Geneviève Bordet et Elsa Pic
Universités Paris Diderot et Sorbonne Nouvelle

Recension

Alex Boulton, Shirley Carter-Thomas and Elizabeth Rowley-Jolivet (eds.). **2012.** *Corpus-informed Research and Teaching in ESP*. Amsterdam /Philadelphia: John Benjamins Publishing Company, 306 pages, ISBN 978-9027-0357-1.

Just like John Swales, who wrote the preface to the book, we have every reason to welcome the publication of *Corpus-informed Research and Learning in ESP*, the 52nd volume in the John Benjamins "Studies in Corpus Linguistics" series, co-edited by Alex Boulton, Shirley Carter-Thomas and Elizabeth Rowley-Jolivet. The names of the editors as well as the names of the other authors in the book will sound familiar to our readers as they are well-known for their research in ESP. So not only does this volume belie the sometimes supposed weakness of ESP research in France (compared to Italy or Spain, for instance), but, as one reads through it, it paints a true portrait of the "French way" of doing research in ESP. We develop this point in the concluding remarks to this review.

The book is made up of ten individual contributions, distributed into three well-balanced parts: "ESP corpora for language research"; "ESP corpora for genre-based approaches"; "ESP corpora for language teaching and learning". We review each contribution in turn before assessing the book as a whole, particularly in what it reveals about the specificity of ESP research in France.

Shirley Carter-Thomas and Angela Chambers – "From text to corpus"

This contrastive study of the use of personal pronouns in economics articles written in French and English aims not only at exploring the way writers build their authorial stance in their introductions but also at demonstrating that corpus linguistics can provide the basis for discourse analysis as applied to scientific writing.

In a text-to-corpus approach, the preferred patterns accompanying pronoun/verb combinations are first studied in a single text for each language. Their connection with Swales' CARS model is established, shedding light on the process used first to determine a territory, then occupy a niche of research. This first step provides the basis for the study of a bilingual corpus comprised of fifty introductions of research articles in economics, taken from the KIAP corpus (KIAP 2005). The study of concordances of combined verbs and pronouns leads to a classification and an analysis of the range of possibilities in choosing an authorial stance: author as a writer, as a researcher, as an arguer, based on the KIAP research team find-

ings. It also shows that, while uses of pronouns in French and English LSP are very similar, the first person singular is much more frequent in English, contrary to what is usually taught to ESL learners.

This paper goes far beyond contributing to the understanding of the array of techniques available to build a writer's stance according to the chosen genre, since it offers an in-depth analysis of the conditions under which findings based on corpus analysis can provide support for data-driven learning. Two stages are considered, the first one being the study of pre-selected texts and concordances, which is by far the easiest one to set up. The second step consists in giving learners direct access to data and use of concordancing. To be successful in helping the student build his/her own learning, this pedagogical approach requires a shift in both teacher and learner's roles, where the teacher stops being an expert to become a "research organiser" while the learner takes an active stance and defines his/her own queries. This reflection on corpus-driven learning thus paves the way for an interactive approach of the links between corpus study and pedagogical applications, where research choices determine pedagogical strategies.

Anthony Saber – "Phraseological patterns in a large corpus of biomedical articles"

Admittedly, biomedical articles have been the subject of extensive research, their IMRAD structure and terminology having been investigated by and large. Still Saber's contribution turns out to be both original and informative. Saber tackles the highly constrained genre of OBRAs (original biomedical research articles) through the study of phraseological patterns in relation to the article section and rhetorical steps they occur in.

Readers will appreciate the crystal-clear methodology section explaining the compilation of a 375 article corpus and the selection of key words refined into "section-specific salient words" around which relevant phraseological patterns are identified within each section. For instance, Saber identifies strong and stable phraseological patterns in the Results sections of OBRAs: description of clinical results by researchers appear to be highly standardised, salient words such as "observed" tending to cluster with the same words within the same kind of constructions. Similar predictions can be made within each step.

This contribution suggests very useful ways in which the outcome of corpus-driven studies could be used in EAP classes – medical students or novice researchers writing their papers in English will find in Saber's contribution a mine of information allowing them to combine the right phraseology with the right rhetorical step.

François Maniez – "A corpus-based study of adjectival *vs* nominal modification in medical English"

In this slightly more linguistics-oriented work, Maniez tries to shed light on the use of nouns or relational adjectives as noun modifiers in medical English, for instance, when to use "heart" vs. "cardiac", or "liver" vs. "hepatic".

After a welcome review of the literature on relational adjectives, Maniez draws attention on the pitfall they represent for ESL learners. But as Maniez himself points out, the predominance of those denominal adjectives in medical English is well-known and has been the subject of numerous research projects. The originality of the present contribution therefore lies in the different corpora used to try and identify the difference between nominal and adjectival modification: a comparison between general language corpora (CoCA; web-as-corpus: Google) and specialised corpora (EMEA; the academic medical subset of CoCA). The author contends that only the use of both types of corpora by the learners is a reliable way of selecting the appropriate modifier.

Natalie Kübler and Alexandra Volanschi – "Semantic prosody"

Kübler and Volanschi offer a study of the implications of semantic prosody phenomena on "pragmatic" or specialised translation i.e., the way in which the choice of negatively or positively connoted collocates modifies the perception of a term's environment. Their approach is doubly contrastive, comparing general and specialised language on one side, French and English on the other side. This implies the use of a doubly comparable corpus, comprised of the BNC and the CoCA corpus for English general language, and the Leipzig corpus Français for French, together with articles from the French paper *Le Monde*. Earth Sciences academic texts written in English and in French, collected at the University Paris Diderot, were used as a comparable specialised corpus. The use of this corpus gives the authors the opportunity to provide an interesting reflection as to the specific difficulties involved in the collection of a parallel or a specialised corpus, when one of the two languages (English) is dominant and the other is a more "peripheral" scientific language, such as French. The study focuses on two very common verbs in the scientific domain, "to commit" ("*commettre*" in French) and "to cause" ("*causer*" in French), and compares the type of semantic prosody they imply, in each language and in specialised and general language uses. An extensive study of concordances shows that despite the strong pervasiveness of semantic prosody, its effects tend to be smoothed in ESP, which can be deceptive, particularly for non-native speakers, tempted to transfer general language semantic uses to specialised language. Therefore, the authors insist on the necessity of offering scientific writers and learners a terminological database that includes collocational and prosodic aspects.

Beyond this local description, these findings open interesting perspectives for a reflection over the mutual influence between general and specialised languages, between French and English uses of scientific language and questions the way this mutual influence can and should be taken into account for an appropriate translation considering genre-specific objectives.

Elizabeth Rowley-Jolivet – "Oralising slides"

This chapter aims both at offering a multimodal analysis of a corpus comprised of complementary oral and text material and providing help to researchers for the preparation of scientific presentations. The analysis, based on a systemic functional approach, uses the concepts of grammatical metaphor and communicative metafunctions to consider the differences, similarities and complementarities in the way presenters address the audience, convey the informational content and build a two-folded cohesion, combining oral and textual information. The corpus studied is based on filmed presentations. The slides and commentaries are considered as two synchronous parallel discourses. Results show that while slides focus mainly on ideational content, the oral presentation mostly performs the textual and interpersonal functions. However, presenters are faced with a choice of various rhetorical strategies as to the type of complementarities between written and oral information, which requires a high level of flexibility in the use of lexicogrammatical resources.

One of the striking points of that study is that it succeeds in combining a theoretical and a pragmatic approach, confirming the relevance of the systemic functional framework as applied to naturally occurring data, and providing support to ESP learners and researchers for the promotion of their research. Of specific interest is the demonstration of the importance for a scientific presentation of mastering the handling of grammatical metaphors, both in their congruent and non-congruent versions, in other words in the capacity of alternatively "packing" and "unpacking" scientific metaphors in the parallel and combined written and oral presentations.

Céline Poudat and Peter Follette – "Corpora and academic writing – a contrastive analysis of research articles in biology and linguistics"

This contribution somehow stands out from the others, since it explicitly aims at presenting the methods and tools developed in France for the quantitative analysis of corpora. French readers might already be familiar with the text statistics software and techniques mentioned; otherwise this article constitutes a very good introduction to the state of the art in this country. The interest and efficiency of those tools and methodology (hypergeometric distributions, correspondence factor analysis) are then illustrated through a contrastive analysis of biology and linguistics articles, which (expectedly?) reveals important differences between an experimental laboratory discipline and a discipline belonging to the humanities. But Poudat and Follette insist on the larger "cultural" differences and traditions underlying the two disciplines and how these play an important role in the acquisition of the lan-

guage used, which leads them to question the very possibility of teaching "general" EAP. Given the differences across disciplines and, within disciplines, across genres, teaching would have to be very specific to be efficient.

Dacia Dressen-Hammouda – "Construction of discoursal expertise"

This corpus-based genre analysis offers an innovating method to investigate the way researchers' uses of genre-specific norms evolve over time as their discoursal expertise builds up. It addresses the issue of norm and genre, i.e., the common range of implicit values which determine the reception of a text by a specific discourse community, and the extent of acceptable variations.

The study compares a diachronic corpus comprised of "field accounts" produced by five geologists over a ten-year period after their PhD, with another larger "control" corpus of the same part-genre. The latter corpus is used to identify thirteen variables which can be considered to be typical of the part-genre. These variables and their evolution over time are then studied in the diachronic corpus. The variables are classified along three types of cues: "personalization" (e.g., first person pronouns); "doing-the-work" (e.g., metric measures or locational adverbs); "disciplinarity" (e.g., field descriptors or technical verbal adjectives"). They are then used to measure the evolution of "standard deviation" from the norm, as determined from the "control" corpus, each variable being studied as to its range of variation according to the writer's level of experience. The method used allows for a factual evaluation of the influence of the writer's level of expertise on his/her interpretation of the norm. Results show that experienced writers tend to have more flexibility in bending the norms and often use it to highlight their personal role and build their "voice of expertise".

Besides shedding light on the decisive and difficult issue of interactions between genre norms and individual creativity, this research will help pedagogical practitioners understand how to support L1 and L2 learners in raising their awareness of the type of discoursal shift they will have to achieve to get their voice heard in the central scientific publishing world.

Geoffrey Williams – "Bringing data and dictionary together – real science in real dictionaries"

Geoffrey Williams's contribution concentrates on how general language dictionaries could be improved thanks to corpus data, that is, how entries could be extended ("tweaked") using prototype theory to take into account specialised usage. To complement such "organic" dictionaries (i.e., corpus-driven ones) for encoding purposes, phraseology will be taken into account as well.

Indeed, most advanced learner's dictionaries based on static, outdated wordlists are of no use for production purposes. Williams convincingly shows how the adoption of prototype theory and mind-mapping for the design of on-line dictionaries could

result in much more accessible and interactive tools, with "overlapping" entries rather than alphabetical ones.

Susan Birch-Bécaas and Ray Cooke – "Raising collective awareness of rhetorical strategies"

This chapter describes an on-line writing tool designed to provide assistance to non-native speakers in writing a scientific paper in English. The difficulties resulting from that situation are first analysed, pointing out the contradiction between the necessity of reproducing conventional scientific style and avoiding pure plagiarism. The conception of the writing tool is based on the idea that learners first have to be aware of the existence of rhetorical strategies and moves. This can be achieved through exposition to a variety of genres (e.g., case reports, letters of requests, research papers) in various disciplines, and their successive drafts, thanks to an entirely free and open on-line tool. This chapter presents a specific pedagogical sequence where TYOS ("Type your own script") is used to help French doctoral students draft the introduction of a scientific research paper in English.

In a first step, students are encouraged to identify and discuss disciplinary discourse variations. This is the opportunity for a first approach of fundamental scientific tenets, such as the complementarities of discipline-specific lexis and lexico-grammatical items, presented as "bricks" and "cement". Students are then introduced to move analysis, through a comparative approach of texts. "Fill in the gap" strategies are used to help them build their own introduction, taking into account the knowledge acquired in the previous steps. Students' "writing propositions" are checked using the TYOS corpus and concordancer, as well as literature in their own discipline found on the Web. The conclusion emphasizes the importance of providing students with attractive and convenient support for controlled information transfer from existing authoritative literature to their own production.

Besides offering a pragmatic and directly efficient tool for teachers and learners of scientific writing, this research provides interesting leads for the underlying discussion about the type of help that should be given to learners, specifically non-native learners, confronted with the necessity of producing a specialised discourse, in a foreign language. Should these students be taught to mimic the existent mainstream discourse, or should they be given the possibility to create their own academic voice without erasing their individual and cultural specificities? The authors claim that students should "have at their disposal bedrock of occurrences and clusters that they may learn so as to be able to branch out later and be more linguistically adventurous once the basics are in place". This is at least a stimulating though controversial proposition for an approach of the dilemma between gaining acceptance from the desired community and getting one's specific voice heard by that community: a dilemma which is not specific but even more difficult to face for non-native speakers.

Alex Boulton – "Corpus consultation in ESP: a review of empirical research"

Boulton's chapter focuses on the assessment of data-driven learning, or how language learners use corpora. Boulton immediately admits that no conclusion can be drawn from his survey of twenty different studies representing empirical evaluations of corpus use. Indeed, the studies differ widely, be it in the number and background of participants or in the type of skills under scrutiny, some being centred on vocabulary acquisition, others on writing skills. So no meta-analysis is possible, which is somewhat frustrating for the reader. But when looking at the short and long-terms benefits felt by the students, it is comforting to see that most are generally favourable to the use of corpora.

All in all, the book lives up to its promise and the contents do reflect what the title announces, except that ESP could easily have been replaced by **EAP**, given that all the papers deal with scientific/academic writing. This could be due to the fact that the focus is set on learning and teaching at *university*, which, in France, does not teach technologies or give professional training.

Still, what we found really interesting were the invisible threads connecting the papers around central notions in today's ESP research.

All the papers deal with corpus linguistics; most combine it with an approach of **texts** as part of a specific discourse and **genre**, and highlight their importance in the choice of corpora (see Kübler & Volanschi for instance, on the difficulties in collecting comparable corpora in two languages, including English). Most authors are also interested in **rhetorical strategies**, be it in relation to phraseology, lexico-grammar or stance-building.

All the contributors acknowledge the importance of combining a **fine-grained analysis** of the lexicogrammatical texture (the "cement" in Birch-Bécaas & Cooke's terms) and the wider scope of the **functions** of discourse. Several shed light on the connection between the selection of specific lexicogrammatical traits and the writer's building of his/her stance as a researcher.

The importance given to **verbs** and their role as "cement" between the "bricks" of specialised lexicon and functional terms is striking (Carter-Thomas & Chambers, Rowley-Jolivet, Kübler & Volanschi, Williams).

In most articles, the **porosity between ESP and general uses** of language (Kübler & Volanschi, Williams) is a salient feature. Some are original enough to consider the mutual influence between general and specialised discourse rather than focusing on differences alone.

Last, **the study of NNS learning difficulties** is viewed as a contribution to a better comprehension of the nature of genres and of specialised discourses. This resonates with the impossible choice for ESL teachers: should they teach students to conform to the dominant type of discourse and gain acceptance in the community; or should

they try to provide a type of support which "creatively informs the authors' own choices in writing" (Carter-Thomas & Chambers)? The authors of the book share this questioning about non-native and learner writers' peripheral status as compared to mainstream centres of knowledge: should ESP research focus on helping learners to imitate the dominant discourse or is it possible to use the findings of combined corpus linguistics and discourse analysis to provide them with tools to get their own voice heard while respecting its cultural specificities? This involves an in-depth reflection on interactions between genre and norm (Hammouda), between lexicogrammatical choices and rhetorical strategies (Carter-Thomas & Chambers, Rowley-Jolivet, Birch-Bécaas & Cooke, Kübler & Volanschi). Therefore, clarifying the question of the integration in the central world of research of non-native speakers of English might strongly contribute to the understanding of the nature of ESP. This could be the most important if somewhat implicit message of a book which offers both stimulating issues and practical material to researchers and teachers of ESP in native and non-native languages.

Electronic web site references

EMEA. European Medicines Agency. Undated. Opus. Retrieved August 2012 <http://opus.lingfil.uu.se/EMEA.php>.

KIAP Corpus. 2005. University of Bergen. Retrieved August 2012 <http://kiap.uib.no/KIAPCorpus.htm>.