Contents lists available at ScienceDirect

Journal of English for Academic Purposes

journal homepage: www.elsevier.com/locate/jeap

Phrases in EAP academic writing pedagogy: Illuminating Halliday's influence on research and practice

David Oakey

Department of English, University of Liverpool, 19 Abercromby Square, Liverpool, L69 7ZG, United Kingdom

ARTICLE INFO

ABSTRACT

This paper looks at how phrases are conceptualized in English for Academic Purposes (EAP) research and practice. It presents examples of phrases identified in academic discourse by EAP researchers, and compares these with examples of the phrases taught in EAP writing course books. The paper highlights the ways in which the forms and functions of phrases recommended for teaching by EAP researchers are different from the forms and functions recommended for teaching in EAP writing materials. This paper illustrates this apparent divergence between EAP research and practice through reviewing the concept of 'function' as applied to phrases in both academic discourse research and EAP pedagogy. It then compares the forms and functions categories for six sets of phrases: two produced by researchers and intended for pedagogy, and four found in teaching materials. The paper suggests that, in this area of EAP at least, Halliday's work is more directly influential on current research than on current practice, and that EAP phraseology is one area where the growing gap between EAP research and practice can in future be quantified.

1. Introduction

The past two decades have seen regular contributions to the debate over the relevance of research from areas of applied linguistics to language teaching practice (Borg, 2009; Crookes, 1998; Lightbown, 2000, 2002; Mackay, 2018; Maley, 2015; Medgyes, 2017; Nassaji, 2012; Paran, 2017; Pica, 1997; Sheen, 2002; Tomlinson, 2013). In recent years aspects of this debate have similarly been discussed in relation to EAP. It has been argued that a gap has opened whereby the current research discourses of EAP have diverged too far from EAP classroom practice. Some published applied linguistic research work, in consequence, seems to lack relevance to EAP teachers. Hamp-Lyons (2015, p. A3) pointed out that "many of the research articles that we publish have strong connections to teaching and learning in academic contexts; but some seem to grow out of theoretical interest in how academic language works with little attention to context." She finds that, in the first 13 years of the *Journal of English for Academic Purposes,* the most downloaded articles were on practical topics close to the classroom, such as the EAP learner and the teaching of various skills. By contrast, studies of language in academic discourse, particularly corpus-based studies, tended to be downloaded less, and cited more. It is reasonable to presume that these citations were by other EAP researchers, rather than by EAP teachers. Hamp-Lyons suggested that this difference reflected a relevance gap between research and practice, where on one side researchers were interested in revealing ever more specialised uses of academic English, while on the other practitioners preferred to read about practical concerns they felt were closer to their classroom experience. She wondered "how far these differences, captured as they may be in fine detail through corpus analyses, can be

(and are) applied in instructional materials used by academics in the discipline within their teaching materials for students" (Hamp-Lyons (2015), p. A2), and asked whether the EAP profession should "take a more critical view of the salience of various phenomena" found in such language-focused research.

This paper is a response to the questions raised by Hamp-Lyons; it argues that phraseology is an area of EAP which has salience for both corpus linguistic researchers, who search corpora for frequent phrases in academic discourse, and classroom-based EAP practitioners, who help their students acquire and use phrases for appropriate academic writing. Phraseology is therefore an ideal lens through which to view the nature of the gap between EAP research and practice, and this paper resonates with the theme of this special issue of JEAP by illustrating M. A. K. Halliday's varying influence on these two aspects of the field.

Over the last 50 years, the central role of phraseology in language in general, and in academic discourse in particular, has been revealed in increasing detail by the adoption of corpus techniques by language researchers. Well before the turn of the 21st century, EAP researchers collected and studied large corpora of academic writing across texts, genres, registers, and disciplines in order to identify and list frequently occurring phrases.[1] These researchers have then examined the way these phrases function in the discourse, and grouped them in various functional categories. These functional categories for phrases correspond to ideational, interpersonal, and textual metafunctions (Christie, 2017; Halliday & Hasan, 1989; Halliday & Matthiessen, 2014), reflecting a clear Hallidayan influence on corpus-linguistic EAP research in this area.

In EAP practice, lists of phrases can similarly be found in course books and websites for teaching academic writing. These are intended to provide students with language which they can adapt for use with their own ideas so that their written expression is phrased appropriately. Again, each of these phrases is also associated with a particular function for use at a specific point in a text.

Yet on closer inspection the functions specified for phrases in these materials for EAP writing classroom *practice* do not obviously fit into the Hallidayan categories for phrases universally adopted in EAP *research*. The functions assigned to phrases in teaching are often instead pragmatic phrases specific to a particular genre or text type or for a particular rhetorical purpose. It seems odd that while EAP researchers unanimously apply Halliday's categories to the phrases they recommend for EAP teaching, the writers of EAP materials, who are very often teachers themselves, view the function of their phrases so differently.

This paper therefore aims to illustrate and give examples, in an impressionistic way, of this apparent divergence between EAP research and practice. It first summarizes Halliday's early contributions to EAP and phraseology and connects these to work on phrases in academic discourse research and EAP pedagogy. It then reviews the concept of 'function' as applied to these phrases in academic discourse research and EAP pedagogy, orienting rhetorical functions in relation to what Tribble (1996, 2009, 2015) terms the "social/genre" academic writing tradition.

The paper then presents a discussion[2] of a sample of six sets of phrases proposed for the teaching of EAP writing — two from research and four from EAP teaching materials. It then reviews a sample of six sets of phrases proposed for the teaching of EAP writing: two produced by researchers and specifically intended for pedagogy (Liu, 2012; Simpson-Vlach & Ellis, 2010), and four from a variety of resources for teaching phrases in EAP writing (Barros, 2016; Godfrey, 2013; Morley, 2018; OPAL, 2019). The paper describes the forms of these phrases and the functions given to them and illustrates the breadth and limits of Halliday's influence in this area.

2. Halliday: phases and functions in EAP

This first section summarizes Halliday's early contributions to EAP and phraseology and connects these to the work on phrases in academic discourse research and EAP pedagogy illustrated in this paper. Halliday's work on phraseology and EAP in the 1960s, many years before his first full articulation of Systemic Functional Linguistics in *Introduction to Functional Grammar* in 1985, anticipates much subsequent computerised research on phrases in academic corpora and its application to EAP pedagogy. During the second half of the 20th century, when English became widespread as the global language of communication in the fields of science, technology, business and academia (Crystal, 2003; Dudley-Evans & St John, 1998; Pennycook, 1994; Phillipson, 1988) there was a corresponding increase in the number of learners of English who needed to use it to communicate in these settings, and a lack of suitable materials from which they could learn. Halliday led efforts to meet this growing demand from learners of academic English by calling for the principled study of academic discourse to inform pedagogy. In *The Linguistic Sciences and Language Teaching*, for example, Halliday, McIntosh, and Strevens (1964) argued that materials for these new learners should be based on sound linguistic descriptions of the language used in these situations: "detailed studies of restricted language and special registers [should be] carried out on the basis of large

---

[1] The term 'phrase' is used here as a superordinate label for any linguistic item identified in the EAP research and teaching literature as somehow formulaic, pre-fabricated, chunk-like or otherwise phraseological in nature. Specific terms used by researchers and materials writers are used in the discussion where relevant. See Liontas (2019), pp59-62) for a recent survey of different terms.

[2] An empirical study is currently being conducted (Oakey, Hughes, & Zhang, in preparation) which employs an "approximate string matching" technique using Jaccard Similarity (Jaccard, 1912), a statistic which measures similarities and differences between data sets and which can be applied to these lists of phrases in order to quantify the extent to which they overlap.

samples of the language used by the particular persons concerned" (1964, p. 190). In thus setting out the research agenda for English for Specific Purposes (ESP) Halliday also laid the foundations for corpus linguistic research in EAP.

Around the same time, Halliday was also working on collocation, the phraseological area of linguistic theory developed by his tutor J. R. Firth, which sought to account for the tendency for particular words to combine more often with some words rather than others. Halliday improved on Firth's (1957) somewhat unclear theoretical statements on the nature of collocation (Robins, 1961, p. 198) by proposing a definition which specified both linear and statistical elements and thus could be empirically tested:

> the syntagmatic association of lexical items, quantifiable, textually, as the probability that there will occur, at n removes (a distance of n lexical items) from an item x, the items a, b, c … . Any given item thus enters into a range of collocation, the items with which it is collocated being ranged from more or less probable.
>
> (Halliday, 1961, p. 276 cited in Oakey, 2009, p. 141)

An item, such as a word, is therefore very likely to occur with some words, and much less likely to co-occur with other words. Halliday pointed out that such empirical investigation of collocational relations between items "would require the study of very large samples of text" (Halliday, 1966, p. 159), a methodological approach which his student John Sinclair and others carried forward into the use of ever-larger corpora for lexicography and English Language Teaching (COBUILD, 1996; Daley, Jones, & Sinclair, 1971/2004; Sinclair, 1987, 1991) and which was then more widely adopted by EAP researchers to study and identify phrases for teaching academic writing.

2.1. SFL and metafunctions

Halliday, later joined by many other collaborators, eventually developed the SFL framework (Halliday, 1973, 1978, 1994; Halliday & Matthiessen, 2004, 2014) discussed in the papers in the current issue of this journal. SFL formulated the systemic relationship between language and Malinowski's notions of meaning as 'context of situation' and 'context of culture' (Malinowski, 1923) by introducing the notion of 'metafunctions' of language. In this model, language is not only a semiotic system itself, but is also the means by which meaning is expressed through two other semiotic systems, namely 'register' and 'genre'. The term 'register' in this model refers to the systematic relationship between language and Malinowski's context of situation, while 'genre' corresponds to Malinowski's broader notion of context of culture (Martin, 2010, p. 17). Genre acts on register by constraining the possible combinations of different categories of register, and register in turn constrains language by limiting the linguistic choices available to the participants.

The relationship between language and meaning as context of situation is the correspondence of these three categories of register to three aspects of meaning, termed 'metafunctions': the 'field' of discourse, i.e. the area of operation of language activity, corresponds to the 'ideational' metafunction of language, in which human experience is construed through identifying and categorising the world. The 'tenor' of discourse, the relationship between discourse participants, relates to the 'interpersonal' metafunction of language which takes account of participants and communicative circumstances, while the 'mode' of discourse, primarily spoken or written, relates to the 'textual' metafunction of language, which organises and creates cohesion in the discourse (Halliday, 1994, p. xiii; Halliday & Matthiessen, 2014, pp. 30−31).

Academic discourse is particularly amenable to analysis in terms of SFL metafunctions since it occurs within clearly identifiable and describable contexts of culture and situation. Writing takes place on recognised textual platforms, such as research articles and course books, which are subject to ideational constraints from the subject matter, and to interpersonal and textual constraints from the established conventions of the discourse community. Metafunctions thus relate linguistic form to what the writer is writing about, who the writer is writing for, and the textual platform on which the writing takes place. Systemic linguists "use grammatical evidence about the nature of field, mode and tenor at the same time as it gives them a way of explaining why language has the shape it does in terms of the way in which people use it" (Martin, 2010, p. 18).

It will be seen in this paper that these three Hallidayan metafunctional components have been most influential in EAP research on phrases in academic discourse: they were initially adopted by Biber et al., whose framework then influenced the functional categorisations of Simpson-Vlach and Ellis (2010) and Liu (2012). Approaches to teaching EAP writing, however, reflect a different understanding of 'function', and these differences will be looked at in the following section.

2.2. Functions in the social/genre approach to EAP writing pedagogy

Since the adoption of the communicative language teaching paradigm in the 1970s, many English language course books have contained phrases associated with a particular communicative function. Both Wilkins in *Notional Syllabuses* (1976) and Munby in *Communicative Syllabus Design* (Munby, 1978) stressed that a learner's communicative competence includes knowing how to say or write the phrases appropriate to the relevant communicative context (Widdowson, 1989). English language teaching course books such as *Functions of English* (Jones, 1981) were organised around the idea that students "must learn which structures are appropriate to the situation they are in" (Jones, 1981, p. 1), and should acquire a repertoire of phrases for functions such as "asking for information", "refusing to do something", and "giving an opinion" (Jones, 1981). Such a teaching approach was criticised for focusing too much on the interpersonal functional aspect of language, placing "an overemphasis on language as an instrument of social interaction and a neglect of its equally important conceptualising function" (Doff, Jones, & Mitchell, 1984, p. 12), i.e. its ideational aspect. This type of approach was also faulted for its simplistic

Jordan (1980, 1990, 1999)	Bunting et al (2012)
Unit 1 Structure and Cohesion Unit 2 Description: Process and Procedure Unit 3 Description: Physical Unit 4 Narrative Unit 5 Definitions Unit 6 Exemplification Unit 7 Classification Unit 8 Comparison and Contrast Unit 9 Cause and Effect	PART 1 Cause and Effect PART 2 Comparison and Contrast PART 3 Narrative PART 4 Classification and Definition PART 5 Problem - Solution PART 6 Summary - Response and Persuasion

Fig. 1. Comparison of functional headings in academic writing text books in the 'social/genre' tradition 1980−2013.

pedagogy: "the student is simply presented with pairings of form and function which he (sic) can only commit to memory" (Doff, Jones, & Mitchell, 1984).

Phrases associated with particular functions have continued nonetheless to figure prominently in EAP writing course books. The concept of 'function' in the teaching of EAP writing encompasses a broad range of meanings depending on the approach taken. In the approach to teaching writing common at north American universities, referred to by Tribble (1996, 2009, 2015) as the "intellectual/rhetorical" tradition, all students, whether L1 or L2 English speakers, take classes in what Tribble (2009, p. 411) terms "essayist literacy," i.e. personal, rather than evidence-based, writing. This sort of teaching is less genre- or register-based in the Hallidayan ESP tradition and more aimed at producing students who can argue a case clearly and persuasively (Tardy & Jwa, 2016). An influential phrase book in this intellectual/rhetorical tradition is *"They Say/I Say": The Moves That Matter in Academic Writing* (Birkenstein & Graff, 2018) which is discussed elsewhere in a wider empirical survey of phrases in academic writing (Oakey, Hughes & Zhang, in preparation). Being developed for the intellectual/rhetorical tradition, writing materials such as Birkenstein and Graff (2018), as Tribble (2009, p. 402) points out, are not always understood well outside north America.[3]

Rather than look at phrases as taught in the intellectual/rhetorical tradition, therefore, this paper instead focuses on phrases taught in what Tribble (2009) calls the 'social/genre' tradition. This tradition was informed, as has already been seen, by work arising from Halliday et al. (1964)'s call for descriptions of academic contexts of situation and culture, and the language shaped by them, to be "the starting points for any pedagogic solutions that are developed to meet learners' needs" (Tribble, 2009, p. 401). The 'social/genre' tradition prioritises written genres used in academic contexts; teaching materials focus on functions that realise different genres or particular stages of a genre, such as the influential "Create a Research Space" (CARS) model (Swales, 1990, 2004; Swales & Feak, 2012) for teaching rhetorical moves in Introduction-Methodology-Results-Discussion (IMRD) research article writing. A more general set of rhetorical functions is based on writing purpose, i.e. by describing types of processes, writing definitions, or describing cause-and-effect relationships. Students on their degree courses are unlikely to be required to write complete texts solely based on just one of these functions. Nonetheless, functions such as these have long offered a useful way of organising academic writing teaching textbooks, as can be seen in the similarities between the contents pages from Jordan (1980, 1990, 1999) and Bunting, Diniz and Reppen (2012) in Fig. 1.

It is clear from Fig. 1 that concepts such as 'definition', 'classification', 'comparison and contrast', and 'cause and effect' - appearing as they do as chapter headings in books written three decades apart - are long established functions in the 'social/ genre' tradition of teaching EAP writing, and, as will be seen below, this is one of the key continuing influences of Halliday et al. (1964) on current EAP teaching. Having briefly surveyed the notions of "function" in relation to the broad field of EAP writing research and pedagogy, we now look more closely at the forms and then the functions of the phrases investigated in this paper.

3. Forms of phrases in EAP research and teaching materials

Phrases of various kinds have long been seen as important in many areas of first and second language theory and practice: in first and second language acquisition, processing, and production (McCauley & Christiansen, 2014; Pawley & Syder, 1983; Peters, 1973; Theakston & Lieven, 2017; Wray, 2012); cognitive linguistics (Culicover, Jackendoff, & Audring, 2017; Ellis & Ferreira-Junior, 2009; Goldberg, 2006); empirical pragmatics (Bardovi-Harlig, 2012) lexicology and lexicography (Cowie, 1981, 1992, 1998; Cruse, 1986); literary stylistics (MacKenzie, 2000; MacKenzie & Kayman, 2016; Oakey & González-Díaz, 2019), and even self-improvement and etiquette guides (Kleiser, 1917; Sullivan, 2016; Werner, 1985). In English language teaching they have long been a part of both grammar and vocabulary teaching (Hornby, 1948; Mittins, 1950; Palmer, 1933) and in EAP since corpus research was applied to academic discourse following the suggestions by Halliday et al. (1964). This section accordingly reviews the various forms of phrases to be found in published research and EAP teaching materials.

---

[3] In my own experience, I have noticed that EAP writing teachers from the UK working in Turkey and China find that their North American colleagues have a very different understanding of what is meant by "academic writing".

Lexical Bundles (Biber, Conrad, & Cortes 2004)	Academic Formulas (Simpson-Vlach & Ellis 2010)	Multi-Word Constructions (Liu 2012)
the nature of the	in terms of	such as (det + N)
in the case of	at the same time	for example
in terms of the	from the point of view	as well as (det + N)
as a result of	in order to	NP suggest that
on the basis of	as well as	according to (det + N)
in the absence of	part of the	(be) based on (det + N)
the way in which	the fact that	there be det + N
the extent to which	in other words	there be no NP
in the presence of	the point of view of	a/the (large/small) number of
at the same time	there is a	out of (det + N)

Fig. 2. Representative examples of phrases from corpus-informed EAP research.

3.1. Forms of phrases in EAP research

As has already been mentioned, large-scale corpus work on phrases in academic discourse began in the 1990s and, over the next two decades, a number of research studies resulted in the publication of lists of phrases. These phrases took various forms, and the forms were given different labels. As well as 'lexical bundles' (Biber, 2006; Biber, Conrad, & Cortes, 2004; Biber, Johansson, Leech, Conrad, & Finegan, 1999; Hyland, 2008 and numerous subsequent papers by these and other researchers), lists were published of 'academic formulas' (Simpson-Vlach & Ellis, 2010), and 'multi-word constructions' (Liu, 2012).[4] Fig. 2 shows representative examples of the forms of each of these phrases.

Lexical bundles (Biber et al., 1999) were originally described exclusively in terms of their structure, rather than being associated with functions, and were used as a way of illustrating differences between 'registers' (in Biber et al.'s non-SFL meaning of the term). Lexical bundles take the form of fixed strings of words which show "a statistical tendency to co-occur" (Biber et al., 1999, p. 989) and therefore "commonly go together in natural discourse" (Biber et al., 1999, p. 990). The lexical bundle is a 'n-gram', i.e. a contiguous string of words *n* words long, without gaps. To be classed as a lexical bundle, an n-gram needs to occur in at least 10 different texts in a corpus a minimum number of times. This minimum frequency whereby an n-gram can be classed as a lexical bundle, originally 40 occurrences per million words, is set by different researchers at a "somewhat arbitrary" level (Biber, 2006, p. 134; Biber & Barbieri, 2007, p. 267; Biber et al., 2004, p. 376; Hyland, 2008, p. 8). The choice of the minimum frequency threshold setting appears to be driven by the need to keep the number of items in the analysis down to a manageable level, although it has recently been pointed out that even small alterations in the minimum frequency threshold have the potential to significantly affect which bundles are identified and which functions they perform (Samraj, 2018).

A notable initial finding by Biber et al., 1999 was that lexical bundles in the register of academic prose were largely made up of noun phrases, or fragments thereof, and those in the register of conversation usually contained verb phrases or were parts of clauses (Biber et al., 1999, p. 992). It can be seen from Fig. 2 that lexical bundles are often fragments of more traditional linguistic structures, such as a noun phrase followed by part of a post-modifying prepositional phrase as in *the nature of the*. Whether this fragmented nature of phrases identified through corpus analysis affects their salience for EAP teachers and learners is food for thought. As will be seen below, lexical bundles were later assigned various discourse functions to become form-function units, together with recommendations by researchers that these phrases were suitable for use in teaching EAP writing.

The academic formulas list (Simpson-Vlach & Ellis, 2010) was an attempt to improve on phrases such as lexical bundles which had been identified purely by frequency and which seemed "neither terribly functional nor pedagogically compelling" (Simpson-Vlach & Ellis, 2010, p. 493). Academic formulas were instead intended to be more pedagogically useful for academic writing while still being empirically identified. These phrases were identified by triangulating between corpus frequency and judgements, by practising EAP teachers, of their meaning and teachability. Mutual Information (MI) scores were preferred to raw frequency for identifying these formulas, since a higher MI score suggests a stronger association between the words in the phrase and that their co-occurrence is less likely to be due to chance. Corpus data was taken from Hyland's (2008) research article corpus and the academic portion of the online British National Corpus (BNC) (Davies, 2004-). As can be seen from Fig. 2, academic formulas such as *there is a* and *part of the* still appear fragmented compared to traditional linguistic structures, but the key difference from lexical bundles is their validation as meaningful and teachable (and by implication more salient) by practising EAP teachers.

---

[4] Other lists of phrases published in the last decade include 'grammatical collocations' (Durrant, 2009), 'phrasal expressions' (Martinez & Schmitt, 2012), 'academic collocations' (Ackerman & Chen, 2013) and 'idioms' (Miller, 2020), but since the phrases in these lists were not assigned functions in academic discourse they are not considered in this paper. The recent 'academic English collocation list' (Lei & Liu, 2018) mentions only two discourse functions specifically and so was also not included.

Finally, the list of multi-word constructions (Liu, 2012) was an attempt to present phrases for teaching that were more complete in form than lexical bundles or academic formulas. Liu was dissatisfied with the fragmented nature of these previous phrases, and instead adapted an item from cognitive linguistics, the 'construction' (Culicover et al., 2017; Ellis & Ferreira-Junior, 2009; Goldberg, 2006). Constructions combine cognitive and social perspectives on language, and are seen as "basic units of the linguistic system, accepted as convention in the speech community and entrenched as grammatical knowledge in the speaker's mind" (Ellis & Ferreira-Junior, 2009, p. 370). Constructions are in opposition to the traditional binary model according to which "each person's brain contains a lexicon of words and the concepts they stand for (a mental dictionary) and a set of rules that combine the words to convey relationships among concepts (a mental grammar)" (Pinker, 1994, p. 85). Constructions underlie both the form and function of language: they "specify the morphological, syntactic, and lexical form of language and the associated semantic, pragmatic, and discourse functions … and it is their communicative functions that motivate their learning" (Ellis & Ferreira-Junior, 2009, p. 370) and so for Liu the construction from cognitive linguistics was a useful item on which to base the multi-word constructions for EAP pedagogy. Unlike other types of phrases, constructions are not necessarily specified as strings of words; while they can be 'filled' and appear like fully-formed phrases, such as *for example* in Fig. 2, they also can be partially filled with a word class, as in *such as (det + N)*. The 'partially-filled' element means each item is a 'complete' construction rather than a fragment and thus more learnable: "presenting structurally-incomplete lexical bundles as partially-filled constructions enables us to represent them as complete constructions" (Liu, 2012, p. 28) and thus by implication make these phrases more salient for teachers and learners. Corpus data for Liu's study was again taken from the academic portion of the BNC and also the much larger academic subsection of the Corpus of Contemporary American English (COCA) (Davies, 2008-). Unlike lexical bundles and academic formulas, however, multi-word constructions were not automatically extracted items: a list of 559 items was drawn up from a variety of previous studies and learner dictionaries and then searched for manually in the corpora (Liu, 2012, p. 29).

As we have seen in this section, corpus linguistic research on phrases produces novel language items which could not easily have been identified without the use of software. Their fragmentary, incomplete or partially-filled nature makes them different in form from traditionally taught language items. In spite of this, the researchers who compiled these lists of phrases believe them to be useful for pedagogy in a similar way to Lewis (2000), who maintains that incomplete phrases like *as a function of,* and *in the case of* and so on are "precisely the kind of language … which is likely to be invisible to learners," (Lewis, 2000, p. 147) who are much more likely to focus on familiar, 'salient' language items like lexical words. Having described the forms of the phrases in EAP research, therefore, we now move on to surveying the phrases which *are* visibly presented to learners directly in EAP teaching materials.

3.2. Forms of phrases in EAP pedagogy

Phrases were a feature of EAP writing teaching course books long before the adoption of corpus linguistic research methods. An example of the types of phrases in academic writing textbooks can be found in Jordan (1980) who presents "impersonal verb phrases often associated with conclusions" as in Fig. 3:

It can be seen that these phrases pair form with function. In terms of their form and structure they are 'extraposed *that* clauses' or 'introductory *it* patterns' (Oakey, 2002; Hewings & Hewings, 2002; Larsson, 2017; Peacock, 2007), while in terms of their function, they have a discourse organising or textual function, i.e. signalling the discourse structure to the reader, in this case the conclusion to a text or part of a text. Some also appear to have an interpersonal function: the choice of verb indicates the degree of commitment to the proposition made by the writer, i.e. his or her 'stance', and that this item, rather than an alternative formulation such as *many people think that*, is a more appropriate academic writing style.

A decade after Jordan's phrases, Nattinger and DeCarrico (1992) provided lists of phrases they called 'lexical phrases,' which they defined as collocations which have a pragmatic function in discourse. Lexical phrases can be discontinuous strings of words with slots or letters which can be replaced by noun phrases, such as *according to _____, X is Y* (Nattinger & DeCarrico, 1992, p. 165) to which Nattinger and DeCarrico assign a "topic priming" function. They can also be continuous strings of words containing upper-case letters that can be replaced with time adverbials or clauses, such as *For a long time X, it has been the case that X*, to which they also assign a "topic priming" function. They argued that a combination or "mosaic" (Nattinger & DeCarrico, 1992, p. 166) of appropriate lexical phrases constitutes the framework for the content of a text which adheres to the constraints imposed by the genre. Nattinger and DeCarrico accordingly specified the framework for a first year composition essay written in the 'intellectual/rhetorical' tradition (Tribble, 2009, p. 402). Nattinger and DeCarrico were an influence on the Lexical Approach to English language teaching (Lewis, 1993, 1997, 2000) and this approach in turn influenced approaches to using phrases in the teaching of academic writing. Lewis recommended that students writing in "a particular

It has been suggested that

It is generally agreed that

It is widely accepted that

It is now generally recognised that   (these indicate stronger evidence)

Fig. 3. "Impersonal verb phrases" in *Academic Writing Course* (Jordan, 1980).

genre such as academic English [develop] an awareness of the sentence heads and frames typical of the genre." (Lewis, 1996, p. 15).

In the current paper, two lists of phrases are examined which come from academic writing phrase books published independently by practising EAP teachers: *The Only Academic Phrasebook You'll Ever Need: 600 Examples of Academic Language* (Barros, 2016) and *The Academic Phrasebank* (Morley, 2018).

The inclusion of these two books in this study reflects Tribble's observation that teachers have

> a growing independence from published course materials. As teachers come to a fuller understanding of the specificity of the needs of their students, they become less happy with published coursebooks and more confident in their capacity to develop materials which will be specifically relevant to the needs of their learners
>
> (Tribble, 2009, pp. 401–402)

Tribble's observation, in tandem with that of Hamp-Lyons referred to in the Introduction above, feeds into the debate regarding the usefulness of EAP research for language teaching practice: if EAP researchers continue to identify and recommend language features which teachers and learners do not see as salient, then teachers will produce other materials themselves. These two books were selected by the "Customers who bought this item also bought …" algorithm on Amazon.co.uk, which suggests they are highly relevant in terms of content and consumer demand to other mainstream phrase books. The two other lists of phrases come from materials produced by 'mainstream' publishers: the first is the highest-selling academic writing phrase book in the *Student Life* category on Amazon.co.uk: *The Student Phrase Book: Vocabulary for Writing at University* (Godfrey, 2013) while the *Oxford Phrasal Academic Lexicon* (OPAL) is part of the widely used *Oxford Advanced Learner's Dictionary* website (OPAL, 2019).

The lists of phrases in these different sets of materials were prepared using a variety of methods and data sets. The OPAL phrases, according to the website, were taken using data from the Oxford Corpus of Academic English (OCAE), "using a scientific method called 'keyword analysis' to identify the words and phrases that are the most important in an academic setting" (OPAL, 2019). Morley's phrases originally came from 100 postgraduate dissertations completed at the University of Manchester (Morley, 2004) and have since been expanded from academic articles drawn from a wide range of academic disciplines (Morley, 2018, p. 4). Godfrey's phrases were "collected over 24 years of teaching writing to university students" (Godfrey, 2013, p. viii). Barros's phrase list originated in the phrases he noticed in his own MA reading and which he collected to help him in his own academic writing and which subsequently became a popular EAP blog post before publication as a book (Barros, 2016, pp. 69–70). These phrases have a particular face validity for teaching EAP writing since they were evidently salient features for a non-native speaker learning to write academic English. Representative examples of these phrases are shown in Fig. 4 and Fig. 5:

In terms of structure, the forms of the phrases in OPAL stand out as being the most similar to the lexical bundles, academic formulas, and multi-word constructions described in the previous section. This stands to reason, since these phrases were similarly identified by researchers from a corpus, even if the identification criteria are less fully described. The most notable difference between the OPAL phrases and the phrases in the other three sets of materials is the prevalence of clausal elements in the latter. We have already seen that lexical bundles in academic prose tend to consist more of nominal fragments than parts of clauses. The phrases in the published EAP materials, by contrast, are made up much largely of clauses. This can either

OPAL (Oxford 2019)	Phrases (Godfrey, 2013) (emphasis in the original)
in terms of the idea of it is important to note that there is no doubt that it has been argued that result(s) in depend(s) on the manner in which the influence of have a significant impact on on the other hand there are a number of reduction in the number of in this study in recent years the emergence of	The question of whether to legalise all drugs is regularly debated in the media. The (principal) purpose / objective / aim / goal of this report is to identify solutions to the company's problems. The participants were allocated / assigned randomly to one of five groups. A large number of / A small number of students returned the questionnaire. Our results suggest that student numbers will level off / level out / reach a plateau over the next few years. The chemical inhibits / restricts / curtails / limits plant growth. Many people think / hold the opinion / are of the opinion that using animals for experimentation is inhumane. A carving depicting / showing / representing a mammoth has recently been found on a bone fragment in Florida. This survey provides fresh insights into / gives us a better understanding of why people smoke. Picasso is widely regarded as / viewed as / thought of as / considered to be a major influence in twentieth-century art.

Fig. 4. Representative examples of phrases in published EAP course materials (use of formatting and ellipsis is retained from the original lists).

Sentence templates (Barros, 2016)	Phrases (Morley, 2018)
A number of recent studies have looked into the effects of _____ on _____ . Few studies have investigated the impact of _____. This study critically reviews _____. The aim of this study is/was to discuss the extent to which _____. This paper provides an overview of recent developments in _____. Much previous work on _____ has focused on _____. While Smith's findings indicate that _____, it has not yet been demonstrated that _____. This study used a combination of qualitative and quantitative analysis tools. Means and standard deviations were determined through _____. Future studies will have to investigate the role of _____.	X has received considerable scholarly attention in recent years … The issue of X has been a controversial and much disputed subject within the field of ... Despite the importance of X, there remains a paucity of evidence on … Previous studies have failed to consider the important role played by … The approach to empirical research adopted for this study was one of ... The remaining part of the paper proceeds as follows: ... The most striking result to emerge from the data is that ... A note of caution is due here since … The findings of this research provide insights for … The present study was subject to a number of potential methodological weaknesses.

Fig. 5. Representative examples of phrases in self-published EAP course materials (use of formatting and ellipsis is retained from the original lists).

be a full clause with an incomplete adverbial phrase, as in "Means and standard deviations were determined through _____" (Barros, 2016) or complete sentences, as in "The present study was subject to a number of potential methodological weaknesses." (Morley, 2018). This difference is so marked that it again brings to mind Hamp-Lyons's comment regarding salience referred to in the Introduction above. These more complete structures such as clauses and sentences appear to be more salient to EAP materials producers in identifying phrases as important for EAP learners, and less so to producers of corpus linguistic phrase lists who as we have seen use frequency and other statistical measures to identify their phrases.

That said, some might argue that it is debateable whether complete sentences such as those in Morley's and Barros's lists can be considered phraseological at all, unless they are seen as "fully-filled" rather than "partially-filled" constructions. Davis and Morley (2015) asked academics in different disciplines what was considered acceptable re-use of academic phrases. Long, complete sentences containing general lexical items, such as "These results therefore need to be interpreted with caution" and "Much of the research up to now has been descriptive in nature" were regarded as acceptable by a high proportion of respondents, while short sentences containing content-specific lexical items, such as "Dawkins is deaf to theology" were regarded as unacceptable by respondents.

After this survey of the forms of phrases in both research and EAP pedagogy, and the discussion of the differences between them, the paper now focuses on the functions these phrases have been assigned.

4. Functions assigned to phrases in EAP research and teaching materials

As discussed in section 2.2 above, communicative approaches to language teaching have long associated language forms with functions. The functions of the phrases reviewed in this paper also show a marked difference between those assigned by researchers and those assigned by materials writers. The aim of this section is not to propose new functional categories for phrases, but to bring together existing categories proposed by researchers and practitioners and reveal similarities and differences as a way of illustrating the gap between EAP research and practice in this area.

4.1. Functions assigned to phrases in EAP writing research

The functions assigned to the two sets of phrases identified by researchers, academic formulas (Simpson-Vlach & Ellis, 2010, p. 503) and multi-word constructions (Liu, 2012, p. 30) are based on those of Biber et al. (2004). Biber et al.'s framework is in turn based on Halliday's meta-functions reviewed in section 2.1, and so it is relevant to discuss it first. It will be seen that the Hallidayan influence extends to all the phrases identified by corpus linguistic research which have been assigned functions: lexical bundles, academic formulas, and multi-word constructions, despite their substantial differences in form, are all ascribed similar referential (ideational), stance (interpersonal), and discourse organising (textual) functions in academic writing.

The functional framework for lexical bundles in Biber et al. (2004) was developed from an initial taxonomy, greatly influenced by SFL, proposed by Cortes (2001). The definition of 'function' refers to "the meanings and purposes of the

Referential bundles perform an ideational function; they help writers structure their experience and determine their way of looking at things. Expressions in this category are time, place, or text markers, such as *at the beginning of, the end of the, or at the same time.*

Text organizers are word combinations used to express textual functions which are concerned with the meaning of the sentence as a message in relation to the surrounding discourse. Some of the functions performed by these expressions are contrast (e.g. *on the other hand*), inference (e.g. *as a result of*), or focus (e.g. *it is important to*)

Stance bundles and interactional bundles perform interpersonal functions. Stance bundles express attitudes that frame some other proposition; expressions such as (e.g., *I don't know why, are more likely to*). Interactional bundles are conversational word combinations used to express politeness or to report, as in *thank you very much, and I said to him*.

Fig. 6. Functional categories of lexical bundles (Cortes, 2004, p. 41).

language, functions that try to provide texture or organise the discourse according to situations or contexts" (Cortes, 2004, p. 400). This definition explicitly connects the pragmatic meaning of lexical bundles, i.e. their discourse functions, to the context of situation. The original framework proposed by Cortes has three core categories of function: 'referential', 'text', and 'stance', shown in Fig. 6, which correspond to Halliday's metafunctional categories of register discussed in section 2.1, i.e. the ideational, textual, and interpersonal:

Cortes' framework has subsequently evolved according to the data in which the bundles have been found, for example in the more detailed study of academic register by Biber et al. (2004), which in turn was further revised by Biber (2006). Since this last study focused solely on academic discourse and did not make any comparisons with other registers, some bundles occurred more frequently or less frequently than in previous studies, and so the categories into which bundles were grouped either became more detailed and contained further subcategories, or were subsumed into superordinate categories. For example, some of the "interactional bundles" mentioned above by Cortes, such as *thank you very much* and *I said to him*, occur often in the register of conversation, but seldom occur in the register of academic discourse.

Another functional framework proposed by a corpus linguistic researcher for phrases in academic discourse was for lexical bundles by Hyland (2008), in which he suggests an alternative functional framework for lexical bundles based on his register corpus of academic prose (this is a different corpus from Hyland's previous journal article corpus (Hyland, 1998)). Although his lexical bundle framework was again developed from that of Cortes (2004) and Biber et al. (2004), Hyland also found that it was un-necessary to use the interactional categories arising from conversation data when classifying lexical bundles in academic discourse. Biber's data from "service encounters, institutional texts, and so on … seems to have yielded far more personal, referential, and directive bundles than my more research-focused genres" (Hyland, 2008, p. 13) and so the titles of his functional categories, 'research', 'text' and 'participant-oriented' "specifically reflect the concerns of research writing" (Hyland, 2008). Fig. 7 summarizes and presents Biber's and Hyland's functional frameworks for lexical bundles side by side with those of Ellis and Simpson-Vlach's academic formulas and Liu's multi-word constructions.[5]

Fig. 7 clearly shows the overlap between their superordinate categories and their alignment with the SFL metafunctions. All frameworks have a roughly "interpersonal" category which is used to show a writer's attitude towards and stance on what is being written about, evaluating the propositions made, and including the audience. While "stance" is superordinate in the frameworks of Biber, Simpson-Vlach and Ellis, and Liu, it is a subcategory in Hyland's 'participant-oriented' category.

All frameworks have a superordinate "textual" category, in which bundles are used to organise the discourse, although Hyland's equivalent 'text-oriented' category includes some functions, particularly those relating to 'framing', which Biber, Simpson-Vlach and Ellis, and Liu all place in the ideational category. Biber's 'discourse organisers' is a single category, while Hyland's text-oriented category is sub-divided into transitional, resultative, structuring, and framing signals. Simpson-Vlach and Ellis and Liu include 'topic introduction' under this grouping. Lastly all four studies have an ideational or referential category involving phrases that structure writers' experience of the real world; this includes functions like quantification, time and location.

The functions assigned to these different frameworks match to a remarkable degree, and illustrate how influential Halliday's ideas have been on EAP research, yet when the functions assigned to phrases in EAP writing course books are examined, these seem to have much less in common. This will be discussed in the next section.

---

[5] A fuller version is contained in Appendix A.

Functions of Common Lexical Bundles in Academic Textbooks (Biber, 2006, pp. 166-168)	Functions of Lexical Bundles in Research Articles, Dissertations and MA Theses (Hyland, 2008, pp. 13-14)	Functions of Academic Formulas (Simpson-Vlach & Ellis 2010)	Functions of Multi-word Constructions (Liu 2012)
Stance bundles	Participant-oriented	Stance Expressions	Stance / Interpersonal / Impersonal Expressions
Discourse Organizers	Text-oriented	Discourse Organizing Functions	Discourse/Textual Organizers
Referential Bundles	Research-oriented –	Referential Expressions	Referential/Ideational

Fig. 7. Comparison of functional categories for phrases found in corpus-informed EAP research.

4.2. Functions assigned to phrases in EAP teaching materials

The functional categories for phrases in EAP teaching materials show little direct evidence of the influence of Halliday's three metafunctional categories as a basic organising principle, although there is an underlying suggestion of implicit influence. The functional categories for the four sets of phrases can be seen in Appendices B (Barros, 2016), C (OPAL, 2019), D (Morley, 2018), and E (Godfrey, 2013). Barros's categories have the least in common with the SFL categories: they are based more on Swales's genre-specific CARS model mentioned above and are organised around the structure of the IMRD paper.

At times the other three sets of phrase functions resemble the contents pages of traditional academic writing course books in the social/genre tradition described in Fig. 1, featuring purpose-led functions like 'making contrasts' and 'comparing' (OPAL), 'comparing and contrasting' (Morley), and 'similarity and difference' (Godfrey), as well as 'causes and effects' (OPAL), 'explaining causality' (Morley) and 'cause and effect' (Godfrey). Other functional categories reflect sets of genre-based text-specific functions similar to those of Barros: 'expressing aims' (OPAL), 'introducing work' (Morley) and 'talking about aim and proposition '(Godfrey). Some functional categories actually seem to be less like functions than formal categories, as in 'Phrases with *(the) (noun) of*' such as *the quality of* and *the study of* (OPAL).

There is, however, some overlap with the SFL-derived functional categories for phrases proposed by EAP researchers described in section 4.1. Ideational headings can be seen such as 'referring to time' (OPAL), 'writing about the past' (Morley), and 'structure, time, sequence and frequency' (Godfrey) as well as 'expressing the existence/non-existence of something', and 'describing quantities: presence or absence' (Godfrey). Similarly, some sets of phrases are grouped according to textual functions such as 'referring to the text' (OPAL), and 'signalling transition' (Morley), and all sets contain phrases with interpersonal functions such as 'hedging and expressing degrees of certainty' (OPAL), 'indicating shared knowledge or understanding' (Morley), and 'importance' (Godfrey). Godfrey's categories are notable in combining genre-based categories like 'methodology and method, findings' with associated ideational functions like 'size, amount, level and proportion'. All these reveal some underlying Hallidayan influence on these functional categories, even though they are not explicitly organised along SFL lines.

5. Discussion and conclusion

This paper has surveyed the forms and functions of phrases in EAP research and practice and illustrated how they both reflect Halliday's influence to differing degrees. It will now explore the reasons why the considerable amount of work by corpus linguistic researchers on phrases in academic discourse does not seem to have percolated far into the EAP teaching materials surveyed here. The systematic, comprehensive lists of phrases produced by corpus linguistic researchers certainly indicate that we have moved on from the type of EAP corpus studies that Swales (2002, p. 152) tactfully suggested were "enthusiastic yet fragmented efforts" constituting "banked intellectual resources whose pedagogical time has yet to come" (Swales, 2002, p. 159). Yet, as this paper has hopefully illustrated, there remains a seeming reluctance on the part of course book writers and teachers to incorporate these phrases in their materials, and a corresponding enthusiasm for producing their own lists.

The sheer amount of detail yielded by corpus results does not seem in itself to be the cause of the problem. The lists of phrases from both research and teaching materials comprise several thousand items, and so incorporating so many phrases in teaching practice is a formidable task regardless of where they were identified.

As suggested at the start of this paper, the gap between EAP research and practice may arise from a problematic relationship between corpus research on academic discourse and its practical application to EAP pedagogy. As Hamp-Lyons (2015) pointed out, it may be that phrases found through corpus linguistic research, although attested examples of language use in authentic academic contexts, somehow lack salience to teachers and learners, leading to the situation whereby

writers and teachers are filling the gap with materials based on their own experience, and/or pedagogical corpora they have collected themselves.

A possible reason for the lack of uptake of research phrases may be the dearth of suggestions by researchers how teachers can use the phrases identified in their practice. While word lists (Coxhead, 2000; Thorndike & Lorge, 1938; West, 1953; Xue & Nation, 1984), have a long history in English language teaching and are widely incorporated into teaching materials, researchers seem to intend their lists to inform pedagogy in a similar way, regardless of the fact these are list of *phrases* rather than single words. The "implications for teaching" sections of the articles introducing these lists urge syllabus and materials designers to incorporate the phrases into teaching materials, but with little space to give much more than a general indication of how this might be done. Simpson-Vlach & Ellis are clear about the importance of academic formulas in academic discourse for learners, for example, but, despite their inclusion of practising EAP teachers in the phrase identification process, their paper ultimately has little space for practical advice on incorporating these phrases into teaching, other than the observation that "functional linguistic classification and the organization of constructions according to academic needs and purposes is essential in turning a list into something that might usefully inform curriculum or language testing materials" (Simpson-Vlach & Ellis, 2010, p. 510). Liu is similarly brief on this point but is more practical, suggesting a learner-centred data-driven approach (Johns, 1991a; 1991b) by which learners bypass teachers and materials designers altogether: "students may search and examine the examples of the constructions in a concordancing format and in the process they can often identify their usage patterns" (Liu, 2012, p. 33). There still appears to be a need for practising teachers with current experience in the classroom to become more involved in conducting this kind of corpus linguistic research to better ensure that phrases are more classroom-friendly.

Another reason for the gap between research and practice may be that the decontextualised presentation of these phrases makes them unappealing to teachers (Flowerdew, 2015, p. 106). Functions are often assigned to phrases by researchers through the inspection of concordance lines, which allow little context to be visible for each example of the use of a phrase. Some lexical bundle function categories, for example, seem to equate the function of a lexical bundle with its lexical content: the lexical bundle *the number of*, for example, is termed a referential "quantifier" (Biber's referential category 2a in Appendix A below) because it can be seen from the concordance line that it contains the word 'number'. The category of "stance bundles of importance" (Biber's stance category 2b in Appendix A below) similarly includes bundles such as *of the most important*. It would be difficult, however, using concordance lines alone, to assign a function to a phrase which contains no obviously lexical words, such as *is that it is.* This is one of Simpson-Vlach & Ellis's formulaic sequences which, as part of their identification methodology, was judged to be salient and teachable by their EAP informants, and yet does not appear to carry much functional meaning when decontextualised and presented in a list. When researchers extract phrases from their context and present them in a list, it may make these corpus linguistic phrases opaque, and thus less salient, to learners and teachers. It was clear from section 3.2 that the phrases in EAP teaching materials contain more clausal features, giving the impression that practitioners preferred to highlight phrases where a discourse function could clearly be identified from the context, rather than fragmented snippets that they felt were unmeaningful.

A related salience issue may be that the arcane nature of some of the functions suggested by researchers for their phrases, while a necessary part of a comprehensive theoretical framework, makes them unsuitable for incorporation into teaching materials because they are conceptually too challenging for teachers and learners. Functional classifications for phrases such as "epistemic stance" and "intangible framing attributes," for example, might be cognitively challenging for some EAP teachers to teach, and for their students to learn about. It is not hard to imagine that the idea of functions in general - as a "meta" concept beyond the surface of texts - is difficult for teachers and learners to grasp alongside the myriad other features they have to contend with.

To sum up, based on the above comparison of the forms and functions of phrases for teaching EAP writing, it seems safe to say that Halliday has had more direct influence on EAP *research* in this area than on EAP *practice*. Practitioner producers of phrase lists for EAP writing pedagogy, by and large, seem to be less concerned with adopting the work of corpus linguistic researchers, and have focused on more pragmatic, text-focused phrases they feel are more salient to their learners.

Author statement

David Oakey: Conceptualization; Methodology; Investigation; Resources; Writing - Original Draft Preparation; Writing - Review & Editing; Visualization; Supervision; Project administration. Figures for "Phrases in EAP academic writing pedagogy: illuminating Halliday's influence on research and practice".

Appendix A. Comparison of the functional frameworks of Biber (2006), Hyland (2008), Simpson-Vlach and Ellis (2010), and Liu (2012)

Functions of Common Lexical Bundles in Academic Textbooks (Biber, 2006, pp. 166–168)	Functions of the Top 50 Lexical Bundles in Research Articles,	Functions of Academic Formulas (Simpson-Vlach & Ellis, 2010)	Functions of Multi-word Constructions (Liu, 2012)

(*continued*)

Functions of Common Lexical Bundles in Academic Textbooks (Biber, 2006, pp. 166–168)	Functions of the Top 50 Lexical Bundles in Research Articles, Dissertations and MA Theses (Hyland, 2008, pp. 13–14)	Functions of Academic Formulas (Simpson-Vlach & Ellis, 2010)	Functions of Multi-word Constructions (Liu, 2012)
	Dissertations and MA Theses (Hyland, 2008, pp. 13–14)		
Stance bundles express attitudes that frame some other proposition 1 Epistemic stance – impersonal *are more likely to, by the fact that* 2 Attitudinal/Modality stance a) Ability/effort: Impersonal *it is difficult to, it is possible to* b) Importance: Impersonal *it is important to, of the most important*	**Participant-oriented** 1 Stance features – convey the writer's attitudes and evaluations *are likely to be, it is possible that* 2 Engagement features – address readers directly *it should be noted that, as can be seen*	**Stance Expressions** (1) Hedges: *(more) likely to (be), [it/there] may be, to some extent* (2) Epistemic stance: *assumed to be, be seen as, be considered as* (3) Obligation and directive *(it should) be noted, need not be* (4) Expressions of ability and possibility: *can be used (to)* (5) Evaluation: *the importance of, important role in* (6) Intention/volition, prediction: *to do so, we do not*	**Stance/Interpersonal/Impersonal Expressions** 3.1. Epistemic stance: *NP(I/We) argue that, NP (I/We) believe that* 3.2. Attitudinal/modality stance: *be able to VP, tend to VP, be (more) likely to VP*
Discourse Organizers express textual functions which are concerned with the meaning of the sentence as a message in relation to the surrounding discourse *on the other hand, at the same time*	**Text-oriented** – concerned with the organisation of the text and its meaning as a message or argument: 1 Transition signals – establishing additive or contrastive links between elements: *on the other hand, in contrast to the* 2 Resultative signals – mark inferential or causative relations between elements *as a result of, it was found that* 3 Structuring signals – text-reflexive markers which organise stretches of discourse or direct reader elsewhere in text *in the present study, as shown in figure* 4 Framing signals – situate arguments by specifying limiting conditions *in the case of, with respect to the*	**Discourse Organising Functions** (1) Metadiscourse and textual reference: *as shown in, in the next section, (in) this paper (we)* (2) Topic introduction and focus: *For example [if/in/the], what are the* (3) Topic elaboration (a) non-causal *are as follows, in more detail, see for example* (b) cause and effect *[a/the] result of, due to the, so that the* (4) Discourse markers: *as well as, at the same (time)*	**Discourse/Textual Organizers** 2.1. Linking: *such as (det + N), for example, as well as, in addition (to)* 2.2. Topic introduction: *as to wh-clause/NP, let us/me + infinitive*
Functions of common lexical bundles in academic textbooks (Biber, 2006, pp. 166–168)	Functions of the top 50 lexical bundles in RAs, dissertations and MA theses (Hyland, 2008, pp. 13–14)	Functions of Academic Formulas (Simpson-Vlach & Ellis, 2010)	Functions of Multi-word Constructions (Liu, 2012)
Referential Bundles help writers structure their experience 1. Identification *is one of the, is known as the* 2. Specification of attributes a) Quantity/mathematical expression *of the number of, the magnitude of the* b) Predicative *is equal to the, is given by the* c) Tangible framing attributes *the size of the, in place of the* d) Intangible framing attributes *in the case of, as a result of* 3. Time/place/text/other reference a) Place or institution reference *the united states and, in the united states* b) General location reference or framing *in the same direction, the surface of the* c) Text deixis *as shown in figure, in this chapter we*	**Research-oriented** – help writers to structure their activities and experiences of the real world: 1. Location – indicating time/place: *at the beginning of, at the same time* 2. Procedure: *the use of the, the role of the* 3. Quantification: *the magnitude of the, a wide range of* 4. Description: *the structure of the, the size of the* 5. Topic – related to the field of research: *in the Hong Kong, the currency board system*	**Referential Expressions** (1) Specification of attributes (a) Intangible framing attributes *[a/the] form of, (in) such a (way)* (b) Tangible framing attributes *(as) part of, [a/the] the change in* (c) Quantity specification: *a list of, [a/large/the] number of* (2) Identification and focus: *a variety of, is for the, it is not* (3) Contrast and comparison: *and the same, different from the, is much more, (the) difference between (the)* (4) Deictics and locatives *a and b, the real world, of the system* (5) Vagueness markers *and so on*	**Referential/Ideational** 1.1. Framing/intangible framing attributes: *according to (det + N)* 1.2. Tangible framing attributes: *(in) the form of (det/N), (as/in) the size of (det + N)* 1.3 Identification/identity specification: *there be no NP, there be det + N* 1.4. Qualifying: *at all, as a whole* 1.5. Quantity specification: *a/the (large/small) number of, a lot of* 1.6. Referential place/text/direction: *in this article/chapter, in public, see table X* 1.7 Referential time/sequence: *(at/by) the end of (det + N), (at) the time of (det + N)* 1.8. Reporting/description/interpretation: *NP suggest that, NP show that* 1.9. Multifunctional: *the*

Functions of Common Lexical Bundles in Academic Textbooks (Biber, 2006, pp. 166–168)	Functions of the Top 50 Lexical Bundles in Research Articles, Dissertations and MA Theses (Hyland, 2008, pp. 13–14)	Functions of Academic Formulas (Simpson-Vlach & Ellis, 2010)	Functions of Multi-word Constructions (Liu, 2012)
			use of (det + N), (be) associated with (det + N)

Appendix B. Functional Framework for *The Only Academic Phrasebook You'll Ever Need: 600 Examples of Academic Language* (Barros, 2016)

ESTABLISHING A RESEARCH TERRITORY
DESCRIBING RESEARCH GAPS
STATING YOUR AIMS
DESCRIBING THE SCOPE AND ORGANIZATION OF YOUR PAPER
GENERAL LITERATURE REVIEW
REFERENCING
SAMPLING AND DATA COLLECTION
DATA ANALYSIS AND DISCUSSION

Appendix C. Functional Framework for the *Oxford Phrasal Academic Lexicon* (Oxford 2019)

Specifying topics and relations between ideas
Drawing attention to something or focusing on it
Hedging and expressing degrees of certainty
Explaining and defining
Giving examples or presenting evidence
Expressing aims, causes and effects
Making contrasts
Comparing
Adding
Expressing quantity/degree and increase/decrease
Expressing the existence/non-existence of something
Referring to the text and to other texts
Referring to time
Phrases with (the) (noun) of
Miscellaneous

Appendix D. Functional Framework for the *Academic PhraseBank* (Morley, 2018)

Introducing Work
Reviewing the Literature
Describing Methods
Reporting Results
Discussing Findings
Writing Conclusions
Being Cautious
Being Critical
Classifying and Listing
Comparing and Contrasting
Defining Terms
Describing Trends and Projections
Describing Quantities
Explaining Causality
Giving Examples as Support
Signalling Transition
Indicating Shared Knowledge or Understanding
Writing about the Past

Appendix E. Functional Framework for the *Phrase Book* (Godfrey, 2013)

Introducing, defining and classifying
Talking about aim and proposition
Structure, time, sequence and frequency
Methodology and method, findings, size, amount, level and proportion
Movement and change, getting better or worse, allowing or preventing and eliminating
Circumstance, advantage or presence or absence, and importance
Cause and effect, dependency, similarity and difference
Analysing and evaluating ideas
Drawing your own conclusions, stating your own position and summarising your ideas

---

References

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*, 235—247.

Bardovi-Harlig, K. (2012). Formulas, routines, and conventional expressions in Pragmatics research. *Annual Review of Applied Linguistics, 32*, 206—227.

Barros, L. O. (2016). *The only academic Phrasebook you'll ever need: 600 examples of academic language*. CreateSpace Independent Publishing Platform.

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. Amsterdam: John Benjamins.

Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes, 26*, 263—286.

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.

Birkenstein, C., & Graff, G. (2018). *They say/I say": The moves that matter in academic writing* (4th ed.). New York: W. W. Norton & Company, Inc.

Borg, S. (2009). English language teachers' conceptions of research'. *Applied Linguistics, 30*, 358—388.

Bunting, J. D., Diniz, L., & Reppen, R. (2012). *Grammar and Beyond 4*. Cambridge: Cambridge University Press.

Christie, F. (2017). Genres and institutions: Functional perspectives on educational discourse. In S. Wortham, D. Kim, & S. May (Eds.), *Discourse and education. Encyclopedia of language and education* (3rd ed., pp. 29—40). Cham: Springer.

COBUILD. (1996). *COBUILD grammar patterns 1: Verbs*. London: HarperCollins.

Cortes, V. (2001). *Lexical bundles in context: A new taxonomy*. In A. Z. Flagstaff (Ed.). Northern Arizona University.

Cortes, V. (2004). Lexical bundles in published and student writing in history and biology. *English for Specific Purposes, 23*, 397—423.

Cowie, A. P. (1981). The treatment of collocations and idioms in learners' dictionaries. *Applied Linguistics, 2*, 223—235.

Cowie, A. P. (1992). Multiword lexical units and communicative language teaching. In P. J. L. Anaud, & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 1—12).

Cowie, A. P. (1998). Phraseological dictionaries: Some east-west comparisons. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and application* (pp. 209—243). Oxford: Clarendon Press.

Coxhead, A. (2000). A new academic word list. *Tesol Quarterly, 34*, 213—238.

Crookes, G. (1998). On the relationship between second and foreign language teachers and research. *TESOL Journal, 7*, 6—11.

Cruse, D. A. (1986). *Lexical semantics*. Cambridge: Cambridge University Press.

Crystal, D. (2003). *English as a global language* (2 ed.). Cambridge: Cambridge University Press.

Culicover, P. W., Jackendoff, R., & Audring, J. (2017). Multiword constructions in the grammar. *Topics in Cognitive Science*, 552—568.

Daley, R., Jones, S., & Sinclair, J. M. (1971/2004). *English collocation studies - the OSTI report*. London: Continuum.

Davies, Mark (2004). *British national corpus*. (from Oxford University Press). Available online at https://www.english-corpora.org/bnc/.

Davies, M. (2008). The corpus of contemporary American English (COCA): 560 million words, 1990-present. Available online at: https://corpus.byu.edu/coca/.

Davis, M., & Morley, J. (2015). Phrasal intertextuality: The responses of academics from different disciplines to students' re-use of phrases. *Journal of Second Language Writing, 28*.

Doff, A., Jones, C., & Mitchell, K. (1984). *Meanings into words upper-intermediate: Teacher's book*. Cambridge: Cambridge University Press.

Dudley-Evans, A., & St John, M.-J. (1998). *Developments in ESP*. Cambridge: Cambridge University Press.

Durrant, P. (2009). Investigating the viability of a collocation list for students of English for academic purposes. *English for Specific Purposes, 28*, 157—169.

Ellis, N. C., & Ferreira-Junior, F. (2009). Construction learning as a function of frequency, frequency distribution, and function. *The Modern Language Journal, 93*, 370—385.

Firth, J. R. (1957). A synopsis of linguistic theory 1930-1955. In J. R. Firth (Ed.), *Studies in linguistic analysis* (pp. 1—32). Oxford: Basil Blackwell.

Flowerdew, L. (2015). Corpus-based research and pedagogy in EAP: From lexis to genre. *Language Teaching, 48*, 99—116.

Godfrey, J. (2013). *The student phrase book: Vocabulary for writing at university*. London: Palgrave Macmillan.

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Halliday, M. A. K. (1961). Categories of the theory of grammar. *Word, 17*, 241—292.

Halliday, M. A. K. (1966). Lexis as a linguistic level. In C. E. Bazell, J. C. Catford, M. A. K. Halliday, & R. H. Robins (Eds.), *In memory of J. R. Firth* (pp. 148—162). London: Longmans.

Halliday, M. A. K. (1973). *Explorations in the functions of language*. London: Edward Arnold.

Halliday, M. A. K. (1978). *Language as social semiotic*. London: Edward Arnold.

Halliday, M. A. K. (1994). *An introduction to functional grammar* (2 ed.). London: Edward 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., & Matthiessen, C. M. I. M. (2004). *An introduction to functional grammar* (3rd ed.). London: Hodder Arnold.

Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's introduction to functional grammar* (4 ed.). Abingdon: Routledge.

Halliday, M. A. K., McIntosh, A., & Strevens, P. (1964). *The linguistic sciences and Language teaching*. Harlow: Longman.

Hamp-Lyons, E. (2015). The future of JEAP and EAP. *Journal of English for Academic Purposes, 20*, A1—A4.

Hewings, M., & Hewings, A. (2002). It is interesting to note that…. In *A comparative study of anticipatory 'it' in student and published writing* (Vol 21, pp. 367—383). English for Specific Purposes.

Hornby, A. S. (1948). *Oxford advanced learner's dictionary*. Oxford: Oxford University Press.

Hyland, K. (1998). Boosting, hedging and the negotiation of academic knowledge. *Text, 18*, 349—382.

Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes, 27*, 4—21.

Jaccard, P. (1912). The distribution of the flora in the alpine zone. *New Phytologist, 11*, 37—50.

Johns, T. F. (1991a). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. In T. Odlin (Ed.), *Perspectives on pedagogical grammar* (pp. 293—313). Cambridge: Cambridge University Press.

Johns, T. F. (1991b). Should you be persuaded: Two examples of data-driven learning. *ELR Journal, 4*, 1−16.

Jones, L. (1981). *Functions of English: Teacher's book*. Cambridge: Cambridge University Press.

Jordan, R. R. (1980). *Academic writing course*. London: Collins ELT.

Jordan, R. R. (1990). *Academic writing course* (2nd). London: Nelson/Longman.

Jordan, R. R. (1999). *Academic writing course* (3rd). London: Longman.

Kleiser, G. (1917). *Fifteen thousand useful phrases: A practical handbook of pertinent expressions, striking similes, literary, commercial, conversational, and oratorical terms, for the embellishment of speech and literature, and the improvement of the vocabulary of those persons who read, write, and speak English*. New York: Funk & Wagnalls.

Larsson, T. (2017). A functional classification of the introductory it pattern: Investigating academic writing by non-native-speaker and native-speaker students. *English for Specific Purposes, 48*, 57−70.

Lei, L., & Liu, D. (2018). The academic English collocation list. *International Journal of Corpus Linguistics, 23*, 216−243.

Lewis, M. (1993). *The lexical approach: The state of ELT and a way forward*. Hove: Language Teaching Publications.

Lewis, M. (1996). Implications of a lexical view of language. In J. Willis, & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 10−16). Oxford: Heinemann.

Lewis, M. (1997). *Implementing the lexical approach*. Hove: Language Teaching Publications.

Lewis, M. (2000). *Teaching collocation: Further developments in the lexical approach*. Hove: Language Teaching Publications.

Lightbown, P. M. (2000). Anniversary article. Classroom SLA research and second language teaching. *Applied Linguistics, 21*, 431−462.

Lightbown, P. M. (2002). The role of SLA research in L2 teaching: Reply to Sheen. *Applied Linguistics, 23*, 529−535.

Liontas, J. I. (2019). Teaching idioms and idiomatic expressions across the second language curriculum. In E. Hinkel (Ed.), *Teaching Essential Units of Language: Beyond Single-word Vocabulary* (pp. 55−105). London: Routledge.

Liu, D. (2012). The most frequently-used multi-word constructions in academic written English: A multi-corpus study. *English for Specific Purposes, 31*, 25−35.

Mackay, J. (2018). Introduction. In J. Mackay, M. Birello, & D. Xerri (Eds.), *ELT research in action - bridging the gap between research and classroom practice* (pp. 15−21). Faversham: IATEFL.

MacKenzie, I. (2000). Improvisation, creativity, and formulaic language. *Journal of Aesthetics and Art Criticism, 58*, 173−179.

MacKenzie, I., & Kayman, M. A. (Eds.). (2016). *Formulaicity and Creativity in Language and Literature*. London: Routledge.

Maley, A. (2015). 'More research is needed' − a mantra too far? *Humanising Language Teaching Magazine, 1*, 1−5. April 2015.

Malinowski, B. (1923). The problem of meaning in primitive languages. In C. K. Ogden, & I. A. Richards (Eds.), *Meaning of meaning: A study of the influence of language upon thought and of the science of symbolism* (pp. 451−510). London: Kegan Paul, Trench, Trubner & Company.

Martin, J. R. (2010). Language, register and genre. In C. Coffin, T. Lillis, & K. A. O'Halloran (Eds.), *Applied linguistics methods: A reader* (pp. 12−32). New York: Routledge.

Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics, 33*, 299−320.

McCauley, S. M., & Christiansen, M. H. (2014). Acquiring formulaic language: A computational model. *The Mental Lexicon, 9*, 419−436.

Medgyes, P. (2017). The (ir)relevance of academic research for the language teacher. *ELT Journal, 71*, 491−498.

Miller, J. (2020). The bottom line: Are idioms used in English academic speech and writing? *Journal of English for Academic Purposes, 43*(100810).

Mittins, W. H. (1950). *Pattern in English*. London: George Allen & Unwin.

Morley, J. (2004). An initiative to develop Phrasebank: A university-wide online writing resource. In *Phraseology and metaphor in EAP' professional issues meeting of the British association of lecturers in English for academic purposes (BALEAP)*. University of Birmingham, 19 June 2004.

Morley, J. (2018). *Academic Phrasebank - a compendium of commonly used phrasal elements in academic English*. Manchester: The University of Manchester.

Munby, J. (1978). *Communicative syllabus Design*. Cambridge: Cambridge University Press.

Nassaji, H. (2012). The relationship between SLA research and language pedagogy: Teachers' perspectives. *Language Teaching Research, 16*, 337−365.

Nattinger, J. R., & DeCarrico, J. (1992). *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.

Oakey, D. J. (2002). Formulaic language in English academic writing: A corpus-based study of the formal and functional variation of a lexical phrase in different academic disciplines in English. In R. Reppen, S. Fitzmaurice, D. Biber, et al. (Eds.), *Using Corpora to Explore Linguistic Variation* (pp. 111−130). Amsterdam: John Benjamins.

Oakey, D. J. (2009). Fixed collocational patterns in isolexical and isotextual versions of a corpus. In P. Baker (Ed.), *Contemporary Corpus Linguistics* (pp. 142−160). London: Continuum.

Oakey, D. J., & González-Díaz, V. (2019). "Incredibly wild and monstrously hideous …": Adjectival degree intensification in H.P. Lovecraft stories. *Paper presented at the 39th Poetics and Linguistics Association Conference "Stylistics Without Borders", University of Liverpool, 11 July 2019*.

Oakey, D. J., Hughes, L., & Zhang, Y. (In preparation). Using Jaccard Similarity to investigate the extent to which phrases identified by corpus research overlap with phrases in academic writing materials.

OPAL. (2019). Oxford Phrasal Academic Lexicon.

Palmer, H. E. (1933). *Second interim report on English collocations: Submitted to the tenth annual conference of English teachers*. Tokyo: The Institute for Research in English Teaching.

Paran, A. (2017). 'Only connect': Researchers and teachers in dialogue. *ELT Journal, 71*, 499−508.

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards, & R. W. Schmidt (Eds.), *language and communication* (pp. 191−226). London: Longman.

Peacock, M. (2007). A comparative study of introductory 'it' in research articles across eight disciplines. In *Corpus linguistics 2007 conference, 28 july 2007*. University of Birmingham.

Pennycook, A. (1994). *The cultural politics of English as an international language*. London: Longman.

Peters, A. (1973). *The units of language acquisition*. Cambridge: Cambridge University Press.

Phillipson, R. (1988). *Linguistic imperialism*. Oxford: Oxford University Press.

Pica, T. (1997). Second language teaching and research relationships: A north American view. *Language Teaching Research, 1*, 48−72.

Pinker, S. (1994). *the language instinct*. London: Penguin.

Robins, R. H. (1961). John rupert Firth. *Language, 37*, 191−199.

Samraj, B. T. R. (2018). Using lexical bundles research findings in writing instruction: Some questions. In *American asssociation for applied linguistics. Chicago, IL*.

Sheen, R. (2002). A response to lightbown's (2000) 'anniversary article: Classroom SLA research and second language teaching'. *Applied Linguistics, 23*, 519−528.

Simpson-Vlach, R. C., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics, 31*, 487−512.

Sinclair, J. M. (1987). *Looking Up: An account of the COBUILD Project in lexical computing*. London: Collins ELT.

Sinclair, J. M. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Sullivan, J. (2016). *Simply said: Communicating better at work and beyond*. Hoboken, NJ: Wiley.

Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Swales, J., M. (2002). Integrated and fragmented worlds: EAP materials and Corpus Linguistics. In J. Flowerdew (Ed.), *Academic Discourse* (pp. 150−164). Harlow: Longman.

Swales, J. M. (2004). *Research genres: Explorations and applications*. Cambridge: Cambridge University Press.

Swales, J. M., & Feak, C. B. (2012). *Academic writing for graduate students: Essential tasks and skills* (3rd ed.). Ann Arbor: University of Michigan Press.

Tardy, C. M., & Jwa, S. (2016). Composition studies and EAP. In K. Hyland, & P. Shaw (Eds.), *The routledge handbook of English for academic purposes* (pp. 56–58). Oxford: Routledge.

Theakston, A., & Lieven, E. (2017). Multiunit sequences in first language acquisition. *Topics in Cognitive Science, 9*, 588–603.

Thorndike, E. L., & Lorge, I. (1938). *The teacher's word book of 30,000 words.* New York: Teachers College.

Tomlinson, B. (2013). Introduction: Applied linguistics and materials development. In *Applied linguistics and materials development* (pp. 1–8). London: Bloomsbury.

Tribble, C. (1996). *Writing.* Oxford: Oxford University Press.

Tribble, C. (2009). Writing academic English – a survey review of current published resources. *English Language Teaching Journal, 63*, 400–417.

Tribble, C. (2015). Writing academic English further along the road. What is happening now in EAP writing instruction? *English Language Teaching Journal, 69*, 442–462.

Werner, M. (1985). *The bride's thank-you note handbook* (Revised ed.). New York: Simon and Schuster.

West, M. (1953). *A general service list of English words.* New York: Longmans, Green and Co.

Widdowson, H. G. (1989). Knowledge of language and ability for use. *Applied Linguistics, 10*, 129–137.

Wilkins, D. A. (1976). *Notional Syllabuses.* Oxford: Oxford University Press.

Wray, A. (2012). What do we (think we) know about formulaic language? An evaluation of the current state of play. *Annual Review of Applied Linguistics, 32*, 231–254.

Xue, G., & Nation, I. S. P. (1984). A university word list. *Language Learning and Communication, 3*, 215–229.

David Oakey is a lecturer in TESOL and Applied Linguistics at the University of Liverpool, UK. He has taught undergraduate and graduate classes in EAP writing, grammar, lexis, and corpus linguistics at universities in China, Turkey, and the USA, and has published research on the phraseology of disciplinary and interdisciplinary discourse.