# Language learning as discourse analysis: Implications for the LSP learning environment

## Angela Chambers

This text was automatically generated on 19 April 2019.

# Language learning as discourse analysis: Implications for the LSP learning environment

**Angela Chambers**

## 1. Introduction

1   From the mid-twentieth century onwards, three parallel developments in language studies have contributed to shaping the current landscape in Languages for Specific Purposes (LSP). Firstly, discourse analysis has become an increasingly influential area of research within a number of disciplines, including sociology, sociolinguistics, psychology, anthropology, and applied linguistics. Secondly, the communicative approach to language learning and teaching, with its emphasis on practical communicative activities, has led to a need for genuine examples of language use, both written and spoken, to be made available to language learners. The Web provides an excellent source of such material. Thirdly, technological advances have also made developments in corpus linguistics potentially available to all the actors in the language-learning process, including the producers of materials and resources, teachers, and even learners. This may appear to represent an ideal learning and teaching environment for those involved in specialised language use, with easy access to individual examples of specialised texts (used here to refer to both written and spoken language), readily and freely available corpora permitting the analysis of discourse patterns across texts, and publications on language learning and teaching as discourse analysis providing guidance to teachers who are not experienced discourse analysts.

2   However, two key phrases in the above lines, "parallel developments" and "potentially available", point to the fact that these three developments have not yet been integrated into the language-learning and teaching environment. While the communicative approach has fundamentally altered this environment, the integration of discourse analysis and corpus linguistics has been a slower development.

3    This paper aims to investigate this situation in two ways. Firstly a number of publications from the late 1960s onwards will be examined to show how the study of individual texts as discourse can be of particular relevance in the LSP context. This will then be extended to the study of discourse across texts referring to data from a number of corpora in areas such as academic writing, economics, and business discourse. Finally this will enable us to envisage the type of learning environment which could facilitate the development of language learning as discourse analysis.

## 2. Discourse and LSP

4    Within linguistics and applied linguistics the research literature on discourse and discourse analysis is vast, but a small number of seminal texts lay the foundations for the development of a discourse-based approach to LSP. The first is Foucault's claim that one can no longer

> traiter les discours comme des ensembles de signes (d'éléments signifiants renvoyant à des contenus ou à des représentations) mais comme des pratiques qui forment systématiquement les objets dont ils parlent. (1969: 67)

5    For those involved in teaching and researching academic discourse, the author of a study of Foucault's work inadvertently asks a fundamental question.

> *Les mots et les choses* [...] repérait les mécanismes de pouvoir à l'œuvre dans les discours scientifiques, en exhumant les règles auxquelles ceux-ci se trouvent contraints d'obéir : mais de quelle sorte « d'obéissance » s'agit-il ici ? Sans doute peut-on soutenir que parler français, c'est « obéir » aux règles de la grammaire française, mais cette « obéissance » est-elle du même ordre que celle du soldat qui exécute les ordres d'un officier, et quel sens y a-t-il, autre que métaphorique, à interpréter ces deux situations en termes de pouvoir ? Or, le cas du discours scientifique s'apparente beaucoup plus à celui de la langue qu'à celui du soldat. (Bouchard 2003: 494-495)

6    The theme of GERAS 2007, and the text accompanying and problematising it in the conference documentation, make it clear that speaking or writing English or any other language entails much more than obeying the rules of grammar. Underlying this study, and arguably all studies of language learning in general and LSP in particular, is thus the question of whether, in teaching our students to master a specific discourse, we are treating them like soldiers trained to obey orders, or tacticians able to work out the best way to deal with the various discourses which they will encounter in their future careers. This issue will be raised again later in this study.

7    The link between LSP, discourse analysis, and corpus linguistics developed slowly but steadily throughout the second half of the twentieth century. Works such as Munby's *Communicative Syllabus Design* (1978) emphasised that the content of language teaching should focus on the needs of the learners. The link between the teaching and learning of LSP and research on discourse was also developing from the 1960s. In *Episodes in ESP* (1988a), Swales included what he considered to be the key texts in the area. Although a small number of the chapters focus on discourse and genre (see, for example, Swales 1988b; Tarone *et al.* 1988), it would be fair to conclude that discourse plays what can best be described as an emerging role in the development of ESP from 1962 until the first publication of Swales's book by Pergamon in 1985. Indeed in the conclusion, Swales (1988a: 211-212) notes the importance of a text-based approach, but goes beyond this to emphasise context as a direction for future development, citing publications which focus

on academic, disciplinary, professional, and occupational cultures and sub-cultures. As we shall see, his prediction was certainly well-founded.

8    In 1990 another publication by Swales was to play a leading role in linking language learning and discourse, particularly in the context of LSP, namely *Genre Analysis: English in Academic and Research Settings.* His definitions of the characteristics of a discourse community include participatory mechanisms of intercommunication, mastery of a specific lexis, and the use of one or more genres. More importantly, for Swales a defining characteristic of a discourse community is "a threshold level of members with a suitable degree of relevant content and discoursal expertise" (Swales 1990: 27). This raises two fundamental questions: what is "discoursal expertise" and how does one teach it to nonnative speakers who are novices in the discourse community? Is it only training in obedience to the norms in the target language community, accepting them as models created by others (i.e., native speaker experts) or can the nonnative speaker participate fully as a member of the discourse community and contribute to the process of the development of those norms? Clearly the answer is yes, as many researchers publish extensively and successfully in languages other than their native language, principally in English. For the nonnative speaker, however, the period of apprenticeship is more difficult, and the status of nonnative speaker may inhibit confidence and creativity. When recently choosing the title of a corpus of academic writing in French, for example, I followed the advice of native speaker colleagues from France and did not yield to the temptation to use the expression "*discours académique*", despite more than 10,000 Google results, many in a Canadian context. As a perhaps cowardly nonnative speaker I chose " *articles de recherche*" (Chambers & Le Baron 2006), which corresponds to the norm in the European discourse community. But was I perhaps also motivated by fear of my choice being interpreted as a nonnative speaker error or, worse still, a native English speaker's preference for an Anglicism? The nonnative speaker's or language learner's progress as a member of a discourse community can be situated on a continuum on which the following milestones can be recognised:

- Introduction to the characteristics of the discourse community: lexis, genres and participatory mechanisms
- Familiarity with those characteristics and ability to use them following the accepted norms
- Awareness of the choices to be made when operating within the discourse community and of the implications of those choices
- Confidence to choose a non-standard element while otherwise conforming to the norms of the participatory mechanism in question.

9    If one accepts the existence of this continuum, "discoursal expertise" entails not only the ability to use the participatory mechanisms and genres within a discourse community but also the ability to understand aspects of the reasons for their existence and the implications of changing them.

10   In LSP research and practice the discourse-based approach developed in two main ways, the qualitative study of one or a limited number of texts, and the combined qualitative and quantitative study made possible by developments in corpus linguistics. Firstly the study of individual texts or groups of texts to inform the teaching and learning activity was already in evidence from the 1960s onwards, but the study of classroom applications gained momentum in the late '80s and in the '90s.[1] In particular the needs of learners clearly influenced the selection of the types of discourse being analysed, and genre-based studies of language use in specialised settings multiplied (see, for example, Sager *et al.*

1980; Gunnarsson 1993; Henry & Roseberry 2001; Flowerdew 2002; Hyland 2002). The first issues of *ASp* appeared at this time, focusing on the one hand on the educational environment (see, for example, Mémet 1993; Bertin 1994; Chénik 1994), and on the other on the analysis of scientific texts (Petit 1993; Corbisier 1994). This double focus is still evident in LSP research in general and in the context of *ASp* in particular, with the analysis of specialised genres playing an important role (Carter-Thomas & Rowley-Jolivet 2003; Resche 2003; Laffont 2006; Percebois 2006). It is interesting to note the continuity in the focus on this area of research, as the title of Régent's (1994) article is particularly appropriate to the decision to focus on the cultural environment in the theme of the 2007 conference, "*L'article scientifique : un produit culturel*".

11    In several of the publications on genre mentioned above (see, for example, Carter-Thomas & Rowley-Jolivet 2003; Resche 2003), the analysis of a collection of texts or a corpus has clearly formed the resource on which the study is based. In other words, the discourse-based approach to language learning and teaching involves not only the study of individual texts but also discourse across texts. This does not include corpus-based analysis using concordancing software, and thus belongs in the first of the two categories as defined above, namely the qualitative study of one or a limited number of texts. As this study focuses on the implications of research for the language-learning environment, the question arises as to how the LSP teacher can apply the findings of this research? One possibility is to use them to inform the study of one or a very limited number of individual texts in class. For example Carter-Thomas & Rowley-Jolivet (2001; 2003) contrast the syntactic choices of the authors of spoken and written scientific texts, conference presentations and research articles. They note the preference for passive constructions in the research article and the relative absence of these in the conference presentation.

| Article | Presentation |
|---|---|
| It is possible to define | We can define |
| It can be implied that | We know |

(Carter-Thomas & Rowley-Jolivet 2003: 62)

12    As the articles and presentations used in the corpus are listed, a teacher could use selected texts in class and encourage the learners to discover these preferences, or indeed use other written and spoken texts corresponding to the specific field of study of the learners. By encouraging the learners to become aware of the characteristics common to texts in a specific genre, the teacher is thus encouraging in them the awareness of discourse variety and thus by implication the ability to transfer the knowledge acquired to the study of other types of discourse and to acquire the competence to perform in those genres. As Thompson (2001: 33) points out,

> students should be encouraged to view genres as potential forms, but not as models for writing; the analysis of exemplars of a genre should be considered in terms of the rhetorical choices that a writer had to make and the forms of language and organization that were available to help the writer to achieve the purpose.

13    This corresponds closely with the tendencies evident in current research on language learning and teaching, well summarised by Hyland (2002: 120) when he refers to "the

need to find ways for students to take more active, reflective and autonomous roles in their learning".

# 3. Discourse across texts: empowering the learner

14  Alongside this type of research, developments in corpus linguistics have made it possible for individual teachers and learners to consult a corpus directly, becoming independent researchers. The question thus arises: has the integration of corpus linguistics and language learning enjoyed the same success as the integration of discourse analysis and language learning? This is not an easy question to answer. It is initially tempting to reply that yes, it has developed along similar lines, perhaps a little later than the linking of discourse analysis and language learning. In 1985, for example, when Swales was publishing *Episodes in LSP*, he was able to include a number of chapters and references to work linking ESP and the concepts of discourse and genre, and to confidently predict that this connection would become more important in the future. He was understandably not able to predict that access to corpora by teachers and learners would be an important future development.

15  The main reason for this delay is undoubtedly that corpus linguistics was seen for several decades as a methodology for language research, not teaching. The early corpora, such as the one-million word Brown and LOB corpora were thus created as resources for researchers, and it is now commonplace for dictionaries, such as *The Oxford-Hachette French Dictionary* (2001) and the *Collins Cobuild* dictionaries and reference books to be based on vast corpora of several million words. The development of the use of corpora as a resource with learners, however, was a slow one. The first achievement was the influence of corpora on the content of language teaching, on what was taught. Carter and McCarthy (1988: Ch. 3) stressed the importance of ensuring that the language taught corresponded closely to actual language use, in parallel with the emerging role of corpora as the basis for dictionaries and grammars. In later publications they emphasise that spoken language has its own grammar, that it is not an inferior variety of written language, and that, for example, "Anybody want soup?" is not bad English for "Does anybody want soup?" (2006: 182). Interestingly the earliest attested use of corpora in the classroom was in the LSP context. According to McEnery and Wilson (1997: 12), "Peter Roe started to use LSP corpora in teaching at Aston University in 1969". Johns's article in *System* (1986) and Tribble and Jones's (1990) resource book for teachers, re-edited in 1997, were early landmarks in the introduction of corpora into the classroom environment. Tim Johns's web page also played an important role in disseminating information on the ways in which corpora could be used to create materials for language learning. A substantial number of publications deal with the potential of using corpus data in the classroom (see, for example, Hunston 2002; Sinclair 2004; Wichmann *et al.* 1997), and a smaller but still significant number of publications exist which report on the findings of quantitative and qualitative studies of corpus consultation by learners. Examples of studies in the LSP context include Stevens 1991; Cobb 1997; Johns 1997; Gaskell & Cobb 2004; O'Sullivan & Chambers 2006; Lee & Swales 2006. There is a lot of similarity between the findings of these studies, particularly in relation to the design of the experiments and the evaluation of the learners. Three aspects are particularly relevant in the present context. Firstly the learners have very similar positive and negative reactions to the experience of corpus consultation. Secondly the majority of the researchers, although

working in the context of English where a significant number of large corpora are freely available, choose to create their own smaller corpora. Finally, and perhaps most significantly in the context of the LSP learning environment, the small corpora created by the researchers are generally not publicly available. These aspects of the studies of corpus consultation by learners deserve more detailed attention here.

16    The majority of the learners in these studies reacted positively to corpus consultation as a learning activity, appreciating two aspects in particular. Firstly they liked having access to a large number of authentic examples of the aspect of language use which they were studying, (Cheng *et al.* 2003: 181; Yoon & Hirvela 2004: 275; Chambers 2005: 117). Secondly they enjoyed the exploratory nature of the activity, illustrating the phrase which Johns used to describe learners working with corpus data: "Every learner a Sherlock Holmes". (Johns 2002: 108). Johns was not referring to direct consultation of the corpus by learners, as in these studies, but to indirect corpus consultation, where the learner analyses corpus data provided by the teacher. McEnery and Wilson (1997: 6) point out that this is already "active participation in the process of learning", but it is nonetheless limited in the level of autonomy it allows the learners to exercise, in that they cannot decide what searches to undertake in the corpus. One learner wrote of the process of directly consulting a corpus: "I discovered that achieving results from my concordance was a highly motivating and enriching experience. I've never encountered such an experience from a textbook" (Chambers 2005: 120). It is important to note, however, that this enjoyment of corpus consultation was tempered by the negative reactions which are also evident in a number of studies (Cheng *et al.* 2003: 182-3; Yoon & Hirvela 2004: 274; Chambers 2005: 120), namely that the learners found the activity of analysing the corpus data time-consuming, laborious and tedious. As we shall see, one possible solution to this is to integrate the consultation of corpus data with the study of individual texts.

17    The second characteristic shared by the studies of learner consultation of corpora is that the majority of the teachers/researchers created corpora specifically for their students. Notable exceptions are Bernardini (2000), Cheng *et al.* (2003), and Lee & Swales (2006), who advocate large corpus concordancing by learners, although it is important to note that they have at their disposal a significant amount of class time to train the students in corpus consultation. Johns (1997: 103) created a three-million word corpus which he describes as follows:

> a 3-million word corpus chosen to give a rough reflection of an overseas student's life in Birmingham, both on and off the campus, including lectures, scientific journals, newspapers, 'domestic' texts on childcare and household management.

18    This corpus includes the types of texts typically found in these small corpora, namely familiar texts of a general nature and texts belonging to one specialised genre, although Johns includes both in his corpus, while most of the other researchers choose one. Stevens (1991), for example, used texts from the course books of his physics students. While these corpora are not accessible, there is now a trend to create small corpora for learners and make them publicly available, either by directly creating them, as in the case of the *ELISA* corpus (2004) of individuals from several English-speaking countries speaking about their lives, or by extracting them from larger corpora, as in the case of *BNC Baby* (Burnard 2004), a four-million word corpus extracted from the British National Corpus and including four sub corpora of academic writing, journalistic discourse, literary texts, and spoken language. The *ELISA* corpus is freely available on the web, while *BNC Baby* is available for a modest price.

19    The easy availability of corpora such as these has important implications for the language-learning environment in general and the LSP learning environment in particular. It is a necessary first step in popularising corpus consultation by language teachers and learners. It means that teachers who wish to incorporate corpus data in their teaching do not immediately come up with the obstacle of the lack of availability of relevant corpora. On the negative side, however, the number of such corpora, what Braun (2005: 47) terms "pedagogically relevant corpora" in contrast to corpora created specifically for research, remains very limited, with the result that, while a teacher can easily adopt a discourse-based approach to the study of individual texts, complementing this with a corpus-based study of discourse across texts will still immediately encounter the problem of the lack of easy access to corpora suitable for learners, particularly in the LSP context. In their EAP course, Lee & Swales (2006) provided a substantial training course to their postgraduate students on the creation of genre-specific corpora specially tailored to meet their own needs. A similar practice would clearly meet the needs of many LSP teachers, although it is perhaps unrealistic to expect large numbers of teachers who are not corpus linguists to undertake the necessary training to enable them to create and subsequently update appropriate corpora.

20    An example concerning the use of corpora to investigate academic writing will illustrate both the usefulness of a small corpus of this nature and the problem of the lack of easy access to appropriate corpora. In an effort to improve the level of academic writing in French of my language students on the BA in Applied Languages in the University of Limerick, I have created a one-million word corpus of academic French, the *Corpus Chambers-Le Baron d'articles de recherche en français (CLAREF)* (Chambers & Le Baron 2006), which will be available via the Oxford Text Archive. The corpus contains 159 research articles from 20 journals in ten areas within the Humanities, including economics, history, ICT, law, linguistics and language learning, literature, media/culture, philosophy, social anthropology, and social sciences. Weary of seeing extensive use of the first-person singular in student essays ("*Comme j'ai déja dit plus tôt*", for example), I have examined the use of the first-person plural in the 147 single-authored articles in the corpus and prepared worksheets on the use of *nous, notre* and *nos.* This provides learners with a wealth of information on the metadiscourse of academic writing, including not only the variety of verbs used with *nous* (*nous analyserons, nous nous pencherons sur*, etc.), but also expressions such as "*dans un premier temps*", "*dans un deuxième temps*", and many other lexicogrammatical patterns which can be observed including *nous.*

21    Randomly selecting 150 of the 3,117 occurrences of *nous* makes it possible to classify its use in the following way:

- *nous* referring to the author alone
- *nous* referring to the author and the readers
- *nous* referring to the author and a wider public.

22    Examples are listed below.

Nous avons fait ce choix parce que nous voulions obtenir des données comparables pour les deux périodes étudiées
Pour conclure, nous reviendrons sur la méthodologie employée
En conclusion, nous pouvons nous interroger sur les raisons profondes de cette pluralité
Nous aurons l'occasion de revenir plus en détail sur la structuration communale
Burkina Blues nous entraîne alors dans le voyage de la liberté
Nous savons aujourd'hui qu'ils incarnent les vestiges de grandes civilisations

23    A detailed analysis of the use of *nous* in the corpus is beyond the scope of this article. It is important to note, however, that simple concordances based on this corpus can be used in a variety of ways. A search of "*comme nous*" for example, reveals 67 occurrences (three of which can be omitted as they are not part of the metadiscourse of the articles), and a much richer variety of expressions than the simple examples which I had proposed to the students, namely "*comme nous l'avons vu plus haut*" and "*comme nous le verrons plus loin*". The intention is that the "massed concordance examples", to borrow Cobb's phrase (1997: 303) will promote "noticing" by the students, in the sense in which Schmidt (1993: 217) uses the term in the language-learning context. In the 47 references to the past, the eight occurrences of *voir* are easily outnumbered by a variety of other verbs, such as *souligner* (4 occurrences), *montrer* (3), *expliquer* (2), *indiquer, exposer, signaler, tenter de le montrer* (one occurrence each). It was thus easy for me to produce a handout with a concordance of "*comme nous*", illustrating the variety of ways in which native speakers of French use it in academic writing to refer to earlier and later parts of research articles. A few examples are listed below.[2]

    Comme nous l'avons indiqué plus haut, les premières critiques furent formulées
    Or, comme nous l'avons vu, ces formes de consommation coexistent
    Comme nous l'avons constaté, la participation d'un nombre croissant d'organismes
    Le gouvernement américain [...] prend, comme nous l'avons vu, des mesures
    Comme nous le verrons par la suite, il est difficile d'attribuer à chacune de
    Il s'agit d'une décennie qui a débuté, comme nous le verrons, avec un accroissement

24    Concordance data such as these also have the advantage that they not only enable some learners to correct basic errors, but also provide useful examples of interest to learners who have already mastered the basic conventions. In addition, as these students have some limited training in corpus consultation, they can be encouraged to search for other uses of *nous* in the corpus, such as the verbs which accompany it, the tenses used, etc.

25    While individual corpora such as this can provide a valuable resource for learners, the limitations of the current environment quickly become clear if one decides to encourage learners to undertake a comparative study of language use in one disciplinary area in the native language of the learners and the target language, in this case articles on economics in English and French. While *BNC Baby* includes a subcorpus of academic writing, it does not include articles which could be used as the basis of a comparison with the *CLAREF* articles. To provide the basis of a comparison of articles on economics in English and French, a number of articles from two online journals similar to those in *CLAREF*, *Economics and Politics*[3] and *The American Journal of Economics and Sociology,*[4] were used as a corpus.[5] This provided the possibility of comparing language use in approximately 100,000 words in each language. It is important to note that this is not recommended practice (Wynne 2005) as it does not ensure the reusability of the resource if Web references change or access is no longer available. It is nonetheless a strategy which, in the absence of available resources, can be used by those who wish to use corpus data in their teaching and cannot easily access the resources which they require.

26    A search for "we" in the single-authored articles among the English texts (approximately 75,000 words) reveals that it is indeed much less common than in a similar number of words in French, 64 occurrences (excluding 29 in quotations) as opposed to 134 in French (excluding one in a quotation). Unlike the French data, the majority of the uses of the first person in English are inclusive in nature, referring to the author and the readers, or a wider public, as in the examples below.

> Redirection in major public policy seldom occurs until we are in a crisis
> If as a nation we are concerned that such high dependance is far too risky
> we face not so much a question of technology policies as of geopolitical
> As we will see below, the international arena provides considerable evidence
> To answer this question, we must be more precise about what "new" is

27 Indeed, there are only five clear examples of "we" used to represent a single author.

> In the following analysis we focus upon these eight countries
> we now move to analyzing more specific materials
> sharing customs and practices to initiate exchange, just as we have described
> We pursue this issue further here for our four ICT-intensive growth countries
> Here, we will analyze the Kula Ring gift exchange

28 In a further two occurrences, the use of "we" is ambiguous and could refer either to the single author or to the author and the readers.

> We considered evidence demonstrating that heterogeneous individuals
> even if it was methodologically sound, evidence that we have received here

29 A comparison of these small collections of texts can thus be used by teachers and students to illustrate aspects of the use of the first-person plural in academic writing in French and English, not as a model to follow but rather as a way of exploring how it is used. While the limited number of words makes it impossible to draw firm conclusions about language use in economic writing, the data nonetheless provide learners and teachers with substantial numbers of attested examples to inform their development of discoursal expertise in a given discourse community.

30 Even a small collection of articles of only 100,000 words can be used in a large variety of ways with learners. For example, concordances of a number of nouns selected from the frequency list of the economic articles in English reveals a very high number of occurrences of nominal compounds. Of the 79 occurrences of "budget" for example, only thirteen involve the use of the noun on its own. Nominal compounds are much more common, as shown in Table 1.

**Table 1: Nominal compounds with "budget"**

| Nominal compound | Occurrences |
| --- | --- |
| Budget rule/s | 25 |
| Budget constraint/s | 14 |
| Budget cap | 3 |
| Budget deficits | 3 |
| Budget requirements | 3 |
| Budget size | 2 |
| Budget passage | 2 |
| Budget process | 2 |

| Budget laws, restrictions, years, expansions, limits, data, officers, legislation, disputes | One each |
|---|---|
| TOTAL | 63 |

31 In addition the concordance includes the expressions "off-budget accounts", "budget-making" and "budget-constrained".

32 A teacher could use this concordance to create worksheets on nominal compounds involving "budget". Examples are provided below.

> States with stricter balanced budget requirements were predicting deficits that were 20% lower as a percent
> To determine what constitutes a strict, enforceable balanced budget rule requires some background.
> at first glance there appears to be little difference between, say, a balanced budget rule that allows deficits to be carried over and a no-carryover rule
> This dummy variable is then modified to account for budget passage and veto rules that affect the impact of the governor
> to consider the effects of an executive with veto authority, a key feature of budget-making
> institutions on government spending is evident in recent estimates of state budget deficits
> While there is an extensive literature linking state budget rules to deficit spending, there is relatively little work probing the

33 Alternatively a learner with a basic command of corpus consultation could produce the concordance of 79 occurrences and use the expressions as an aid in writing. Similarly interesting results can be obtained from concordances of a number of other common nouns. Of the 95 occurrences of "resource", for example, 85 are nominal compounds. Examples are given below.

> even mass privatization in transition countries – one of the largest shocks to resource allocation in economic history –
> First, technical progress may be resource augmenting (possibly stimulated by rising resource prices)
> Osborne saw the possibilities of a technology-based concept of natural resource availability,
> contain the key elements of Zimmermann's technology-based concept of natural resource availability.
> Recent advances in green national accounting mean that estimates of domestic resource depletion are readily available; (Morse 1963: 75)

34 The texts from which these examples are taken were put together as an illustration. It is easy to envisage a small corpus of research articles in economics created to correspond specifically to the content of courses in a number of universities, thus providing teachers and learners with easily available resources without the need for each individual teacher to acquire the necessary training and create his or her own corpora. As Gavioli (2005: 55) puts it, ESP and small, specialised corpora are "a happy marriage". As we have seen, Lee & Swales (2006) are already exploring the potential of this type of corpus creation for nonnative speaker researchers.

35 Although the creation of what Tribble (1997) has termed "quick and dirty" resources such as the English texts referred to above is not recommended practice, it is at least possible. Easily available corpus resources for spoken language are much more difficult to acquire

for a number of reasons. Firstly, it is much more labour intensive to create a spoken corpus because of the time necessary for transcription. Secondly, as a result of this it is more expensive. Thirdly, as a result of these two factors the creators of spoken corpora are less inclined to make their corpora easily available to others than their counterparts who create written corpora from texts which are already digitised. Thus teachers and learners interested in the use of the first-person plural in spoken business discourse will face insurmountable obstacles in acquiring data, although they exist. In a doctoral thesis recently completed in the University of Nottingham, Handford (2007) analyses CANBEC, a one-million word corpus of business meetings in English, created as part of a large project being undertaken by Cambridge University Press, namely the creation of the Cambridge International Corpus. The thesis includes an analysis of the use of "we" which was identified as one of the keywords in the corpus. Handford's analysis reveals, for example, that it is more frequent in external meetings than internal meetings, and that it is used in a number of ways: inclusive referring to all those present in a meeting, or to both companies; and exclusive referring to only one group of people at a meeting or to one of the companies represented at the meeting (2007: 109-115). He also identifies "we need to" and "we gotta" as common ways of expressing commands rather than the imperative mood (2007: 264). While the findings of this project will be available to teachers and researchers in the form of the thesis and publications, and while the data will no doubt influence the content of dictionaries, grammars and course books published by Cambridge University Press, the type of cognitive learning described above will not be possible.

## 4. Conclusion

36    For the corpus data which are easily available, the question arises as to how they can best be integrated into the learning environment. In the current environment, the absence of suitable resources is not the only obstacle. As we have seen earlier, students evaluating the activity of corpus consultation found the data analysis time-consuming, laborious and tedious. Römer (2006: 127) notes that, "we have to consider whether the tools that are currently available are easy enough to use for learners and teachers who, obviously, have not had the same training as the corpus-linguistic researcher". For a number of teachers/ researchers, the answer lies in providing students with corpus data prepared by the teacher. Tim Johns went a stage further by making such data available on the Web. The potential of activities such as this in the LSP context is clear, particularly in a collaborative context. It is important to note, however, that this is not a plea for exclusive concentration on concordance-based exercises. Braun (2005: 60) recommends combining the study of whole texts with corpus data added to provide greater detail on individual aspects of language use. She cites the following text as a good example for the study of the use of tenses in English, and recommends adding concordance data to provide a greater number of examples of the use of specific tenses.

> I'm the owner of the Broken Saddle Riding Company, have been for the last eleven years. I used to be in the horse racing industry back in New Jersey, worked around horse racing. I wasn't making any money, and the woman I was seeing at the time decided that she wanted to come to Santa Fe. So I decided to come along with her, because I was very much in love. ... I started this eleven years ago. Came up with the name Broken saddle, because when I started the business, my saddle broke. And I've been doing it now full time for the last nine years. It took me about two years to get

it going and it's been just a lot of fun. For the last nine years we've been riding in the hills. Silver, turquoise, mines, the canyons.

37    When working with corpora, it is easy to find extensive texts which illustrate a single aspect of language use. It is thus easy to envisage an environment where LSP teachers can adopt a discourse-based approach to teaching, studying individual texts in detail as has traditionally been the case, and complementing this with corpus data to make the learners aware of recurrent patterns of use. As Gavioli (2005: 141) observes, "awareness of what is typical may provide the learner with more autonomy to be untypical". This combination of integrated text-based and corpus-based language teaching is, however, still rare in research publications, although it is clear that the LSP learning environment would have much to gain from it.

## BIBLIOGRAPHY

Bernardini, Silvia. 2000. "Systematising serendipity: Proposals for concordancing large corpora with language learners". In Lou Burnard and Tony McEnery (eds), *Rethinking Language Pedagogy from a Corpus Perspective*. Frankfurt am Main: Peter Lang, 225-234.

Bertin, Jean-Claude. 1994. "L'enseignant, le professionnel et l'apprenant : confrontation des cultures et choix des matériaux pédagogiques". *ASp* 5/6, 69-78.

Bouchard, Guy. 2003. "Michel Foucault : unité ou dispersion de l'œuvre ?" *Laval Théologique et Philosophique* 59/3, 485-502. Accessed 25 June 2007 <http://www.erudit.org/revue/ltp/2003/v59/n3/008791ar.pdf>.

Braun, Sabine. 2005. "From pedagogically relevant corpora to authentic language learning contents". *ReCALL* 17/1, 47-64.

Burnard, Lou (ed.) 2004. *BNC Baby* [CD-Rom]. Oxford: Research and Technology Service, Oxford University. Accessed 20 March 2007 <http://www.natcorp.ox.ac.uk/babyinfo.html>.

Carter, Ronald A. and Michael J. McCarthy. 1988. *Vocabulary and Language Teaching*. Harlow: Longman.

Carter, Ronald and Michael McCarthy. 2006. *Cambridge Grammar of English*. Cambridge: Cambridge University Press.

Carter-Thomas, Shirley and Elizabeth Rowley-Jolivet. 2001. "Syntactic differences in oral and written scientific discourse: the role of information structure". *ASp* 31/33, 19-37.

Carter-Thomas, Shirley and Elizabeth Rowley-Jolivet. 2003. "Analysing the scientific conference presentation (CP). A methodological overview of a multimodal genre". *ASp* 39/40, 59-72.

Celce-Murcia, Marianne and Elite Olshtain. 2000. *Discourse and Context in Language Teaching*. Cambridge: Cambridge University Press.

Chambers, Angela. 2005. "Integrating corpus consultation in language studies". *Language Learning & Technology* 9/2, 111-125.

Chambers, Angela and Florence Le Baron (eds). 2006. *Le Corpus Chambers-Le Baron d'articles de recherche en français / The Chambers-Le Baron Corpus of Academic Writing in French.* Oxford, University of Oxford: Oxford Text Archive <http://ota.ahds.ac.uk/headers/2527.xml>.

Cheng, Winnie, Martin Warren and Xu Xun-feng. 2003. "The language learner as language researcher: Putting corpus linguistics on the timetable". *System* 31/2, 173-186.

Chénik, Nicole. 1994. "The use of new technologies in teaching ESP". *ASp* 4, 1-16.

Cobb, Thomas. 1997. "Is there any measurable learning from hands on concordancing?" *System 25/*3, 301-315.

Cook, Guy. 1989. *Discourse.* Oxford: Oxford University Press.

Corbisier, Agnès. 1994. "Metaphor not as a stylistic but as a cognitive device: English-language information technology manuals as a test case". *ASp* 4, 43-59.

*ELISA English Language Interview Corpus as a Second-Language Application.* Accessed 25 June 2007 < http://www.uni-tuebingen.de/elisa/html/elisa_index.html>.

Flowerdew, John (ed.). 2002. *Academic Discourse.* London: Pearson Education.

Foucault, Michel. 1969. *L'Archéologie du savoir.* Paris: Gallimard.

Gaskell, Delian and Thomas Cobb. 2004. "Can learners use concordance feedback for writing errors?" *System 32/3,* 301-319.

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

Gunnarsson, Britt-Louise. 1993. "Pragmatic and macrothematic patterns in science and popular science: A diachronic study of articles from three fields". In Mohsen Ghadessy (ed.), *Register Analysis. Theory and Practice.* London & New York: Pinter Publishers, 165-179.

Handford, Michael. 2007. The genre of the business meeting: a corpus-based study. Unpublished doctoral thesis. University of Nottingham.

Henry, Alex and Robert L. Roseberry. 2001. "Using a small corpus to obtain data for teaching a genre". In Mohsen Ghadessy, Alex Henry and Robert L. Roseberry, *Small Corpus Studies and ELT. Theory and Practice.* Amsterdam: John Benjamins, 93-133.

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

Hyland, Ken. 2002. *Teaching and Researching Writing.* London: Pearson Education.

Johns, Tim. 1986. "Micro-concord: a language learner's research tool". *System* 14 /2, 151-162.

Johns, Tim. 1997. "Contexts: The background, development, and trialling of a concordance-based CALL program". In Anne Wichmann, Steven Fligelstone, Tony McEnery and Gerry Knowles (eds), *Teaching and Language Corpora.* London & New York: Longman, 100-115.

Johns, Tim. 2002. "Data-driven learning: The perpetual challenge". In Bernhard Kettemann and Georg Marko (eds), *Teaching and Learning by Doing Corpus Analysis. Amsterdam &* New York: Rodopi, 107-117.

Johns, Tim. Data-driven Learning page. Accessed 25 June 2007 <http://www.eisu.bham.ac.uk/ johnstf/index.html>.

Laffont, Hélène 2006. "Un ingénieur est-il bien *an engineer* ?" *ASp* 49/50, 9-21.

Lee, David and John Swales. 2006. "A corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self-compiled corpora". *English for Specific Purposes* 25/1, 56-75

McCarthy, Michael. 1991. *Discourse Analysis for Language Teachers.* Cambridge: Cambridge University Press.

McEnery, Tony and Andrew Wilson. 1997. "Teaching and language corpora". *ReCALL* 9/1, 5-14.

Mémet, Monique. 1993. "Contraintes, libertés : le premier DEA de 'langue anglaise des spécialités scientifiques et techniques'". *ASp* 2, 229-257.

Munby, John. 1978. *Communicative Syllabus Design.* Cambridge: Cambridge University Press.

Nunan, David. 1993. *Introducing Discourse Analysis.* London: Penguin.

Olshtain, Elite and Marianne Celce-Murcia. 2001. "Discourse Analysis and Language Teaching". In Deborah Schiffrin, Deborah Tannen and Heidi E. Hamilton (eds), *The Handbook of Discourse Analysis* . Malden, Mass.; Oxford: Blackwell, 707-724.

O'Sullivan, Íde and Angela Chambers. 2006. "Learners' writing skills in French: Corpus consultation and learner evaluation". *Journal of Second Language Writing* 15, 49-68.

Percebois, Jacqueline. 2006. "Les économistes universitaires français et leurs homologues anglais et américains : aspects institutionnels, culturels et linguistiques". *ASp* 49/50, 79-91.

*Oxford-Hachette French Dictionary (The).* 2001. Oxford: Oxford University Press.

Petit, Michel. 1993. "Contraintes et libertés dans le passage à l'anglais du discours mathématique". *ASp* 2, 93-106.

Régent, Odile. 1994. "L'article scientifique : un produit culturel". *ASp* 5/6, 55-60.

Resche, Catherine. 2003. "Décryptage d'un genre particulier : les communiqués de presse de la banque centrale américaine". *ASp* 39/40, 21-35.

Römer, Ute. 2006. "Pedagogical applications of corpora: Some reflections on the current scope and a wish list for future developments". *Zeitschrift für Anglistik und Amerikanistik 54/2,* 121-134.

Sager, Juan C., David Dungworth and Peter F. McDonald. 1980. *English Special Languages: Principles and Practice in Science and Technology.* Wiesbaden: Brandstetter.

Schmidt, Richard. 1993. "Awareness and second language acquisition". *Annual Review of Applied Linguistics* 13, 206-226.

Scott, Mike. 2003. *Wordsmith Tools* Version 4.0. Oxford: OUP.

Sinclair, John. McH. 2004. *How to Use Corpora in Language Teaching.* Amsterdam: John Benjamins.

Stevens, Vance. 1991. "Concordance-based vocabulary exercises: a viable alternative to gap-fillers". In Tim Johns and Philip King (eds), *Classroom Concordancing: English Language Research Journal 4.* Centre for English Language Studies: University of Birmingham.

Swales, John. 1988a. *Episodes in ESP.* London: Prentice Hall.

Swales, John. 1988b. *"Writing Scientific English".* In John Swales, *Episodes in ESP,* 37-44.

Swales, John. 1990. *Genre Analysis: English in Academic and Research Settings.* Cambridge: CUP.

Tarone, Elaine, Sharon Dwyer, Susan Gillette and Vincent Icke. 1988. "On the use of the passive in two astrophysics journal papers". In John Swales, *Episodes in ESP,* 188-207.

Thompson, Paul. 2001. "A pedagogically-motivated corpus-based examination of PhD theses: Macrostructure, citation practices and uses of modal verbs". Unpublished doctoral thesis, University of Reading, UK.

Tribble, Chris. 1997. "Improvising Corpora for ELT: Quick and Dirty Ways of Developing Corpora for Language Teaching". In Barbara Lewandowska-Tomaszczyk and Patrick James Melia (eds), *Proceedings of the First International conference on Practical Applications in Language Corpora*. Accessed 25 June 2007 <http://www.eisu.bham.ac.uk/johnstf/palc.htm>.

Tribble, Chris and Glyn Jones. 1990. *Concordancing in the Classroom.* Harlow: Longman.

Wichmann, Anne, Steven Fligelstone, Tony McEnery and Gerry Knowles (eds). 1997. *Teaching and Language Corpora*. London & New York: Longman.

Wynne, Martin. 2005. "Archiving, preservation and distribution". In Martin Wynne (ed.), *Developing Linguistic Corpora: A Guide to Good Practice*. Oxford: Oxbow. Accessed 25 June 2007 < http://ahds.ac.uk/creating/guides/linguistic-corpora/index.htm>.

Yoon, Hyunsook and Alan Hirvela. 2004. "ESL student attitudes towards corpus use in L2 writing". *Journal of Second Language Writing* 13, 257-283.

## APPENDIXES

**Articles used as a corpus**

*Economics & Politics.* Accessed 26 June 2007 <http://www.blackwellpublishing.com/journal.asp?ref=0954-1985>

Falaschetti, Dino. "Aggregate health expenditures, national income, and institutions for private property". *Economics & Politics* 17/3, 393-431.

Gehlbach, Scott. 2007. "Revenue traps". *Economics & Politics* 19/1, 73-96.

Hanson, Gordon H., Kenneth Scheve, & Matthew J. Slaughter. 2007. "Public finance and individual preferences over globalization strategies". *Economics & Politics* 19/1, 1-33.

Karayalçin, Cem and Kathryn McCollister. 2005. "Income distribution, sovereign debt, and public investment". *Economics & Politics* 17/3, 351-365.

Primo, David M. 2006. "Stop us before we spend again: institutional constraints on government spending". *Economics & Politics* 18/3, 269-312.

Treisman, Daniel. 2006. "Fiscal decentralization, governance, and economic performance: a reconsideration". *Economics & Politics* 18/ 2, 219-235.

*The American Journal of Economics and Sociology*. Accessed 26 June 2007 <http://www.blackwellpublishing.com/journal.asp?ref=0002-9246>.

Alley, Sean and John Marangos. 2006. "A comparative political economy approach to farming interest groups in Australia and the United States". *The American Journal of Economics and Sociology* 65/3, 497-524.

Forsyth, Grant D. 2006. "Special interest protectionism and the antebellum woolen textile industry. A Contemporary Issue in a Historical Context". *The American Journal of Economics and Sociology* 65/5, 1025-1058.

Leeson, Peter T. 2006. "Cooperation and conflict. Evidence on self-enforcing arrangements and heterogeneous groups". *The American Journal of Economics and Sociology* 65/4, 891-907.

Lyson, Thomas, A. 2006. "Big business and community welfare. Revisiting a classic study by C.Wright Mills and Melville Ulmer", *The American Journal of Economics and Sociology* 65/5, 1001-1023.

Parsons, Stephen D. 2006. "Max Weber and economic sociology. A response to Peukert". *The American Journal of Economics and Sociology* 65/5, 1111-1124.

Peacock, Mark S. 2006. "The moral economy of parallel currencies. An analysis of local exchange trading systems". *The American Journal of Economics and Sociology* 65/5, 1059-1083.

Perraton, Jonathan. "Heavy constraints on a "weightless world"? Resources and the new economy". *The American Journal of Economics and Sociology* 65/3, 641-691.

Stringham, Edward and Ilkay Pulan. 2006. "Evaluating economic justifications for alcohol restrictions". *The American Journal of Economics and Sociology* 65/4, 971-990.

Vaughn, Gerald F. 2006. "The resource economics of Grover Pease Osborne, author of America's first textbook on resource economics". *The American Journal of Economics and Sociology* 65/3, 587-603.

## NOTES

**1.** In addition, in the more general context of language teaching and learning, several publications on language learning as discourse analysis (Cook 1989; McCarthy 1991; Nunan 1993; Carter & McCarthy 1994; Celce-Murcia & Olshtain 2000; Olshtain & Celce-Murcia 2001) provide guidance to teachers on the development of a discourse-based approach.
**2.** *Wordsmith Tools* (Scott 2003) was used in the creation of the concordances in this article.
**3.** <http://www.blackwell-synergy.com/journal.asp?/ref=0954-1985>.
**4.** http://www.blackwellpublishing.com/journal.asp?ref=0002-9246>.
**5.** I am indebted to my colleague Florence Le Baron for identifying these journals as comparable to the subcorpus of economic articles in *CLAREF*. The articles are listed in an appendix.

## ABSTRACTS

The teaching of Languages for Specific Purposes (LSP) in the second half of the twentieth century has been greatly influenced both by the communicative approach and by developments in research in discourse analysis, particularly with reference to genre. The integration of research in corpus linguistics in the language-learning environment has been a more recent phenomenon, with publications on the use of concordance data by language teachers and learners multiplying in the 1990s and the early years of the twenty-first century. This article aims to bring together the view of language learning as discourse analysis on the one hand, and on the other the contribution which corpus data can make to the development of "discoursal expertise" (Swales 1990: 27) by learners. Examples are included from areas such as academic writing, economics and business discourse. The article focuses not only on the potential of corpora but also on the obstacles to their integration in the language-learning environment.

Au cours de la deuxième moitié du vingtième siècle, l'enseignement des langues de spécialité a été fortement influencé par l'approche communicative et par la recherche en analyse du discours, surtout dans le domaine du genre. L'intégration de la recherche en linguistique de corpus dans l'apprentissage des langues est un phénomène plus récent. À partir de 1990 les publications sur l'emploi des concordances par les professeurs et les apprenants de langues se sont multipliées. Le but de cette étude est de rapprocher d'une part l'apprentissage des langues, vu comme analyse du discours, et, d'autre part, l'apport potentiel des corpus à l'acquisition par les apprenants de ce que Swales (1990 :27) nomme l'expertise du discours (« *discoursal expertise* »). Des exemples sont tirés de plusieurs domaines, dont l'article de recherche, les sciences économiques et la communication en entreprise. L'auteur se penche non seulement sur le potentiel des corpus mais aussi sur les obstacles qui entravent la réussite de leur intégration dans l'apprentissage des langues.

## INDEX

## AUTHOR

**ANGELA CHAMBERS**

Angela Chambers is Professor of Applied Languages at the University of Limerick in Ireland. Her research publications deal with the role of corpora in language learning. She is a member of the Irish Research Council for the Humanities and Social Sciences. Angela.Chambers@ul.ie