39†	16	5.33	15.96****	9.56†	1	0.33	66.25****	59.85†

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = ≥ positive evidence against H0

Table A12 (cont.)

Pattern frequencies for three subcorpora and result of cross-corpora comparison for appear

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
8	[[Anything]] appears [to-Infinitive] <Anything seems to/inf [verb]>	119	39.67	6.05*	-0.35	160	53.33	9..51**	3.12†	220	73.33	30.55****	24.16†
9	{it} appears [that-clause] <that-[CLAUSE] is probably true>	20	6.67	1.34	-5.06	28	9.33	0.07	-6.32	26	8.67	0.78	-5.61
10	[[Entity 1]] or [[Eventuality 1]] appears (as) [[Entity 2]] or(as) [[Eventuality 2]] <Entity 1 or Eventuality 1 is perceived as or assumed to be Entity 2 or Eventuality 2>	16	5.33	7.71**	1.31	4	1.33	0.68	-5.72	2	0.67	12.4***	6†
11	[[Human]] appears in or before [[Institution]] or before [[Human]] <Law: court procedure. Human presents Self formally in Court or before Judge as a Defendant to be tried in respect of Accusation>	2	0.67	2.77	-3.62	0	0.00	0	-6.4	0	0.00	2.77	-3.62
12	[[Abstract Entity]] appear in [[Document]] <Document mentions abstract entity>	2	0.67	2.77	-3.62	0	0.00	0	-6.4	0	0.00	2.77	-3.62
13	[[Human]] appears as {an MP}	1	0.33	1.39	-5.01	0	0.00	0.00	-6.4	0	0.00	1.39	-5.01
	TOTAL	300				300				300			

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = ≥ positive evidence against H0

Table A13

Pattern frequencies for three subcorpora and result of cross-corpora comparison for apply

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
1	[[Human]] or [[Institution]] applies [[Concept]] or [[Process]] (to [[State of Affairs]]) <Human or Institution makes use of Concept or Process in a particular situation or State_of_Affairs>	163	54.33	7.18**	0.78	215	71.67	1.3	-5.1	192	64.00	2.37	-4.03
2	[[Concept]] or [[Process]] applies (to [[State of Affairs]] or [[Activity]]) <Concept or Process is relevant to State_of_Affairs or Activity>	91	30.33	2.22	-4.18	72	24.00	21.22****	14.83†	27	9.00	36.65****	30.26†
3	[[Human]] applies for [[Benefit]] benefit or injunction or extension or admission or readmission or {[NP] to be [V]} <Human 1 formally asks Human 2 or Institution to give Benefit to Human 1>	13	4.33	2.64	-3.76	6	2.00	8.32**	1.92	0	0.00	18.02****	11.62†
4	[[Human]] or [[Device]] applies [Fluid]] or [[Stuff]] to [[Surface]] <Human (uses Device to) spread Fluid or Stuff on Surface, typically in order to heal, improve, protect Surface or conduct test>	8	2.67	11.09***	4.69†	0	0.00	97.04****	90.64†	70	23.33	56.54****	50.15†
5	[[Human]] applies [[Word]] to [[Anything]] <Human uses Word to suitably refer to Anything>	17	5.67	6.92**	0.52	5	1.67	0.09	-6.31	6	2.00	5.48*	-0.91
6	[[Human]] or [[Institution]] applies {pressure} <Human tries to influece decision or action of another human or institution>	7	2.33	2.94	-3.45	2	0.67	0.68	-5.72	4	1.33	0.83	-5.57
7	[[Human]] applies [[Self]] to [[Activity]] <Human makes great effort to conduct activity>	1	0.33	1.39	-5.01	0	0.00	0	0	0	0.00	1.39	-5.01
8	[[Human]] applies {caution} <Human take a cautious approach to [[Anything]]>	0	0.00	0.00	0	0	0.00	1.39	-5.01	1	0.33	1.39	-5.01
	TOTAL	300				300				300			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = ≥ positive evidence against H0

Table A14

Pattern frequencies for three subcorpora and result of cross-corpora comparison for call

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
1	[[Anything]] is called [[Name]] <Anything is referred to as Name>	76	25.33	5.59*	-0.8	108	36.00	74.6****	68.2†	274	91.33	118.92****	112.52†
2	[[Human]] OR [[Institution]] calls [[Anything]] [[Name]] <Human or Institution invents or uses Name to refer to Anything>	119	39.67	0.35	-6.04	110	36.67	101.19****	94.8†	9	3	112.31****	105.91†
3	[[Human]] or [[Institution]] calls [[Event]] <Human or Institution instructs people to cause Event = Meeting or Action to happen immediately>	4	1.33	1.93	-4.47	1	0.33	1.39	-5.01	0	0	5.55*	-0.85
4	[[Human]] or [[Institution]] or [[Document]] calls {for} [[Action]] or {for} [[State of Affairs]] <Human or Institution or Document says that other people should do Action or create [State_of_Affairs]>	38	12.67	1.39	-5.0	49	16.33	30.34****	23.95†	9	3	19.25****	12.85†
5	[[Human]] or [[Institution]] calls [[Human 2]]{for} [[Activity]] <Human 1 or Institution officially invites or instructs Human 2 to take part in Activity>	1	0.33	1.39	-5.01	0	0.00	0	-6.4	0	0	1.39	-5.01
6	[[Human 1]] calls ((in or round) (on [[Human 2]]) at [[Location]]) <Human 1 goes to Location in order to meet Human 2>	3	1.00	4.16*	-2.24	0	0.00	0	-6.4	0	0	4.16*	-2.24

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A14 (cont.)
Pattern frequencies for three subcorpora and result of cross-corpora comparison for call

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
7	[[Human 1]] calls [[Human 2]] <Human 1 shouts to Human 2, typically in order to ask Human 2 to come to them>	6	2.00	8.32**	1.92	0	0.00	0	-6.4	0	0	8.32**	1.92
8	[[Human]] calls [[Speech_Act]] (out) <Human says Speech_Act or QUOTE in a loud, clear voice>	3	1.00	4.16*	-2.24	0	0.00	0	-6.4	0	0	4.16*	-2.24
9	[[Human 1]] calls [[Human 2]] or [[Institution]] or [[Number]] <Human 1 contacts Human 2 or Institution on the telephone by selecting Number = Telephone Number>	1	0.33	2.91	-3.49	5	1.67	6.93**	0.53	0	0	1.39	-5.01
10	[[Human 1]] or [[Institution 2]] or [[Document]] calls {on} or {upon} [[Human 2]] OR [Institution 2]] [TO-INFINITIVE] <Human 1 or Institution 1 or Document asks Human 2 or Institution 2 to [verb]>	23	7.67	2.81	-3.58	13	4.33	9.01**	2.62†	2	0.67	20.72****	14.32†
11	[[Human 1]] or [[Institution 1]] calls [[Human 2]] or [[Institution 2]] {in} <Human 1 or Institution 1 asks Human 2 or Institution 2 to come and give help or advice>	2	0.67	0.20	-6.2	3	1.00	4.16*	-2.24	0	0	2.77	-3.62
12	[[Human]] calls [[Activity]] {off} <Human cancels Activity>	3	1.00	1.05	-5.35	1	0.33	1.39	-5.01	0	0	4.16*	-2.24
13	[[Human]] calls [[Information]] or [[Image]] {up} <Human causes Information or Image to be displayed on a computer screen>	1	0.33	1.39	-5.01	0	0.00	0	-6.4	0	0	1.39	-5.01

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A14 (cont.)

Pattern frequencies for three subcorpora and result of cross-corpora comparison for call

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
14	[[Human]] calls {attention to} [[Event]] <Human asks people to notice Event>	5	1.67	0.00	-6.4	5	1.67	2.91	-3.49	1	0.33	2.91	-3.49
15	[[Human]] or [[Anything] calls [[Proposition]] {into question} <Human or Anything provides a reason for doubting whether Proposition is correct>	6	2.00	1.02	-5.38	3	1.00	0.51	-5.89	5	1.67	0.09	-6.31
16	[[Human]] calls {upon} [[Human]] <Human 1 asks Human 2 for assistance>	3	1.00	0.20	-6.2	2	0.67	2.77	-3.62	0	0	4.16*	-2.24
17	[[Human 1]] calls [[Anything]] {down upon} [[Human 2]] <Human evokes Anything to affect Human 2>	2	0.67	2.77	-3.62	0	0.00	0	0	0	0	2.77	-3.62
18	[[Human]] is called to {celibacy} <Human chooses to remain celibate>	1	0.33	1.39	-5.01	0	0.00	0	0	0	0	1.39	-5.01
19	[[Concept]] is called into {being} <Human creates concept to explain or deal with Eventuality>	1	0.33	1.39	-5.01	0	0.00	0	0	0	0	1.39	-5.01
20	[[Institution]] calls [[Human]] {up} < Institution = Government requires Human to undertake military service>	2	0.67	2.77	-3.62	0	0.00	0	0	0	0	2.77	-3.62
	TOTAL	300				300				300			

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A15

Pattern frequencies for three subcorpora and result of cross-corpora comparison for conduct

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
1	[[Human]] or [[Institution]] conducts [[Activity]] <Human or Institution carries out Activity>	141	94.00	0.29	-5.82	298	99.33	0.02	-6.09	147	98.00	0.13	-5.58
2	[[Human 1]] conducts [[Human 2]] [Direction] <Human 1 accompanies and guides Human 2 Direction>	3	2.00	6.59*	0.48	0	0.00	0	-6.11	0	0.00	4.16*	-1.54
3	[[Human]] conducts [[Self]] [Manner] <Human behaves in a particular way>	5	3.33	4.23*	-1.88	2	0.67	1.62	-4.49	0	0.00	6.93**	1.23
4	[[Metal]] or [[Fluid]] conducts [[Energy]] <Energy is transferred via Metal>	1	0.67	2.20	-3.91	0	0.00	0	-6.11	0	0.00	1.39	-4.32
5	[[Route]] conducts [[Fluid]] <Stuff passes through Aperture>	0	0.00	0.00	-6.11	0	0.00	2.2	-3.91	1	0.67	1.39	-4.32
6	[[MB Entity]] conducts [[Process]] <MB entity carries out Process = Microbiological>	0	0.00	0.00	-6.11	0	0.00	4.39*	-1.71	2	1.33	2.77	-2.93
	TOTAL	150				300				150			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A16

Pattern frequencies for three subcorpora and result of cross-corpora comparison for explain

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
1	[[Human]] or [[Institution]] or [[Document]] or [[Proposition]] or [[Eventuality 1]] explains [[Eventuality 2]] (in terms of [[Eventuality 3]]) (to [[Human 2]]) <Human or Institution formulates (in conversation with Human 2 or in a Document intended for Human 2 to read) a Proposition concerning the cause or effect of an Eventuality 2>	171	57.00	9.27**	2.87†	232	77.33	0.67	-5.72	250	83.33	14.91***	8.52†
2	[[Human]] or [[Institution]] or [[Document]] or [[Proposition]] explains (to [[Human 2]]) <Human or Institution formulates Proposition in Document concerning the cause or effect of an Eventuality that is expressed in wh- [clause]>	62	20.67	3.87*	-2.53	42	14.00	0.18	-6.22	46	15.33	2.38	-4.02
3	[[Human]] or [[Institution]] explains [That-Clause] (to [[Human 2]]) <Human 1 or Document tells Human 2) that [clause], as a way of accounting for Eventuality (often implied rather than stated explicitly)>	37	12.33	22.44****	16.04†	7	2.33	2.94	-3.45	2	0.67	38.29****	31.89†
4	[[Human]] or [[Institution]] explains [[Quote]] <Human says [QUOTE] as a way of accounting for Eventuality (often implied rather than stated explicitly)>	27	9.00	1.40	-5	19	6.33	19.79****	13.39†	1	0.33	30.19****	23.79†
5	[[Human]] or [[Institution]] explains [[Eventuality]] {away} <Human 1 formulates Proposition as a way of accounting for Eventuality = Inconvenient>	3	1.00	4.16*	-2.24	0	0.00	1.39	-5.01	1	0.33	1.05	-5.35
	TOTAL	300				300				300			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A16.1

Reanalysed pattern frequencies for three subcorpora and result of cross-corpora comparison for explain

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Frequency	%			Frequency	%			Frequency	%		
1	[[Human]] or [[Institution]] or [[Document]] or [[Proposition]] or [[Eventuality 1]] explains [[Eventuality 2]] (in terms of [[Eventuality 3]]) (to [[Human 2]]) <Human or Institution formulates (in conversation with Human 2 or in a Document intended for Human 2 to read) a Proposition concerning the cause or effect of an Eventuality 2>	171	57.00	9.27**	2.87†	232	77.33	0.67	-5.72	250	83.33	14.91***	8.52†
2a	[[Human]] or [[Institution]] or [[Document]] or [[Proposition]] explains (to [[Human 2]]) <Human or Institution formulates Proposition in Document concerning the cause or effect of an Eventuality that is expressed in wh- [clause]>	126	42.00	17.61****	11.21†	68	22.67	3.1	-3.3	49	16.33	35.07****	28.67†
5	[[Human]] or [[Institution]] explains [[Eventuality]] {away} <Human 1 formulates Proposition as a way of accounting for Eventuality = Inconvenient>	3	1.00	4.16*	-2.24	0	0.00	1.39	-5.01	1	0.33	1.05	-5.35
	TOTAL	300				300				300			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = ≥ positive evidence against H0

Table A17

Pattern frequencies for three subcorpora and result of cross-corpora comparison for fail

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Frequency	%			Frequency	%			Frequency	%		
1	[[Human]] or [[Institution]] fails [[to-infinitive]] <Human or Institution does not do something that they intended to do or were expected to do>	95	63.33	9.6**	3.9†	57	38.00	44.54****	38.83†	7	4.67	90.39****	84.68†
2	[[Human]] or [[Institution]] fails (in[[Activity]]) <Human or Institution attempts to do Activity but does not produce the expected or intended result>	10	6.67	3.99*	-1.71	21	14.00	22.36****	16.66†	1	0.67	8.55**	2.84†
3	[[Activity]] or [[System]] fails <Activity or System does not produce the expected or intended result>	27	18.00	0.08	-5.63	25	16.67	5.59*	-0.11	11	7.33	6.95**	1.25
4	[[Activity]] or [[System]] or [[Abstract_Entity]] fails [[to-infinitive]] <Activity or System or Abstract_Entity does not produce the expected or intended result>	14	9.33	16.3****	10.59†	44	29.33	0	-5.5	44	29.33	16.3***	10.59†
5	[[Artifact]] or [[Plant]] or [[Body_Part]] or [[MB Unit]] fails <Artifact or Body_Part or Plant does not function effectively>	0	0.00	1.39	-4.32	1	0.67	1.05	-4.66	3	2.00	4.16*	-1.54
6	[[Artifact]] or [[Body_Part]] or [[MB Entity]] or [[Plant]] fails [[to-infinitive]] <Artifact or Body_Part or Plant does not work or perform in the way that it is intended or expected to>	2	1.33	0.34	-5.36	1	0.67	102.87****	97.19†	81	54.00	96.21****	90.5†

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A17 (cont.)
Pattern frequencies for three subcorpora and result of cross-corpora comparison for fail

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Frequency	%			Frequency	%			Frequency	%		
7	[[Human 1]] fails [[Activity]] <Human does not meet the standard required to be successful in Activity = Test>	0	0.00	0.00	-5.7	0	0.00	4.16*	-1.54	3	2.00	4.16*	-1.54
8	[[Human 1]] or [[System]] or [[Institution 1]] fails [[Human 2]] or [[Institution 2]] <Human 1 or System or Institution 1 fails Human 2 or Institution 2>	1	0.67	0.00	-5.7	1	0.67	1.39	-4.32	0	0.00	1.39	-4.32
9	[[Weather_Event]] fails <Weather_Event = Desirable does not occur in the way that is expected or hoped for>	1	0.67	1.39	-4.32	0	0.00	0	-5.7	0	0.00	1.39	-4.32
	TOTAL	150				150				150			

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A18
Pattern frequencies for three subcorpora and result of cross-corpora comparison for follow

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Frequency	%			Frequency	%			Frequency	%		
1	[[Human 1]] or [[Animal 1]] or [[Vehicle 1]] or [[MB Entity 1]] follows [[Human 2]] or [[Animal 2]] or [[Vehicle 2]] or [[MB Entity 2]] <Human 1 or Animal 1 or Vehicle 1 moves in the same direction as that selected by Human 2 or Animal 2 or the driver of Vehicle 2>	16	5.33	9.77**	3.73†	3	1.00	0	-6.4	3	1	9.77**	3.37†
2	[[Human 1]] follows [[Human 2]] <Human 1 = Student or Disciple studies and is influenced by or tries to practice the teachings of Human 2 = Teacher>	19	6.33	13.7***	7.31†	49	16.33	67.93****	61.53†	0	0.00	26.34****	19.94†
3	[[Human_Group 1]] or [[Institution 1]] follows [[Human_Group 2]] or [[Institution 2]] <Human_Group 1 or Institution 1 is influenced by and thus copies the behaviour of Human Group 2 or Institution 2>	9	3.00	4.82*	-1.58	2	0.67	2.77	-3.62	0	0.00	12.48***	6.08†
4	[[Human]] or [[Institution]] or [[Concept]] or [[Process]]follows [[Command]] or [[Rule]] or [[Plan]] or [[Document]] <Human or Institution acts in accordance with Command or Rule or Plan (expressed in Document)>	84	28.00	1.74	-4.65	102	34.00	31.61****	25.22†	37	12.33	18.75****	12.35†

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A18 (cont.)
Pattern frequencies for three subcorpora and result of cross-corpora comparison for follow

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Frequency	%			Frequency	%			Frequency	%		
5	[[Human]] or [[Institution]] follows [[Event]] <Human or Institution takes an interest in Event = Unfolding>	12	4.00	10.97***	4.57†	1	0.33	22.36****	15.97†	21	7.00	2.49	-3.91
6	[[Human]] follows [[Proposition]] or [[wh-clause]] <Human is able to understand Proposition or wh-clause>	0	0.00	1.39	-5.01	1	0.33	1.39	-5.01	0	0.00	0	-6.4
7	[[Human]] or [[Institution]] follows {lead} or {line of enquiry} <Human or Institution pursues an investigation based on specific information>	0	0.00	2.77	-3.62	2	0.67	2.77	-3.62	0	0.00	0	-6.4
8	[[Human]] or [[Institution]] or [[Event]] follows {trend} <Human or Institution performs in line with the expectations dictated by trend >	3	1.00	1.05	-5.35	1	0.33	1.93	-4.47	4	1.33	0.14	-6.25
9	[[Event 1]] follows (Event 2) <Event 1 happens after and typically as a consequence of Event 2>	103	34.33	27.59****	21.19†	41	13.67	73.43****	67.03†	158	52.67	11.68***	5.28†
10	[[Eventuality 1]] follows from [[Eventuality 2]] <Eventuality 1 is a necessary consequence of Eventuality 2>	8	2.67	2.72	-3.68	16	5.33	12.4***	6†	2	0.67	3.85*	-2.54

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A18 (cont.)

Pattern frequencies for three subcorpora and result of cross-corpora comparison for follow

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Frequency	%			Frequency	%			Frequency	%		
11	{it} follows [[that-clause]] (from Eventuality 1) <Eventuality 2 expressed by [CLAUSE] is a necessary consequence of Eventuality 1 (either expressed in a from-phrase or one that was previously stated)>	5	1.67	2.31	-4.09	11	3.67	3.4	-3	4	1.33	0.11	-6.29
12	[[Anything 1]] follows [[Anything 2]] <Anything 1 = Item in List comes after Anything 2 = Item in List in a list or sequence>	13	4.33	0.04	-6.36	14	4.67	8.59**	2.2†	34	11.33	9.72**	3.33†
13	[[Human]] follows {up} [[Eventuality]] <Human monitors Eventuality and finds out more about it>	5	1.67	1.33	-5.07	2	0.67	0.34	-6.06	1	0.33	2.91	-3.49
14	as follows <as will now be stated here>	13	4.33	24.12****	17.72†	51	17.00	6.79**	0.4	28	9.33	5.62*	-0.78
15	[[Human 1]] or [[Institution 1]] follows {suit} <Human 1 or or Institution 1 does the same as Human 2 or Institution 2 has done>	4	1.33	0.68	-5.72	2	0.67	2.77	-3.62	0	0.00	5.55*	-0.85

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A18 (cont.)
Pattern frequencies for three subcorpora and result of cross-corpora comparison for follow

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Frequency	%			Frequency	%			Frequency	%		
16	[[Human 1]] follows in footsteps {of [[Human 2]]} <Human 1 engages in similar activity to one previously engaged in by Human 2>	5	1.67	6.93**	0.53	0	0.00	0	-6.4	0	0.00	6.93**	0.53
17	[[Human]] follows up with [[Speech Act]] <Human adds additional details to information already given>	1	0.33	1.39	-5.01	0	0.00	0	-6.4	0	0.00	1.39	-5.01
18	[[MB Entity]] follows [[Route]] <MB Entity follows an established route or conduit>	0	0.00	0.00	-6.4	0	0.00	11.09***	4.69†	8	2.67	11.09***	4.69†
19	[[Human]] follows {through on commitment} <Human fulfills a commitment>	0	0.00	2.77	-3.62	2	0.67	2.77	-3.62	0	0.00	0	-6.4
	TOTAL	300				300				300			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A19

Pattern frequencies for three subcorpora and result of cross-corpora comparison for handle

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro / Hist LL	BIC
		Frequency	%			Frequency	%			Frequency	%		
1	[[Human]] handles [[Physical_Object]] <Human uses his/her hands to hold, manipulate, and/or use Physical_Object>	15	20.27	16****	10.95†	1	1.23	25.37****	20.74†	10	47.62	4.03*	-0.52
2	[[Human]] or [[Institution]] or [[Computer]] handles [[Eventuality]] (Manner) <Human or Institution or Computer uses intelligence or an established procedure to deal with Eventuality as required or in a particular manner>	32	43.24	5.88*	0.83	59	72.84	15.92****	11.29†	2	9.52	6.81**	2.26†
3	[[Human]] or [[Institution]] or[[Device]] or [[MB Entity]] handles [[Entity]] or [[Stuff]] <Human or Institution or Device processes or deals with Entity or Stuff>	16	21.62	0.01	-5.04	18	22.22	1.49	-3.14	8	38.10	1.59	-2.96
4	[[Human 1]] handles [[Human 2]] or [[Human_Group]] (Manner) <Human 1 manages or copes with Human 2 or Human_Group (in a particular manner)>	8	10.81	2.83	-2.21	3	3.70	1.38	-3.24	0	0.00	4	-0.56

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A19 (cont.)

Pattern frequencies for three subcorpora and result of cross-corpora comparison for handle

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro / Hist LL	BIC
		Frequency	%			Frequency	%			Frequency	%		
5	[[Human]] handles [[Artifact]] <Human receives Artifact = that should not legally have in his/her possesion >	2	2.70	2.96	-2.09	0	0.00	0	-4.62	0	0.00	1	-3.55
6	[[Human]] handles [[Self]] (manner) <Human behaves in a particular manner>	1	1.35	1.48	-3.56	0	0.00	0	-4.62	0	0.00	0.5	-4.05
7	[[Human]] handles [[Animal]] <Human cares for and commands animal>	0	0.00	0.00	-5.04	0	0.00	3.16	-1.46	1	4.76	3.02	-1.54
	TOTAL	74				81				21			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A20

Pattern frequencies for three subcorpora and result of cross-corpora comparison for lead

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Frequency	%			Frequency	%			Frequency	%		
1	[[Eventuality 1]] leads to [[Eventuality 2]] <Eventuality 1 is the cause of Eventuality 2>	158	52.67	14.83***	8.43†	234	78.00	6.81**	0.41	295	98.33	42.09****	35.69†
2	[[Eventuality 1]] leads {up to} [[Eventuality 2]] <Eventuality 1 = PLURAL precedes but may not be the cause of Eventuality 2>	5	1.67	6.93**	0.53	0	0.00	0	-6.4	0	0.00	6.93**	0.53
3	[[Eventuality]] leads [[Human]] or [[Institution]] [[to-infinitive]] <Eventuality causes, enables, or encourages Human or Institution to/INF [verb]>	47	15.67	0.01	-6.39	46	15.33	41.44****	35.04†	4	1.33	42.66****	36.26†
4	[[Human]] or [[Institution 1]] leads [[Human_Group]] or [[Institution 2]] <Human or Institution 1 organizes, directs, or provides a model for Activity of Human_Group or Institution 2>	44	14.67	32.62*	26.23†	6	2.00	8.32**	1.92	0	0.00	61****	56.6†
5	[[Human]] or [[Institution]] leads [[Activity]] <Human or Institution organizes, directs, or provides a model for the Activity of Human_Group>	13	4.33	1.83	-4.57	7	2.33	5.06*	-1.33	1	0.33	12.2***	5.81†

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A20 (cont.)

Pattern frequencies for three subcorpora and result of cross-corpora comparison for lead

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Frequency	%			Frequency	%			Frequency	%		
6	[[Human 1]] leads [[Human 2]] or [[Human_Group]](Direction to Location) <Human 1 organizes and directs the movement of Human 2 or Human_Group to Location by accompanying Human 2 or Human Group and showing the way>	12	4.00	4.19*	-2.21	4	1.33	5.55*	-0.85	0	0.00	16.64****	10.24†
7	[[Route]] leads [[Direction]] to [[Location]] <Route is a path or road Direction or to Location>	12	4.00	10.97***	4.57†	1	0.33	1.39	-5.01	0	0.00	16.64****	10.24†
8	[[Human]] or [[Human_Group]] leads (Activity) <Human = Competitor or Human_Group = Competitor is in first place in Activity = Competition or Race)>	1	0.33	0.00	-6.4	1	0.33	1.39	-5.01	0	0.00	1.39	-5.01
9	[[Human]] or [[Institution]] leads the way(in-ING) <Human or Institution is the best at doing something and shows how it can best be done>	2	0.67	2.77	-3.62	0	0.00	0	-6.4	0	0.00	2.77	-3.62

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A20 (cont.)

Pattern frequencies for three subcorpora and result of cross-corpora comparison for lead

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Frequency	%			Frequency	%			Frequency	%		
10	[[Human]] or [[Animate]] leads MOD {life} or {existence} <Human or Institution is the best at doing something and shows how it can best be done>	5	1.67	2.91	-3.49	1	0.33	1.39	-5.01	0	0.00	6.93**	0.53
11	[[Human 1]] or [[Eventuality]] leads [[Human 2]] on <Human 1 or Eventuality encourages Human 2, typically to do or believe something stupid or reprehensible>	1	0.33	1.39	-5.01	0	0.00	0	-6.4	0	0.00	1.39	-5.01
	TOTAL	300				300				300			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A21

Pattern frequencies for three subcorpora and result of cross-corpora comparison for maintain

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro / Hist LL	BIC
		Freq	%			Freq	%			Freq	%		
1	[[Human]] or [[Process]] maintains [[State_of_Affairs]] <Human takes action to ensure or Process has the effect that State_of_Affairs remains unchanged>	211	70.33	3.98*	-2.41	254	84.67	0	-6.4	254	84.67	3.98*	-2.41
2	[[Human]] or [[Institution]] or [[Document]] maintains [[that-clause]] <Human or Institution or Document asserts strongly and consistently over time that [clause] or [Proposition]>	64	21.33	17.68****	11.29†	25	8.33	19.75****	13.35†	3	1.00	68.38****	61.98†
3	[[Human]] or [[Institution]] maintains [[Document]] <Human or Institution regularly and systematically updates Document>	1	0.33	8.55**	2.15†	10	3.33	5.82*	-0.57	2	0.67	0.34	-6.06
4	[[Human]] maintains [[Building]] or [[Vehicle]] or [[Artifact]] <Human takes action to ensure that Building or Vehicle or Artifact is kept in good working order>	19	6.33	3.65*	-2.75	9	3.00	7.36**	0.96	1	0.33	19.79****	13.39†
5	[[Human 1]] maintains [[Self]] or [[Human 2]] <Human 1 provides food and drink for Self or Human 2, enabling Self or Human 2 to stay alive>	5	1.67	1.33	-5.07	2	0.67	0	-6.4	2	0.67	1.33	-5.07
6	[[MB_Entity]] is maintained in [[Location]] or on [[Surface]] <MB Entity is place in Location or on Surface for the purpose of a test>	0	0.00	0.00	-6.40	0	0.00	52.68****	46.28†	38	12.67	52.68****	46.28†
	TOTAL	300				300				300			

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A22

Pattern frequencies for three subcorpora and result of cross-corpora comparison for manage

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Frequency	%			Frequency	%			Frequency	%		
1	[[Human]] or [[Institution]] or [[MB Entity]] manages [to-infinitive] <Human or Institution is able to/inf [verb], despite difficult circumstances>	90	60.00	120.76****	114.66†	16	5.33	0.44	-5.33	2	9.09	13.59***	8.44†
2	[[Human]] manages <Human is able to deal with difficult circumstances successfully>	0	0.00	0.81	-5.3	1	0.33	0.14	-5.63	0	0.00	0	-5.15
3	[[Human]] or [[Institution]] [[MB Entity]] manages [[Action]] or [[State_of_Affairs]] <Human or Institution is able to perform Action = Desirable or achieve State_of_Affairs = Desirable, despite difficult circumstances>	1	0.67	0.85	-5.26	5	1.67	6.22*	0.45	3	13.64	8.11**	2.97†
4	[[Human]] manages [[Artifact]] <Human is able to lift, carry, or use Artifact without difficulty>	0	0.00	0.81	-5.3	1	0.33	0.14	-5.63	0	0.00	0	-5.15
5	[[Human]] or [[Institution 1]] manages [[Human 2]] or [[Institution 2]] <Human or Institution 1 is responsible for and controls Institution 2>	11	7.33	5.94	-0.17	47	15.67	0.08	-5.7	4	18.18	2.07	-3.08

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A22 (cont.)

Pattern frequencies for three subcorpora and result of cross-corpora comparison for manage

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Frequency	%			Frequency	%			Frequency	%		
6	[[Human]] or [[Institution]] or [[MB Entity]] manages [[System]] or [[Activity]] <Human or Institution is responsible for and controls System or Activity = Ongoing>	32	21.33	41.78****	35.67†	191	63.67	4.06*	-1.71	7	31.82	0.84	-4.31
7	[[Human]] or [[Institution]] manages [[Resource]] <Human or Institution is responsible for Resource>	16	10.67	0.24	-5.87	37	12.33	1.36	-4.42	1	4.55	0.89	-4.26
8	[[Human]] or [[Institution]] manages [[Location]] <Human or Institution is responsible for the use and maintenance of Location>	0	0.00	1.62	-4.49	2	0.67	0.28	-5.49	0	0.00	0	-5.15
9	[[Human]] manages [[Disease]] <Human alleviates symptoms of Disease with drugs or Therapy>	0	0.00	0.00	-6.11	0	0.00	26.84****	21.06†	5	22.73	20.56****	15.42†
	TOTAL	150				300				22			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A23

Pattern Frequencies for Three Subcorpora and Result of Cross-corpora Comparison for need

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro / Hist LL	BIC
		Freq	%			Freq	%			Freq	%		
1	[[Human]] or [[Institution]] needs [[Eventuality]] or [[Entity]] <Human or Institution requires that Eventuality or Entity must be realized or available, in order to accomplish some goal>	122	40.67	4.85*	-1.55	90	30.00	1.18	-5.21	76	25.33	10.79**	4.93†
2	[[Entity 1]] or [[Eventuality 1]] needs [[Entity 2]] or [[Eventuality 2]] <Entity 1 or Eventuality 1 is an essential precondition for or attribute of Entity 2 or Eventuality 2>	29	9.67	0.76	-5.64	36	12.00	26.08****	19.68†	93	31.00	35.31****	28.92†
3	[[Entity]] or [[Eventuality]] needs [[to-infinitive]] <An essential precondition for the realization of Eventuality is that [verb] must be realized typically to be>	62	20.67	0.03	-6.37	64	21.33	2.71	-3.69	84	28.00	3.33	-3.07
4	[[Human]] or needs [[to-infinitive]] <Human must do [verb]>	83	27.67	3.79	-2.61	110	36.67	46.93****	40.53†	31	10.33	24.62****	18.22†
5	[[Plant]] or [[Animate]] or [[MB Entity]] needs [[Eventuality]] or [[Stuff]] <Plant or Animate must have Eventuality or Stuff in order to survive and flourish>	4	1.33	5.55*	-0.85	0	0.00	22.18****	15.78†	16	5.33	7.71**	1.31
	TOTAL	300				300				300			

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A24

Pattern frequencies for three subcorpora and result of cross-corpora comparison for note

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro / Hist LL	BIC
		Frequency	%			Frequency	%			Frequency	%		
1	[[Human]] notes [[Information]] or [[Eventuality]] or [[Visible_Feature]] <Human notices and pays particular attention to Information about Eventuality or Visible_Feature>	104	34.67	0.18	-6.22	98	32.67	6.5**	0.11	137	45.67	4.53*	-1.86
2	[[Human]] notes [[quote]] or [[that-clause]] or [[wh-clause]] <Human notices and pays particular attention to Eventuality or Visible_Feature>	173	57.67	0.80	-5.6	190	63.33	4	-2.4	153	51.00	1.23	-5.17
3	[[Human]] notes [[Information]] (on or in [[Document]]) (under [[Document_Part]]) <Human makes a written note of Information (on | in Document) (under Document_Part = Heading)>	23	7.67	3.52	-2.88	12	4.00	2.97	-3.43	5	1.67	12.54****	6.14†
4	[[MB Entity]] is noted [[Name]] <MB Entity is referred to as Name>	0	0.00	0.00	-6.4	0	0.00	6.93**	0.53	5	1.67	6.93**	0.53
	TOTAL	300				300				300			

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A25

Pattern Frequencies for Three Subcorpora and Result of Cross-corpora Comparison for overcome

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro / Hist LL	BIC
		Freq	%			Freq	%			Freq	%		
1	[[Human]] or [[Institution]] or [[Activity]] or [[MB Entity]] overcomes [[Eventuality]] <Human or Institution or Activity successfully deals with Eventuality = Problem>	122	88.41	1.03	-4.64	150	100.00	0	-5.57	112	100.00	0.43	-4.98
2	[[Eventuality]] overcomes [[Human]] <Human fails to deal successfully with Eventuality = Problem>	5	3.62	7.36**	1.69	0	0.00	0	-5.57	0	0.00	0	-5.41
3	[[Human 1]] or [[Human_Group 1]] overcomes [[Human 2]] or [[Human_Group 2]] <Human 1 or Human_Group 1 defeats Human 2 or Human Group 2 in a fight, war, or contest>	11	7.97	16.19****	10.52†	0	0.00	0	-5.57	0	0.00	0	-5.41
	TOTAL	138				150				112			

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A26

Pattern frequencies for three subcorpora and result of cross-corpora comparison for prescribe

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro / Hist LL	BIC
		Frequency	%			Frequency	%			Frequency	%		
1	[[Human]] prescribes [[Drug]] or [[Activity]] ((for Human 2) (for Illness)) <Human = Medical Professional advises the use of Drug or Activity = Treatment to treat Human 2 = Patient) (for Illness))>	16	61.54	0.02	-4.18	24	58.54	2.62	-1.52	21	95.45	1.77	-2.10
2	[[Human]] or [[Institution]] or [[Rule]] prescribes [[State_of_Affairs]] or [[Activity]] <Formal. Human or Institution or Rule authoritatively sets out the terms and conditions necessary for State_of_Affairs or Activity to occur>	10	38.46	0.04	-4.18	17	41.46	8.99**	4.84†	1	4.55	7.12**	3.25†
	TOTAL	26				41				22			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A27

Pattern Frequencies for Three Subcorpora and Result of Cross-corpora Comparison for propose

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro / Hist LL	BIC
		Freq	%			Freq	%			Freq	%		
1	[[Human]] or [[Institution]] or [[Document]] proposes [[Action]] or [[Plan]] or [[Proposition]] <Human or Institution suggests, for consideration by others, Plan or Proposition as a basis for Action>	94	62.67	0.00	-5.7	94	62.67	7.57**	1.86	60	40.00	7.57**	1.86
2	[[Human]] or [[Institution]] proposes [[Entity]] <Human or Institution suggests, as a plan for consideration by others, that Entity should be created>	4	2.67	5.55*	-0.16	0	0.00	20.79****	15.09†	15	10.00	6.78**	1.08
3	[[Human]] or [[Document]] proposes [[that-clause]] <Human suggests [CLAUSE] as a proposition for consideration by others>	24	16.00	8.74**	3.03†	49	32.67	2	-3.71	64	42.67	18.87****	13.16†
4	[[Human]] or [[Institution]] or [[Document]] proposes [[to-infinitive]] <Human or Institution intends to [verb]>	23	15.33	8.99**	3.29†	7	4.67	0.9	-4.81	11	7.33	4.33*	-1.38
5	[[Human 1]] proposes [[Human 2]] for or as [[Human_Role]] <Human 1 formally suggests that Human 2 should be elected or appointed to undertake Human_Role>	4	2.67	5.55*	-0.16	0	0.00	0	-5.7	0	0.00	5.55*	-0.16
6	[[Human 1]] proposes {marriage} <Human 1 asks Human 2 to marry him or her.>	1	0.67	1.39	-4.32	0	0.00	0	-5.7	0	0.00	1.39	-4.32
	TOTAL	150				150				150			

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A28

Pattern Frequencies for Three Subcorpora and Result of Cross-corpora Comparison for replicate

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro / Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
1	[[Human]] replicates [[Artifact]]	5	16.67	6.73**	1.57	4	2.78	5.71*	0.03	0	0.00	17.92****	12.72†
2	<Human creates a copy of Artifact> [[Human]] or [[Institution]] or [[Eventuality 1]] replicates [[Eventuality 2]] <Human or Institution or Eventuality 1 recreates Eventuality 2>	19	63.33	0.71	-4.44	73	50.69	88.46****	82.77†	2	1.33	55.61****	50.41†
3	[[Human]] or [[Activity]] replicates [[Eventuality]] or {results findings} <Science. Human = Scientist or Activity = Study repeats Eventuality = Experiment in order to investigate whether this produces the same results or findings>	0	0.00	24.22****	19.06†	64	44.44	70.9****	65.22†	3	2.00	1.09	-4.10
4	[[MB_Entity]] replicates (Self) <MB Entity = DNA or Cell produces an exact copy of Self>	1	3.33	1.12	-4.04	1	0.69	164.65****	158.97†	130	86.67	39.24****	34.05†
5	[[Concept]] or [[Institution]]replicates ([[Self]]) <Concept reoccurs in different context>	5	16.67	9.96**	4.8†	2	1.39	2.86	-2.83	0	0.00	17.92****	12.72†
6	[[MB_Entity 1]] or [[Process]] replicates [[MB_Entity 2]] <[[MB_Entity 1]] or [[Process]] produces copy of [[MB_Entity 2]]>	0	0.00	0.00	-5.16	0	0.00	20.19****	14.5†	15	10.00	5.47*	0.28
	TOTAL	30				144				150			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A29

Pattern frequencies for three subcorpora and result of cross-corpora comparison for see

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
1	[[Human]] or [[Animal]] sees [[Physical_Object]] or [[Stuff]] <Human or Animal perceives or observes Physical_Object or Stuff with his or her eyes>	29	9.67	14.44***	8.04†	7	2.33	12.56	6.16†	27	9	0.07	-6.9
2	[[Human]] sees [[Event]] <Human is a witness to Event [NO ADVL]>	45	15.00	22.15****	15.75†	11	3.67	35.11****	28.72†	58	19.33	1.65	-4.75
3	[[Human]] sees [[Document]] or [[Document_Part]] <Human refers to Document or Document_Part for further information>	1	0.33	176.72****	170.32†	135	45.00	12.34***	5.94†	199	66.33	264.67****	258.27†
4	[[Human]] sees [[Location]] <Human refers to Document_Part as a reference>	25	8.33	1.67	-4.72	35	11.67	48.52****	42.12†	0	0.00	34.66****	28.26†
5	[[Human]] sees [[Proposition]] or [[Concept]] <Human achieves an understanding of Proposition or Concept>	81	27.00	76.94****	70.54†	6	2.00	0.29	-6.11	8	2.67	69.57****	63.18†

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A29 (cont.)

Pattern frequencies for three subcorpora and result of cross-corpora comparison for see

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
6	[[Human]] sees [[Anything]] [[Manner]] <Human regard Anything Manner, i.e., in a particular way>	8	2.67	54.28****	47.81†	68	22.67	94.27****	87.87†	0	0.00	11.09***	4.69†
7	[[Human]] sees [[Eventuality]] or [[Physical_Object]] as [[Anything]] <Human classifies Eventuality or Physical_Object as being Anything>	63	21.00	12.87***	6.47†	29	9.67	40.2****	33.81†	0	0.00	87.34****	80.94†
8	[[Human]] sees [[that-clause]] or [[wh-clause]] <Human achieves an understanding of Proposition or Concept embodied in that [clause] or wh- [clause]>	27	9.00	37.43****	31.03†	0	0.00	6.93**	0.53	5	1.67	16.62****	10.23†
9	[[Human]] or[[Institution]] sees [[Emotion]] or [[Attitude]] or [[Responsibility]] or [[Obligation]] <Human or Institution experiences an Emotion or Attitude or Obligation or Responsibility>	1	0.33	1.39	-5.01	0	0.00	0	-6.4	0	0.00	1.39	-5.01
10	[[Human 1]] sees [[Human 2]] <Human 1 goes to and spends some time with Human 2, typically for social reasons>	9	3.00	3.14	-3.26	3	1.00	4.16*	-2.24	0	0.00	12.48***	6.08†

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A29 (cont.)

Pattern frequencies for three subcorpora and result of cross-corpora comparison for see

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Freq.	%			Freq.	%			Freq.	%		
11	[[Human 1]] sees [[Human 2]] <Human 1 consults Human 2 = Expert or Doctor in order to get expert advice or treatment>	0	0.00	1.39	-5.01	1	0.33	1.39	-5.01	0	0.00	0	-6.4
12	[[Action]]is seen [[to-infinitive]] <Action is considered to/inf [verb], where the clause identifies an effect of the Action>	2	0.67	1.33	-5.07	5	1.67	0.51	66,67†	3	1.00	0.2	-6.2
13	[[Time_Period]] or [[Time_Point]] sees [[Event]] <Time_Period or Time_Point is the time when Event occurs>	6	2.00	8.32**	1.92	0	0.00	0	-6.4	0	0.00	8.32**	1.92
14	[[Human]] sees {fit} <Human considers it appropriate to/inf [verb]>	2	0.67	2.77	-3.62	0	0.00	0	-6.4	0	0.00	2.77	-3.62
15	[[Human1]] sees {the writing on the wall} <Human realises that Event=negative will happen soon>	1	0.33	1.39	-5.01	0	0.00	0	-6.4	0	0.00	1.39	-5.01
	TOTAL	300				300				300			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A30

Pattern frequencies for three subcorpora and result of cross-corpora comparison for specify

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro/ Hist LL	BIC
		Frequency	%			Frequency	%			Frequency	%		
1	[[Human]] or [[Institution]] or [[Concept]] or [[Document]] specifies [[Anything]] <Human or Institution or Document states clearly and precisely that Anything is a necessary condition for some activity, process, or concept>	56	80.00	1.82	-4.09	291	97.00	15.53****	9.66†	25	46.30	5.49*	0.67
2	[[Human]] or [[Institution]] or [[Concept]] or [[Document]] specifies [[that]] <Human or Institution or Document states clearly and precisely the conditions implied by [that [CLAUSE] as being necessary for some activity, process, or concept>	14	20.00	19.61****	13.69†	9	3.00	2.98	-2.89	0	0.00	16.01****	11.19†
3	[[MB Entity 1]] specifies [[MB Entity 2]] <MB Entity 1 contains the instructions necessary for the creation of MB Entity 2>	0	0.00	0.00	-5.91	0	0.00	109.06****	103.19†	29	53.70	48.22****	43.39†
	TOTAL	70				300				54			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = ≥ positive evidence against H0

Table A31

Pattern frequencies for three subcorpora and result of cross-corpora comparison for treat

PT #	Pattern <implicature>	History		Hist/ Man LL	BIC	Management		Man/ Micro LL	BIC	Microbiology		Micro / Hist LL	BIC
		Freq	%			Freq	%			Freq	%		
1	[[Human 1]] or [[Institution 1]] or [[Animal 1]] treats [[Human 2]] or [[Animal 2]] or [[Entity]] or [[Event]] Manner <Human 1 or Institution 1 or Animal 1 behaves toward Human 2 or Animal 2 or Entity or Event in the Manner specified>	122	81.33	0.76	-4.94	136	90.67	125.59****	119.89†	11	7.33	108.48****	102.78†
2	[[Human 1]] or [[Process]] or [[Drug]] treats [[Human 2]] or [[Animal]] or [[Disease]] or [[Injury]] <Human 1 = Health Professional applies a Drug or Process = Medical to Human 2 =Patient for the purpose of curing the patient`s Disease or Injury>	9	6.00	0.06	-5.64	8	5.33	43.87****	38.17†	59	39.33	41.11****	35.41†
3	[[Human]] or [[Device]] treats [[Inanimate]] (with [[Stuff]] or by [[Process]]) <The chemical or other properties of Inanimate are improved or otherwise changed by Process or the application of Stuff>	0	0.00	4.16*	-1.54	3	2.00	3.98*	-1.73	10	6.67	13.86***	8.16†
4	[[Human 1]] treats [[Human 2]] or [[Self]] to [[Eventuality]] <Human 1 gives or pays for Eventuality = Good as a benefit for Human 2 or Self>	1	0.67	1.39	-4.32	0	0.00	0	-5.7	0	0.00	1.39	-4.32
5	[[Human]] treats [[Anything]] <Human discusses Anything = Topic>	15	10.00	8.73**	3.03†	3	2.00	0	-5.7	3	2.00	8.73**	3.03†
6	[[Human 1]] treats with [[Human 2]] <Human 1 reaches a peace agreement with Human 2>	3	2.00	4.16	-1.54	0	0.00	0	-5.7	0	0.00	4.16*	-1.54
7	[[Human]] treats [[MB entity 1]] or [[Animal]] with [[MB entity 2]] or [[Drug]] <Human indtroduces drug or MB entity 2 to MB Entity 1=cell or Animal for experimental purposes>	0	0.00	0.00	-5.7	0	0.00	92.88****	87.18†	67	44.66	92.88****	84.18†
	TOTAL	150				150				150			

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Appendix D: Results of intradisciplinary comparison

Table A32

Pattern frequencies for sample a and b and result of intradisciplinary comparison for accept

PT #	Pattern <implicature>	History A		History B		A / B LL	BIC
		Frequency	%	Frequency	%		
1	[[Human]] or [[Institution]] accepts [[Proposition]] or [[Concept]] or [[Eventuality]] <Human or Institution agrees that Proposition or Concept or Eventuality is correct and does not need to be contested>	106	70.67	107	71.33	0.00	-5.71
2	[[Human]] or [[Institution]] accepts <Human or Institution agrees that that [clause] is true or correct>	8	5.33	13	8.67	1.2	-4.5
3	[[Human]] or [[Institution]] accepts {responsibility} (for Eventuality or for Entity) <Human or Institution agrees that he, she, or it is or will be responsible for Eventuality or Entity>	5	3.33	2	1.33	1.33	-4.38
4	[[Human 1]] accepts [[Entity]] or [[Money]] <Human 1 consents to receive Entity = Valuable or Money as a gift (or bribe) or as part of buisiness transaction from Human 2>	7	4.67	6	4.00	0.08	-5.63
5	[[Human]] accepts [[Offer]] <Human 1 agrees to act on the Offer of Human 2>	4	2.67	1	0.67	1.93	-3.78
6	[[Human 1]] accepts {resignation} <Human 1 = Authority receives and consents to resignation of Human 2 = Employee>	0	0.00	0	0.00	0	-5.7
7	[[Human]] accepts [[Human Role]] <Human agrees to undertake the work specified as required to fulfil appointment or job>	7	4.67	5	3.33	0.33	-5.37
8	[[Human 1]] or [[Institution 1]] accepts [[Human 2]] or [[Institution 2]] as [[Human_Role]] or as [[Institution_Role]] <Human 1 or Institution 1 acknowledges that Human 2 or Institution 2 is suitable for Human_Role or Institution_Role>	9	6.00	8	5.33	0.06	-5.7

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A32 (cont.)

Pattern frequencies for sample a and b and result of intradisciplinary comparison for accept

PT #	Pattern <implicature>	History A		History B		A / B LL	BIC
		Frequency	%	Frequency	%		
9	[[Human_Group 1]] or [[Human 1]] accepts [[Human 2]] or [[Human_Group 2]] <Human_Group 1 or Human 1 agrees to allow Human 2 or Human Group 2 to become a member of or be respected by Human Group 1>	5	3.33	7	4.67	0.33	-5.37
10	[[Location]] accepts [[Human Group]] <Location is suitable for Human Group>	0	0.00	1	0.67	1.39	-4.32
11	[[MB Entity]] accepts {task} <MB Entity= part of cell undertakes task in cell>	0	0.00	0	0.00	0	-5.7
12	[[MB Entity1]] accepts [[MB Entity 2]] <MB Entity 1 takes MB Entity 2 which becomes an integral part of MB Entity 1>	0	0.00	0	0.00	0	-5.7
13	[[MB_Entity 1]] accepts [[MB_Entity 2] as [[MB_Role]] <MB_Entity 2 plays a role in the normal functioning of MB_Entity 1>	0	0.00	0	0.00	0	-5.7
14	[[Device]] accepts [[Stuff]] <Device is capable of processing Stuff = biological sample>	0	0.00	0	0.00	0	-5.7
15	[[MB Entity]] accepts {topology} <MB Entity=gene occurs in a given spacial configuration>	0	0.00	0	0.00	0	-5.7
16	[[Concept 1]] accepts [[Concept 2]] or [[Eventuality]] <Concept 1 is compatible with Concept 2 or Eventuality>	0	0.00	0	0.00	0	-5.7
17	[[Human]] or [[Institution]] accepts [[Document]] <Human or Institution agrees that Document meets the standards required for publication>	0	0.00	0	0.00	0	-5.7
	TOTAL	150		150			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A33

Pattern frequencies for sample a and b and result of intradisciplinary comparison for account

PT #	Pattern <implicature>	Management A		Management B		A / B LL	BIC
		Frequency	%	Frequency	%		
1	[[Entity 1]] or [[Eventuality 1]] accounts for [[Numerical_Value]] of [[Entity 2]] or of [[Eventuality 2]] <Entity 1 or Eventuality 1 makes up a certain Numerical_Value = Percentage of the total number of Entity 2 or of Eventuality 2>	19	12.67	21	14.00	0.10	-5.6
2	[[Eventuality 1]] accounts for [[Eventuality 2]] <Eventuality 1 was a factor in causing Eventuality 2 to happen>	65	43.33	60	40.00	0.2	-5.5
3	[[Human]]or [[Concept]] accounts for [[Eventuality]] <Human or Concept explains or justifies Eventuality>	62	41.33	68	45.33	0.28	-4.43
4	[[Human]] accounts for [[Entity]] <Human is able to explain where Entity is or how it was used>	4	2.67	1	0.67	1.93	-3.78
5	[[MB Entity]] accounts for [[Eventuality]] <the presence of MB_Entity explains Eventuality>	0	0.00	0	0.00	0	-5.7
	TOTAL	150		150			

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A34

Pattern frequencies for sample a and b and result of intradisciplinary comparison for act

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
1	[[Human]] or [[Institution]] or [[Animal]] or [[Machine]] or [[MB unit]] acts <Human or Institution or Animal or Machine = Agent performs a motivated Action>	27	18.00	23	15.33	0.32	-5.38	22	14.67	29	19.33	0.96	-4.74	21	14.00	23	15.33	0.09	-5.61
2	[[Human]] or [[Institution]] or [[MB unit]] acts [[Manner]] <Human or Institution behaves in the manner specified>	55	36.67	60	40.00	0.22	-5.49	72	48.00	62	41.33	0.75	-4.96	34	22.67	41	27.33	0.65	-5.05
3	[[Entity 1]] acts as or like [[Anything]] <In a particular situation, Entity 1 performs the role or function specified>	50	33.33	49	32.67	0.01	-5.69	41	27.33	42	28.00	0.01	-5.69	74	49.33	62	41.33	1.06	-4.64
4	[[Human 1]] or [[Institution 1]] acts for [[Human 2]] or for [[Institution 2]] <Human 1 = Lawyer or Agent or Institution 1 is employed to perform certain tasks for Human 2 or Institution 2>	0	0.00	2	1.33	2.77	-2.93	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A34 (cont.)

Pattern frequencies for sample a and b and result of intradisciplinary comparison for act

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
5	[[Human 1]] or [[Institution 1]] acts on behalf of [[Human 2]] or on behalf of [[Institution 2]] <Human 1 or Institution 1 performs a motivated Action for the benefit and/or at the request of Human 2 or Institution 2>	2	1.33	4	2.67	0.68	-5.02	1	0.67	1	0.67	0	-5.7	0	0.00	0	0.00	0	-5.7
6	[[Human]] or [[Institution]] acts on [[Eventuality]] <Human or Institution behaves in accordance with Eventuality=Motivation>	10	6.67	5	3.33	1.7	-4	6	4.00	9	6.00	0.6	-5.1	0	0.00	0	0.00	0	-5.7
7	[[Human]] or [[Institution]] acts under [[Rule]] or under [[Command]] <Human or Institution behaves in accordance with Rule or Command>	0	0.00	1	0.67	1.39	-4.32	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7
8	[[Entity 1]] acts on or upon [[Entity 2]] <Entity1 has a particular effect on Entity 2>	3	2.00	5	3.33	0.51	-5.2	5	3.33	5	3.33	0	-5.7	21	14.00	23	15.33	0.09	-5.61

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A34 (cont.)

Pattern frequencies for sample a and b and result of intradisciplinary comparison for act

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
9	[[Human]] acts (Role) (in Performance) <Human plays Role = Theatrical (in Performance)>	1	0.67	0	0.00	1.39	-4.32	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7
10	Phrasal verb. [[Human]] acts [[Event]] or [[Human_Role]] or [[Emotion]] out <Human performs Role, not necessarily sincerely, or behaves as if feeling Emotion>	2	1.33	0	0.00	2.77	-2.93	0	0.00	1	0.67	1.39	-4.32	0	0.00	0	0.00	0	-5.7
11	[[Human]] acts [[ADJ]] <Human behaves in the manner specified by ADJ>	1	0.67	0	0.00	1.39	-4.32	1	0.67	0	0.00	1.39	-4.32	0	0.00	0	0.00	0	-5.7
12	[[Drug]] acts <Drug has an effect>	0	0.00	0	0.00	0	-5.7	1	0.67	0	0.00	1.39	-4.32	0	0.00	0	0.00	0	-5.7
13	[[Process]] acts <Process exerts an influence>	0	0.00	0	0.00	0	-5.7	1	0.67	1	0.67	0	-5.7	0	0.00	1	0.67	1.39	-4.32
	TOTAL	150		150				150		150				150		150			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A35

pattern frequencies for sample a and b and result of intradisciplinary comparison for appear

PT #	Pattern <implicature>	History A		History B		A / B LL	BIC	Man. A		Man. B		A / B LL	BIC	Micro. A		Micro. B		A / B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
1	[[Physical Object] appears <Physical_Object becomes visible>	4	2.67	4	2.67	0.00	-5.7	1	0.67	1	0.67	0.00	-5.7	2	1.33	1	0.67	0.34	-5.36
2	[[Anything]] appears [ADJ] <Anything seems to be ADJ>	21	14.00	25	16.67	0.35	-5.36	30	20.00	39	26.00	1.18	-4.53	13	8.67	19	12.67	1.13	-4.57
3	[[Abstract Entity]] or [[State of Affairs]] appears <Abstract_Entity or State_of_Affairs becomes noticeable or comes into existence>	6	4.00	10	6.67	1.01	-4.69	9	6.00	10	6.67	0.05	-5.65	3	2.00	2	1.33	0.2	-5.5
4	[[Stuff]] OR [[Physical Object]] appears [[Time Period]] <Stuff or Physical_Object comes into use or becomes available Time_Period>	6	4.00	4	2.67	0.4	-5.3	0	0.00	0	0.00	0	-5.7	5	3.33	0	0.00	6.93	1.23
5	[[Human]] OR [[Animal]] or [[MB Entity]] appears ([[Location]]) <Human or Animal arrives at Location>	2	1.33	1	0.67	0.34	-5.36	0	0.00	2	1.33	2.77	-2.93	2	1.33	4	2.67	0.68	-5.02

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A35 (cont.)
pattern frequencies for sample a and b and result of intradisciplinary comparison for appear

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
6	[[Human]] appears in [[Performance]] on [[TV Programme]] on [[Radio Programme]] or on [[Stage]] <Human takes part in Performance>	2	1.33	1	0.67	0.34	-5.36	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7
7	[[Document 1]] OR [[Image]] appears (in [[Document 2]] or [[Document Part]]) <Document 1 or Image is published (in Document 2 or Document Part)>	25	16.67	29	19.33	0.3	-5.41	7	4.67	9	6.00	0.25	-5.45	0	0.00	1	0.67	1.39	-4.32
8	[[Anything]] appears [To-Infinitive] <Anything seems to/inf [verb]>	58	38.67	61	40.67	0.08	-5.63	81	54.00	79	52.67	0.03	-5.68	110	73.33	110	73.33	0	-5.7
9	{It} appears [[That Clause]] <that-[CLAUSE] is probably true>	12	8.00	8	5.33	0.81	-4.9	14	9.33	14	9.33	0	-5.7	13	8.67	13	8.67	0	-5.7

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A35 (cont.)
pattern frequencies for sample a and b and result of intradisciplinary comparison for appear

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
10	[[Entity 1]] or [[Eventuality 1]] appears (as) [[Entity 2]] or(as) [[Eventuality 2]] <Entity 1 or Eventuality 1 is perceived as or assumed to be Entity 2 or Eventuality 2>	10	6.67	6	4.00	1.01	-4.69	3	2.00	1	0.67	1.05	-4.66	2	1.33	0	0.00	2.77	-2.93
11	[[Human]] appears in or before [[Institution]] or before [[Human]] <Law: court procedure. Human presents Self formally in Court or before Judge as a Defendant to be tried in respect of Accusation>	1	0.67	1	0.67	0	-5.7	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7
12	[[Abstract Entity]] appear in [[Document]] <Document mentions abstract entity>	2	1.33	0	0.00	2.77	-2.93	5	3.33	0	0.00	6.9	1.23	0	0.00	0	0.00	0	-5.7
13	[[Human]] appears as {an MP} <Human undertakes the role of MP>	1	0.67	0	0.00	1.39	4.32	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7
	TOTAL	150		150				150		150				150		150			

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A36

Pattern frequencies for sample a and b and result of intradisciplinary comparison for apply

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
1	[[Human]] or [[Institution]] applies [[Concept]] or [[Process]] (to [[State of Affairs]]) <Human or Institution makes use of Concept or Process in a particular situation or State_of_Affairs>	84	56	79	52.67	0.15	-5.55	106	70.67	109	72.67	0.04	-5.66	99	66.00	93	62.00	0.19	-5.52
2	[[Concept]] or [[Process]] applies (to [[State of Affairs]] or [[Activity]]) <Concept or Process is relevant to State_of_Affairs or Activity>	43	28.67	48	32.00	0.27	-5.43	36	24.00	36	24.00	0	-5.7	13	8.67	14	9.33	0.04	-5.67
3	[[Human]] applies for [[Benefit]] benefit or injunction or extension or admission or readmission or {[NP] to be [V]} <Human 1 formally asks Human 2 or Institution to give Benefit to Human 1>	5	3.33	8	5.33	0.7	-5.01	4	2.67	2	1.33	0.68	-5.02	0	0.00	0	0.00	0	-5.7

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A36 (cont.)

Pattern frequencies for sample a and b and result of intradisciplinary comparison for apply

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
4	[[Human]] OR [[Device]] applies [Fluid]] or [[Stuff]] to [[Surface]] <Human (uses Device to) spread Fluid or Stuff on Surface, typically in order to heal, improve, protect Surface or conduct test>	5	3.33	3	2.00	0.51	-5.2	0	0.00	0	0.00	0	-5.7	33	22.00	37	24.67	0.23	-5.48
5	[[Human]] applies [[Word]] to [[Anything]] <Human uses Word to suitably refer to Anything>	9	6.00	8	5.33	0.06	-5.64	2	1.33	3	2.00	0.2	-5.5	2	1.33	4	2.67	0.68	-5.02
6	[[Human]] or [[Institution]] applies {pressure} <Human tries to influece decision or action of another human or institution>	3	2.00	4	2.67	0.14	-5.56	2	1.33	0	0.00	2.77	-2.93	2	1.33	2	1.33	0	-5.7
7	[[Human]] applies [[Self]] to [[Activity]] <Human makes great effort to conduct activity>	1	0.67	0	0.00	1.39	-4.32	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7
8	[[Human]] applies {caution} <Human take a cautious approach to [[Anything]]>	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7	1	0.67	0	0.00	1.39	-4.32
	TOTAL	150		150				150		150				150		150			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A37

Pattern frequencies for sample a and b and result of intradisciplinary comparison for call

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
1	[[Anything]] is called [[Name]] <Anything is referred to as Name>	37	24.67	39	26.00	0.05	-5.65	54	36.00	54	36.00	0.00	-5.7	136	90.67	138	92.00	0.01	-5.69
2	[[Human]] OR [[Institution]] calls [[Anything]] [[Name]] <Human or Institution invents or uses Name to refer to Anything>	61	40.67	58	38.67	0.08	-5.63	53	35.33	57	38.00	0.15	-5.56	6	4.00	3	2.00	1.02	-4.68
3	[[Human]] or [[Institution]] calls [[Event]] <Human or Institution instructs people to cause Event = Meeting or Action to happen immediately>	2	1.33	2	1.33	0	-5.7	1	0.67	0	0.00	1.39	-4.32	0	0.00	0	0.00	0	-5.7
4	[[Human]] or [[Institution]] or [[Document]] calls {for} [[Action]] or {for} [[State of Affairs]] <Human or Institution or Document says that other people should do Action or create [State_of_Affairs]>	17	11.33	21	14.00	0.42	-5.28	25	16.67	24	16.00	0.02	-5.68	4	2.67	5	3.33	0.11	-5.59
5	[[Human]] or [[Institution]] calls [[Human 2]] {for} [[Activity]] <Human 1 or Institution officially invites or instructs Human 2 to take part in Activity>	1	0.67	0	0.00	1.39	-4.32	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A37 (cont.)
Pattern frequencies for sample a and b and result of intradisciplinary comparison for call

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
6	[[Human 1]] calls ((in or round) (on [[Human 2]]) at [[Location]]) <Human 1 goes to Location in order to meet Human 2>	1	0.67	2	1.33	0.34	-5.36	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7
7	[[Human 1]] calls [[Human 2]] <Human 1 shouts to Human 2, typically in order to ask Human 2 to come to them>	1	0.67	5	3.33	2.91	-2.79	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7
8	[[Human]] calls [[Speech_Act]] (out) <Human says Speech_Act or QUOTE in a loud, clear voice>	1	0.67	2	1.33	0.34	-5.36	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7
9	[[Human 1]] calls [[Human 2]] or [[Institution]] or [[Number]] <Human 1 contacts Human 2 or Institution on the telephone by selecting Number = Telephone Number>	1	0.67	0	0.00	1.39	-4.32	2	1.33	3	2.00	0.2	-5.5	0	0.00	0	0.00	0	-5.7
10	[[Human 1]] or [[Institution 2]] or [[Document]] calls {on} or {upon} [[Human 2]] OR [Institution 2]] [TO-INFINITIVE] <Human 1 or Institution 1 or Document asks Human 2 or Institution 2 to [verb]>	13	8.67	10	6.67	0.39	-5.31	7	4.67	6	4.00	0.08	-5.63	2	1.33	0	0.00	2.77	-2.93

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A37 (cont.)
Pattern frequencies for sample a and b and result of intradisciplinary comparison for call

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
11	[[Human 1]] or [[Institution 1]] calls [[Human 2]] or [[Institution 2]] {in} <Human 1 or Institution 1 asks Human 2 or Institution 2 to come and give help or advice>	1	0.67	1	0.67	0	-5.7	3	2.00	0	0.00	4.16	-1.54	0	0.00	0	0.00	0	-5.7
12	[[Human]] calls [[Activity]] {off} <Human cancels Activity>	3	2.00	0	0.00	4.16	-1.54	1	0.67	0	0.00	1.39	-4.32	0	0.00	0	0.00	0	-5.7
13	[[Human]] calls [[Information]] or [[Image]] {up} <Human causes Information or Image to be displayed on a computer screen>	0	0.00	1	0.67	1.39	-4.32	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7
14	[[Human]] calls {attention to} [[Event]] <Human asks people to notice Event>	2	1.33	3	2.00	0.2	-5.5	1	0.67	4	2.67	1.93	-3.78	1	0.67	0	0.00	1.39	-4.32
15	[[Human]] or [[Anything] calls [[Proposition]] {into question} <Human or Anything provides a reason for doubting whether Proposition is correct>	2	1.33	4	2.67	0.68	-5.02	2	1.33	1	0.67	0.34	-5.36	1	0.67	4	2.67	1.93	-3.78
16	[[Human 1]] calls {upon} [[Human 2]] <Human 1 asks Human 2 for assistance>	2	1.33	1	0.67	0.34	-5.36	1	0.67	1	0.67	0	-5.7	0	0.00	0	0.00	0	-5.7

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A37 (cont.)
Pattern frequencies for sample a and b and result of intradisciplinary comparison for call

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
17	[[Human 1]] calls [[Anything]] {down upon} [[Human 2]] <Human evokes Anything to affect Human 2>	2	1.33	0	0.00	2.77	-2.93	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7
18	[[Human]] is called to {celibacy} <Human chooses to remain celibate>	1	0.67	0	0.00	1.39	-4.32	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7
19	[[Concept]] called into {being} <Human creates concept to explain or deal with Eventuality>	0	0.00	1	0.67	1.39	-4.32	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7
20	[[Institution]] calls [[Human]] {up} <Institution = Government requires Human to undertake military service>	2	1.33	0	0.00	2.77	-2.93	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7
	TOTAL	150		150				150		150				150		150			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A38

Pattern frequencies for sample a and b and result of intradisciplinary comparison for conduct

PT #	Pattern <implicature>	Management A		Management B		A/B LL	BIC
		Frequency	%	Frequency	%		
1	[[Human]] or [[Institution]] conducts [[Activity]] <Human or Institution carries out Activity>	149	99.33	149	99.33	0.00	-5.7
2	[[Human 1]] conducts [[Human 2]] [Direction] <Human 1 accompanies and guides Human 2 Direction>	0	0.00	0	0.00	0.00	-5.7
3	[[Human]] conducts [[Self]] [Manner] <Human behaves in a particular way>	1	0.67	1	0.67	0.00	-5.7
4	[[Metal]] or [[Fluid]] conducts [[Energy]] <Energy is transferred via Metal>	0	0.00	0	0.00	0.00	-5.7
5	[[Route]] conducts [[Fluid]] <Stuff passes through Aperture>	0	0.00	0	0.00	0.00	-5.7
6	[[MB Entity]] conducts [[Process]] <MB entity carries out Process = Microbiological>	0	0.00	0	0.00	0.00	-5.7
	TOTAL	150		150			

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A39

Pattern frequencies for sample a and b and result of intradisciplinary comparison for explain

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
	[[Human]] or [[Institution]] or [[Document]] or [[Proposition]] or [[Eventuality 1]] explains [[Eventuality 2]] (in terms of [[Eventuality 3]]) (to [[Human 2]])																		
1	<Human or Institution formulates (in conversation with Human 2 or in a Document intended for Human 2 to read) a Proposition concerning the cause or effect of an Eventuality 2> [[Human]] or [[Institution]] or [[Document]] or [[Proposition]] explains (to [[Human 2]])	82	54.67	89	59.33	0.29	-5.42	113	75.33	119	79.33	0.16	-5.55	122	81.33	128	85.33	0.14	-5.56
2	<Human or Institution formulates Proposition in Document concerning the cause or effect of an Eventuality that is expressed in wh- [clause]> [[Human]] or [[Institution]] explains [That-Clause] (to [[Human 2]])	30	20.00	32	21.33	0.06	-5.64	21	14.00	21	14.00	0	-5.7	25	16.67	21	14.00	0.35	-5.36
3	<Human 1 or Document tells Human 2) that [clause], as a way of accounting for Eventuality (often implied rather than stated explicitly)>	24	16.00	13	8.67	3.32	-2.38	3	2.00	4	2.67	0.14	-5.56	2	1.33	0	0.00	2.77	-2.93

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A39 (cont.)
Pattern frequencies for sample a and b and result of intradisciplinary comparison for explain

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
4	[[Human]] or [[Institution]] explains [[Quote]] <Human 1 or Document tells Human 2) that [clause], as a way of accounting for Eventuality (often implied rather than stated explicitly)> [[Human]] or [[Institution]] explains [[Eventuality]] {away}	13	8.67	14	9.33	0.04	-5.67	13	8.67	6	4.00	2.64	-3.06	0	0.00	1	0.67	1.39	-4.32
5	<Human 1 formulates Proposition as a way of accounting for Eventuality = Inconvenient>	1	0.67	2	1.33	0.34	-5.36	0	0.00	0	0.00	0	-5.7	1	0.67	0	0.00	1.39	-4.32
	TOTAL	150		150				150		150				150		150			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A40

Pattern frequencies for sample a and b and result of intradisciplinary comparison for follow

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
1	[[Human 1]] or [[Animal 1]] or [[Vehicle 1]] or [[MB Entity 1]] follows [[Human 2]] or [[Animal 2]] or [[Vehicle 2]] or [[MB Entity 2]] <Human 1 or Animal 1 or Vehicle 1 moves in the same direction as that selected by Human 2 or Animal 2 or the driver of Vehicle 2>	7	4.67	9	6.00	0.25	-5.45	0	0.00	3	2.00	4.16	-1.54	2	1.33	1	0.67	0.34	-5.36
2	[[Human 1]] follows [[Human 2]] <Human 1 = Student or Disciple studies and is influenced by or tries to practice the teachings of Human 2 = Teacher>	10	6.67	9	6.00	0.05	-5.65	23	15.33	26	17.33	0.18	-5.52	0	0.00	0	0.00	0	-5.7
3	[[Human_Group 1]] or [[Institution 1]] follows [[Human_Group 2]] or [[Institution 2]] <Human_Group 1 or Institution 1 is influenced by and thus copies the behaviour of Human Group 2 or Institution 2>	4	2.67	5	3.33	0.11	-5.59	2	1.33	0	0.00	2.77	-2.93	0	0.00	0	0.00	0	-5.7
4	[[Human]] or [[Institution]] or [[Concept]] or [[Process]]follows [[Command]] or [[Rule]] or [[Plan]] or [[Document]] <Human or Institution acts in accordance with Command or Rule or Plan (expressed in Document)>	40	26.67	42	28.00	0.05	-5.65	53	35.33	49	32.67	0.16	-5.55	16	10.67	21	14.00	0.68	-5.03

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A40 (cont.)
Pattern frequencies for sample a and b and result of intradisciplinary comparison for follow

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
5	[[Human]] or [[Institution]] follows [[Event]] <Human or Institution takes an interest in Event = Unfolding>	6	4.00	6	4.00	0	-5.7	1	0.67	0	0.00	1.39	-4.32	8	5.33	13	8.67	1.2	-4.5
6	[[Human]] follows [[Proposition]] or [[wh-clause]] <Human is able to understand Proposition or wh-clause>	0	0.00	0	0.00	0	-5.7	0	0.00	1	0.67	1.39	-4.32	0	0.00	0	0.00	0	-5.7
7	[[Human]] or [[Institution]] follows {lead} or {line of enquiry} <Human or Institution pursues an investigation based on specific information>	0	0.00	0	0.00	0	-5.7	1	0.67	1	0.67	0	-5.7	0	0.00	0	0.00	0	-5.7
8	[[Human]] or [[Institution]] or [[Event]] follows {trend} <Human or Institution performs in line with the expectations dictated by trend >	2	1.33	1	0.67	0.35	-5.36	1	0.67	0	0.00	1.39	-4.32	2	1.33	2	1.33	0	-5.7
9	[[Event 1]] follows (Event 2) <Event 1 happens after and typically as a consequence of Event 2>	49	32.67	54	36.00	0.24	-5.46	23	15.33	18	12.00	0.61	-5.09	86	57.33	72	48.00	1.24	-4.46
10	[[Eventuality 1]] follows from [[Eventuality 2]] <Eventuality 1 is a necessary consequence of Eventuality 2>	5	3.33	3	2.00	0.51	-5.2	10	6.67	6	4.00	1.01	-4.69	1	0.67	1	0.67	0	-5.7

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A40 (cont.)
Pattern frequencies for sample a and b and result of intradisciplinary comparison for follow

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
11	{it} follows [[that-clause]] (from Eventuality 1) <Eventuality 2 expressed by [CLAUSE] is a necessary consequence of Eventuality 1 (either expressed in a from-phrase or one that was previously stated)> [[Anything 1]] follows [[Anything 2]]	5	3.33	0	0.00	6.93	1.23	5	3.33	6	4.00	0.09	-5.61	4	2.67	0	0.00	5.55	-0.16
12	<Anything 1 = Item in List comes after Anything 2 = Item in List in a list or sequence> [[Human]] follows {up} [[Eventuality]]	6	4.00	7	4.67	0.08	-5.63	7	4.67	7	4.67	0	-5.7	16	10.67	18	12.00	0.12	-5.59
13	<Human monitors Eventuality and finds out more about it>	2	1.33	3	2.00	0.2	-5.5	1	0.67	1	0.67	0	-5.7	1	0.67	0	0.00	1.39	-4.32
14	as follows <as will now be stated here> [[Human 1]] or [[Institution 1]] follows {suit}	7	4.67	6	4.00	0.08	-5.63	21	14.00	30	20.00	1.6	-4.11	11	7.33	17	11.33	1.3	-4.41
15	<Human 1 or or Institution 1 does the same as Human 2 or Institution 2 has done> [[Human 1]] follows in footsteps {of [[Human 2]]}	2	1.33	2	1.33	0	-5.7	2	1.33	0	0.00	2.77	-2.93	0	0.00	0	0.00	0	-5.7
16	<Human 1 engages in similar activity to one previously engaged in by Human 2>	4	2.67	1	0.67	1.93	-3.78	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table 40 (cont.)
Pattern frequencies for sample a and b and result of intradisciplinary comparison for follow

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
17	[[Human]] follows up with [[Speech Act]] <Human adds additional details to information already given>	1	0.67	0	0.00	1.39	-4.32	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7
18	[[MB Entity]] follows [[Route]] <MB Entity follows an established route or conduit>	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7	3	2.00	5	3.33	0.51	-5.2
19	[[Human]] follows {through on commitment} <human fulfills a commitment>	0	0.00	0	0.00	0	-5.7	0	0.00	2	1.33	2.77	-2.93	0	0.00	0	0.00	0	-5.7
	TOTAL	150		150				150		150				150		150			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A41

Pattern frequencies for sample a and b and result of intradisciplinary comparison for lead

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
1	[[Eventuality 1]] leads to [[Eventuality 2]] <Eventuality 1 is the cause of Eventuality 2>	72	48.00	86	57.33	1.24	-4.46	117	78.00	117	78.00	0.00	-5.7	148	98.67	147	98.00	0.00	-5.7
2	[[Eventuality 1]] leads {up to} [[Eventuality 2]] <Eventuality 1 = PLURAL precedes but may not be the cause of Eventuality 2>	1	0.67	4	2.67	1.93	-3.78	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0.00	-5.7
3	[[Eventuality]] leads [[Human]] or [[Institution]] [[to-infinitive]] <Eventuality causes, enables, or encourages Human or Institution to/INF [verb]>	22	14.67	25	16.67	0.19	-5.51	24	16.00	22	14.67	0.09	-5.62	2	1.33	2	1.33	0.00	-5.7
4	[[Human]] or [[Institution 1]] leads [[Human_Group]] or [[Institution 2]] <Human or Institution 1 organizes, directs, or provides a model for Activity of Human_Group or Institution 2>	27	18.00	17	11.33	2.29	-3.41	2	1.33	4	2.67	0.68	-5.02	0	0.00	0	0.00	0.00	-5.7
5	[[Human]] or [[Institution]] leads [[Activity]] <Human or Institution organizes, directs, or provides a model for the Activity of Human_Group>	7	4.67	6	4.00	0.08	-5.63	5	3.33	2	1.33	1.33	-4.38	0	0.00	1	0.67	0.00	-5.7

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A41 (cont.)
Pattern frequencies for sample a and b and result of intradisciplinary comparison for lead

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
6	[[Human 1]] leads [[Human 2]] or [[Human_Group]](Direction to Location) <Human 1 organizes and directs the movement of Human 2 or Human_Group to Location by accompanying Human 2 or Human Group and showing the way>	8	5.33	4	2.67	1.36	-4.34	2	1.33	2	1.33	0	-5.7	0	0.00	0	0.00	0.00	-5.7
7	[[Route]] leads [[Direction]] to [[Location]] <Route is a path or road Direction or to Location>	6	4.00	6	4.00	0	-5.7	0	0.00	1	0.67	1.39	-4.32	0	0.00	0	0.00	0.00	-5.7
8	[[Human]] or [[Human_Group]] leads (Activity) <Human = Competitor or Human_Group = Competitor is in first place in Activity = Competition or Race)>	0	0.00	1	0.67	1.39	-4.32	0	0.00	1	0.67	1.39	-4.32	0	0.00	0	0.00	0.00	-5.7
9	[[Human]] or [[Institution]] leads the way(in-ING) <Human or Institution is the best at doing something and shows how it can best be done>	2	1.33	0	0.00	2.77	-2.93	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0.00	-5.7

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A41 (cont.)
Pattern frequencies for sample a and b and result of intradisciplinary comparison for lead

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
10	[[Human]] or [[Animate]] leads MOD {life} or {existence} <Human or Institution is the best at doing something and shows how it can best be done>	4	2.67	1	0.67	1.93	-3.78	1	0.67	0	0.00	1.39	-4.32	0	0.00	0	0.00	0.00	-5.7
11	[[Human 1]] or [[Eventuality]] leads [[Human 2]] on <Human 1 or Eventuality encourages Human 2, typically to do or believe something stupid or reprehensible>	1	0.67	0	0.00	1.39	-4.32	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0.00	-5.7
	TOTAL	150		150				150		150				150		150			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A42

Pattern frequencies for sample a and b and result of intradisciplinary comparison for maintain

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
1	[[Human]] or [[Process]] maintains [[State_of_Affairs]] <Human takes action to ensure or Process has the effect that State_of_Affairs remains unchanged>	104	69.33	107	71.33	0.04	-5.66	120	80.00	134	89.33	0.77	-4.93	126	84.00	128	85.33	0.02	-5.69
2	[[Human]] or [[Institution]] or [[Document]] maintains [[that-clause]] <Human or Institution or Document asserts strongly and consistently over time that [clause] or [Proposition]>	35	23.33	29	19.33	0.56	-5.14	15	10.00	10	6.67	1.01	-4.7	3	2.00	0	0.00	4.16	-1.54
3	[[Human]] or [[Institution]] maintains [[Document]] <Human or Institution regularly and systematically updates Document>	0	0.00	1	0.67	1.39	-4.32	7	4.67	3	2.00	1.65	-4.06	1	0.67	1	0.67	0	-5.7
4	[[Human]] maintains [[Building]] or [[Vehicle]] or [[Artifact]] <Human takes action to ensure that Building or Vehicle or Artifact is kept in good working order>	9	6.00	10	6.67	0.05	-5.65	7	4.67	2	1.33	2.94	-2.76	0	0.00	1	0.67	1.39	-4.32

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A42 (cont.)
Pattern frequencies for sample a and b and result of intradisciplinary comparison for maintain

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
5	[[Human 1]] maintains [[Self]] or [[Huaman 2]] <Human 1 provides food and drink for Self or Human 2, enabling Self or Human 2 to stay alive>	2	1.33	3	2.00	0.2	-5.5	1	0.67	1	0.67	0	-5.7	2	1.33	0	0.00	2.77	-2.93
6	[[MB_Entity]] is maintained in [[Location]] or on [[Surface]] <MB Entity is place in Location or on Surface for the purpose of a test>	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7	18	12.00	20	13.33	0.11	-5.6
	TOTAL	150		150				150		150				150		150			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A43

Pattern frequencies for sample a and b and result of intradisciplinary comparison for manage

PT #	Pattern <implicature>	Management A		Management B		A / B LL	BIC
		Frequency	%	Frequency	%		
1	[[Human]] or [[Institution]] or [[MB Entity]] manages [to-infinitive] <Human or Institution is able to/inf [verb], despite difficult circumstances>	5	3.33	11	7.33	2.31	-3.4
2	[[Human]] manages <Human is able to deal with difficult circumstances successfully>	0	0.00	1	0.67	1.39	-4.32
3	[[Human]] or [[Institution]] [[MB Entity]] manages [[Action]] or [[State_of_Affairs]] <Human or Institution is able to perform Action = Desirable or achieve State_of_Affairs = Desirable, despite difficult circumstances>	1	0.67	4	2.67	1.93	-3.78
4	[[Human]] manages [[Artifact]] <Human is able to lift, carry, or use Artifact without difficulty>	1	0.67	0	0.00	1.39	-4.32
5	[[Human]] or [[Institution 1]] manages [[Human 2]] or [[Institution 2]] <Human or Institution 1 is responsible for and controls Institution 2>	31	20.67	16	10.67	4.87*	-0.83
6	[[Human]] or [[Institution]] or [[MB Entity]] manages [[System]] or [[Activity]] <Human or Institution is responsible for and controls System or Activity = Ongoing>	96	64.00	95	63.33	0.01	-5.7
7	[[Human]] or [[Institution]] manages [[Resource]] <Human or Institution is responsible for Resource>	16	10.67	21	14.00	0.68	-5.03
8	[[Human]] or [[Institution]] manages [[Location]] <Human or Institution is responsible for the use and maintenance of Location>	0	0.00	2	1.33	2.77	-2.93
9	[[Human]] manages [[Disease]] <Human alleviates symptoms of Disease with drugs or Therapy>	0	0.00	0	0.00	0	-5.7
	TOTAL	150		150			

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A44

Pattern frequencies for sample a and b and result of intradisciplinary comparison for need

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A / B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
1	[[Human]] or [[Institution]] needs [[Eventuality]] or [[Entity]] <Human or Institution requires that Eventuality or Entity must be realized or available, in order to accomplish some goal>	59	39.33	63	42.00	0.13	-5.57	45	30.00	45	30.00	0.00	-5.7	37	24.67	39	26.00	0.05	-5.65
2	[[Entity 1]] or [[Eventuality 1]] needs [[Entity 2]] or [[Eventuality 2]] <Entity 1 or Eventuality 1 is an essential precondition for or attribute of Entity 2 or Eventuality 2>	15	10.00	14	9.33	0.03	-5.67	17	11.33	19	12.67	0.11	-5.59	43	28.67	50	33.33	0.53	-5.18
3	[[Entity]] or [[Eventuality]] needs [[to-infinitive]] <An essential precondition for the realization of Eventuality is that [verb] must be realized typically to be>	30	20.00	32	21.33	0.06	-5.64	30	20.00	34	22.67	0.25	-5.45	44	29.33	40	26.67	0.19	-5.51
4	[[Human]] or needs [[to-infinitive]] <Human must do [verb]>	45	30.00	38	25.33	0.59	-5.11	58	38.67	52	34.67	0.33	-5.38	16	10.67	15	10.00	0.03	-5.67
5	[[Plant]] or [[Animate]] or [[MB Entity]] needs [[Eventuality]] or [[Stuff]] <Plant or Animate must have Eventuality or Stuff in order to survive and flourish>	1	0.67	3	2.00	1.05	-4.66	0	0.00	0	0.00	0	-5.7	10	6.67	6	4.00	1.01	-4.69
	TOTAL	150		150				150		150				150		150			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A45

Pattern frequencies for sample a and b and result of intradisciplinary comparison for note

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
1	[[Human]] notes [[Information]] or [[Eventuality]] or [[Visible_Feature]] <Human notices and pays particular attention to Information about Eventuality or Visible_Feature>	53	35.33	51	34.00	0.04	-5.67	48	32.00	50	33.33	0.04	-5.66	62	41.33	75	50.00	1.24	-4.47
2	[[Human]] notes [[quote]] or [[that-clause]] or [[wh-clause]] <Human notices and pays particular attention to Eventuality or Visible_Feature>	86	57.33	87	58.00	0.01	-5.7	95	63.33	95	63.33	0	-5.7	82	54.67	71	47.33	0.79	-4.91
3	[[Human]] notes [[Information]] (on or in [[Document]]) (under [[Document_Part]]) <Human makes a written note of Information (on | in Document) (under Document_Part = Heading)>	11	7.33	12	8.00	0.04	-5.66	7	4.67	5	3.33	0.33	-5.37	3	2.00	2	1.33	0.2	-5.5
4	[[MB Entity]] is noted [[Name]] <MB Entity is referred to as Name>	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7	3	2.00	2	1.33	0.2	-5.5
	TOTAL	150		150				150		150				150		150			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A46

Pattern frequencies for sample a and b and result of intradisciplinary comparison for see

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
1	[[Human]] or [[Animal]] sees [[Physical_Object]] or [[Stuff]] <Human or Animal perceives or observes Physical_Object or Stuff with his or her eyes>	15	10.00	14	9.33	0.03	-5.67	4	2.67	3	2.00	0.14	-5.56	15	10.00	12	8.00	0.33	-5.37
2	[[Human]] sees [[Event]] <Human is a witness to Event [NO ADVL]>	23	15.33	22	14.67	0.02	-5.68	3	2.00	8	5.33	2.36	-3.35	27	18.00	31	20.67	0.42	-5.28
3	[[Human]] sees [[Document]] or [[Document_Part]] <Human refers to Document or Document_Part for further information>	0	0.00	1	0.67	1.39	-4.32	75	50.00	60	4.00	1.67	-4.03	100	66.67	99	66.00	0.01	-5.7
4	[[Human]] sees [[Location]] <Human refers to Document_Part as a reference>	15	10.00	10	6.67	1.01	-4.7	16	10.67	19	12.67	0.26	-5.45	0	0.00	0	0.00	0	-5.7

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A46 (cont.)
Pattern frequencies for sample a and b and result of intradisciplinary comparison for see

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
5	[[Human]] sees [[Proposition]] or [[Concept]] <Human achieves an understanding of Proposition or Concept>	37	24.67	44	29.33	0.61	-5.1	2	1.33	4	2.67	0.68	-5.02	5	3.33	3	2.00	0.51	-5.2
6	[[Human]] sees [[Anything]] [[Manner]] <Human regard Anything Manner, i.e., in a particular way>	5	3.33	3	2.00	0.51	-5.2	32	21.33	36	24.00	0.24	-5.47	0	0.00	0	0.00	0	-5.7
7	[[Human]] sees [[Eventuality]] or [[Physical_Object]] as [[Anything]] <Human classifies Eventuality or Physical_Object as being Anything>	33	22.00	30	20.00	0.14	-5.56	16	10.67	13	8.67	0.31	-5.39	0	0.00	0	0.00	0	-5.7
8	[[Human]] sees [[that-clause]] or [[wh-clause]] <Human achieves an understanding of Proposition or Concept embodied in that [clause] or wh-[clause]>	13	8.67	14	9.33	0.04	-5.67	0	0.00	0	0.00	0	-5.7	1	0.67	4	2.67	1.93	-3.78

Note. *= $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$, † = at least positve evidence against H_0

Table A46 (cont.)
Pattern frequencies for sample a and b and result of intradisciplinary comparison for see

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
9	[[Human]] or[[Institution]] sees [[Emotion]] or [[Attitude]] or [[Responsibility]] or [[Obligation]] <Human or Institution experiences an Emotion or Attitude or Obligation or Responsibility>	1	0.67	0	0.00	1.39	-4.32	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7
10	[[Human 1]] sees [[Human 2]] <Human 1 goes to and spends some time with Human 2, typically for social reasons>	4	2.67	5	3.33	0.11	-5.59	2	1.33	1	0.67	0.34	-5.36	0	0.00	0	0.00	0	-5.7
11	[[Human 1]] sees [[Human 2]] <Human 1 consults Human 2 = Expert or Doctor in order to get expert advice or treatment>	0	0.00	0	0.00	0	-5.7	1	0.67	0	0.00	1.39	-4.32	0	0.00	0	0.00	0	-5.7

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A46 (cont.)
Pattern frequencies for sample a and b and result of intradisciplinary comparison for see

PT #	Pattern <implicature>	History A		History B		A/B LL	BIC	Man. A		Man. B		A/B LL	BIC	Micro. A		Micro. B		A/B LL	BIC
		Freq.	%	Freq.	%			Freq.	%	Freq.	%			Freq.	%	Freq.	%		
12	[[Action]]is seen [[to-infinitive]] <Action is considered to/inf [verb], where the clause identifies an effect of the Action>	0	0.00	2	1.33	2.77	-2.93	2	1.33	3	2.00	0.2	-5.5	2	1.33	1	0.67	0.34	-5.36
13	[[Time_Period]] or [[Time_Point]] sees [[Event]] <Time_Period or Time_Point is the time when Event occurs>	3	2.00	3	2.00	0	-5.7	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7
14	[[Human]] sees {fit} <Human considers it appropriate to/inf [verb]>	1	0.67	1	0.67	0	-5.7	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7
15	[[Human1]] sees {the writing on the wall} <Human realises that Event=negative will happen soon>	0	0.00	1	0.67	1.39	-4.32	0	0.00	0	0.00	0	-5.7	0	0.00	0	0.00	0	-5.7
	TOTAL	150		150				150		150				150		150			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0

Table A47

Pattern frequencies for sample a and b and result of intradisciplinary comparison for specify

PT #	Pattern <implicature>	Management A		Management B		A/B LL	BIC
		Frequency	%	Frequency	%		
1	[[Human]] or [[Institution]] or [[Concept]] or [[Document]] specifies [[Anything]] <Human or Institution or Document states clearly and precisely that Anything is a necessary condition for some activity, process, or concept>	147	98.00	144	96.00	0.03	-5.67
2	[[Human]] or [[Institution]] or [[Concept]] or [[Document]] specifies [[that]] <Human or Institution or Document states clearly and precisely the conditions implied by [that [CLAUSE] as being necessary for some activity, process, or concept>	3	2.00	6	4.00	1.03	-4.68
3	[[MB Entity 1]] specifies [[MB Entity 2]] <MB Entity 1 contains the instructions necessary for the creation of MB Entity 2>	0	0.00	0	0.00	0	-5.7
	TOTAL	150		150			

Note. *= p < .05, ** = p < .01, *** = p < .001, **** = p < .0001, † = at least positve evidence against H_0