What can linguistic approaches bring to English for Specific Purposes?

*Quel est l'apport des approches linguistiques à l'anglais de spécialité?*

Christopher Gledhill and Natalie Kübler
What can linguistic approaches bring to English for Specific Purposes?

Quel est l’apport des approches linguistiques à l’anglais de spécialité?

Christopher Gledhill and Natalie Kübler

1. Introduction

What is a ‘linguistic approach’ in the field of English for Specific Purposes (ESP)? And what can contemporary linguistic approaches contribute to the theory and practice of ESP? A linguistic approach to ESP attempts to answer the following broad questions: (1) What qualifies as ESP, and what textual forms can it take? (2) What are most relevant language skills needed to learn ESP? (3) What features of language do we typically find in ESP?

1.1. What qualifies as ESP, and what textual forms can it take?

ESP is the study of a specific configuration of the English language that is used by a group of specialists for that group's particular purposes (whatever their language origins, whatever their degree of expertise). For each type of specific English, it is important to specify which one we are discussing (for example ‘English for Academic Purposes’, ‘English for Business Purposes’ and so on). For many linguists, the study of ESP involves the study of a conceptual domain (economics, the law, medicine, etc.). In this paper, we contrast this ‘conceptual’ approach with our own ‘contextual’ point of view, which assumes that each type of ESP can be most usefully apprehended in context, i.e. in the form of texts (whether spoken, written or in a hybrid form). For example in English for Academic Purposes, the core linguistic activities (genres) are typically represented by a whole ‘ecology’ of text types, such as university-level science lectures, textbooks, PhD theses, conference presentations, research articles, and so on. Most linguists interested in
this question (conversation analysts, ethnographers of communication, genre analysts, etc.) are also interested in the social and pragmatic factors which lead to the development of an ESP among a particular speech or discourse community, and they turn to the experts or practitioners in order to decide what is a core or a peripheral text type in their particular domain.

1.2. What are most relevant language skills needed to learn ESP?

This question can be framed either in terms of communication skills (oral/ written comprehension, oral/written production, etc.), or in terms of language forms and systems (terminology, lexis, grammar, rhetorical structure, etc.). This question necessarily leads to a debate about methodology, materials, testing (see for example Bloor & Bloor 1986) or ‘needs analysis’ (the design of a language syllabus on the basis of learners’ or practitioners’ needs, as discussed in Basturkmen & Elder 2006). In this paper, we leave these issues to the ‘didacticians’ (as they are called in the French ESP context). In the following sections, we limit our discussion to the forms and systems of language, and we argue that it is necessary to adopt a systematic framework when describing the architecture of language (such as the Systemic Functional model). In addition, a key tenet of the context-oriented approach is that both students and teachers of ESP need direct experience of working not only with authentic texts as exemplars of a particular type of discourse, but also with large ‘bodies’ of interrelated texts (hence the term ‘corpus’).

1.3. What features of language do we typically find in ESP?

Again, this question can only effectively be answered in relation to a particular type of ESP. For example, in written scientific discourse, linguists prototypically find: specialised technical vocabulary, phraseological paradigms such as ‘cells express genes, genes express proteins’, a preference for grammatical structures such as the passive, conventionalised rhetorical moves such as ‘create a research space’, and so on. This type of description looks straightforward, for what else could one find in a typical ESP text apart from specialised words (terminology) and specialised combinations of words (phraseology)?

However, here, we suggest that the advent of corpus-based analysis has changed the ways in which we identify and categorise different ESPs (as discussed for example in Biber et al. 2010; Groom et al. 2015; Hoffmann et al. 2015). Following Rundell & Stock (1992), Granger (1994), Hanks (2012) and others, we find it appropriate to talk about a ‘corpus revolution’, especially in relation to many areas of applied and descriptive linguistics. This paradigm has changed the names of the features we can observe in relation to their naturally-occurring environment. Thus for example, whereas a discourse analyst or grammarian might analyse a passive clause such as “The gene was expressed in middle ear mucosa” (cited in Gledhill 2011b) as a ‘thematicisation’ of an affected object gene in initial position in the clause, the corpus analyst also sees a ‘collocation’ involving a co-dependence between the lexical items gene + express as well as a ‘colligation’ or an ‘extended lexical pattern’ (involving the co-selection of the passive and a prepositional expression of location). Furthermore, the widespread use of corpora has also changed the way we discuss the status and typicality of these features in relation to the language system: it is no longer enough to look at one particular occurrence in one particular ESP text – it has become now possible to compare similar instances across the larger ‘ESP corpus’ and then an even
larger ‘reference corpus’ (see, for example, large-scale comparative cross-disciplinary studies such as Biber 2010; Biber et al. 2010; Grey 2015).

The first two questions mentioned above (What qualifies as ESP? What are the most relevant skills in ESP?) are crucial to the theory and practice of ESP. However most linguists would agree that it is the third issue, with its focus on the forms of language, which lies at the heart of the linguistic approach to ESP. In the following pages, the authors of this paper set out our own personal view of ESP, and we admit that this is a very personal, subjective perspective. It is also important to bear in mind that we are both applied linguists and teachers of specialised translation at a French university. We therefore appreciate the various issues involved in, for example, ESP syllabus design. However, we also believe that by providing systematic corpus-based descriptions of ESP texts, and by training our students to do the same, we are not only helping them to learn a specialised variety of English, we are also introducing them to an important transferable skill-set which will serve them in their careers as language specialists, namely: the ability to do research in descriptive linguistics.

It follows from what we have just said that, in this paper, we are not attempting to provide a comprehensive survey of all of the different linguistic approaches that have so far made significant contributions to the field of ESP. We have however instead conducted a survey (presented in Appendices 1-3) of the main linguistic topics and linguistic approaches that have been addressed in the present journal (ASp – for Anglais de spécialité) since 1993. While clearly not a complete overview of the field, these data do provide some interesting indications about where linguistic approaches stand in relation to ESP as an academic field. We comment briefly on the main themes that emerge from the Appendices later on in this paper (section 4).

However, before looking at these data, in section 2 we present what we believe to be the two main linguistic schools of thought which have had a major influence on to ESP studies: the concept-oriented approach and the context-oriented approach. Both represent very different schools of thought in linguistics. However, more recently both traditions have begun to converge in a number of interesting ways, most notably through the widespread use of corpora and the rise of computational methods in formal linguistics.

In the final parts of this paper, we examine how ESP may be shaped in the future by new concepts arising in linguistics. We cannot claim to represent the majority of linguists; all we can hope to do is to set out our own perspective on what may be fruitful lines of enquiry for future research in ESP. In the first instance (section 3.1), we make the case for the analysis of lexico-grammatical patterns in ESP texts. Then (section 3.2), we examine how teaching students of specialised translation to become corpus analysts can help them to become familiar with the technical terminology and phraseology of their chosen specialist domains.

2. What are the main linguistics approaches to ESP?

There are two basic approaches to the study of ESP (also known more generally as ‘Languages for Specialised purposes’, LSP), namely: (1) the concept-oriented approach, and (2) the context-oriented approach.
This distinction is based on:

a) The differences in theoretical orientation and methodology which can usually can be
observed in any research paper published in the field of ESP;

b) Our own observation of linguistics papers published in this journal (ASp), and set out in
the Appendices and summarised in section 4 of this paper.

In the following discussion, it is important to realise that both the concept-oriented and
context-oriented approaches can sometimes be found within the same study. However, it
occurs to us that it is still useful to make a basic distinction between at least two main
perspectives on ESP which also happen to coincide with a distinction often made between
structural or formal linguistics on the one hand and discourse-analytic or functional
linguistics on the other.

2.1. The concept-oriented approach

A concept-oriented approach to ESP/LSP focuses on the conceptual structures of
specialist domains. We mention three typical examples of this kind of study below
in the technical words, notions, definitions, symbols, semantic networks and ‘spaces of
knowledge’ with which specialists and other practitioners represent and construe their
specialist domain. We would suggest that there are at least three main subtypes of
concept-oriented study in ESP. Below we give some prototypical examples of these
subtypes, as represented by studies published in the journal ASp.

a) **Metaphor analysis.** A typical example of this is English (1998), who adopts the terms
and methodology of metaphor theory (including notions such as assumptive frameworks,
metaphorical description and explanation, blending, etc.) to examine the processes by
which experts and learners use metaphors to name and understand new concepts in
mathematics.

b) **Domain analysis.** For example, Resche (2013) underlines the importance of
disciplinary labels and definitions in the conceptual construction and historical
development of economics (including early terms such as ‘social philosophy’ and sub-
branches such as regeneration economics, gender economics, etc.). Resche is careful to point
out the extent to which domain labels can involve heterodox as well as conventional
status; she also points out that a dynamic domain undergoes the multiple influences of
other disciplines, as evidenced by blends such as ‘neurofinance’ and ‘econophysics’.

c) **Ontology analysis.** For example, Peraldi (2012) uses large-scale corpus analysis as well
as the traditional techniques of componential analysis to explore the conceptual domain
of organic chemistry: she points in particular to the fact that the syntax of multiply-
modified nominal groups can be used to build up a complex ontology in this domain.

It has to be recognised that here are many different perspectives in the concept-oriented
approach, as demonstrated by the different studies mentioned above. However, we would
suggest that many of these studies share some core assumptions. In particular, we would
point to a position paper by Pierre Lerat (1997), “Linguistic approaches to specialised
languages”. In the following paragraphs, we describe the main positions taken by Lerat
(1997), and we make the case for seeing some of his assumptions as fairly typical of an
approach which places a specialised discipline and the semantic networks of a discipline
(in terms of terminology) as central components of ESP.
The starting point of Lerat’s paper is to argue that specialised languages (he uses the abbreviation LSP) are “nothing more than specialised uses of natural languages” (1997: 1 [our translation]). He later adds that “Paradoxically, specialised languages are relatively simple, because they are subject to a high degree of lexical, semantic and pragmatic constraints” (1997: 5). Clearly this is a point of contention, and we will return to this later.

For Lerat, the main differences between LSP and LGP (language for general purposes) can be found in what he calls ‘lexico-semantic’ and ‘lexico-syntactic’ features. As far as semantic specificity goes, he has three features in mind: a) monosemy (LSPs reduce the meanings of polysemic words to single, specialised meanings), b) denomination (the most typical features of LSPs are nominalised expressions; significantly Lerat states that these “can only be validated as denominations by professionals” [Lerat 1997: 2]), and c) distortion of the link between signifier and signified, or as Lerat puts it: “the link between the technical term and its root tends to be stretched” (1997: 2). Turning to syntactic specificities, Lerat points briefly to a general tendency for LSPs towards a more abstract, more impersonal, and more explicit style. It could be argued that these are essentially semantic rather than formal features of syntax. However, Lerat does not elaborate on this; rather at this point in his article, he turns more specifically to the analysis of language above the level of the sentence:

At a higher level of analysis, that of the text, the linguistic tools are still very limited: ‘text-grammar’ has obtained few results – apart from work on anaphora, the typology of speech acts still lacks rigour, and the notion of cognitive environment is necessary but lacks technical accuracy for the moment; currently this has more to do with logic than with syntax proper. (Lerat 1997: 3)

As we discuss below, this point would be contested by many linguists. Nevertheless, according to Lerat, it is at the lower levels of linguistic analysis, that of the group or word, that theoretical linguistics has most to contribute to the description of LSPs (1997: 4). He cites in particular G. Gross’s (1996) notion of the “class of objects”; that is to say the set of arguments which are typically used with each particular predicate in the LSP, thus revealing the typical semantic relations of a specialised domain (Lerat gives examples from the medical domain such as to catch + contagious disease vs. to contract + non-contagious disease, etc.). Thus linguists who conduct this kind of analysis can contribute to the study of LSP, since they can “provide a better conception of the whole by lexicalising syntax and establishing the typical conceptual relations between words [...]”. By ‘lexicalising syntax’, Lerat means ‘establishing the distinct terminological units of the LSP’ [our reformulation], thus enabling the linguist to study the conceptual relations of the given domain. According to Lerat, by examining these relations linguists can contribute practically to LSP by “creating the best kind of database, that is, one which can combine linguistic as well as encyclopedic knowledge” (Lerat 1997: 1).

Finally, it is worth mentioning Lerat’s views on the relationship between linguistic research and LSP as a field, especially in relation to translation. Lerat mentions that the principal aim of LSP is “to assist translation, technical writing, documentation and learning” (1997: 1). It occurs to us that this corresponds to a very negative view of applied linguistics: essentially, Lerat considers that these applied activities can benefit from linguistic research, but he does not mention the possibility that research in these areas might be of any interest to theoretical linguistics – in this respect, Lerat is echoing Corder, who once stated that “the applied linguist is a consumer or user, not a producer of theories” (1973: 10). Later on in the paper, Lerat states that: “because specialised
translation supposes a good command of both languages, most of the difficulties to be encountered will be lexical." (1997: 6). It is because of this comment that we characterise Lerat’s approach as essentially 'conceptual': i.e. when Lerat uses the word 'lexical', he is not talking about the behaviour of a lexical item as used in a given text or context, he is referring to its semantics. Similarly, for the conceptual linguist, the term ‘lexis’ refers primarily to a structured system of concepts, i.e. an ontology or taxonomy (as discussed, for example in Thelen 2002).

Let us take Lerat’s comments about translation as a starting point for some critical discussion. It is true that for many professional translators, one of the central problems they encounter is the search for conceptual and terminological equivalence between source and target text. This is certainly the assumption behind university courses which train translators to use terminological databases (including our own). However, we would suggest that for translators who are experts in their own domain, or translators who are not writing in their mother tongue (despite having ‘a good command of both languages’), terminology may not be the only problem, or even the main one; we would suggest that these translators need help with phraseology. This is particularly the case of our own students of translation – mostly native French speakers – who are asked to translate technical texts into English (on this point see Gleddhill 2011a; Volanschi & Kübler 2011, and section 3.2 below). And of course for all translators and revisers, there is the issue of clarity and the desirability of producing reliable, readable copy. Thus in our view, Lerat’s comment about the centrality of lexical (i.e. semantic) problems appears to be rather limited. More generally, such an approach assumes that meaning is entirely dependent on the abstract network of concepts, with little or no consideration for the phraseological constructions which make up a text. Yet at the same time, Lerat’s comment does raise an interesting research question, and one which may be tested empirically: do lexical or terminological problems really represent a core difficulty for technical writers or translators?

We now return to the assumption underlying Lerat’s paper that ESP/LSP can somehow be discussed as a homogenous notion, or that it is possible to identify and characterise certain features of language as (purely) LGP. Many readers of ASp will be familiar with the debate about the all-embracing or overlapping status of LGP in relation to LSP, but it seems appropriate here to review the issue again in the light of our distinction between concept-oriented and context-oriented approaches. For Lerat and many other linguists, it seems to be important to make a distinction between a LGP on the one hand and LSP. However, these linguists are also often careful to point out that any one type of LSP corresponds to a selection of potential resources in relation to the LGP, as Resche (1999) puts it:

The idea that “specialised language” should be closely linked to general language is now commonly accepted. Thus this expression is preferred today in order to emphasise the idea of a continuum, not that of a break [...]. (Resche 1999:121 [our translation])

Yet the idea of a continuum between LSP and LGP raises the question of whether it is possible to identify any elements of language which might characterise the LGP. One possible position would be to claim that, as Bloor & Bloor once put it

[...] the Common Core does not exist as a language variety. It is as much an idealization as Chomsky’s “ideal speaker-listener in a completely homogenous speech-community”. (Bloor & Bloor 1986: 19)
For these linguists (see also Maingueneau 2002), the notion of LSP as variety can be replaced by a more porous notion of ‘mode of discourse’, i.e. the idea that any type of language is necessarily adapted to a specific discourse context, and as such may involve distinctly specific rhetorical functions (such as: reporting, recommending, exploring, expounding, etc.). It follows that, for each register, there is a specific repertoire of different discourse functions, and these are realised by specific lexical and grammatical phrases which are derived from, but also constitutive of, the repertoire of all potential lexico-grammatical resources. As we suggest below, this view is also that of the ‘register’ approach put forward by linguists such as Halliday (Halliday & Matthiessen 2014).

So far we have discussed a handful of questions arising from Lerat (1997), but the main technical point of his paper is that linguistic analysis can contribute to our understanding of ESP/LSP by developing conceptual ontologies on the basis of argument-predicate patterns. On this particular issue, there is no doubt that both theoretical and applied linguists have made useful contributions to the understanding of how conceptual and terminological structures operate within the context of various types of ESP/LSP. It will be sufficient for our purposes here to mention a sample of studies from our own circle of colleagues and associates, such as Volanschi & Kübler (2011) on the relationship between recurring collocational frameworks and the expression of key metaphors in biology, Pecman (2012) on the recurring lexical patterns associated with particular terms and the interpretation of these patterns as definitions in the multilingual terminological database ARTES, or Sánchez-Cárdenas & Buendía-Castro (forthcoming) on the observation of specialised collocations and their role in determining the argument-predicate structure of a particular domain relating to water disposal. These studies represent an interesting development of the traditional approach to terminology, in that they involve corpus analysis, the statistical observation of specialised word combinations (collocations), and the building of extensive, user-oriented databases. To this extent, Lerat’s predictions about the ways in which the language sciences could contribute the practice of ESP/LSP were entirely well-founded.

2.2. The context-oriented approach

The ‘context-oriented approach’ is a term that we have temporarily adopted (see also Williams 2003b) in order to gather together various kinds of linguistic analysis. One common feature of these studies, in our view, is a concern for the analysis of language in its social or cultural context, as well as a focus on language forms as they occur in actual text (hence an overlap between the terms ‘context’ and ‘co-text’). We would suggest that the context-oriented approach includes at least three subtypes, as typified by the following studies previously published in ASp.

a) Socio-cultural analysis. For example Banks (2009) describes the historical development of two of the earliest scientific journals in English and French, with a description of the scientific debates, technical topics, textual sub-genres and other factors which have defined the social and ideological context in which each journal issue evolved over time.

b) Discourse and genre analysis. An example of this is Resche (2003), who analyses a specific text type – Press Releases from the US Federal Reserve – in terms of their salient lexical, syntactic and phraseological features, as well as their rhetorical move structure.
Resche is careful to emphasise the extent to which this is a ‘living genre’, involving productive features which evolve over time.

c) **Language feature analysis.** This category ranges from small-scale manual studies of single texts, to large-scale computational comparisons of whole text collections. In some cases, the linguistic features in question lend themselves to small-scale analysis. Thus for example, Banks (1998) examines vague quantification and hedging in three hard-science texts and three ‘not-so-hard’ texts. On the other hand, large-scale analyses can be subdivided into ‘corpus-based’ studies, in which a corpus has been built in order to target specific features of language. For example Pic, Furmaniak & Hugou (2013) compare the use of three reformulation markers (*that is, i.e.* and *namely*) in two different corpora of scientific research and popular science articles (over a million words), themselves subdivided into five sub-corpora corresponding to different scientific disciplines. Another category of study involves the ‘corpus-driven’ approach: in such studies, the linguistic features of interest emerge from statistical data analysis. In Biber et al.’s (2010) method of ‘multifactorial’ analysis, for example, lexical and grammatical similarities are first observed across a large range of different text types, and it is only after this comparison that functional differences are assigned to each major ‘register’ (type of discourse). Looking at more terminological types of analysis, Williams (2003a) describes how to build a specialised ESP dictionary on the basis of concordances (derived from the analysis of reference corpora involving several hundreds million words) as well as collocational networks (derived from the statistical correlation of key lexical items).

Admittedly, the approaches set out above are very different in scope. However, we would argue that most of these studies share a degree of overlap, in that they often simultaneously combine two or more complementary methodologies. Thus as Poole & Samraj point out, “most studies of the discourse of target communities employ some form of genre analysis” (2010: 129). And as Biber et al. have claimed:

> [...] English for specific purposes and English for academic purposes have been especially influenced by corpus research, so that nearly all articles published in these areas employ some kind of corpus analysis. (Biber, Reppen & Friginal 2010: 559)

Aside from methodological considerations, we would add that many of these studies share some or all of the following theoretical assumptions (as argued especially in Gledhill 2000 & Kübler 2003).

1) ESP can only be understood properly in relation to **situational and social context**; most notably in terms of the speech communities or networks of experts who are responsible for each particular ESP; in other words the conditions in which any type of ESP text is produced and received determine its form.

2) The **semantics of a domain** (its conceptual frameworks and cognitive potential for meaning) can only be observed through language, thus from the point of view of the context-oriented analyst (but not necessarily all concept-oriented linguists), the knowledge structure of a domain is in effect equivalent to the language of the community of specialists, *i.e.* a particular type of ESP (or LSP). It follows from this logo-centric perspective that the systematic analysis of ESP should proceed from the bottom-up, *i.e.* from the analysis of language in particular texts.

3) The task of the applied descriptive linguist is not to analyse artificial examples of invented data, but rather to analyse authentic language phenomena in context, that is to say the **use of language in a particular co-text** (by ‘co-text’ we mean the language
features in the immediate textual environment of a particular example, and by ‘text’ we mean spoken, written or otherwise transmitted units of coherent communicative activity; this can include such disembodied items as headwords in a dictionary, diagrams in mathematical proofs, non-verbal semiotic gestures, etc.).

4) Any authentic example of a language form has a relevant context of use, and thus belongs to at least one genre or register. As mentioned above, the context-oriented approach questions the relevance of a ‘core language’, and is suspicious of examples which are claimed to be representative of ‘the LGP’. It is perhaps worth noting that this argument has most often been adopted by proponents of the Systemic Functional approach which, unlike many other linguistic theories (Generative Grammar, Dependency Grammar, Lexical-Functional Grammar, the Frames approach to Semantic analysis, etc.), defines itself as a theory of language in context:

[Systemic Functional Grammar is] a theory of functional variation in the genre system correlated with contextual variation. (Matthiessen 1993: 223)

5) Second-language learners of ESP have to learn the most typical patterns of usage of the ESP much in the same way native speakers have to learn it, i.e. learning ESP by doing ESP. This principle is expressed by Bloor & Bloor in the following terms:

LSP employs teaching strategies (in particular the “Deep End Strategy”) that are incompatible with certain theories of language acquisition. We explain the success of these strategies by claiming not only that learners need to be involved in the use of language appropriate to their needs, but also that by its nature this language must be “specific”. It is only by exposure to “specific” language that learners can learn the appropriate grammatical and lexical dependencies [...]. A language learner is as likely to acquire “the language” from one variety as from another, but the use of language, being geared to situation and participants, is learned in appropriate contexts. (Bloor & Bloor 1986: 28)

The notion of ‘context’ thus emerges as a central theme for many linguists who have worked on ESP/LSP. Of course, the term itself is rather slippery, and we have used it here rather opportunistically, in order to group together a variety of studies under the same label. But it seems clear to us that many linguistic studies do simultaneously attempt to provide an account of ESP in terms of both its extralinguistic context (context of situation) and its discursive context (i.e. by studying the communicative function of language forms as they occur in actual texts). Nevertheless, although we have attempted to bring together the various strands of the context-oriented approach, there are many dissenting voices and contrasting visions. We discuss some of these details in the following sections of this paper. However, before moving on, we would like to add a final assumption, which is not shared by all context-oriented linguists, but which appears to have been observed by several other analysts (see for example Sockett 2011), and which certainly underlies our own approach to ESP.

6) The typical linguistic features of ESP cannot be characterised as a list of discreet items (technical terminology, the passive, hedging, impersonal expressions, etc.), rather the most typical features of ESP texts are chains of meaningful interlocking lexical and grammatical structures, which we have called lexico-grammatical patterns (Gledhill 2011b; Kübler & Volanschi 2012). As mentioned below, these are not the same as fixed collocations or other phraseological units (such as formulae or idiomatic expressions), which may occur from time to time in ESP texts, but which are not, as we claim, the basic building-blocks of ESP.
In the following sections of this paper, we make the case for the ‘lexico-grammatical pattern’, and suggest ways in which this approach to language description may contribute very positively to the field of ESP in the future.

3. What can linguistic approaches bring to ESP now, and what can they bring in the future?

In the following sections, we make the case for two simple ideas which we believe have shown considerable potential in the development of ESP studies, namely: 1) the basic building blocks of ESP: extended lexico-grammatical patterns, 2) research skills and the ESP learner: how corpus linguistics can help trainee translators become experts in their chosen specialist domain.

3.1. The building blocks of ESP: extended lexico-grammatical patterns

As we suggested above, the development of corpus analysis has been one of the main contributions from contemporary linguistics to the field of ESP. And one of the main consequences of the widespread use of corpora in linguistics has been the development of increasingly sophisticated corpus tools which have not only allowed linguists to provide more detailed and statistically accurate descriptions of different ESPs, but have also enabled learners and practitioners to explore their own ESPs. We explore this important ‘autonomising’ feature of corpus work in the following section. However, another important consequence of the spread of corpus analysis has been the development of new ways of analysing language.

In linguistics the introduction of new methods generally involves a focus on different linguistic forms, as well as the invention of a new metalanguage. For example, in the era of generative grammar (1960s-1970s), linguists were preoccupied with grammatical paraphrase, derived structures and movement rules, while during the development of text linguistics (1970s-1980s) linguists focused on features of cohesion, discourse markers and information structure. It is not surprising then that with the availability of large-scale machine-readable corpora since the 1990s, corpus analysts have similarly developed their own analytical apparatus. From our point of view, the most significant concept to emerge from large-scale corpus analysis has been the notion of ‘phraseology’, as defined succinctly by Stubbs:

Phraseology: the identification of typical multi-word units of language use and meaning. (Stubbs 2006: 115)

This definition is based on the approach developed by John Sinclair and the lexicographers of the Cobuild dictionary (Sinclair 1987; Hunston 2002; Williams 2003b). As we have argued elsewhere (Gledhill 2000, 2011a, 2011b), this context-oriented perspective to phraseology has profoundly altered the way in which linguists view phraseology and phraseological phenomena in ESP. Whereas it is possible to analyse proverbs and idiomatic expressions out of context, using intuition for example, it is practically impossible to identify ‘multi-word units’ without using a corpus and the appropriate corpus software. Similarly, from the point of view of the corpus linguist, it may be meaningful to say ‘a typical ESP text contains few idioms’, but it would not be meaningful
to claim ‘this ESP text has no collocations’, because corpus analysis has demonstrated
time and time again that regularities of expression can always be found in any type of text
(as long as a representative corpus can be analysed).

45 But even if we accept this very inclusive view of phraseology, there are many different
ways of identifying phraseological phenomena within a corpus. It is not our purpose here
to review this topic in detail, but we suggest the following overall typology, which covers
many of the commonly described types referred to in the literature.

46 - Lexico-grammatical patterns (a formal category which covers: bound collocations,
bundles, clusters, colligations, collocational frameworks, n-grams, phrasal schemas, etc.): the
definitions for each of these terms vary, but basically these are all sequences of
recurrently co-occurring word forms (whether lexical or grammatical) which range from
very fixed sequences (for example, a 3-gram such as proceed as follows is a fixed sequence
which is three orthographic words in length) to highly variable sequences (collocational
frameworks, for example, involve two or more fixed pivotal items surrounded by more
variable elements, as in: Failure to do (so / this) + will result in (death / serious injury / property
damage, etc.).

47 - Semantic patterns (a functional category which includes: discourse prosody, extended
lexical units, semantic preference, stabilised expressions, etc.): again, the details vary, but
generally these are regularities of expression which correspond to conventionalised units
of meaning within a text, whether this meaning corresponds to a well-defined semantic
field (semantic preference) or a marker of speaker attitude (semantic prosody, as
discussed in the next section). When semantic patterns correspond to lexico-grammatical
patterns, they signal something about the conventional meaning or communicative
function of the expression (e.g. Proceed as follows = INSTRUCTION, Failure to do so =
WARNING, etc.). But semantic patterns can be more diffuse: they can correspond to
abstract networks or paradigms of meaning which are built up over time, either within a
single text or across a whole discourse (for example the lexical item ‘express’ in the ESP
of biology and biochemistry which conventionally signals a transitive chain of
relationships between cells, genes and proteins, as in: cells express genes, genes express
proteins, etc.)

48 - Register patterns (a contextual category which includes such notions as Biber et al.’s
(1990) rhetorical moves, etc.): we use this term to refer to various types of high-level
regularities, which for a particular register (a variety of language corresponding to a
specific context), correspond to the overall configuration of all of the relevant lexico-
grammatical and semantic patterns which are conventionally (or potentially) associated
with that register among a particular community of language users.

49 The categories above clearly cannot do justice to a very rich and complex field, but at
least by setting them out in this way we can see how various related concepts may appear
to fit together in a unified scheme. These categories were chosen intentionally; they
correspond to the main strata of language posited in the systemic functional model
(Halliday & Matthiessen 2014), namely: Lexicogrammar, Semantics and Context. We
would suggest that what emerges from all of these different categories is the notion of
‘pattern’ or ‘regularity of expression’. We would also suggest that the analysis of one type
of phraseological phenomenon (e.g. lexico-grammatical pattern) systematically implies the
analysis of the others (semantic pattern, register pattern), to such an extent that it in effect becomes redundant to specify at which point one has entered the analytical system.

In order to demonstrate this point, in the following discussion we examine two related lexico-grammatical patterns in order to make a more general point about their communicative functions (semantics) and their typical context of use (register).

The following examples were first identified as part of a Master’s level exercise in corpus linguistics (as reported in Gledhill & Kübler 2015). As part of their training in corpus linguistics, we ask our students to analyse five ‘generic collocations’, that is to say a selection of five lexico-grammatical patterns used in a corpus of French or English specialised texts. In the examples we examine below, one student (initials FS) collected an unannotated corpus of 94 English-language technical manuals (a typical title being *Yale Electric Wire Rope Winch: Installation and Operator’s Manual*), all roughly corresponding to around 20 different makes of electric winch (all of these machines are used to launch gliders, so the products and their technical manuals are subject to very stringent safety regulations). According to the AntConc4 wordlist tool (Anthony 2014), the Technical Manual corpus includes 568,998 words.

Not all students use ‘n-grams’ to look for generic collocations, but early in his project FS decided to use the AntConc clusters tool to look for fixed sequences of words in the corpus. For example, if we look for 4-gram sequences, the tool finds over 230,000 occurrences (although over 60% of these are ‘hapaxes’: sequences which occur only once). Many at the top of the list represent noise or meaningless examples. Others are meaningful, but not obviously related to other sequences, as in:

> West Artesia Blvd Compton, Refer to Dwg MHP, please call Smittybilt West, REDUCTION WORM GEAR DRIVES...

Finally, a third category of n-grams are meaningful, especially when they are associated with other sequences, such as:

Owner’s manual for, The wire rope is, And replace as necessary, Failure to observe these, Instructions could lead to, Instructions could result in, Observe these instructions could, Exceed the wire rating, On the wire rope, Serious injury or death...

Evidently, the search for n-grams does not provide the analyst with immediate results, and so some ‘manual’ interpretation of the evidence is still necessary. Still, such a tool can be used to reveal some of the more statistically salient patterns in this type of text. Looking at the above results, FS found that one of the most statistically salient patterns to emerge in this text type is ‘failure to follow instructions could result in...’. Although many of our students (including FS in this case) consider this type of occurrence to be an exemplar (a pattern in itself), it is clear that such examples have to be analysed in terms of a more generic pattern (involving potential variations), but also across a more extended stretch of text, as in the following schema (based on 230 occurrences):

Pattern 1. FAILURE TO + V (where V = comply with, follow, heed, observe, understand, etc.) MAY/CAN + LEAD TO / RESULT IN + N (where N = death, fire, injury, property damage, etc.)

The following examples show a variety of different possible instantiations of this overall pattern (NB we usually try to present examples in the same way: pivotal elements – the stable elements of the pattern – are in bold, while paradigms – regular but more variable elements – are underlined):
(1a) Failure to comply with safety precautions described in the manuals supplied with the winch, this manual or any of the labels and tags attached to the winch is a safety violation that may result in death, serious injury or property damage.

(1b) WARNING! Read all instructions. Failure to follow all instructions listed on pages 4 to 7 may result in fire, serious injury and/or DEATH.

(1c) Failure to heed these instructions may result in personal injury and/or property damage.

(1d) WARNING RISK OF BURNS. Failure to observe these instructions could lead to serious injury or death.

(1e) Failure to understand the proper operation of this product can result in serious injury and/or property damage.

56 These examples demonstrate a number of general principles about lexico-grammatical patterns:

57 - the pattern is non-linear (it can be interrupted by other lexical material);

58 - the pattern can extend over several phrases and even clauses (and one pattern appears to lead on to or ‘cascade’ into another);

59 - the pattern involves several permanent or ‘pivotal’ elements of grammatical structures as well as lexical items (nominal group: failure, embedded clause: to follow instructions, main verb group: can result in, etc.) as well as a ‘paradigm’ of more variable elements (lead to / result in + death / injury, etc.);

60 - the pattern involves lexical items which belong to related semantic fields (for example, the verb introduced after failure refers to behaviour: comply, follow, heed, observe, or in a minority of cases cognition: understand, while the complement refers to the text itself: instructions, safety precautions);

61 - the pattern has a regular textual function (it constitutes a formal legalistic warning) as well as a regular context (it is often used in close proximity to a textual signal such as ‘Warning!’).

62 However, it is necessary to examine another set of examples in order to demonstrate a further general principle: similar patterns are usually contrastive, and do not necessarily have the same contexts of use. Thus, for example, if we look in the Technical Manual corpus for variants of pattern 1, such as failure to do so... /failure to do this..., we find that the semantic range of the pattern is more restricted. In this case, while the embedded verb refers to a generic process to be retrieved in the preceding text (to do so, to do this), the complement of the main verb (results in) refers exclusively to undesired technical consequences—but not as bad as the dire consequences we saw for pattern (1), giving the following schema (based on 45 occurrences):

Pattern 2. FAILURE TO DO SO / THIS (MAY/CAN etc.) + (modal) + V (where V = result in) + N (where N = the cable parting, damage to the cable, pressure buildup, electric shock, equipment damage, etc.).

(2a) Failure to do this could result in the cable parting from the drum under load.

(2b) CAUTION: ... Failure to do so will result in the outer wraps pressing against the inner wraps resulting in the damage of the cable.

(2c) Failure to do so could result in pressure buildup which can cause the gearbox to leak or damage the equipment.

(2d) All electrical work must be performed by a licensed electrician. Failure to do so could result in electric shock or poor winch operation.

(2e) Remove port caps from exhaust ports for operation. Failure to do so may result in equipment damage and limit performance.
In this section, we have identified one of the most typical multi-word units of language in technical manuals. More generally, we would argue that this kind of analysis represents a relatively simple but systematic method for identifying the most typical lexico-grammatical patterns – in other words the ‘building blocks’ – of ESP and indeed of any specific type of language. Of course, it has to be borne in mind that much of this analysis is dependent on human interpretation, and requires the ability to identify a representative corpus of relevant texts. It is for this reason that corpus analysis is increasingly seen as a key feature of the contemporary language syllabus (for a discussion, see Lenko-Szymańska & Boulton 2015).

3.2. Research skills and the ESP learner: how corpus linguistics can help trainee translators become experts in their chosen specialist domain

The spread of corpus analysis mentioned above has also triggered new approaches in translation studies, and especially in specialised or pragmatic translation training. Aston (1999) was one of the first to explain how corpus linguistics could be applied to the training of translators in this area. In his analysis of the different corpora that can be of use in specialised translation, such as monolingual corpora in the target language, parallel corpora and comparable corpora, he underlines the advantages of comparable specialised corpora that are very often 'ad hoc' corpora, i.e. compiled for a specific use in a specific field. Following Aston’s work, a series of conferences entitled Corpus Use and Learning to Translate (CULT) led to several publications in this area (Bernardini & Zanettin 2000; Zanettin et al. 2003; Beeby et al. 2009). More recently, other researchers (Frankenberg-Garcia 2009; Stewart 2009; Gledhill 2011a, 2011b; Loock 2014) have emphasised the advantages of using corpora and corpus linguistics in translation training, although in France not many courses have implemented this approach. In this section, we describe a corpus-based approach that has been adopted and effectively applied at our university over many years (Kübler 2003) in the training of specialised translators.

The translation process is presented to trainees as including three phases: 1) a documentation phase, 2) a translation phase, and 3) a revision phase. Corpora play different roles in all these three phases. We will not go into details here, as the three phases have been described elsewhere (Kübler 2011), but we will give a few examples to make our point on what our theoretical approach can bring to linguistics and translation training.

- **Documentation phase**

The analysis of genre and register for technical writing can also be applied to the text to be translated, as a thorough understanding of the linguistic features of the source text is necessary to achieve a high quality translation. Subsequently, translators have to explore the domain and understand concepts that are difficult or unknown. As mentioned above in our discussion of the concept-oriented approach, the technique of querying specialised parallel and comparable corpora in order to understand a domain and to find out its terminological system (in both source and target languages) has been widely used for many years in the training of translators (Maia 2003 for example). Moreover, there has been much research on linguistic markers, which allow the translator to come up with
definitional contexts (Pearson 1998) in the source and target language. Markers of course differ from one language to another. In English, for example, a very common marker appears in the following definition, as shown in example (3):

(3) The motion of continents relative to the Earth’s spin axis may be either due to the drift of individual continents or due to a rotation of the entire Earth relative to its spin axis; the latter is called True Polar Wander (TPW)

Here is called signals a definition followed by the name of the defined term (in italics). In French, our students’ target language, a similar marker is on dira (literally ‘one will say’):

(4) Si l’arête est la diagonale intérieure d’un quadrilatère concave, on dira que cette arête est flippable

‘If the edge corresponds to the internal diagonal of a concave quadrilateral, this is called a flippable edge.’

Not only does this example represent a regular lexico-grammatical pattern, it can be also be used as a template for the observation of regular patterns of meaning. In this case, we can infer that the term arête (edge or ‘arête’ in English) is the hyperonym of arête flippable, whose definition appears before the actual lexical item is actually named. Generally speaking, such examples show how the trainee translator can come to understand the conceptual structure of a domain by the systematic observation of a corpus, and without any specific computational skills (as mentioned above in our discussion of the concept-oriented and context-oriented approaches). The advantage of querying a corpus for domain-specific knowledge lies in the fact that the corpus yields contexts and definitions that are not in specialised dictionaries, and gives information on the semantic relationships of the terms in the domain (hyperonyms, hyponyms, meronyms...).

Apart from finding terms in the source language and their equivalents in the target language, which is what specialised corpora are widely used for today, our approach also involves the detection of important lexico-grammatical information in the second learning phase mentioned above, namely, the translation process.

- Translating phase

One of the most interesting problems which can arise in the translation process is the recognition, and then translation, of semantic prosody. Louw defines semantic prosody as “a consistent aura of meaning with which a form is imbued by its collocates” (1993: 157), while Sinclair, who originally coined the term, points out that it is attitudinal and on the “pragmatic side of the semantics /pragmatics continuum” (1996: 87). As discussed in Kübler & Volanschi (2012),

Semantic prosody is thus situated on the highest level of abstraction in the extended description of the lexical unit: collocation (co-occurrence of a node with a specific word), colligation (co-occurrence of a node with a set of words belonging to a specific syntactic category), semantic preference (co-occurrence of a node with a set of words belonging to the same semantic class, i.e. sharing a set of semantic features), and semantic prosody. (Kübler & Volanschi 2012: 104-105)

Thus, in brief, semantic prosody can be defined as the contextually-determined (or ‘pragmatic’) meaning of a word shown by its combination with a series of collocates that have a positive or a negative connotation: the word is then said to have a negative or a positive semantic prosody. The phenomenon is difficult to detect for non-native speakers, and as discussed by Stewart (2009), it may account for a major difficulty in translation. As translators usually translate into their mother tongue, from a second language, semantic prosody constitutes a collocational pitfall for French-speakers translating from English into French. Moreover, as many researchers (Partington 1998; Berber Sardinha 2000)
have shown, semantic prosody can differ from one language to the other, and that an equivalent in the target language does not necessarily have the same semantic prosody as the initial word in the source language. It has also been demonstrated that semantic prosody can differ when used in LSPs (Hunston 2007; Louw & Château 2010; Kübler & Volanschi 2012).

We think that trainee translators should be trained to spot semantic prosody in a text, as they will have to choose among different equivalents. The verb *give rise to* for example has a negative semantic prosody in contexts in which human beings or their productions or artefacts are concerned, and no semantic prosody at all in other cases. This ‘semantic pattern’ was first identified by Hunston (2002). The presence (but also absence) of this prosody can be illustrated in the following examples:

(5a) widespread changes in erosion rates and sediment flux can have important repercussions and *give rise to* significant socio-economic and environmental problems
(6a) Such a process may *give rise to* explosive eruptions, leading to important human and material damages, and may have a major impact on the global Earth climate
(7a) the relatively planar topography of the caldera may *give rise to* misinterpretations
(8a) Such a phylogenetic distribution in association with the apparent textural stability of nacre throughout the geological times *give rise to* two fundamental questions
(9a) however, it appears unable to *give rise to* TTG-like magmas
(10a) which, upon addition of hyaluronan, *give rise to* negatively charged composite vesicles

Examples (5) to (7) show that *give rise* has a negative semantic prosody, as the pattern *give rise to N* contains the nouns *problem* or *damages* or *misinterpretations*, which have a negative connotation. Moreover in example (5) the presence of *important repercussions* and of the evaluative adjective *significant* reinforce the negative connotation. In all three sentences, the consequences have an impact on human beings. In contrast, sentence (8) has no semantic prosody (or it is neutral), and the pattern *give rise to question(s)* appears to belong to more general academic discourse. Finally, in sentences (9) and (10), *give rise to* TERM belongs to the ESP of geology, since it is used with other terms in this domain (L’Homme 1998), namely *magma* and *vesicle*. Again, in these examples, there is no observable semantic prosody.

If we look up the translation of *give rise to* in the Wordreference dictionary, we find several potential translation equivalents in French:

- déclencher, provoquer, engendrer, donner lieu à
- produire, soulever, générer, déchainer, donner suite à, enfanter, prêter à confusion, prêter à rire, etc.

Even though it is possible, by relying on our ‘Chomskyan’ (native speaker’s) intuition, to sort out some of the suggested translations, choosing the correct one remains a problem without the help of any context, and without understanding the phenomenon of semantic prosody. As mentioned, the translation equivalents for the first three examples (5, 6, 7) should also have a negative semantic prosody. The only way of finding out which suggested equivalent (or even other possibilities) is adapted in French is to look up each verb in a French specialised corpus, and then to check, for each individual noun, which verb fits best, taking into account the negative or neutral semantic prosody of the verb in French. Here the French corpus suggests:
(5b) widespread changes in erosion rates and sediment flux can have important repercussions and give rise to significant socio-economic and environmental problems => engendrer, entraîner des problèmes environnementaux et socio-économiques importants

(6b) Such a process may give rise to explosive eruptions, leading to important human and material damages, and may have a major impact on the global Earth climate => provoquer des éruptions explosives

(7b) the relatively planar topography of the caldera may give rise to misinterpretations => conduire à des interprétations erronées

(8b) Such a phylogenetic distribution in association with the apparent textural stability of nacre throughout the geological times give rise to two fundamental questions => soulever deux questions fondamentales

(9b) however, it appears unable to give rise to TTG-like magmas => générer des magmas de type TTG (Tonalite, Trondjhemite, Granodiorite)

(10b) which, upon addition of hyaluronan, give rise to negatively charged composite vesicles => amènent à la formation de vésicules composites chargées négativement

78 The approach adopted in this translation phase can also be applied to the revision phase, once the trainee’s translations have been revised by the trainers. The revision process is a critical and necessary stage in any professional translation (Mossop 2007); in our case, it is complicated by the fact that the trainers are often neither native speakers of English, nor specialists in the domain. Thus we would suggest that corpus analysis, once again, can be extremely useful when revising a translation to check for potential equivalents and lexico-grammatical patterns.

79 In this section we have picked out just two features in the complex process of training specialist translators: the identification of lexico-grammatical patterns relating to definitions, and the analysis and comparison of semantic patterns in both source and target languages. This approach not only requires that the trainee translators learn how to use a corpus, but also – and more importantly – what to look for in a corpus. This is why it is crucial for them to receive specific training in the analysis of monolingual and comparable corpora.

4. The history of linguistic contributions to the journal ASp

80 So far we have set out our own particular perspective on linguistic approaches to ESP. However, it would be instructive to finish this article by exploring the different ways in which linguistics has contributed to the field of ESP in the particular context of this journal. Despite its international perspective and team of contributors, ASp can still be fairly described as a ‘national’ journal (its readership and editors have traditionally been based mostly in France). So a good case can be made to see ASp as representative of a particular community of specialists. A further argument for this is to consider that many contributors to the journal attend the regular GERA S conferences. This is of course a subjective observation, but we would suggest that there are likely to be closer professional and social bonds between ASp’s editorial team, contributors and readership than for some other comparable international journals in this field. Given this degree of cohesion, it would be interesting to observe what articles on linguistics have been submitted to ASp, and how the various linguistic approaches to ESP have evolved in the journal.
Overall, we have found 174 linguistics-oriented articles out of 483 papers published in ASp between 1993-2015. In the early years of ASp (1993-2000), articles were separated topically into four areas (Linguistics, Didactics, Culture, Technology). This made our selection fairly easy before 2000. After 2000, the division of papers into topics was replaced by a more synthetic approach. In Appendix 1 we have included all papers which before 2000 were either categorised as Linguistics (or a specific branch of linguistics), while after 2000 we have identified linguistics articles as those which correspond to one of the concept-oriented or context-oriented approaches discussed above (see sections 2.1 and 2.2). In some cases, a paper was not considered to be linguistic even though it might have looked like one (citing several extracts of text, for example, or studies which adopted an essentially literary or ‘cultural’ approach to language: this was particularly the case of studies on ‘FASP’: a French term for ESP in fictional contexts).

Generally speaking, Appendix 1 shows that the proportion of papers adopting a linguistic approach in ASp has been gradually increasing. Between 1993-1997, the proportion was approximately 20%, but this percentage rose regularly between 1998-2008 with scores of around 30-40%, despite one or two dips. Between 2008-2015, the score has risen even more sharply, and has been regularly above 50%. It is not clear to us why this might be. It may be that the apparent growth of linguistics in ASp corresponds to a particular change in the way ESP is taught or researched in the context of French higher education. But perhaps another explanation lies in a relative expansion in the range and diversity of topics. As we mentioned above, the context-oriented approach and the pervasive use of corpora have multiplied the different number of ways in which linguistic analysis can be conducted on different types of ESP. Some support for this may be found in Appendix 2. Although the range is very broad, it can be seen that among the 40 topics listed, the most frequent categories of linguistic contribution to ASp focus on genre analysis (the language of specific text types or disciplines), or discourse analysis (specific discourse modes or discourse markers). Although such studies are typically context-oriented, there have also been many concept-oriented contributions, most notably looking at metaphor, terminology and the morphological aspects of term-formation. It is also interesting to note that although most traditional features of syntax (verbal groups, nominal groups, pronouns, etc.) are represented, these topics do not attract as much attention as higher-level or textual features (Theme/Rheme analysis, cohesive markers, etc.).

When we turn to Appendix 3, we see a similar state of affairs. In this table, we have attempted to categorise the theoretical background of each of the linguistic contributions to ASp. This has not been very systematic; in most cases we simply looked at the main references in the bibliography, or the keywords for each article (or failing that, we have made an ‘educated guess’ about the overall approach adopted). As can be seen, the number of ‘approaches’ is not the same as the number of articles, for the simple reason that in many cases, no single approach can be identified. One paper for example (Resche 2006) covers many different linguistic approaches, as can be seen in the keywords: definition, economics textbook, market, orthodoxy and heterodoxy, paradigmatic and syntagmatic axes. Notwithstanding this complexity, we would suggest that the most frequently occurring contributions to ASp appear, perhaps unsurprisingly, to adopt a context-oriented approach (genre analysis, discourse analysis, descriptive grammar, systemic functional linguistics, etc.). In addition, of the 16 studies in terminology, many of these also happen to be context-oriented, in that many of these studies often involve corpus analysis, or
adopt a lexicographic or discourse-analytic approach to the definition of terms (see for example, Resche 1998, 2001).

5. Conclusion

So, what can linguistic approaches bring to English for Specific Purposes (ESP)? We began this paper with a critical look at the concept-oriented tradition. It is undeniable that concept-oriented linguistics has shaped the field of ESP fundamentally, not least by emphasising the central role of technical terminology in the conceptualisation of specialist domains and the construal of specialised knowledge. We would not like to give the impression that we are critical of all concept-oriented studies, although in section 2.1, we did react to many of the assumptions made by Lerat (1997), whose approach seems to be representative of a certain kind of linguistics which sees LSP or ESP as being restricted to the language (and in particular terminology) of a specialised discipline.

In contrast, it should be clear from section 2.2 above that we see ourselves as context-oriented linguists. We would claim that the context-oriented school has also fundamentally shaped the field of ESP, in that it has brought with it a plethora of new methods and research questions, all of them determined not by theoretical considerations, but by the development of new techniques in the observation of language data and changes in the way that we interact with and perceive the practitioners of ESP. However, as we suggested above (sections 2.1 and 2.2), in many ways, the eclectic, empirical approach adopted in context-oriented studies does not stand in total opposition to the concept-oriented approach; rather there has come to emerge a hybrid kind of linguistic analysis which combines both perspectives. Thus a serious in-depth study of a particular ESP might contain a survey of the knowledge structure of the domain, an analysis of ethnographic context, a thorough description of the ecology of genres, a critical discourse analysis of several key texts in the field, and all of this triangulated by the systematic analysis of one or (preferably several) comparative corpora (not to mention possible diachronic analysis...). In some ways such studies have existed since the beginning of the corpus revolution (1980-1990), but it is not hard to find more recent and certainly more systematic studies which appear to be moving in this direction (to name but one example: Grey 2015).

Regardless of whether one accepts or rejects our distinction between concept-oriented and context-oriented approaches, it seems undeniable that the pervasive use of corpus analysis has changed the rules of the game for most linguists interested in ESP. Of course, the corpus revolution has not been adopted by all linguists (see for example the review of current linguistic approaches in Heine & Narrog 2010). And indeed it would be a mistake to assume that all analysts use corpus-informed research in the ways that we have been discussing (in sections 3.1 and 3.2 above). Nevertheless, we would claim that for many linguists working in the field of ESP, academic practice has changed immeasurably in recent years: in the pre-corpus era, the expectation was that all of the lexical forms and phrases of a language variety could be analysed out of context, in terms of an abstract grammar; now these forms can be seen in the light of multi-word patterns (features of text which can only be systematically appreciated by observing corpus data). Similarly, the world of the language practitioner has also changed. For example, in the pre-corpus era technical translators relied on translation equivalences set out in static, authoritative dictionaries, now most technical translations are carried out using computer-assisted...
translation tools which make use of multi-word units, and as a consequence translators almost obligatorily need training in how to manage corpus tools.

This final point leads us to mention very briefly an important issue which we unfortunately did not have space to discuss in the main body of this paper: the indirect benefits of linguistics. By ‘indirect’ we mean all of the techniques, tools and theories that have come into the field of ESP (and LSP) without necessarily being associated with linguistic analysis. For example, on-line electronic resources such as the analytical toolbox SketchEngine\textsuperscript{11} (Kilgarriff et al. 2004), or terminology databases such as the Grand Dictionnaire Terminologique\textsuperscript{12} or ARTES\textsuperscript{13} (Pecman 2012) are of immense use to language-users, but especially so to professionals and practitioners in ESP/LSP. We would argue that such tools, among many hundreds of others, are an indirect benefit of linguistics – mostly because they originated as fundamental research in (applied) linguistics.

We have no space here to go into this issue in any detail, but perhaps the best way we can frame it is by echoing the famously tongue-in-cheek question from the film *Life of Brian*: “Well, just what have the Romans (or in this case ‘the Linguists’) ever done for us?” To this question we could very well reply, as in the film: “Well, apart from online dictionaries and term-banks, keyword searches, concordances and corpora, grammar checkers, style checkers and readability indexes, computer-assisted translation, subtitling and voice recognition software, not to mention knowledge engineering, the semantic web, the use of online grammars, training modules and other corpus-informed materials for the language classroom and teacher training – in other words all the basic theories, tools and techniques that language learners, professional translators, language engineers, technical communicators and for that matter pretty much all literate people in the educated world who need to deal with language – it’s true, what have ‘the Linguists’ ever done for us?”

*The authors of this paper are grateful to two anonymous reviewers for their valuable suggestions and comments.*

---

**BIBLIOGRAPHY**


GREY, Bethany. 2015. Linguistic Variation in Research Articles: When Discipline only Tells Part of the Story. Amsterdam: John Benjamins.


APPENDIXES


<table>
<thead>
<tr>
<th>Year</th>
<th>Authors (listed alphabetically)</th>
<th>Linguistics Articles</th>
<th>Total Articles</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993</td>
<td>Boughedaoui, Coulardeau, Diana, Fries-Verdeil, Miller, Monin, Perry, Petit, Rybar, Videlenc (a), Videlenc (b)</td>
<td>11</td>
<td>61</td>
<td>18%</td>
</tr>
<tr>
<td>1994</td>
<td>Carter-Thomas, Corbisier, Dudley-Evans, Fade, Magnet, Miller, Osborne, Petit, Schuwer</td>
<td>9</td>
<td>47</td>
<td>19%</td>
</tr>
<tr>
<td>1995</td>
<td>Banks, Busuttil, Coulardeau, English, Partington, Phillips, Resche, Villez</td>
<td>8</td>
<td>36</td>
<td>22%</td>
</tr>
<tr>
<td>1996</td>
<td>Bondi-Paganelli, Monin, Osborne, Resche, Tassard, Van der Yeught, Wolosin</td>
<td>7</td>
<td>32</td>
<td>22%</td>
</tr>
<tr>
<td>1997</td>
<td>Bachschmidt, Banks, Beaufrère-Bertheux, Birch-Bécaasas, Boughedaoui, Carnet, Chukwu, Coulardeau, Crosnier, Galonnier, Gledhill, Labrosse, Lacic, Lerat, Martin, Mathis, Resche, Sionis (a), Sionis (b), Sionis (c), Petit, Rézeau, Webber</td>
<td>23</td>
<td>35</td>
<td>66%</td>
</tr>
<tr>
<td>1998</td>
<td>Banks, Dressen-Hammouda, English, Greenstein, Resche (a), Resche (b), Salager-Meyer, Trouillon</td>
<td>8</td>
<td>22</td>
<td>36%</td>
</tr>
<tr>
<td>Year</td>
<td>Authors</td>
<td>Documents</td>
<td>Citations</td>
<td>Coverage</td>
</tr>
<tr>
<td>------</td>
<td>------------------------------------------------------------------------</td>
<td>-----------</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>1999</td>
<td>Bachschmidt, Carter-Thomas, Cotte, Fries-Verdeil, Magnet, Monin, Mullen, Percebois, Resche, Rouveyrol, Rowley-Jolivet, Schuwer</td>
<td>12</td>
<td>25</td>
<td>48%</td>
</tr>
<tr>
<td>2000</td>
<td>Arleo, Banks, Boulton, Galonnier, Labrosse, Magnet, Resche, Rowley-Jolivet</td>
<td>8</td>
<td>22</td>
<td>36%</td>
</tr>
<tr>
<td>2001</td>
<td>Banks, Carter-Thomas &amp; Rowley-Jolivet, Dechesne, Eason, Fries-Verdeil, Resche, Spalding Andréolle, G. Williams.</td>
<td>8</td>
<td>20</td>
<td>40%</td>
</tr>
<tr>
<td>2002</td>
<td>Banks, Busch-Lauer, Carnet &amp; Magnet, Divasson &amp; León, Fontaine &amp; Kodratoff, Hilton, Maniez, Resche, Rotgé, Sionis, G. Williams Sionis &amp; Boucher</td>
<td>11</td>
<td>16</td>
<td>69%</td>
</tr>
<tr>
<td>2003</td>
<td>Baud, Resche, Rowley-Jolivet &amp; Carter-Thomas, Trouillon, G. Williams</td>
<td>5</td>
<td>25</td>
<td>20%</td>
</tr>
<tr>
<td>2004</td>
<td>Banks, Duchet &amp; Chaulet, Negroni &amp; Humbley, Percebois, Van der Yeught</td>
<td>5</td>
<td>19</td>
<td>26%</td>
</tr>
<tr>
<td>2005</td>
<td>Resche</td>
<td>1</td>
<td>8</td>
<td>13%</td>
</tr>
<tr>
<td>2006</td>
<td>Banks, Boulton &amp; Wilhelm, Resche, Richard</td>
<td>4</td>
<td>10</td>
<td>40%</td>
</tr>
<tr>
<td>2007</td>
<td>Chambers, Hay, Maniez, Moirand &amp; Tréguer-Felten</td>
<td>4</td>
<td>9</td>
<td>44%</td>
</tr>
<tr>
<td>2008</td>
<td>Humbley, Pennec, Rowley-Jolivet &amp; Carter-Thomas</td>
<td>3</td>
<td>11</td>
<td>27%</td>
</tr>
<tr>
<td>2009</td>
<td>Banks a, Banks b, Davier, Maniez, Robert, Trouillon</td>
<td>6</td>
<td>12</td>
<td>50%</td>
</tr>
<tr>
<td>2010</td>
<td>Banks, Bex, Flettum, Landure &amp; Boulton, Pennec, Pic &amp; Furmaniak</td>
<td>6</td>
<td>12</td>
<td>50%</td>
</tr>
<tr>
<td>2011</td>
<td>Bendinelli, Chapon, Gledhill, Isani, Peynaud, Sockett</td>
<td>6</td>
<td>12</td>
<td>50%</td>
</tr>
<tr>
<td>2012</td>
<td>Banks, Kusyk &amp; Sockett, Peraldi, Pic &amp; Furmaniak (a), Pic &amp; Furmaniak(b)</td>
<td>5</td>
<td>9</td>
<td>56%</td>
</tr>
<tr>
<td>2013</td>
<td>Bardière, Biros, Bordet, Dressen-Hammouda, Gaillat, Lacase, Pic Furmaniak &amp; Hugou, Resche</td>
<td>8</td>
<td>16</td>
<td>50%</td>
</tr>
<tr>
<td>2014</td>
<td>Anesa, Bordet, Debras, Flettum, José-Marin &amp; Rea, McAlister, Meyers, Percebois, Peynaud, Pic &amp; Furmaniak</td>
<td>10</td>
<td>14</td>
<td>71%</td>
</tr>
<tr>
<td>2015</td>
<td>Cloarec, Domenec, Gaillard, Millot, Nesi, Resche</td>
<td>6</td>
<td>10</td>
<td>60%</td>
</tr>
</tbody>
</table>

Linguistic features of a genre/text type (conference presentation, presidential debate, email list announcement, headline, film review, letter to the editor, mission statement, military operation order, press release, research article, etc.) or field (astronomy, economics, finance, the law, medicine, etc.)

<table>
<thead>
<tr>
<th>Topic</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argumentation, discourse modes/sequences, rhetorical moves</td>
<td>10</td>
</tr>
<tr>
<td>Bundles, clusters, collocations, lexical phrases, phraseological units</td>
<td>9</td>
</tr>
<tr>
<td>Metaphor, conceptual metaphor (cognitive approach)</td>
<td>9</td>
</tr>
<tr>
<td>Morphology in terminology (abbreviation, affixing, blending, composition, derivation)</td>
<td>8</td>
</tr>
<tr>
<td>Terminology of a domain (in LSP and/ or LGP)</td>
<td>8</td>
</tr>
<tr>
<td>Theme/Rheme analysis, information / thematic structure</td>
<td>8</td>
</tr>
<tr>
<td>Cohesive markers of reference (‘this, that’, etc.)</td>
<td>7</td>
</tr>
<tr>
<td>Terminology in translation, translation problems (false friends)</td>
<td>7</td>
</tr>
<tr>
<td>Cohesive markers of conjunction, discourse markers (‘that is to say...’, etc.)</td>
<td>6</td>
</tr>
<tr>
<td>Modality</td>
<td>6</td>
</tr>
<tr>
<td>Aspect, tenses</td>
<td>5</td>
</tr>
<tr>
<td>Cohesion across a whole text</td>
<td>5</td>
</tr>
<tr>
<td>Clause-level structure (if-conditionals, clefts, extraposition, coordination)</td>
<td>4</td>
</tr>
<tr>
<td>Lexical association, complexity, density, comprehension, reading level, AWL</td>
<td>4</td>
</tr>
<tr>
<td>Lexical distribution, keywords, “notional networks”</td>
<td>4</td>
</tr>
<tr>
<td>Terminological creativity/neology</td>
<td>4</td>
</tr>
<tr>
<td>Translation processes, specialised translation</td>
<td>4</td>
</tr>
<tr>
<td>Voice (= active, medio-passive, passive, etc.)</td>
<td>4</td>
</tr>
<tr>
<td>Error analysis</td>
<td>3</td>
</tr>
<tr>
<td>Grammatical metaphor</td>
<td>3</td>
</tr>
<tr>
<td>Hedging</td>
<td>3</td>
</tr>
<tr>
<td>Interrogatives (direct/indirect)</td>
<td>3</td>
</tr>
<tr>
<td>Multimodal aspects of oral and written interaction (images, symbols, non-verbal gestures, etc.)</td>
<td>3</td>
</tr>
</tbody>
</table>
Rhetoric, style, voice (= strategy of communication) 3
Verbal group syntax, causatives, phrasal verbs 3
Verbal semantics (process types – participant roles, transitivity) 3
Adverbs of intensity, modal adjuncts 2
Aphoristic phrases, formulaic language, speech acts, sequences with specific rhetorical functions 2
Denomination, proper nouns, names 2
Nominal group, adjectival group (compound pre-modifiers) 2
Pronouns, relative pronouns 2
Quantifiers 2
Reformulation 2
Citations, quotations 1
Collocational networks 1
Intercultural rhetoric, politeness 1
Language learning/Acquisition 1
Non-standard language (slang) 1


Genre analysis (V. Bhatia, C. Bazerman, F. Salager-Meyer, J. Swales) 16
Terminological analysis, specialised lexicography (J. Humbley, L. Hoffmann, M.-C. L’Homme, R. Temmerman, P. Thoiron) 16
Corpus-based, corpus-driven, corpus-informed analysis (D. Biber, S. Granger, J. Sinclair) 14
Discourse analysis (J-M. Adam, P. Charaudeau, M. Halliday, D. Maingueneau) 8
Descriptive grammar(s) (D. Biber, R. Huddleston, P. Larreya, M. Riegel, R. Quirk) 9
Systemic functional approach (M. Halliday, R. Hasan, J. Martin, C. Matthiessen) 7
Concordance analysis (T. Johns, M. Stubbs, C. Tribble) 5
Contrastive/comparative analysis (U. Connor, R. Hartmann, B. Hatim) 5
<table>
<thead>
<tr>
<th>Meta-linguistic discussion of linguistic models</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multimodal analysis (G. Kress, Van Leuwen)</td>
<td>4</td>
</tr>
<tr>
<td>Diachronic analysis (D. Banks, G. Myers)</td>
<td>3</td>
</tr>
<tr>
<td>Lexicological approach, lexical semantics, phraseology (F. Gaudin, I. Mel'čuk, F. Rastier.)</td>
<td>3</td>
</tr>
<tr>
<td>Meta-linguistic discussion of corpus methodology/construction</td>
<td>3</td>
</tr>
<tr>
<td>Meta-linguistic discussion of learning/teaching through linguistic analysis</td>
<td>3</td>
</tr>
<tr>
<td>Meta-linguistic discussion of student productions</td>
<td>3</td>
</tr>
<tr>
<td>Semiotic analysis (G. Kress, Van Leuwen)</td>
<td>3</td>
</tr>
<tr>
<td>Critical Discourse Analysis (N. Fairclough)</td>
<td>2</td>
</tr>
<tr>
<td>Lexicometrics, logometrics, textometrics (B. Habert, L. Rouveyrol)</td>
<td>2</td>
</tr>
<tr>
<td>Meta-linguistic discussion of glossaries, terminological databases</td>
<td>2</td>
</tr>
<tr>
<td>Psycho-systematic theory (G. Guillaume)</td>
<td>2</td>
</tr>
<tr>
<td>Stylistics (D. Crystal, G. Molinié)</td>
<td>2</td>
</tr>
<tr>
<td>Theory of enunciative operations (A. Culioli)</td>
<td>2</td>
</tr>
<tr>
<td>Cognitive approach, mental lexicon (W. Croft, R. Lakoff, P. Thorndike)</td>
<td>1</td>
</tr>
<tr>
<td>Conversation analysis (J. Atkinson, E. Goffman, J. Sacks)</td>
<td>1</td>
</tr>
<tr>
<td>Pragmatics (J. Anscombre, P. Grice, O. Ducrot)</td>
<td>1</td>
</tr>
<tr>
<td>Rhetorical structure theory (W. Mann, S. Thompson)</td>
<td>1</td>
</tr>
<tr>
<td>Transformational grammar (N. Chomsky)</td>
<td>1</td>
</tr>
</tbody>
</table>

**NOTES**

1. In this and following quotations from Lerat (1997), we have translated the quote from French.
2. The original French text: "À un niveau plus global, celui du texte, les instruments linguistiques d’analyse restent très limités : la « grammaire de texte » obtient peu de résultats en dehors du traitement des anaphores, la typologie des actes de langage manque encore de robustesse, la notion d’environnement cognitif est nécessaire mais peu technique pour le moment, et elle relève plus de la logique que de la syntaxe proprement dite." (Lerat 1997: 3)
3. The original French text: “L'idée que la langue spécialisée est ancrée à la langue générale a maintenant fait son chemin et, à juste titre, on préconise désormais cette appellation de préférence à « langue de spécialité », précisément pour souligner l'idée d'un continuum et non d'une rupture...” (Resche 1999: 121)
5. The following examples were taken from the Earth Science Comparable Corpus compiled at University Paris Diderot.


7. We would have to conduct a detailed ethnographic survey to prove this. But, our knowledge of the local context suggests that the contributors and readers of ASp form a fairly cohesive, tightly-knit community within the French-speaking ESP world.

8. GERAS is a regular annual conference which usually takes place in France; the abbreviation stands for ‘Group for Study and Research on English for Specific Purposes’ (see <http://www.geras.fr/>).

9. For a similar study, but which did not concentrate on linguistic studies see Laffont & Trouillon (2013).

10. The editorial team of ASp have informed us that there have been more papers published in ASp in total than the 483 counted here. It is possible that our total is lower because we did not count certain contributions (editorials, work in progress, etc.).

11. SketchEngine is available at <http://www.sketchengine.co.uk>.


ABSTRACTS

This paper explores the various contributions that linguistics has made to English for Specific Purposes (ESP). First, we set out a broad distinction between concept-oriented and context-oriented approaches in linguistics. Although in theory these traditions are largely incompatible, there are signs that the pervasive use of corpus analysis is producing a convergence of the two. Then, we demonstrate a small number of key ideas based on the techniques and theoretical assumptions which we have developed in our own research and teaching: 1) use of corpora to identify “lexico-grammatical patterns”, the basic building-blocks of ESP texts, and 2) the use of phraseological training and databases to enhance the learning process and skill-set of future technical translators. Finally, we look at how other linguists have contributed to the general development of linguistic approaches over the history of the journal ASp.

Cet article explore les différentes contributions apportées par les études en linguistique à l’anglistique de spécialité. Nous établissons en premier lieu une distinction entre les approches conceptuelles et les approches contextuelles en linguistique. En théorie, ces approches sont incompatibles ; cependant, l’emploi de plus en plus répandu des corpus a généré une certaine convergence entre elles. Nous présentons ensuite quelques idées issues des présupposés théoriques et de la méthodologie que nous avons développés dans notre propre recherche et dans notre enseignement : 1) l’utilisation des corpus pour identifier les « schémas lexico-grammaticaux », qui représentent les unités de structure pour la construction des discours spécialisés, et 2) la formation à la phraséologie et à l’utilisation de bases de données comme facteurs d’amélioration, à la fois du processus d’apprentissage, et des compétences des futurs traducteurs spécialisés. Enfin, nous examinons comment d’autres linguistes ont contribué de manière générale au développement d’approches linguistiques dans l’histoire de la revue ASp.
INDEX

Mots-clés: anglais de spécialité, approche conceptuelle, approche contextuelle, formation de traducteurs, linguistique de corpus, phraséologie

Keywords: concept-oriented ESP, context-oriented ESP, corpus linguistics, phraseology, translator training

AUTHORS

CHRISTOPHER GLEDHILL

Christopher Gledhill is Professor of English Linguistics at the Université Paris Diderot (Paris, France) where he is currently co-ordinator of the first year Master’s in Applied Foreign Languages (Mention LEA) and co-coordinator with Natalie Kübler of a second year Master’s in Languages for Specific Purposes, Corpus Linguistics and Translation studies (Master 2 LSCT). He has taught and conducted research in interlinguistics, specialised translation, phraseology, systemic functional grammar and text linguistics. cgl@eila.univ-paris-diderot.fr

NATALIE KÜBLER

Natalie Kübler is Professor of English Linguistics and translation at the Université Paris Diderot (Paris, France). She was one of the first to introduce corpus linguistics applied to translation in a Master’s in specialised translation in France. She has taught and conducted research in corpus linguistics and LSPs, specialised translation, and corpus linguistics applied to language and translation teaching and learning. She is currently the head of the research team CLILLAC-ARP and of the Language Resource Center at Paris Diderot. nkubler@eila.univ-paris-diderot.fr