EAP Materials: Insights from Corpus-Based Studies
Eleni Agathopoulou
Aristotle University of Thessaloniki

Abstract
Corpus-based research can offer valuable insights to teachers of English for Academic Purposes (EAP). Such research may demonstrate typical and less typical patterns of language use and highlight important phenomena commonly not included in conventional EAP syllabuses. Moreover, studies exploring non-native English corpora inform us on learner problems in academic English. To make prominent the pedagogical implications of corpus research for EAP, in the current paper I focus on corpus findings in three areas, namely, compound words, discourse-organising nouns and titles of scientific articles and offer suggestions for the pedagogical exploitation of these findings.
Key words: academic English, corpus research, teaching materials, compound words, discourse-organising nouns, titles of scientific articles

1. Introduction

As well known, teachers of English for Academic Purposes (EAP)[1] often have to design their own courses and to create teaching materials or to supplement existing ones. This endeavour presupposes, among other things[2], awareness of the essential features of the texts relevant with the students' needs and of the problems the students may have in handling these textual features.

In recent years, knowledge of genre-specific features has been aided significantly by research employing electronic corpora and software tools for corpus search. Studies in this methodological domain have yielded informative findings regarding features of academic English, indicatively: rhetorical structure (e.g. Halleck & Connor 2006; Peacock 2002; Pho 2010) grammatico-rhetorical relations (e.g. Pho 2008; Salager-Meyer 1989), grammatico-functional descriptions (e.g. Ferguson 2001), lexico-grammatical patterns (e.g. Hunston & Francis 1999), cohesion (e.g. Bondi 2004, Gray 2009 and references there; Römer & Wulff 2010; Swales 2005), metadiscourse (Harwood 2003; Hyland 2005; Samson 2004), hedging (e.g. Hyland 1998a,b; Poos & Simpson 2002), collocations (e.g. Curado Fuentes 2001; Lúzon 2000; Ward 2007), lexical bundles (e.g. Biber & Barbieri 2007; Csomay & Cortes 2010; Hyland 2008; Wood 2010), idioms (e.g. Simpson & Mendis 2003) and word-frequencies (e.g. Hyland & Tse 2007; Martínez et al. 2009). Corpus-informed studies have also made prominent cross-disciplinary differences in particular academic genres (e.g. Bruce 2010; Conrad 2001; Golebiowski 2009; Harwood 2005a; Hyland 2001, 2009; Samraj 2005; Thompson 2009).

Even more recently, significant insights regarding students' problems in handling English academic genres come from studies drawing on learner corpora usually alongside comparable corpora from native peers or/and from expert native writers (Granger 2003,

---

[1] In this paper EAP will be used as an umbrella term that includes English for General Academic Purposes (EAGP) and/or English for Specific Academic Purposes (ESAP).
[2] For instance, knowledge about language learning and teaching theories, principles of course design and needs analysis models and methods.

2004; for more recent literature reviews, see Guilquin et al. 2007 and Granger 2009). This research has disclosed discrepancies between English native and non-native use regarding, for example, the expression of: author stance/engagement (e.g. Charles 2009; Hasselgärd 2009; Hatzitheodorou & Mattheoudakis 2007, 2008, 2009; Hyland 2009; Neff et al. 2003), hedging (Hinkel 2005), cohesion (e.g. Condrad 1999; Gray 2009; Hinkel 2003), modality (Aijmer 2002; Hyland & Milton 1997), aspect (Leńko-Szymańska 2007), tense, aspect and voice (Hinkel 2004), use of first person pronoun (Martínez 2005) and collocations (Nesselhauf 2005, 2007).

Ideally, EAP teachers should be able to do their own research based on corpora relevant with their students' needs. However, this is may not be feasible due to lack of corpus-searching skills or access to specialized corpora. Hence, the next best thing for EAP teachers is to be aware of corpus-based research, as this research may (a) "highlight important aspects of a text or text collection that may go unnoticed otherwise" (Römer & Wulff 2010: 101), (b) reveal that language descriptions in commercial EAP textbooks are misleading (Barbieri & Eckhardt 2007; Harwood 2005b,c; Hyland 1994; Paltridge 2002; Swales 2002; Sznajder 2010; WonHo Yoo 2009) or underrepresent aspects of academic language use (e.g. Chen 2010; Pho 2008) and (c) confirm or disprove intuitive assumptions regarding learners' language problems (e.g. Rundell & Granger 2007).

While many authors of corpus studies underscore that their findings have important implications for EAP teaching, relatively few of these authors suggest specific classroom activities. The latter point is not meant as a criticism, but simply to underline that incorporating the findings of corpus research is often left to the ability and the imagination of EAP teachers (cf. Conrad 1999: 16). This takes us to the aim of the present brief paper: to bring into focus some insights from corpus research and suggest how these insights can be implemented for the EAP classroom, as elaborated next in Section 2. The paper is rounded off by Section 3 which contains the recommendation of some corpus-based studies that demonstrate how their findings can be translated into teaching practice.

2. Some corpus-based findings and suggestions for classroom implementation

This section is divided into three subsections, each one focusing on a specific area of interest for EAP, namely: compound words (3.1), discourse-organising nouns (3.2) and titles to research articles (3.3).

2.1. Compound words

English compound words (hereafter 'compounds') consist of two or more words or stems (see, e.g. Bauer 1978: 49), for example: *drug research*, *borderline personality disorder*, *cannabinoid-1 receptor blocker*, *filter-feeding bivalve molluscs*, *micro-column liquid chromatography-tandem mass spectrometry*. Corpus research has corroborated the importance of compounds for EAP. Salager's (1984) small corpus analysis revealed that compounds are considerably more frequent in medical than in general English texts (0.87% vs. 9.76% respectively) and in Biber and Gray's (2010) large corpus analysis compounds occurred about three times more in English research articles than in oral discourse. Even more interesting for teachers of English for Specific Academic Purposes

may be Horsella and Perez's (1991) corpus study of chemical English. This study disclosed that the frequency of compounds varied from 12.75% in introductory course texts to 37.17% in more advanced course texts and, impressively, to 50.08% in texts of graduate level and those concerning communication among scientists; also, the higher the text level, the larger was the number of multi-word compounds.

Halliday (1993: 77) points out that among other features that contribute to the high lexical density of scientific English, compounds may be "the hardest examples to process" and second language research has found that EAP learners have difficulty both with the structural form of such compounds[3] and with the selection of proper contextual clues to interpret their meaning appropriately (Olshtain 1981; see also Trimble 1985). EAP/ESP textbooks and articles offer a variety of exercises on compounds, for example, "Form compounds from the following words" and "Find X two- or three-word compounds in the text" (Saslow & Mongillo 1985), "Turn these phrases into noun compounds" (Master 2003), "Choose the correct definition for the noun compound" (Master 2003; Williams 1985) and "Use a compound to name an entity" (e.g. *A device that scans bar codes = a bar code scanner*; Boeckner & Brown 1993)[4]. However, while exercise types that focus on the structure of compounds abound, to my knowledge, there is a scarcity of exercises/activities that may enhance learners' ability to select the proper contextual clues for interpreting compounds. The activity I suggest below aims at the treatment of the latter type of learner problem.

First a slight diversion is in order to explain the theoretical background of this activity. According to the "linear on-going accumulation principle", "each discourse element and its textual presentation at any stage of the ongoing process of discourse is premised upon the collection of all the foregoing elements and builds upon it" (Candlin & Lotfipour-Saedi 1983, cited in Campoy & Carda 1998: 63). To illustrate how this is related with compounds, consider (1)-(5) appearing in this order in a text (excerpted from Halliday ([1999] 2004: 108).

(1)   …the fracture of glass…
(2)   …preexisting cracks . . .
(3)   …Cracks in glass can grow…
(4)   …the slow growth of cracks…
(5)   …the crack growth …

The compound "crack growth" (5) is gradually built upon elements that appear first in a sentence (3) and then in a noun phrase + preposition + noun phrase construction (4). Also, in the beginning of the text there is a synonym for the first constituent noun of the compound ('fracture') which first appears in an adjective + noun phrase (2).

I suggest that the "linear on-going accumulation principle" can be exploited for pedagogical purposes. Learners can be asked to mark phrases or words in a text that relate to the eventual formation of a compound in this text, as demonstrated in the following text excerpts.

---

[3] Learners may give left-headed instead of right-headed interpretations, that is, translate *engine alterations* as "a kind of engine". For a treatment of this type of difficulty in large contexts see Williams (1985).
[4] For a large variety of exercise types on compounds see Master (op. cit.).

(6)   For the first time, researchers have established how much exercise women should be able to do for their age and found that their capacity is slightly lower than men's. It also declines a bit faster than men's as they grow older.

(7)   Women whose exercise capacity was less than 85 percent of what it should be were twice as likely to die within eight years, the researchers found[5].

(8)   In 1930 the death rate from lung cancer among men was less than 5 per 100,000 population per year. By 1950 it had quintupled to more than 20; today it is above 70. The CDC in November 1993 said lung cancer death rates increased from … [6]

(9)   In 1800, the majority of the world's population resided in Asia and Europe, with 65 percent in Asia alone. By 1900, Europe's share of world population had risen to 25 percent, fueled by the population increase that accompanied the Industrial Revolution. Some of this growth spilled over to the Americas, increasing their share of the world total. World population growth accelerated after World War II…[7]

Additional merits of the activity described above are that it fosters awareness of cohesive relations in texts and, on a practical plane, it can be carried out using texts already included in the EAP materials.

2.2. Discourse-organising nouns

This section deals with a category of abstract nouns in noun phrases that refer to "a stretch of discourse as a linguistic act" and are thus termed "metalinguistic labels" or just "labels" (Francis 1994: 83). Some examples are: *argument, assumption, belief, claim, comment, conclusion, concept, contention, controversy, debate, definition, dispute, example, excerpt, idea, illustration, notion, opinion, passage, point, position, principle, process, quotation, rationale, statement, theory, term* and *view*. More general categories of abstract nouns with discourse-organizing function have been termed 'shell nouns' (Aktas & Cortes 2008 and references there) or 'signalling nouns' (Flowerdew 2003, 2006, 2010).

Phrases containing the discussed nouns may have anaphoric or cataphoric reference ('retrospective' or 'advance' labels, respectively, in Francis, ibid.), as demonstrated by the elements in bold type in (10) and (11) respectively.

---

[5] Source: http://health.dailynewscentral.com/content/view/1423/63
[6] Source:
http://www.prb.org/Content/NavigationMenu/PRB/Educators/Human_Population/Population_Growth/Population_Growth.html
[7] Source:
http://www.prb.org/Content/NavigationMenu/PRB/Educators/Human_Population/Population_Growth/Population_Growth.html

(10) Particular concern is expressed regarding the field of R&D, because of the dominance of military expenditure in this area and the scarcity of qualified people. **This point** is made by Glyn and Harrison. . .(from Charles 2003: 317)

(11) Spending on Medicare, the federal health program for the elderly and the disabled, grew 5.7 percent in 2003, to $283.1 billion, compared with an increase of 7.6 percent in 2002. Out-of-pocket payments were the only major source of health spending to increase faster in 2003 than in 2002, the administration said. Federal health officials offered **two reasons**: the number of people without health insurance is increasing, and employers who provide coverage are requiring workers to bear more of the costs[8].

Research employing learner corpora has confirmed a serious underuse of discourse-organising nouns in academic argumentative essays written by university students of upper-intermediate to advanced levels of proficiency in English. Both L1 Chinese students (Flowerdew 2010) and L1 Greek students (Mattheoudakis & Hatzitheodorou (in print)) have been found to use the discussed nouns about three times less than their native English peers[9]. A characteristic example from the latter study is the noun *argument*, which occurred 253 and 19 times in the native and the non-native corpora, respectively. Mattheoudakis and Hatzitheodorou (in print) also found that while the native English writers often modified discourse-organising nouns with adjectives to express stance, for example, *strong/weak argument* or *heated/fiery debate*, the Greek learners did not. Other learner problems identified in this area by Flowerdew's (2006) study, which was based on corpora similar to those in his 2010 study, regard wrong collocations, for example "draw an impact" or "discrimination to" (instead of "discrimination against"), or noun omission after a demonstrative (e.g. "these"), which resulted in unclear cross-sentence reference. Mattheoudakis and Hatzitheodorou (in print) attribute the learners' underuse of discourse-organising nouns to lack of instruction in this area, probably due to that noun phrases containing the discussed nouns can not be easily classified into categories of connecting phrases for teaching purposes, unlike adverb phrases that can be categorized under labels such as "Summarizing" (e.g. all in all), "Making a concession" (e.g. at any rate), "Time reference" (e.g. afterwards) etc. (*Longman English Grammar*1988: 326)[10].

Explicit instruction within EAP courses may be able to help learners to develop effective use of the discussed phrases. To this end, teachers can provide learners with texts rich in discourse-organising nominal phrases and discuss the function of these phrases. For example, excerpt (12) serves as an illustration of both anaphoric and cataphoric functions of the discussed nominals (underlined and in bold, respectively).

(12) Education in global citizenship, or world citizenship, therefore, requires familiarisation with and understanding of the network of relationships and

---

[8] Source:
http://www.cafeoflifepikespeak.com/Research_Articles/US%20HEALTH%20CARE%20SPENDING%20STATS%202003.doc.

[9] Flowerdew (2006, 2010) deals with "signalling nouns", which is a category of abstract discourse-organizing nouns broader than that of "metalinguistic labels". For the purposes of the present paper I focus on the overlap in findings between Flowerdew's and Mattheoudakis and Hatzitheodorou's studies.

[10] This may explain the attested overuse of adverbials for text cohesion by learners; see Mattheoudakis and Hatzitheodorou (in print) and references there.

processes that pertain to a complex and multifarious world, inclusive of the local community to which citizens belong. By engaging in aspects of <u>this ongoing debate</u>, which are pertinent to the discourse for citizenship education and civic culture, **the following argument** will emerge. A world-embracing vision derived from… (Golmohamad, M. 2009)

I also suggest the following activities.
Activity A
Remove the sentences containing these nouns and ask students to find the place of these sentences in the text. To illustrate, this can be done with the sentence marked in bold type in (13).

(13) There is a mistaken belief that literacy development doesn't truly begin until a child starts formal instruction at an educational institution. **This is a misconception**. In fact, literacy development also involves listening and speaking abilities, which begin prior to elementary school. There are many things parents can do to aid in their baby's literacy development.[11]

Activity B
In addition to the original sentence include other sentences with discourse organizing nouns and ask students to choose the sentence they think that best completes the text, as in (14).

(14) Patriotism does tend to cloud or distort civic judgment, and this can have dire consequences for those at the sharp end of unjust national institutions and policies. _____. It is not necessarily the case that political judgments made by people who love their countries will be distorted and it may be that the danger can be somewhat mitigated by making patriots aware of it.

1. But, again, this argument is not decisive. [The original sentence]
2. But, again, this is not a sound argument.
3. But, again, this argument appears to be quite plausible.
4. But, again, this argument is a fallacious one.

Activity C
Draw learners' attention to collocations with discourse-organizing nouns. For instance, a cursory search for collocates of 'this/these argument(s)' in Google Scholar[12] yielded a wealth of such collocates shown below.
§ Examples of verb collocates: *adopt, clarify, counter, display, defend, dispute, elucidate, endorse, expound, make, meet, prove, question, support, take*.

---

[11] Source: Pear, 2005 <<u>www.ehow.com/facts_6064535_literacy-development-babies.html</u> >
[12] Google Scholar (<u>http://scholar.google.gr/</u>) is a searching engine that indexes the full text of scholarly literature in a wide range of research areas. Note that the 'Advanced Scholar Search' allows the option of searching articles only in specific subject areas, such as Social Sciences, Biology etc.

§ Examples of noun phrase + *of* preposition phrase collocates: *the basis of, (a/the) consideration of, the form of, the foundation of, the implication of, the line of, the premise of, the robustness of, the structure o(a/the) consideration of.*

§ Examples of adjective collocates (in premodifying or postmodifying position): *accurate, baseless, cogent, commonsensical, consistent, decisive, explicit, feeble, firm, frail, groundless, inconsistent, implicit, indecisive, persuasive, robust, sound, specious, straightforward, strong, unconvincing, unequivocal, unfounded, unreasonable, unsound, weak.*

Using appropriate texts, teachers can ask learners to pick out collocations with frequently used discourse-organizing nouns and, furthermore, to brainstorm about synonymous collocations. Also, providing lists of such collocations may help learners improve their essays. Importantly, this activity can offer a rich supply of expressions of stance, as some of the verb collocates and many of the adjective collocates of 'argument' shown previously. For further focus on such collocates, learners can be asked whether the authors of a text agree or disagree with an argument, what they think of a specific claim, idea, opinion etc. Alternatively, learners can be directed to trusted websites for individual discovery of the collocations under discussion.

2.3. *Titles of scientific articles*

Given the importance of heading scientific articles appropriately, textbooks on writing scientific articles usually contain a section with advice on titles. Here are some examples of such textbook advice selected for reasons which will be clear later in this subsection.

> "It [=the title] should include a verb, which should be in the present tense. Find an active verb, and aim for the grammatical structure *Subject–Active verb–Object*, i.e., "Viral Interleukin-6 Blocks Interferon Signaling". (Katz 2009: 159)
> "Hanging titles (=titles with a colon) … appear pedantic, often place the emphasis on a general term rather than a more significant term, necessitate punctuation, scramble indexes, and in general provide poor titles. …. titles phrased as questions … become unintelligible, and in my view "question" titles should not be used" (Day 1988: 20-21)."
> "Try to avoid questions in the title, unless the issue remains unsettled or you came up with a clear answer". (Alexandrov & Hennerici 2007: 257)

Let us next turn to relevant findings from two recent corpus studies.

Soler (2007) searched a corpus of 480 research article titles in the biological (medicine, biology and bio-chemistry) and social sciences (linguistics, psychology and anthropology). Titles were categorized into the following four types: *Full-sentence type*, e.g. "Learning induces a CDC2-related protein kinase", *Nominal group type*, e.g. "Acute liver failure caused by diffuse hepatic melanoma infiltration", *Compound type*, e.g. "Romanian nominalizations: case and aspectual structure" and *Question type*, e.g. "Does the Flynn effect affect IQ scores of students classified as LD?" (Soler: 94 & 96). The corpus analysis demonstrated a preponderance of the nominal group (42%-38%) and a scarcity of the question type (0%-2%) in the titles of research articles across both of the disciplines examined. Moreover, full-sentence titles headed articles in the biological

sciences exclusively (38%) and compound-group titles were used much less in the biological than in the social sciences (10% and 38% respectively). Ball's (2009) analysis of 20 million research article titles revealed that from 1966 to 2005 the use of question-type titles increased very little in physics (0,25%-0.5%) but significantly in life sciences (0.0025%-2.3%) and even more so in medicine (1%-5%). Given that titles with questions may more easily attract one's attention, in Ball's account, "The widespread use of the question mark clearly correlates with the increasing competition in science and research" (: 11). Also, Ball speculates that in the future there may be a further increase in question-type titles in medical articles.

EAP teachers should therefore be aware that evidence from corpus research may fly in the face of textbook advice recommending full-sentence titles and the avoidance of nominal-group and question-type titles in research articles. Also the corpus-attested differences across disciplines in this area offer another useful insight that can be incorporated in the EAP materials.

I suggest the following activity to enhance learner awareness concerning types of titles in research articles. Ask learners to
§   find the most cited or the most downloaded articles in an academic journal of their specialism at online publisher sites,
§   classify these articles into the four types described previously and
§   discuss which types occur more frequently.

To offer a small illustration, in the beginning of October 2010 my own search of the ten most downloaded articles in the Journal of English for Academic Purposes (JEAP) and the American Journal of Medical Quality (AJMQ) yielded the results exhibited in Table 1.

Table 1. Types of titles in the ten most downloaded articles in JEAP & AJMQ (October 2010)

	JEAP	AJMQ
Full-sentence type	-	1
Nominal group type	4	3
Compound type	6	4
Question type	-	2

Although the above results come from a negligible sample size, note that they generally comply with Soler's (2007) findings regarding cross-disciplinary differences in this area and with the relatively frequent appearance of question-type titles in medical research articles attested in Ball's study.

Last, in view of Alexandrov and Hennerici's (op. cit.) recommendation, a further activity, probably more suitable for postgraduate students, would be to scan the abstract or the discussion section in order to find out whether the author(s) came up with a clear answer or whether the issue remained unsettled.

3. Final notes

The aim of this paper was to provide a short illustration of how EAP practitioners "…can benefit from being wise consumers of corpus-based research, even if they do not do the research themselves" (Conrad 1999: 16). For this purpose, I have offered examples of how corpus findings can be incorporated into EAP materials. A variety of other tasks based on data from research on corpora of academic English can be found in studies that are at the interface of corpus linguistics and language teaching. Indicatively, see Harwood (2003) on person markers and metadiscourse, Charles (2009) on the projection of stance, Hyland (2008), Flowerdew (2002), Johns (1995, 2008), Hüttner et al. (2009), Pang (2002), and Swales and Lindemann (2002) on genre awareness, Suganthi (2009) on the construction of academic identities in dissertations, Hyland (1996) and Hyland and Milton (1997) on hedging, and Curado Fuentes (2001) on collocations; also see Aranha (2009) and Feak and Swales (2010) on genre-based syllabuses and Harwood (2010) on academic citations. Examples of relevant studies that draw on corpora of spoken academic English are Jones and Schmitt (2010) on vocabulary and phrases, and Simpson and Mendis (2003) on idioms. Last, some of the books that offer suggestions for pedagogical materials consonant with corpus-informed data are those by Hyland (2006), Swales (2004), and Swales and Feak (2000)[13].

References

Aijmer, K. (2002). "Modality in advanced Swedish learners' written interlanguage". *In* S. Granger, J. Hung, and S. Petch-Tyson (eds), *Computer learner corpora, second language acquisition and foreign language teaching*. Amsterdam: Benjamins, 55-76.

Aktas, R.N. & V. Cortes (2008). "Shell nouns as cohesive devices in published and ESL student writing". *Journal of English for Academic Purposes* 7(1): 3-14.

Alexandrov A. & M. Hennerici (2007). "Writing good abstracts". *Cerebrovascular Diseases* 23: 256-259.

Aranha, S. 2009. "The Development of a Genre-Based Writing Course for Graduate Students in Two Fields". In C. Bazerman, A. Bonini & D. Figueiredo (eds) Genre in a Changing World. Indiana: Parlor Press and WAC Clearinghouse, 465-482.

Ball, R. (2009). "Scholarly communication in transition: The use of question marks in the titles of scientific articles in medicine, life sciences and physics 1966–2005". *Scientometrics* 79(3): 1-13.

Barbieri, F. & S. Eckhardt (2007). "Applying corpus-based findings to form-focused instruction: The case of reported speech". *Language Teaching Research* 11(3): 319-346.

Bauer, L. (1978). *The Grammar of Nominal Compounding*. Odense: Odense University Press.

Bhatia V. K. (1993). *Analysing Genre - Language Use in Professional Settings*. London: Longman.

Biber, D. & F. Barbieri (2007). "Lexical bundles in university spoken and written registers". *English for Specific Purposes* 26(3): 263-286.

Biber D. & B. Gray (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. Journal of English for Academic Purposes 29(1): 2-20.

Boeckner, K. & P.C. Brown (1993). *Oxford English for Computing*. Oxford: Oxford University Press.

Bondi, M. (2004). "The discourse function of contrastive connectors in academic abstracts". In K. Aijmer & A.-B. Stenström (eds), *Discourse patterns in spoken and written corpora* (pp. 139-156). Amsterdam: John Benjamins.

Bruce, I. (2010). "Textual and discoursal resources used in the essay genre in sociology and English". *Journal of English for Academic Purposes* 9(3): 153-166.

Campoy M.C. & L. Carda (1998). "Lexical and syntactic pattern choice in the titles of chemistry research articles". *I Jornades Catalanes sobre Llengües per a Finalitats Específiques 1998*: 61-66.

Candlin, C. N. & K. Lotfipour-Saedi (1983). "Processes of discourse." *Journal of Applied Language Study* 1: 103-133.

Charles, M. (2003). "'This mystery...'": a corpus-based study of the use of nouns to construct stance in theses from two contrasting disciplines". *Journal of English for Academic Purposes* 2(4): 313-326.

---

[13] For an example of a genre-based course that drew on Swales & Feak's book, see Aranha (2009).

Charles, M. (2009). ""Can I use I, should I use we, or do I have to use the passive?" A Pedagogic Approach to Stance in Academic Writing". In K. Katsaboxaki-Hodgets (ed.), *Proceedings of the International Conference "Options and Practices of LSP practitioners"*. Crete: University of Crete Publications, 69-93.

Chen, L. 2010. "An Investigation of Lexical Bundles in ESP Textbooks and Electrical Engineering Introductory Textbooks". In D. Wood (ed.), *Perspectives on Formulaic Language: Acquisition and Communication*. London/New York: Continuum, 107-125.

Conrad, S. (1999). "The importance of corpus-based research for language teachers". *System* 27(1): 1-18. Cortes, V., and E. Csomay (2007), "Positioning lexical bundles in university lectures," in: M. C. Campoy and M.J. Luzon (eds.) *Spoken corpora in applied linguistics*. Bern: Peter Lang, 57–76.

Conrad, S. (2001). Variation among disciplinary texts: A comparison of textbooks and journal articles in biology and history. In S. Conrad & D. Biber (Eds.) Multi-dimensional Studies of Register Variation in English. Harlow: Pearson Education, 94-107.

Csomay, E. & V. Cortes (2010). Lexical bundle distribution in university classroom talk. In Gries, S., S. Wulff, M. Davies (eds), *Corpus Linguistic Applications: Current Studies, New Directions*. Amsterdam/New York: Rodopi V.B., 153-168.

Curado Fuentes, A. (2001). "Lexical behaviour in academic and technical corpora: implications for ESP development". In *Language Learning & Technology* 5/3. <http://llt.msu.edu/vol5num3/pdf/curado.pdf>

Day, R. A. (1994). *How to write and publish a scientific paper*. Cambridge, UK: Cambridge University Press.

Feak, C. B. & J. Swales (2010). "Writing for publication: Corpus-informed materials for postdoctoral fellows in perinatology". In N. Harwood (ed.), *English Language Teaching Materials: Theory and Practice*. Cambridge: Cambridge University Press, 279-300.

Ferguson, G. (2001). "If you pop over there: a corpus-based study of conditionals in medical discourse". *English for Specific Purposes* 20(1): 61-82.

Flowerdew, J. (2002). "Genre in the classroom: a linguistic approach". In A.M. Johns (ed.), *Genre in the Classroom: Multiple Perspectives*. Mahwah: Lawrence Erlbaum Associates, 91-102.

Flowerdew, J. (2003). Register specificity of signalling nouns in discourse. In C. Meyer and P. Leistyna (eds), *Corpus analysis: Language structure and language use*. Amsterdam: Rodopi Publishers, 35-46.

Flowerdew, J. (2006). "Use of signalling nouns in a learner corpus". *International Journal of Corpus Linguistics*, 11(3): 345-362.

Flowerdew, J. (2010). "Use of signalling nouns across L1 and L2 writer corpora". International Journal of Corpus Linguistics 15(1): 36-55.

Francis, G. (1994). "Labelling discourse: an aspect of nominal-group lexical cohesion". In M. Coulthard (ed.), *Advances in Written Text Analysis*. London: Routledge, 83-101.

Golebiowski, Z. (2009). "Prominent messages in Educationand Applied Linguistic abstracts: How do authors appeal to their prospective readers?". *Journal of Pragmatics* 41(4): 753-769.

Golmohamad, M. (2009). "Education for world citizenship: Beyond national allegiance". *Educational Philosophy and Theory* 41 (4): 466-486.

Gray B. (2009). "On the use of demonstrative pronouns and determiners as cohesive devices: A focus on sentence-initial *this/these* in academic prose". *English for Specific Purposes* 9(3): 167-183

Granger, S. 2003. The International Corpus of Learner English: A new resource for foreign language learning and teaching and second language acquisition research. *TESOL Quarterly* 37(3): 538-546.

Granger, S. (2004). "Computer learner corpus research: current status and future prospects." Ulla Connor and Thomas A. Upton, eds. *Applied Corpus Linguistics. A Multi-Dimensional Perspective*. Amsterdam: Rodopi, 123-145.

Granger S. (2009). "The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation". In Aijmer, K. (ed.) *Corpora and Language Teaching*. Amsterdam / Philadelphia: Benjamins, 13-33.

Gray, B. (2009). "On the use of demonstrative pronouns and determiners as cohesive devices: A focus on sentence-initial this/these in academic prose". *Journal of English for Academic Purposes* 9(3): 167-183.

Gilquin, G., Granger, S. & Paquot, M. 2007. "Learner corpora: The missing link in EAP pedagogy". In *Corpus-based EAP Pedagogy,* P. Thompson (ed.). Special issue of *Journal of English for Academic Purposes* 6(4): 319-335.

Halleck, J. B. and U. M. Connor (2006). "Rhetorical moves in TESOL conference proposals". *Journal of English for Academic Purposes* 5(1): 70-86.

Halliday, M.A.K. (1993). "Some grammatical problems in scientific English". In M.A.K. Halliday and J.R. Martin (eds), *Writing Science: Literary and Discursive Power*. London: Falmer, 69-85.

Halliday, M. A. K. ([1999]2004). "The grammatical construction of scientific knowledge: The framing of the english clause". In J. J. Webster (ed.), *The Language of Science, volume 5 of The Collected Works of M. A. K. Halliday*. London: Continuum, 102-134.

Halliday, M.A.K. & Martin, J.R. 1993. *Writing Science: Literacy and Discursive Power*. London: Falmer Press.

Harwood, N. (2003). *Person markers and interpersonal metadiscourse in academic writing: A multidisciplinary corpus-based study of expert and student texts*. Unpublished doctoral dissertation, Canterbury Christ Church University College, Canterbury.

Harwood, N. (2005a). "'We do not seem to have a theory. The theory I present here attempts to fill this gap': inclusive and exclusive pronouns in academic writing". *Applied Linguistics* 26(3): 343-375.

Harwood, N.(2005b). "'I hoped to counteract the memory problem, but I made no impact whatsoever': discussing methods in computing science using *I*'. *English for Specific Purposes* 24(3): 243-267.

Harwood, N. (2005c). "What do we want EAP teaching materials for?". *Journal of English for Academic Purposes* 4(2): 149-161.

Harwood, N. (2010). "Research-based materials to demystify academic citation for postgraduates". In N. Harwood (ed.), *English Language Teaching Materials: Theory and Practice*. Cambridge: Cambridge University Press, 301-321.

Hasselgärd H. (2009). "Thematic choice and expressions of stance in English argumentative texts by Norwegian learners". In K. Aijmer (ed.), Corpora and Language Teaching. Amsterdam / Philadelphia: Benjamins, 121-140.

Hatzitheodorou, A.M. and M. Mattheoudakis. (2007). "The Projection of Stance in the Greek Corpus of Learner English (GRICLE)". In *Proceedings of CorpusLinguistics 2007*: <http://ucrel.lancs.ac.uk/publications/CL2007/>

Hatzitheodorou, A.M. and M. Mattheoudakis. (2008). "Exploring the Marking of Stance in Argumentative Essays Written by EFL Learners and Native Speakers of English". *Proceedings of the 8th Teaching and Language Corpora Conference*. Ana Frankenberg- Garcia (ed) ISLA, Lisboa. 133-141.

Hatzitheodorou, A.M. and M. Mattheoudakis. (2009). "'It is More Than True That Television Reproduces Life': The Effect of Greek Rhetorical Conventions on Greek Learners' Academic Writing in English".*Selected Papers from the 18th International Symposium on Theoretical and Applied Linguistics*. A. Tsangalidis (ed.) Thessaloniki: Monochromia Publishing, 167-176.

Hinkel, E. (2003). "Adverbial markers and tone in L1 and L2 students' writing". *Journal of Pragmatics* 35(7): 1049-1068.

Hinkel, E. (2004). "Tense, aspect and the passive voice in L1 and L2 academic texts". *Language Teaching Research* 8(1): 5-29.

Hinkel, E. (2005). "Hedging, inflating, and persuading in L2 academic writing". *Applied Language Learning* 14(2): 29-54.

Horsella, M. & F. Pérez (1991). "Nominal Compounds in Chemical English Literature: Toward an Approach to Text Typology". *English for Specific Purposes* 10(2): 125-138.

Hüttner, J., U. Smit, B. Mehlmauerlarcher (2009). „ESP teacher education at the interface of theory and practice: Introducing a model of mediated corpus-based genre analysis". *System* 37(1): 99-109.

Hyland K (1994). "Hedging in academic writing and EAP coursebooks". *English for Specific Purposes* 13(3): 239-256.

Hyland, K. (1996), *Nurturing Hedges in the ESP Curriculum. System* 24: 4, 477-490.

Hyland K (1998a). Hedging in scientific research articles. Amsterdam: Benjamins.

Hyland K (1998b). "Boosting, hedging and the negotiation of academic knowledge". *Text* 18(3): 349–382.

Hyland, K. (2001). "Humble Servants of the Discipline? Self-Mention in Research Articles". *English for Specific Purposes* , 20(3): 207-226.

Hyland, K. (2005). *Metadiscourse*. London/New York: Continuum.

Hyland, K. (2006). *English for Academic Purposes: An Advanced Resource Book*. London: Routledge.

Hyland, K. (2008). "Genre and academic writing in the disciplines". Language Teaching 41(4): 543-562.

Hyland, K. (2009). "Corpus informed discourse analysis: the case of academic engagement". In M. Charles, S. Hunston & D. Pecorari (eds), *Academic Writing: At the Interface of Corpus and Discourse*". London: Continuum, 110-128.

Hyland K. & Milton J. (1997). "Qualification and certainty in L1 and L2 students' writing". *Journal of Second Language Writing* 6(2): 183-205.

Hyland K. & P. Tse (2007). "Is there an "Academic Vocabulary"?". *TESOL Quarterly* 41(2): 235-253.

Johns A. M. (1995). "Genre and pedagogical purposes". *Journal of Second Language Writing* 4(2): 181-190.

Johns A. M. (2008). "Genre awareness for the novice academic student: An ongoing quest". *Language Teaching* 41(2): 237–252

Jones M.& N. Schmitt Katz M. J. (2010). "Developing materials for discipline-specific vocabulary and phrases in academic seminars". In N. Harwood (ed.), English Language Teaching Materials: Theory and Practice. Cambridge: Cambridge University Press, 225-248.

Katz M. J. (2009). *From Research to Manuscript: A Guide to Scientific Writing*. New York: Springer Science + Business Media B.V.

Leńko-Szymańska A. (2007). "Past progressive or simple past? The acquisition of progressive aspect by Polish advanced learners of English". In E. Hildalgo, L. Quereda and J. Santana (eds), Corpora in the Foreign Language Classroom. Amsterdam/New York: Rodopi, 253-266.

Luzón, M. J. (2000). "Collocational frameworks in medical research papers: a genre-based study". *English for Specific Purposes* 19(1): 63-86.

Martínez, I. 2005. "Native and non-native writers' use of first person pronouns in the different sections of biology research articles in English". *Journal of Second Language Writing* 14: 174–190.

Martínez, I. A., S. C. Beck & C. B. Panza, (2009). "Academic vocabulary in agriculture research articles: A corpus-based study". *English for Specific Purposes* 28(3): 183-198.

Master, P. 2003. "Noun Compounds and Compressed Definitions". *English Teaching Forum* 41(3): 2-9.

Mattheoudakis, M. & A.M. Hatzitheodorou. (in print)."Structures of Argumentation in the Greek Corpus of Learner English (GRICLE)". *Language Forum: An International Journal of Language and Linguistics,* vol. 35.

Neff, J., Dafouz, E., Herrera, H., Martinez, F., Rica, J., Diez, M., et al. (2003). "Contrasting Learner corpora: the use of modal and reporting verbs in the expression of writer stance". In S. Granger & S. Petch-Tyson (eds), Extending the scope of corpus-based research: New applications, new challenges. Amsterdam/New York: Rodopi.

Nesselhauf, N. (2005). *Collocations in a Learner Corpus*. Amsterdam & Philadelphia: Benjamins.

Nesselhauf, N. (2007). "The path from learner corpus analysis to language pedagogy: some neglected issues." In R. Facchinetti (ed.), *Corpus Linguistics 25 Years On*. Amsterdam & New York: Rodopi, 305-315.

Olshtain, E. (1981). "English nominal compounds and the ESL/EFL reader". In M. Hines and W. Rutherford (eds), *On TESOL 1981/82*. Washington, DC: TESOL, 153-167.

Paltridge, B. (2002). "Thesis and dissertation writing: An examination of published advice and actual practice. *English for Specific Purposes* 21(1): 125-143.

Pang, T.T. (2002). "Textual Analysis and Contextual Awareness Building: A Comparison of Two Approaches to Teaching Genre". In A. M. Johns (ed.), Genre in the Classroom: Multiple Perspectives. Mahwah, NJ: Lawrence, 145-161.

Peacock, M. (2002). "Communicative moves in the discussion section of research articles". *System* 30(4): 479-497.

Pho, P. D. (2008) Research article abstracts in applied linguistics and educational technology: a study of linguistic realizations of rhetorical structure and authorial stance. *Discourse Studies* 10: 231-50.

Pho, P.D. (2010) Linguistic realizations of rhetorical structure: a corpus-based study of research article abstracts and introductions in applied linguistics and educational technology. In Gries, S., S. Wulff, M. Davies (eds.). *Corpus Linguistic Applications: Current Studies, New Directions.* Amsterdam/New York: Rodopi V.B., 135-152.

Poos, D. & R. Simpson (2002). "Cross-disciplinary comparisons of hedging". In R. Reppen (ed.), *Using corpora to explore linguistic variation*. Amsterdam: Benjamins, 3-23.

Römer, U. & S. Wulff (2010). "Applying corpus methods to writing research: Explorations of MICUSP". *Journal of Writing Research* 2(2): 99-127.

Rundell, M. & S. Granger (2007). "From corpora to confidence". *English Teaching Professional* 50: 15-18.

Salager, F. (1984). "Compound Nominal Phrases in Scientific-Technical Literature: Proportion and Rationale". In A.K. Pugh and J.M. Ulijn (eds), *Reading for Professional Purposes: Studies and Practices in Native and Foreign Languages*. London: Heinemann, 136-145.

Salager-Meyer, F. (1989). "Principal Component Analysis and Medical English Discourse: investigation into genre analysis". System 17(1): 21-34.

Samraj, B. (2005). "An exploration of a genre set: Research article abstracts and introductions in two disciplines". *English for Specific Purposes* 24(2): 119-238.

Samson C. (2004). "Interaction in written economics lectures: The meta-discursive role of person markers". In K. Aijmer & A.-B. Stenström (eds), *Discourse Patterns in Spoken and Written Corpora*. Amsterdam: Benjamins, 199-216.

Saslow, J.M. & J.F. Mongillo (1985). *English in Context: Reading Comprehension for Science and Technology*. Englewood Cliffs, New Jersey: Prentice Hall.

Simpson, R. & D. Mendis (2003). "A corpus-based study of idioms in academic speech." *TESOL Quarterly* 37(3): 419-441.

Soler, V. (2007). "Writing titles in science: An exploratory study". *English for Specific Purposes* 26(1): 90-102.

Suganthi J. (2009). "Using the revision process to help international students understand the linguistic construction of the academic identity. In M. Charles, S. Hunston & D. Pecorari (eds), *Academic Writing: At the Interface of Corpus and Discourse*". London: Continuum, 272-290.

Swales, J. M. (2002). "Integrated and fragmented worlds: EAP materials and corpus linguistics". In J. Flowerdew (ed.), *Academic Discourse*. London: Longman, 153-167.

Swales, J. M. (2005). "Attended and unattended "this" in academic writing: A long and unfinished story. *ESP Malaysia*, *11*, 1-15.

Swales, J. M. & Lindemann, S. (2002). "Teaching the Literature Review to International Graduate Students". In A. M. Johns (ed.), Genre in the Classroom: Multiple Perspectives. Mahwah, NJ: Lawrence Erlbaum, 105-119.

Sznajder, S. H. (2010). "A corpus-based evaluation of metaphors in a business English textbook". *English for Specific Purposes* 29(1): 30-42.

Thompson, P. (2009). "Shared disciplinary norms and individual traits in the writing of British undergraduates". In M. Gotti (ed) *Commonality and Individuality in Academic Discourse* Bern: Peter Lang, 53-82.

Trimble, L. 1985. *English for Science and Technology: A Discourse Approach*. Cambridge: Cambridge University Press.

Ward, J. (2007). "Collocation and technicality in EAP engineering". Journal of English for Academic Purposes 6(1): 18-35.

WonHo Yoo, I. (2009). "The English definite article: What ESL/EFL grammars say and what corpus findings show". *Journal of English for Academic Purposes* 8(4): 267-278.

Wood, D. (2010). "Lexical Clusters in an EAP Textbook Corpus". In D. Wood (ed.), *Perspectives on Formulaic Language: Acquisition and Communication*. London/New York: Continuum, 88-106.

Williams, R. (1985). "Teaching Vocabulary Recognition Strategies in ESP Reading". The *ESP Journal* 4: 121-131.

Wulff. S, J. Swales & K. Keller (2009). "We have about seven minutes for questions": The discussion sessions from a specialized conference. *English for Specific Purposes* 28(2): 79-92.