

Chapter 2

CORPUS LINGUISTICS AND ENGLISH FOR ACADEMIC PURPOSES

John M. Swales

The University of Michigan (Ann Arbor, MI, USA)

1. INTRODUCTION

Rather than representing a leading edge of technological applications, I am very much part of that majority of university language teachers who sometimes feel that they are being left further and further behind in the current race to develop LSP-related technologies. Indeed, like many readers of this volume, I remain an LSP teacher, an LSP materials writer, an observer of the academic scene as it variously unfolds in diverse departments and colleges, and an applied discourse analyst of academic and research text and talk. However, since 1997 I have been closely involved with the Michigan Corpus of Academic Spoken English (or MICASE), and my experiences with corpus linguistics constitute the main focus of this article. The chapter is structured as follows. In Section 2, I discuss the advantages of small specialized corpora for LSP purposes, and summarize and evaluate the present status and future prospects of the MICASE project. Section 3 provides an update on an earlier paper (Swales, 2002) originally written in 2000 that examined the advantages and disadvantages of corpus-linguistic work for LSP practitioners. With that as background, I then offer in Section 4 a fragment of what a MICASE-derived grammar of academic speech might look like. In the penultimate section, I discuss some actual teaching materials

based on academic speech data, and after that, offer a few brief closing comments.

2. SPECIALIZED CORPORA AND THE MICASE PROJECT

The 1990s can be seen as the era of large under-differentiated corpora, such as the Bank of English and the British National Corpus. While such corpora continue to have important roles in lexicography, in natural language processing research, and in general grammatical descriptions of whole languages, their immediate use as more specialized resources may be limited. In effect, bigger may not always be better, and size may not win all. Indeed, there are signs that the first decade of this new century will well turn out to be the decade of the small specialized corpus. As Lee notes,

a small specialized corpus has the advantage of more homogeneity across the texts or transcripts in a corpus, which in turn makes the corpus more suitable for genre-based investigations or analyses that can take into account interactional, pragmatic and contextual features in addition to the purely linguistic ones (2001: 37).

As evidence of this trend, we have already seen the emergence and employment of relatively small circumscribed corpora, such as Luzón's collection of medical research articles (2000) or Granger's (1998) corpus of non-native speaker writing.

One of these small corpora is MICASE, available through the web at www.hti.umich.edu/m/micase. This corpus has a number of distinctive features. First, since this is a corpus of transcribed speech it was expensive to put together, and we realized early on that, whatever the fee or license structure might eventually be arrived at, Michigan's English Language Institute would never recoup more than a small fraction of the cost of its investment. The project group thus decided to make it freely available "academicware". Second, after discussions with a statistical consultant, we decided to focus our efforts on only the University of Michigan. As the consultant commented, "You have already got a lot of variables to cope with; don't make it even more complicated". Third, we made a principled decision to not exclude non-native speakers of English (who in fact represent about 15% of all words spoken). In today's increasingly anglophone and increasingly globalized and internationalized research world, the NS-NNS distinction is increasingly hard to maintain (Swales, 2004).

Fourth, we wanted to cover as many disciplines as possible—and the University of Michigan has 19 separate colleges or schools. This ambition has been broadly obtained, even though coverage from professional schools such as business and medicine has turned out to be weaker than planned. Apart from this horizontal dimension, as it were, we were particularly interested in achieving good vertical coverage. In other words, the corpus design was structured in such a way as to include not only undergraduate lectures, but also many other types of speech-events (see Table 1). We also took to heart the comment Michael McCarthy made when he visited Michigan soon after the project began: "Don't only focus on what's easy to get; go for the harder stuff too".

The eventual outcome was 152 speech-events, consisting of around 200 hours of academic speech totaling some 1.7 million words. The speech-events themselves are shown below, preceded by the numbers of each type collected:

Table 1. Breakdown of MICASE speech-events.

31	Large lectures (40+ students)
31	Small lectures (under 40 students)
12	Colloquia (or invited talks; often called "seminars" elsewhere)
11	Student presentation sessions (as part of lecture courses)
9	Discussion sessions (as support of large lecture courses)
8	Advanced, specialized graduate courses (aka "seminars")
8	Student study group sessions (working on homework, etc)
8	Lab sessions (as part of science courses)
8	Office hour sessions (i.e. individual faculty-student meetings)
6	Meetings (4 of research groups; 2 administrative)
5	Advising sessions (discussing possible courses, careers, etc)
4	PhD dissertation defenses
3	Tutorials
3	Interviews
2	Tours (One of the campus; one of the Museum of Art)
2	Service encounters (1 in a Library; 1 in the Science Learning Ctr)

In fact, the MICASE project is not finished and we are currently well embarked on Phase Two. This includes part-of-speech tagging, lemmatization of the vocabulary stock, further manual pragmatic mark-up, and a grammar of academic speech focussing on its distinctive features. If a major grant application is successful, we will also be able to offer the immediate linking of sound file moments to transcribed elements, and a much improved web interface.

One of the fundamental reasons for undertaking the MICASE project was that our understanding of the nature of academic and research speech is much sketchier than that of academic and research prose. Indeed, five years

ago, one of our foundational questions was whether academic speech was "more like" scholarly prose and research writing, or whether it is "more like" everyday conversation. On the one hand, do the pedagogic and intellectual purposes such speech shares with research writing push it toward academic prose? In Halliday's terms, do field and tenor overcome mode? Or, on the other hand, do the interactive nature of speech settings, the immediate sense of audience, and the on-line-processing character of our spoken utterances push academic speaking toward ordinary spoken interaction? Again in Halliday's terms, does mode predominate over field and tenor?

Five years later, and after many partial investigations into the MICASE corpus, the answer to that originally rather naïve question is pretty clear; in general, academic speech is close to ordinary conversation and distant from academic prose. Some of the emerging characteristics of academic speech can be summarized as follows. Such speech employs a much greater use of narrative elements, such as anecdotes and stories and biographical accounts of research activities (see also Rowley-Jolivet, 1999; and Thompson, 2002). Speakers make much greater use of visuals of all kinds than writers. In speech there is much greater stylistic homogeneity across disciplines than in, say, research paper writing. There are several factors that contribute to this lessening of disciplinary variation. One is the pervasive informality (the commonest noun in MICASE is *thing*); a second is that, although visuals vary greatly across disciplines, the *ways* in which they are referred to and discussed seem quite stable ("on the left you can see..."; "over here is the original data"); and a third is the fact that references to the work of others are not realized through the different conventions that separate broad disciplinary areas (such as the presence or absence of superscript numbers), but by talking about what people have done.

Obviously, the interactive nature of many academic speech-events (perhaps especially in the US) leads to much sign-posting of what is to come and, more generally, to a higher level of reflexive metadiscourse (Mauranen, 2001). One clear manifestation of this is that discourse markers do not always come singly (as discourse analytic research would suggest), but can come in fixed doublets or triplets. One of the commonest of these last is the run of "okay so now", which is used by primary speakers to signal the closing of one episode and the opening of the next (Swales & Malcewski, 2001). As in research writing, there is considerable hedging, but it is more likely to take the form of mitigators such as "just" (Lindemann & Mauranen, 2001) and "sorta"/"kinda" (Poos & Simpson, 2002); in addition, there is above average use of the "wobbly" modals *might*, *would* and *could*.

More generally, academic and research speech countenances apologies, admissions to uncertainties and mistakes and confessions in ways that are very rare in research writing. More surprisingly perhaps, the data from the

University of Michigan shows very little of that stereotypic "vicious" academic argument; rather, the general tone is consensual and facilitatory rather than critical and antagonistic. Within this register, humor clearly plays an important role in "defusing" potentially difficult moments as well as in increasing the engagement of the interactants. According to Dyer & Keller-Cohen (2000), demotic elements in academic speech (such as confessions and anecdotes) also go some way to mitigate the power inequalities between, say, lecturers and students.

We could summarize the main differences by saying that we tend to find a *closed style* in research prose, but an *open style* in research talk. All that said, it is not the case that academic speech is exactly the same as casual conversation. One of the remaining differences is that academic speech contains a considerable amount of technical lexis embedded into a loosely co-ordinated sentence structure and surrounded by heavy employment of deictic elements. Other differences include a relatively lighter use of idiom and metaphor, and a level of evaluation that is somewhat less than casual conversation, where of course hyperbole and exaggeration are highly characteristic rhetorical devices. Finally, there are some effects that derive from the didactic nature of many of the interactions, such as the illustrating and glossing of technical terms as they are being introduced.

3. CORPUS LINGUISTICS—A PERSONAL STORY

As a result of the MICASE project, in the 1998-2000 period my research and pedagogical interests took a new turn. I became more interested in studying academic speech than previously, and I began to gain experience in using the *WordSmith Tools* concordancing package (Scott, 1996). At the end of this period, I reflected on this experience in a chapter in the *Academic Discourse* volume edited by Flowerdew and eventually published in 2002. I there argued that the relationship between corpus work and more traditional genre-based EAP approaches could turn out to be more complex—and perhaps more conflicted—than had been generally acknowledged.

As I saw it then, there were four main issues. First, while genre analysts in the 1990s had developed a rich socio-rhetorical conception of genre, corpus linguists had not taken the concept of genre seriously. Second, the procedural differences between the top-down approach adopted by most discourse analysts and the constrained bottom-up approach of corpus linguists were hard to reconcile. Third, there was a strong *incidental* tendency in corpus work that made it hard to see how these small pieces of the puzzle could be assembled to form a general picture. As I wrote at that time, "Certainly, accumulations of such incidental findings provide little in

the way of a *platform* from which to launch corpus-based pedagogical enterprises" (2002: 151). Fourth, the existence of so many corporist options causes many initially promising lines of inquiry to turn out to be either false leads or dead ends. For hard-pressed EAP practitioners, the amount of trial-and-error involved can result in corpus investigations becoming rather time-ineffective.

If I was in 2000 frustrated by my inability to use MICASE (and Hyland's corpus of 80 research articles) to *discover* interesting things about these two registers, I had no doubt that they were excellent for *validating* or *invalidating* statements found in grammar books, or for *finding examples* of phraseological patterns. So, just recently, a member of my undergraduate class asked a question about whether the verb *need* was a full verb or a modal auxiliary. A quick look at the 35 speech-events in the MICASE research sub-corpus showed only four instances of *needn't* but 25 instances using the "do" form. On the other hand, Hyland's RA corpus showed the opposite tendency for writing—22 examples of *need not* but only three using the "do" form. I could go to the next class with a useful answer to the student's question.

Nowadays, these concerns and criticisms (including self-criticisms) have lessened somewhat. On the first issue of genre, there has been a growing rapprochement of attitudes. One important example of this is Lee's construction of an index for the British National Corpus along genre lines. As he says, "It is envisaged that this resource will allow linguists, language teachers, and other users to easily navigate through or scan the huge BNC jungle more easily, to quickly ascertain what is there (and how much) and to make informed selections from the mass of texts available" (2001: 37). The top-down and bottom-up differences have also narrowed (at least for me) because the MICASE web-interface allows immediate access to whole utterances and to whole transcripts. (The third issue of incidentalism will be discussed in the next section.) The trial-and-error issue (number 4) also seems less serious today. Stubbs (p.c.) has convinced me that the "dead ends" problem is no more a feature of corpus linguistics than it is of any other inductive method of proceeding. In fact, it is likely that it only *seems* to be worse with concordancing software because our hopes for outcomes with scrolling and sorting are higher than when simply staring at individual texts and transcripts.

There remain, of course, some difficulties. One is the lack of context in corpus linguistics, which instead has to rely on co-text and inter-text ("repeated occurrences, often a very large number, of similar patterns across different independent texts", Stubbs, 2001: 157). For example, a recorded lecture from a course is stripped of all the lectures that have preceded it and which inevitably have had some contextual effect on the chosen speech-

event. Finally, on a more practical level, there is the problem of how best we can bring concordance lines (especially in speech) into LSP classes consisting of non-language specialists, and how best we can deal with this kind of data to advance the language proficiencies of our students.

4. TOWARD A GRAMMAR—THE CASE OF SUBJECT-INITIAL ELLIPSIS

A doctoral candidate in Dentistry, Ting Wang from China, came to see me to discuss her PowerPoint slides for an upcoming prospectus defense. At one point, referring to one of these slides, she said, "looks kinda busy". This is an example of "subject-initial ellipsis" (henceforth SIE), as opposed to the full form, "this one looks kinda busy". This kind of ellipsis, especially in utterance-initial contexts, is generally thought to be a pervasive feature of English conversation, particularly in casual conversation among intimates and strangers (Wolfson, 1988). The comprehensive Longman Grammar (Biber et al., 1999) deals with this at length in its final chapter entitled "The Grammar of Conversation". *Inter alia*, its authors note that it seems more common in British English than in American English. Their main illustration is as follows:

One example of this trend is the occurrence of *Depends...* (with ellipsis) as contrasted with *It depends...*. In the BrE conversational subcorpus, almost 60 per cent of these constructions had ellipsis of the subject, compared with only about 30 per cent with ellipsis in the AmE subcorpus. (p.1106)

So let us first consider the case of "depends", since it has already been mentioned. In the MICASE research sub-corpus of 36 speech-events, there are 50 instances of *depends*, 29 of which have a pronoun (e.g. "it") or a demonstrative (e.g. "that/this") subject. However, there are only 5 (some 15%) with elliptic or "missing" subjects. Here is an example:

would you be able to communicate with like sign language?
depends on the aphasia

So far then the data for *depend* suggests that SIE can occur in research speech, but less frequently than in ordinary conversation. In order to get a broader picture, we can turn to a number of other verb forms that have some propensity to drop subjects in speech (especially for empty "it"). In Table 2, Column 1 refers to the number of total occurrences in the sub-corpus;

Column 2, the number with SIE; and Column 3, the percentages of SIE utterances.

Table 2. Percentages of SIE with selected verb forms.

<i>Verb</i>	<i>Tokens</i>	<i>SIE Tokens</i>	<i>SIE %</i>
looks	74	6	8.1
turns out	50	4	8.0
sounds	42	2	4.7
got	c.450	8	1.7
seems	140	2	1.4
wanna	c.300	2	0.7

Examples:

looks a lot better

it might work. turns out though it doesn't

sounds good. I like it

got any ideas about that?

seems a bit ironic to me

might wanna back it up a little

The above figures underline the fact that SIE is not a generally prevalent feature of research speech (in which, we should note, there are very few undergraduate speakers). This is confirmed by the absence of tokens like "heard the latest?", "hope it goes well" and "good to see you again", and very few instances of other adjectival uses such as "sorry that I forgot this" or "ready to move on?".

The situation is, however, somewhat different with quantifiers, such as *nothing*, *anything* or *any*. For example, the phrase *anything else* is more likely to have omitted subjects (and verbs):

okay, anything else on that?

okay, anything else about television violence?

in contrast to full forms such as:

does anyone else wanna tease anything else out of that poem?

The canonical case of SIE in research speech occurs with the formulaic phrase "any questions?". In the MICASE sub-corpus, 75% of occurrences used the SIE form, with the remaining quarter having full forms such as "Do you have any questions?". With other nouns, the figures are lower, but SIE remains relatively common:

okay any experiences you want to share?

any last minute thoughts on this particular search?

any any sense of how long it takes for rediversification?

However, SIE is much more likely to be used in these contexts by the primary speakers than by the (mostly) listening participants. One exception is the last of the previous examples ("any any sense...?"), and it may be significant that the questioner here was a senior faculty member in a colloquium.

This preliminary excursion into the grammar of research speech (as represented by the SIE v. non-SIE issue) begins to allow the outline of a grammar to emerge. The excursion itself, I believe, brings home a number of lessons. Most generally, the observant practitioner can use a corpus as successfully to note what does not occur as note what does; in effect, searching for absence may be as relevant and as rewarding as searching for presence. In this particular database, sentence-initial ellipsis (outside of the highly formulaic speaker check-interrogation "any questions on that?") is comparatively infrequent. Its relative rarity can be partly ascribed to the scarcity of those short opening and closing sequences that are much more characteristic of "daily life"—all those occurrences of "glad to meet you", "sorry I'm a bit late", "had a haircut, I see", etc. Its modest occurrence can also be attributed to Wolfson's "Bulge Theory", i.e. spoken discourse tends to be succinct between those who are very close or very distant, but elaborated among acquaintances or in cases where less powerful speakers address more powerful ones. So a graduate student, on meeting her famous head of department in the corridor, would be unlikely to say "got a minute?" but rather something of the order of "Excuse me, I'm just wondering if you could spare me a couple of minutes as I've got a small problem that I'm having difficulty in solving".

However, the limited employment of SIE is well worth bringing to the attention of international students and other interested parties. First, as the immediately previous example attests, SIE is a useful vehicle for raising consciousness about variations in formality and politeness. Second, discussions on productive uses can usefully be directed to what might be appropriate with friends and peers (including in e-mails) but less appropriate for others. Recollect for example, the dentistry student at the beginning of this section who used an SIE in self-criticism, as opposed to other-criticism. But the final word here should belong to a newly-arrived Japanese student who heard me speak on the SIE topic. Then he observed something like the following: "That is very interesting. I have always thought when listening to English speakers that I was not hearing all the words at the beginning of sentences. Now I know that they may not have been there in the first place".

5. A PEDAGOGICAL APPLICATION

In this section I offer a modified version of a discursial-analytic example of teaching materials derived from MICASE and being developed by Sheryl Leicher, myself and others. It deals with a phenomenon that seems common among Michigan undergraduates and junior graduates—complaining about their instructor!

Review and—if possible—act out this dialog

Math study group members complaining about their homework assignment.

Nick: Male senior undergrad

Jane: Female senior undergrad

Amy: Female senior undergrad

Setting: Nesbitt Lounge, East Hall

1. Jane: Alright...Why did she give multiple parts for every single problem we have?
2. Amy: She wanted to torture us.
3. Nick: I think it's cuz like, that last time I don't think we did a lot on the homework, if I remember right.
4. Jane: I really like how she says this problem set contains a total of four problems. It's like no, four times two plus one problems.
5. Nick: Yeah why does she tell us it contains four problems?
6. Jane: Oh I can't see 'em. Oh the printer chewed off, just cut off number four. So, maybe that's why she puts it at the top just in case.
7. Nick: Yeah that is a little quirk. She's trying to trick us into thinking like oh...
8. Jane: Yeah she's like "Oh that's not bad this problem set only has four problems in it." Yeah, whatever.
9. Amy: And there's like two parts for three of 'em.
10. Nick: Alright, so, oh my gosh so this has three parts.
11. Jane: For all of 'em?
12. Amy: Two parts for all of 'em?
13. Jane: Yeah. And number two has three parts.
14. Nick: This has three parts.
15. Amy: Oh you've got to be kidding.
16. Nick: Well at least we only have to do two of the five parts on fourteen.
No, this won't necessarily be that hard.

17. Jane: Did you say two of the five parts on fourteen?
18. Nick: Yeah we only have to do one and two out of five.
19. Jane: Oh, alright.
20. Nick: And on fifteen we only have to do two of, well one out of four but its number two. Why don't we...Fourteen's gonna be easy let's start there. It really is gonna be easy.
21. Jane: Oh yeah alright <Reading: *Find the image of the point...*>

Worksheet for math study group extract

Now do the following—if possible work with a partner.

GENERAL QUESTIONS

1. Would you say this extract is an example of "academic speech"? What are the pros and cons of labeling it so?
2. The given title is "complaining about their homework assignment". This is obviously true on one level. But how else might you describe it? If it is "complaining", what is its underlying purpose or function?
3. The two main speakers are Nick and Jane. How similar or different might their attitudes be to the tasks on hand? What evidence can you bring to support your viewpoint?
4. If you had to divide this extract into two parts, at which turn would you say the second part begins, and why?

SPECIFIC QUESTIONS

5. Jane, in Turn 1, opens the episode by asking a question. Does she really expect an informative answer?
6. She gets two responses. Which one might she prefer, and why?
7. How would you characterize Jane's "I really like..." opening statement in Turn 4?
8. What does Nick's question in Turn 5 tell us?
9. At the end of Turn 8, Jane comments, "Yeah, whatever". How would you explain this use of *whatever* to a newly-arrived international student?
10. In Turn 15, Amy says, "Oh, you've got to be kidding". What other expressions would have got her point across?

COMMENTARY

1. When I first tried this out with an undergraduate class, 23 out of 24 participants were of the firm opinion that this episode was *not* an example of academic speech. The reasons they gave included: "too much slang", "not on an academic topic", and "not concerned with the transmission of knowledge". Subsequent exposures, however, suggest a different view, participants pointing out that a study group is an academic grouping, the episode takes place on campus, and the three students are talking about an academic assignment. And there is doubtless no more "slang" in this extract than in many lectures.
2. Some of the expressions that come to mind are "whining" and "pissing and moaning", a more academic version being "commiserating with each other". The underlying purpose of this activity is to build rapport among the group; a way of recognizing that "we're all in the same boat" in an "us versus them" scenario. Note how the instructor is always referred to as "she"—as some kind of opponent. According to Boxer & Pickering (1995), this kind of complaining is often misunderstood by international students and mis-presented by their ESL instructors.
3. Nick is the "can-do" member of the group; Jane is the main "piss-and-moaner".
4. Nearly everybody so far has opted for Turn 16, when Nick turns the conversation back to actually doing the homework.
5. No, she expects sympathy.
6. Jane clearly prefers Amy's response, which is the complaining one, rather than Nick's attempt to justify the instructor.
7. This is clearly sarcastic. Note that an international student might take it literally and so lose track of the dialogue.
8. It tells us that Nick is now on board, and is on "the same page" as the two women.
9. This kind of *whatever* is used mainly by younger Americans. It is typically dismissive. In this example, it communicates "I don't agree, *whatever* she says" or perhaps "what she says doesn't make any sense". There are 732 examples of *whatever* in MICASE; you might like to look at a selection of them to see how common this dismissive use of *whatever* actually is.
10. How about "you must be joking" or "you can't be serious"?

At this point, you might want to re-read (or re-enact) the dialog, as you should now have a deeper understanding of the situation.

ROLE PLAYS

Imagine there is a fourth member of the study group.

- In Role Play A, the fourth member immediately joins in the complaining and so establishes his or her membership of the embattled group. Act this out.
- In Role Play B, the fourth member is much like Nick, but he is even more concerned to justify the instructor and actually do the homework. Act this out.

6. CLOSING COMMENTS

In this chapter I have attempted to profile the Michigan Corpus of Academic Spoken English as a contemporary instance of a small specialized corpus designed to be made freely available for students, teachers, materials writers and researchers. In effect, the MICASE project offers a time capsule of academic speech at a single large university at the end of the 20th century. As I hope to have made clear, this particular corpus offers exciting descriptive and applied opportunities, if only because academic and research speech is still poorly understood. We know that there are quite a number of people around the world who are using MICASE in their investigations—people in countries like Spain, Canada, Brazil, Taiwan and Finland. I hope others will join this international effort and let us know in Michigan that they have done so. Indeed, the MICASE team is trying to set up a better infrastructure to facilitate these endeavors.

In the third section of this chapter, I tried to communicate something of my own successes and failures with a concordance-line approach to understanding academic speech. Let me give a final mini-illustration. One of the relatively common verbs in MICASE is *feel*, which is only rarely used in its "I feel sick" sense, but rather has a more epistemic meaning. Although the overall uses of the verb *feel* escape me at the time of writing, especially in comparison to the hugely frequent verb *think*, there is a small sub-set about which I think (feel?) I can say something. The prepositional verb *feel about* is primarily interrogative (60+% of the cases):

- what do you guys feel about including in the intro, maybe stereotypes...
- how do you feel about, um, the meaning of the metaphor...
- how do you feel about the decision to run any picture?
- and how do you feel about that?

My sense here—peering through a glass darkly—is that this choice of verb potentially allows participants in the interaction (primarily students) a wider range of responses than the choice of the "straight" epistemic verb *think* would permit. However, the situation is complicated by the fact that these questions are sometimes rhetorical rather than invitational, and at other times they are part of a run of questions designed to engender an audience response. As ESP responds to the increasing use of English for conferences, and the increasing movement of academics across the face of the globe, this kind of small-scale question (*think* versus *feel*) is just one of many similar ones that will increasingly engage our attention in the years ahead.

REFERENCES

- Biber, D., Johansson, S., Leech, G., Conrad, S. and E. Finegan (1999). *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education.
- Boxer, D. and L. Pickering (1995). Problems in the presentation of speech acts in ELT materials: The case of complaints. *ELT Journal* 49, 44-58.
- Dyer, J. and D. Keller-Cohen (2000). The discursive connection of professional self through narratives of personal experience. *Discourse Studies* 2, 283-304.
- Granger, S. (ed.) (1998). *Learner English on Computer*. London: Longman.
- Lee, D.Y.W. (2001). Genres, registers, text-types, domains, and styles: Clarifying the concepts and navigating a path through the BNC jungle. *Language Learning and Technology* 5, 37-72.
- Lindemann, S. and A. Mauranen (2001). It's just real messy: The occurrence and function of *just* in a corpus of academic speech. *English for Specific Purposes* 20, 459-476.
- Luzón, M.J. (2000). Collocation frameworks in medical research papers: A genre-based study. *English for Specific Purposes* 19, 63-86.
- Mauranen, A. (2001). Reflexive academic talk: Observations from MICASE. In R.C. Simpson and J. M. Swales (eds.) *Corpus Linguistics in North America: Selections from the 1999 Symposium*. Ann Arbor, MI: The University of Michigan Press. 165-178.
- Poos, D. and R.C. Simpson (2002). Cross-disciplinary comparisons of hedging: Some findings from the Michigan Corpus of Academic Spoken English. In R. Reppen, S.M. Fitzmaurice and D. Biber (eds.) *Using Corpora to Explore Language Variation*. Amsterdam: Benjamins. 3-23.
- Rowley-Jolivet, E. (1999). The pivotal role of conference papers in the network of scientific communication. *ASP* 23-26, 179-196.
- Scott, M. (1996). *Wordsmith Tools*. Oxford: Oxford University Press.
- Stubbs, M. (2001). Text, corpora, and problems of interpretation: A response to Widdowson. *Applied Linguistics* 22, 149-172.
- Swales, J.M. (2002). Integrated and fragmented worlds: EAP materials and corpus linguistics. In J. Flowerdew (ed.) *Academic Discourse*. Harlow: Longman. 150-164.
- Swales, J.M. (2004). *Research Genres: Explorations and Applications*. Cambridge: Cambridge University Press.
- Swales, J.M. and B. Malcewski (1999). Discourse management and new episode flags in MICASE. In R.C. Simpson and J.M. Swales (eds.) *Corpus Linguistics in North America: Selections from the 1999 Symposium*. Ann Arbor, MI: University of Michigan Press. 145-164.

- Thompson, S. (2002). As the story unfolds: The uses of narrative in research presentations. In E. Ventola, C. Shalom and S. Thompson (eds.) *The Language of Conferencing*. Frankfurt: Peter Lang. 147-167.
- Wolfson, N. (1988). The bulge: A theory of speech behavior and social distance. In J. Fine (ed.) *Second Language Discourse: A Textbook of Current Research*. Norwood, NJ: Ablex. 21-38.

Information Technology in Languages for Specific
Purposes

Issues and Prospects

Arnó Macià, E.; Soler Cervera, A.; Rueda Ramos, C.
(Eds.)

2006, XVII, 267 p., Hardcover

ISBN: 978-0-387-28595-5