NARROWING LEARNERS' LEXICAL COMPETENCE-PERFORMANCE GAP: AN EXPERIMENTAL STUDY ON THE USE OF CORPORA IN THE EFL AND ESP CLASSROOMS[1]

Abstract: This paper presents two different experimental applications of corpora through a common corpus-based approach, based on the authors' experiences respectively in EFL secondary and ESP tertiary settings. The corpus-based approach was not meant to replace other teaching methodologies, but to experiment opportunities of narrowing learners' lexical competence-performance gap through their active involvement in corpora-designed activities.

Keywords: EFL/ESP teaching and learning, lexical competence and performance, corpora, corpus-based pedagogy.

1. Introduction

The theoretical debate on the conceptual dualism between *competence* and *performance* has always oriented ESL approaches in syllabus design and classroom activities. Chomsky's (1965) concern to distinguish competence (a person's knowledge of the rules of a language) from performance (the actual use of language in real situations) has overestimated the value of syntax in language, and determined grammar-oriented language teaching. Hymes (1966) rejected the Chomskyan divide, introducing the notion of *communicative competence* in real-life contexts, which oriented language teaching to the communicative approach. The divide was also rejected by the cognitive-functional school of linguistics, which had an impact on the notional-functional approach in language teaching.

Nowadays, corpus linguistics also runs counter to the Chomskyan view that performance is an imperfect reflection of competence, and challenges the traditional divide between grammar and lexis by investigating samples of authentic texts, or *corpora* of millions of words, which are now stored electronically and analysed computationally. These studies show how actual utterances are highly organized entities, regulated by constraints on word co-occurrence (Stubbs, 2001), which attribute lexis a central role in language teaching (Sinclair, 2004).

This paper is based on the assumption that corpora can be invaluable pedagogic tools in teaching lexis. It presents two different pedagogic applications of corpora through a common corpus-based approach, based on the authors' experience respectively in EFL secondary and ESP tertiary settings. Section 2 of the paper indicates how it may be possible to narrow the gap between lexical competence and performance in the EFL secondary classroom through the use of *general* corpora. It presents a practical example of their use, purposed to enhance both lexical competence and performance in key aspects of the nature of lexis, such as frequency, collocation, and colligation. Along this line, Section 3 positions the corpus-based approach in the ESP tertiary classroom through the use of ready-made specialised corpora. It presents a practical example of their use, targeted to ESP learners' subject-based collocational competence, which is found to widen learners' competence-performance gap if unattended to. Finally, pedagogic implications of the corpus-based approach in these settings are discussed.

[1] Anna Franca Plastina, University of Calabria, Italy, annafranca.plastina@unical.it, Gloria Branca, Liceo Scientifico - IPSIA, Amantea, Italy, gloria_branca@hotmail.com.

2. A Corpus-based Approach in the EFL Secondary Classroom

In general, a corpus-based analysis can investigate almost any language pattern: lexical, structural, lexico-grammatical, discursive, phonological, morphological. Based on their typology, corpora texts are selected following set criteria to capture the regularities of a language, a language variety or a sub-language. Therefore, the use of *general* corpora finds many applications in the EFL secondary classroom. In the following pedagogical experiment, a corpus-based approach was used specifically to develop learners' lexis.

2.1. The Use of General Corpora

Corpora come in many shapes and sizes: they can be monolingual or multilingual, general or specialised, reference corpora with a fixed size (e.g. the British National Corpus), or monitor corpora that are expandable, in that texts are continuously being added (e.g. the Bank of English).

In language teaching, corpus linguistics can find its application in the three main areas of syllabus design, materials development, and classroom activities (Barlow, 2002). According to Leech (1997: 10), corpora can be used in two different ways: *the soft* or *hard version*. The *soft version* simply requires teachers to be skilled in corpus use and in choosing relevant software to print out corpus examples and devise tasks. Corpus examples are usually in the form of a concordance, where the word or structure under study lies in the centre, so that patterns are more easily discernible. The *hard version* requires learners to have direct computer access and corpus facilities, and the necessary skills to use these in two different ways. Teachers can either guide a predetermined investigation, leading to predictable results, or can have students do it on their own with less predictable findings, as they ask questions about language, and draw their own conclusions about language use. Task-based activities can be designed to direct learners to work on single lemma, word pairs, semi-fixed lexical chunks and/or idioms. Activities can be targeted to semantic-based queries, to compare-contrast word pairs (collocates) with results from information dictionaries, notice and find collocation patterns and group them, work on lexical meaning and frequency of occurrence, and on lexical inference, revision and critical examination of grammatical rules, to name a few.

Nevertheless, the effective integration of corpora in EFL teaching appears to be very limited. Corpus linguistics is predominantly research-oriented. For this reason, corpora are often considered too difficult to handle in terms of both the linguistic competence and the technical skills required from secondary-school teachers and learners. A restricted use has thus been made with advanced language learners (CEF B2-C1, Council of Europe, 2001) in other educational settings, such as universities. Technically, corpora interfaces are not appealing to younger learners, compared to the interactive web environment they normally use. Limited access to online and offline corpora and costs of commercial software are other major constraints, which need to be considered. Despite these constraints, a small-scale corpora experimentation was set up in the attempt to narrow the gap between learners' lexical competence and performance through the implementation of corpus-based activities.

2.2. Narrowing the gap: collocational competence for effective lexical performance

Lewis (1997: 8) defines collocation as "the readily observable phenomenon whereby certain words co-occur in natural text with greater than random frequency", and therefore can be valued as an organizing principle of language. The concept itself implies that language is made up of grammaticalised lexis, rather than of lexicalised grammar, and Lewis (1997:204) further explains that "instead of words, we consciously try to think of

collocations, and to present these in expressions. Rather than trying to break things into ever smaller pieces, there is a conscious effort to see things in larger, more holistic, ways". Thus, *multi-word prefabricated chunks* need to be introduced even in the early stages of language learning to facilitate learners' receptive knowledge of words, reduce processing overload in memory retrieval (Ellis, 1997), and to enhance language production. Hill (1999) converges on the issue, claiming that learners' problems with fluency mostly derive from their limited collocational *competence*. Thus, offering learners the opportunity to identify, organise and record collocations as chunks of language appears to be a first step to narrowing the gap.

2.3. Implementation in the EFL classroom: Aims
The experimentation was implemented in an EFL class at a Scientific-Vocational secondary school (*Liceo Scientifico-IPSIA*) in Amantea, Italy. The pedagogical use of general corpora aimed at:
- enhancing learners' lexical competence for more effective lexical performance through three key aspects of lexis: *lexical frequency of occurrence, collocation* and *colligation*;
- guiding EFL learners to identify, organise and record collocations as chunks of language as a first step to narrowing the competence-performance gap
- offering learners the opportunity to engage in the direct use of General Corpora, which are available online
- training learners in simple language investigation through the use of technical tools.

2.3.1. Subjects
Twenty ESL secondary school learners (9 males and 11 females) were involved in the pedagogical experiment of learning English through the application of corpora. Subjects' average age was 16 years old. Overall, they had a CEF A2 language proficiency level, and were highly computer-literate. All subjects were informed about the pedagogical experiment and showed willingness to engage in learning with new technologies.

2.3.2. Procedure
In the preliminary phase, general corpora were chosen, based on their functionality of enabling the investigation of three key aspects of lexis: *lexical frequency of occurrence, collocation and colligation.* Corpora functions were identified for student use with the following objectives:
a. *wordlist* or *keywords* to identify frequency of features in text types and registers, and to understand how content and function words are discriminated;
b. *KWIC* (*Key Word In Context*) *Concordance*, a sort of search engine for language analysis, to enable students to find collocates of a lemma or word form;
c. *POS (part of speech)* functions referring to word class annotation, to allow students to identify word collocations sorted by word classes, and to inquire into colligation, which highlights recurring combinations of lexis and grammar, and typical word patterning in word distribution across grammatical constructions.
In the following phase, learners were guided to use some Web 2.0 tools and were provided with knowledge of basic technical skills to simulate a few simple corpora queries. This preliminary activity was also targeted to make corpora more appealing to students, besides equipping them with basic technical skills to understand and manage processes in observing and investigating language items. Cooperative data-driven learning then took place through pair and group-work. Students were initially given some authentic written

texts from gamers' websites, describing very popular online games, such as *World of Warcraft, Warrock, Farmville* and *Pet Society*, with which they were very familiar. This factor supported their motivational involvement. After some traditional comprehension activities, students were asked to produce short oral texts on their favourite online games, and to report to the class. All students found difficulties in producing fluent utterances in the traditional manner, in which they tried to retrieve single words as separated items to build single sentences. For this reason, they were introduced into four main Web 2.0 digital learning environments: 1. *wordle*; 2. *wordsift*; 3. *wordnik*; 4. *facebook,* where they were required to carry out a number of reflective language activities. These included using wordle and wordsift in order to learn to: a. create and manipulate word clouds; b. visually observe word frequency; c. make hypotheses about frequency of content and function words; d. search online dictionaries; e. create semantic fields and online word lists. Learners then used facebook to share their work, and to report their findings.

In the next phase, two high-frequency lemmas related to the previous topic, i.e., *game and match,* were identified. Students were then guided to access two online English Corpora: *British National Corpus (BNC)* and *Collins Wordbanks Online*[1] to search lemma frequency in relation to text types, concordances and collocates through *POS* queries. In BNC, students were introduced to key query functions displayed on the main interface, and asked to make hypotheses about the frequency of *game* and *match* in text types. Shifting from a visual approach, they were then required to use *Chart*, *the POS list verb* function, and to create concordances and notice collocates. The function *compare* eventually enabled students to observe different common and exclusive collocates of *game* and *match*. After, students were asked to answer some questions, e.g. *Why are crying / birds exclusive collocates of game?* To do so, they sourced their corpora findings, and found support in online dictionaries and simple web-search engines.

This corpus-based activity stimulated: a. observation of authentic language data; b. the development of hypothesis-making skills; c. awareness of the importance of context to disambiguate polysemic lemmas, also through the use of online dictionaries; d. inquiry into collocates (e.g. *crying-game*) through search engines; e. discovery learning (e.g. *crying-game* collocate in the context of the film *The Crying Game* by Neil Jordan).

Furthermore, students developed colligation skills in WORDBANKS through the following corpora-based learning activities: a. group *game* POS collocates - adjectives left and right; b. wordsketch *player* collocates -noun modifier; c. *game* wordsketch- verbs left, verbs right and noun modifier.

3. A Corpus-based Approach in the ESP Tertiary Classroom

A corpus-based approach has shown to offer EFL secondary learners the opportunity to access *general* corpora and to infer generalizations about the language as a whole, or about broad categories of texts (L.Flowerdew, 2004). Hunston, 2002 claims that ESL learners widen and enrich their view of lexical items more through corpora than vocabulary exercises.

At the tertiary level, where learners choose different academic studies in specific subject areas, they need to be guided to perceive a relationship of these studies with language use (Donna, 2000). At this level, the development of lexical competence in learners' domain of use is paramount for professional performance. In a survey related to

---

[1] British National Corpus: http://www.natcorp.ox.ac.uk/.
 Collins WordBanks: http://www.collinslanguage.com/wordbanks/default.aspx.

English for Specific Purposes (ESP), Curado & Rokowski (2002) found that learners perceived semi-technical and specific lexis to be more important than grammar. However, different factors seem to hinder the process of acquiring ESP terminology in tertiary settings. First and foremost, the development of ESP lexis requires the production of sample language which *seems real* (cfr. Sinclair, 1997). ESP teachers, however, often tend to encourage learners to acquire simplified tokens, which are not highly relevant, nor appropriately contextualized in their subject field (McEnery & Wilson, 1996). This may be due to the fact that most instructors are non-native speakers, and do not master the specific domain of knowledge. In this respect, a corpus-based approach converges with the ESP need to study specific terminology (cfr. Pedersen, 1995), overcoming these drawbacks. The integration of *specialised* corpora in ESP classrooms is considered a coherent pedagogic strategy (cfr. L. Flowerdew, 2004, 2005; Gavioli, 2005). Nowadays, although online corpora are available for reference, the number of general corpora outweighs that of specialised ones, so that gaining access to the latter may be more difficult. Nevertheless, it is crucial to use specialised corpora in ESP pedagogy, as these include representative oral and/or written texts, which reflect the kind of language of a specific domain. It could be argued that general corpora may also include specialised sub-corpora. However, these are difficult to access as search fields are not set up for them (Lee, 2001).

At long, there has been some resistance in introducing corpus-driven learning directly in universities (cfr. Gavioli, *op. cit.*). Furthermore, corpus studies have neglected the pedagogic use of corpora, as point out by L. Flowerdew (1998:550): "the implications for pedagogy are not developed in any great detail with the consequence that the findings have had little influence on ESP syllabus and materials design". This is consistent with Blancafort and Gornostay's (2010) online survey findings on terminology and corpora practices, which reveal that teachers spent 10-30 % of their time working with terminology, and that 33.2% used only dictionaries and glossaries as linguistic resources. This reluctance to adopt a corpus-based approach to teaching ESP lexis can be understood to relate directly to lack of competence in the pedagogic use of corpora.

For this reason, the present pedagogical experimentation addresses the need to develop learners' *collocational competence* through the use of specialised corpora, which may offer some insights into a corpus-based approach to ESP.

3.1. The Use of Specialized Corpora: Adopting Criteria for Corpora Choices
In the Web 2.0 age, language teachers can easily access and consult ready-made corpora to develop *awareness* and *understanding* of how to adopt a corpus-based approach, and corpus-based *skills* to use these valuable resources for various pedagogic purposes. At more competent levels, teachers can compile their own customized teacher-made corpora. The first criteria to bear in mind is the level of corpus-based competence, which entails *awareness* of the pedagogic purpose of use, *knowledge* in choosing or compiling appropriate corpora, and methodological and digital skills of use.

In the present experimentation, ready-made corpora were used for the purpose of introducing ESP learners to the nuts and bolts of the corpus-based approach. Furthermore, the decision to use different text-type corpora was made to allow learners to "compare data from different sources, and to discuss language use in relation to different types of text, topic, and genre" (Gavioli & Aston, 2001: 245). Choosing appropriate ready-made corpora requires skills in selecting, evaluating and using appropriate electronic tools, which is understood to be a key competence (Chambers & O'Sullivan, 2004).

Three main criteria were introduced following Aston (1996):

173

1. *corpus specificity:* Although corpus specificity is the key feature in ESP, highly specific ready-made corpora are unlikely to be available online. Conversely, corpora of a variety of specialisations and genres are more likely to be retrievable;

2. *corpus size:* Specificity is also a drawback for corpus size, which relies on the principle of *the larger the better* (100 million or more words). In ESP teaching, smaller corpora may have the advantage of being far more specialised in terms of topics and genres than general larger ones (J. Flowerdew, 1996);

3. *corpus representativeness:* specificity and size limitations tend to decrease corpus representativeness, or the extent to which a corpus aims to capture the most significant discursive tokens.

3.2. Collocational Competence in ESP

Collocational competence in an ESP context refers to learners' knowledge and skills to recognize and use words which co-occur frequently within a specific field. This competence refers to restricted contexts of use, such as professional textbooks, research articles, technical reports, specialised newspapers. For this reason, collocations are *subject-based* as opposed to *general-purpose* ones, and have two peculiar features:

1. they are mostly jargons, which may be incomprehensible to the wider audience;
2. they need longer context of use to be clearly understood.

Lewis (1993: 60) summarises their value, claiming that teachers of ESP in particular should notify the importance of collocations to enable non-native English language users to develop the awareness of the lexical nature of language and thereby to recognize and eventually produce sound language "chunks" themselves.

In the ESP classroom, collocational competence can be enhanced either through a *bottom-up* or a *top-down* corpus approach. In the former, concordances are used for text processing to generate *corpus-driven* bottom-up data out of the original subject-based context; in the latter, applications are used to find *corpus-based* top-down discursive information in the original subject-based context. L. Flowerdew (2009) claims that the top-down approach is more appropriate for ESP collocational competence, as subject-based collocations convey *representativeness* of their context of use. In this respect, a corpus basically needs to show ESP learners the representative features of specific discourse, lexis and subject terminology. A concrete method of identifying representative subject-based collocations as meaningful units is through the use of raw frequency corpus-based data (Danielsson, 2007). These support the design of activities, which draw learner attention to those aspects that may be taken as paradigmatic of a discourse that is only a niche of the language (Swales, 1990).

3.3. Implementation in the ESP classroom: Aims and Pedagogical Tools

The pedagogical experiment was conducted at the University of Calabria in Italy. The main aims were to:

- help learners gain understanding of two main relationships between subject contexts and terminology:
1. context and collocation with focus on *noun + noun* subject-based collocations*;*
2. context and semantic prosody with focus on subject-based collocation in semantic sets in specific genres (Tribble, 2000);
- develop learner awareness of different ESP text types in order to:

174

"compare data from different sources, and to discuss language use in relation to different types of text, topic, and genre" (Gavioli & Aston, 2001: 245).

The choice of these objectives was grounded in Nelson's (2006) findings of Business English *noun + noun* keywords, which occur significantly more often in business contexts than in General English ones. Furthermore, two specialized corpora were introduced as pedagogical tools: PolyU Business Corpus and the British Academic Written English Corpus (BAWE)[1]. These were sourced online specifically for the genres of Business newspaper articles and academic economic texts. These were the object of learning as they were usually used in the traditional ESP classroom.

3.3.1. Subjects

Ten undergraduates (6 females and 4 males), regularly enrolled in Business Administration, and attending a course of Business English, were recruited at the University of Calabria. Subjects' *age ranged from 19 to 21 years old (M=20)*, and had a CEF B1 language proficiency level. Subjects claimed that they had had no previous experience of using corpora, but were already familiar with the genres and text types introduced in the experiment. All subjects volunteered to take part in the pedagogical experiment.

3.3.2. Procedure

Corpora-designed activities were carried out in three ESP pedagogical stages:
I. Micro-structural stage

In this stage, focus was placed on the development of learners' ESP lexical competence. For this purpose, the activity was based on extracting & classifying lexical information according to the following steps:
a. two lexical items *market* and *sales* were investigated for *neighbourhood* collocations (Scott, 1999) of one word before and after (cfr. Hoey, 2005) of the type *noun + noun* in both corpora;
b. learners searched, noticed and categorised the ten most frequent right and left noun collocates of the two given lemmas. This was accomplished by using Sketch Engine functions (http://the.sketchengine.co.uk): *n-modifier* and *modifies*;
c. learners compared and contrasted the use of the two lexical items across the two corpora for lexical variation due to the contextual nature (genre) of the corpus.

In detail, the most frequent n-modifiers found for *market* were *stock, world, job, labour, values, property*, while market +noun right generated *share, value, economy, price, research, condition*. In the case of *sales*, which had no n-modifiers and only nouns left including *manager, director, tax, growth, revenue*, learners were asked to reflect on these findings. They were then induced to compare and contrast the use of the two lexical items across the two corpora for lexical variation due to the contextual nature (genre) of the corpus. In this micro-structural stage, *extracting* and *classifying* lexical information was a preliminary step to acquiring corpus-based lexical knowledge.
II. Macro-structural stage:

In this stage, attention was drawn to learners' lexical performance. For this reason, communicative practice was implemented in the following steps:

---

[1] British Academic Written English Corpus *accessible via* http://the.sketchengine.co.uk ;
PolyU   Business   Corpus:
http://langbank.engl.polyu.edu.hk/engine.aspx?Submit=Search&lang=1&corpus=30

a. learners prepared topic-based *oral reports*, in which they applied the corpora knowledge, acquired in stage 1;
b. learners performed the communicative task, which was supported by the different levels of competence of contextual relationships acquired.
III. Macro-contextual stage

In this stage, the objective was to support learners in bridging corpora knowledge & skills in order to narrow the gap between their lexical competence and performance. This was done according to the following steps:
a. improving coherence and fluency in ESP performance through learners' engagement in activities dealing with semantic prosody and semantic categorisation of keywords;
b. expanding on the micro- and macro structural stages by engaging learners in more subject-oriented tasks in different contexts. For example, learners were required to *conduct a market analysis* by sourcing descriptive texts of corporate companies. In this activity, learners were stimulated to identify *noun + noun* subject-based collocations in a different genre.

In detail, learners first expanded the collocation span to the phrase to find adjective and verb collocates of *market* and *sales*. After, they categorised their findings - right and left - based on the semantic sets: *market of what, where it takes place, macro-level demarcation, positive and negative attributes* as in Nelson (2000). In the following macro-structural stage, learners were challenged by their previous levels of competence of contextual relationships, when engaged in their oral reports. In the last stage, corpora-based activities on semantic prosody, semantic categorisation of keywords supported learners in building a stronger language framework for better ESP performance. Moreover, by expanding on the initial corpora-based activities, learners were working on their collocational competence.

4. Pedagogic Implications for ESL and ESP Teaching and Learning

The theoretical debate on corpus linguistics has led to controversial views on applications of corpora findings. Widdowson (1991) argued that corpus linguistics had no relevance to pedagogy, as these studies are based on the *use* of language, whereas pedagogy deals with the *usefulness* of language. On the other hand, corpus linguists advocate that only corpora language should be introduced in the language classroom. The experiments described in this paper have attempted to overcome this controversy through the design of learner-centred corpus-based activities in their own different ways.

In detail, corpora-based activities can be a starting-point in the EFL classroom for enhancing the development of elementary learners' lexical competence in terms of frequency, collocation and colligation patterns. The learner-centred approach used in the experiment encourages *discovery learning* through independent practice, in which *use* and *usefulness* of language can be integrated to gradually improve lexical performance. Corpus techniques can offer a pedagogic framework, which supports content and language integration through corpus-based tasks.

In the ESP classroom, the experiment focused on corpora searches of subject-based collocations. Then, basic corpora activities on semantic prosody in Business English were introduced to trigger knowledge-based processes, which fostered better content-based communicative performance. However, inexperienced teachers need to receive training in the use specialized corpora for an effective integration of corpora in the ESP classroom. In particular, training should focus on:

176

- equipping teachers with the necessary *knowledge* of how and why specialized corpora should be accessed for pedagogic purposes;
- developing their *skills* in designing corpus-based activities on ESP lexis and content;
- developing their *awareness* of the benefits of corpus-based *discovery learning* processes.

Overall, the present experimental study has shown how corpora can be introduced both at secondary and tertiary English language classrooms. It has also described sample corpus-based activities with the aim of demonstrating that corpora can be used at different educational levels, and for different purposes. The common goal in the present case was to induce the activation of both EFL and ESP learners' cognitive processes in order to facilitate the acquisition of lexical competence leading to better general language performance in the first experiment, and to specialized language performance in the second experiment. These experiments, however, should be considered as basic applications of corpora in English language teaching and learning. Therefore, a lot of ground still needs to be covered both in terms of teacher training on the pedagogical value of corpus linguistics and on the proper integration of suitable corpora in teachers' pedagogical settings. In the meantime, it is worth stressing how corpora have the potential of contributing greatly to language learning processes. An advantage can also be seen in the fact that corpora tend to enrich other teaching methodologies, and therefore do not represent a threat to teachers' usual teaching praxis, nor an attempt to replace their methodologies.

Bibliography
Aston, G., 1996, "What corpora for ESP?", A Conference on ESP at the University of Pavia, November 1996, http://www.sslmit.unibo.it/~guy/pavesi.htm, consulted on October 12th, 2010.
Barlow, M., 2002, "Corpora, concordancing, and language teaching", *Proceedings of the 2002 KAMALL International Conference*, Daejon, Korea.
Blancafort, H., & Gornostay, T., 2010, "Calling Professionals: Help Us to Understand Your Needs!" TTC: Terminology Extraction, Translation Tools, and Comparable Corpora. The Seventh Framework Programme (Grant Agreement no 248005).
Chambers, A., & O'Sullivan, I., 2004, "Corpus consultation and advanced learners' writing skills in French", *ReCALL,* 16 (1), 158–172.
Chomsky, N., 1965, *Aspects of the theory of syntax*, Cambridge, MIT Press.
Council of Europe, 2001, *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, Cambridge, Cambridge University Press.
Curado, A., & Rokowski, P., 2002, "Using Corpus Resources as Complementary Task Material in ESP", *ICAME JOURNAL,* 26, 5-32.
Danielsson, P., 2007, "What constitutes a unit of analysis in language?", *Linguistik online,* 31 (2) 17–24.
Donna, S., 2002, *Teach Business English,* Cambridge, Cambridge University Press.
Ellis, R., 1997, *Second language Acquisition,* Oxford, Oxford University Press
Flowerdew, J., 1996, "Concordancing in language learning", in M. Pennington (ed.), *The Power of CALL*, Houston, Athelstan, pp. 97-113.
Flowerdew, L., 1998, "Corpus-linguistic techniques applied to text linguistics", *System,* 26, 541-552.
Flowerdew, L., 2004, "The argument for using English specialized corpora to understand academic and professional language", in U. Connor & T. A. Upton (eds.), *Discourse in the Professions, Perspectives from Corpus Linguistics,* Amsterdam/Philadelphia, John Benjamins, pp. 11-33.
Flowerdew, L., 2005, "An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP: countering criticisms against corpus-based methodologies", *English for Specific Purposes*, 24, 321-332.

Flowerdew, L., 2009, "Applying corpus linguistics to pedagogy", *International Journal of Corpus Linguistics,* 14:3, 393–417.

Gavioli, L., 2005, *Exploring Corpora for ESP Learning*, Amsterdam, John Benjamins.

Gavioli, L., & Aston, G., 2001, "Enriching reality: language corpora in language pedagogy", *ELT Journal*, 55 (3), 238-246.

Hill, J., 1999, "Collocational competence", *English Teaching Professional,* 11, 3-6.

Hoey, M., 2005, *Lexical priming: a new theory of words and language,* New York, Routledge.

Hunston, S., 2002, *Corpora in Applied Linguistics.* Cambridge, Cambridge University Press.

Hymes, D.H., 1966, "Two types of linguistic relativity", in W. Bright (ed.), *Sociolinguistics.* The Hague, Mouton, pp. 114-158.

Lee, D., 2001, "Genres, registers, text types, domains and styles: clarifying the concepts and navigating a path through the BNC jungle", *Language Learning and Technology,* 5 (3), 37-72.

Leech, G., 1997, "Teaching and language corpora: a convergence", in A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles. (eds.), *Teaching and Language Corpora*, Harlow, Longman, pp. 1-23.

Lewis, M., 1993, *Implementing the Lexical Approach. Putting theory into practice*. Hove, LTP.

McEnery, T., & Wilson, A., 1996, *Corpus Linguistics.* Edinburgh, Edinburgh University Press.

Nelson, M., 2000, "A Corpus-based Study of the Lexis of Business English and Business English Teaching Materials", Unpublished Thesis, Manchester, University of Manchester.

Nelson, M., 2006, "Semantic associations in Business English: A corpus-based analysis", *English for Specific Purposes*, *25*, 217-234.

Pedersen, J., 1995, "The identification and selection of collocations in technical dictionaries", *Lexicographia, 11,* 60-73.

Scott, M., 1999, *WordSmith Tools 3.0.,* Oxford, Oxford University Press.

Sinclair, J., 1997, "Corpus evidence in language description", in A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles. (eds.), *Teaching and Language Corpora*. Harlow, Longman, 27-39.

Sinclair, J., 2004, *How to use corpora in language teaching,* Amsterdam, Benjamins.

Swales, J., 1990, *Genre Analysis*, Cambridge, Cambridge University Press.

Stubbs, M., 2001, *Words and phrases: Corpus studies of lexical semantics*, Oxford, Blackwell.

Tribble, C., 2000, "Genres, keywords, teaching: Towards a pedagogic account of the language of project proposals", in L. Burnard & T. McEnery (eds.), *Rethinking language pedagogy from a corpus perspective: Papers from the Third International Conference on Teaching and Language Corpora.* New York, Peter Lang, pp. 74-90.

Widdowson, H.G., 1991, "The description and prescription of language", in J. Alatis (ed.), *Georgetown University Round Table on Languages and Linguistics 1991.* Washington, D.C., Georgetown UP, pp. 11-24.