10 *Developing materials for discipline-specific vocabulary and phrases in academic seminars*

Martha Jones and Norbert Schmitt

Summary

This chapter reports on the development and piloting of discipline-specific vocabulary materials on a CD-ROM software program entitled *Sound understanding: Listening and language awareness tasks* (Jones 2003a). The language of academic seminars was focused upon and was explored through needs analysis and corpus analysis techniques. After administering a needs analysis questionnaire and conducting interview surveys among international and home students, as well as members of staff in selected departments at the University of Nottingham, a number of spoken genres were identified as important by the respondents. The survey results prompted the compilation of the Needs-Driven Spoken Corpus (NDSC) (Jones 2003b), of which academic seminar data was one element. It was compiled from discourse in three academic departments, and then analyzed quantitatively and qualitatively in order to obtain a list of subject-specific words and phrases frequently used in the seminars. The items on this list were included in the vocabulary section of the CD-ROM program. Three groups of students studying at the Centre for English Language Education were subsequently exposed to the academic seminar materials over a three-week period in teacher-led sessions and independent study to ascertain to what extent this exposure would result in the acquisition of the target words and phrases. Posttests demonstrated considerable acquisition of the target words and phrases, indicating that the CD-ROM program was effective in facilitating the students' learning.

Introduction

Several studies on vocabulary development have suggested that a focus on discipline-specific vocabulary and phrases in the EAP classroom can result

in learning (Gledhill 2000; Mudraya 2006; Nation 2001; Schmitt 2004; Ward 2007). However, the development of appropriate subject-specific teaching materials is essential for this acquisition to take place. This chapter reports on the process Jones followed in producing such materials for the particular subjects of English Studies, International Law, and Business Studies. The materials were corpus-based, and focused on the discipline-specific vocabulary and phrases that occur in academic seminars. The materials were incorporated onto a CD-ROM software program entitled *Sound understanding: Listening and language awareness tasks* (Jones 2003a),[1] and a small-scale exploratory study conducted in order to check the effectiveness of these materials.

Background

A number of research studies have focused on the analysis of academic spoken discourse, based on corpus data with a view to designing materials for the development of speaking skills in an EAP context. Considerable attention has been given to the analysis of initiating acts, components of turns, signaling devices, sequential organization in student-to-student discussion and exchange patterns (Basturkmen 1999, 2002, 2003), seminar structure, and the use of metadiscursive language (Aguilar 2004). Other studies have analyzed language items such as the lexical particle "just" (Lindemann & Mauranen 2001) and hedging devices (Mauranen, 2004; Poos & Simpson 2002).

Understanding and mastering the above linguistic features is important for successful communication in an academic setting, but so is the use of precise, discipline-specific vocabulary. Schmitt (2000: 37) argues that in order to maintain communication in a specific genre, technical vocabulary for specific fields should be taught and suggests the use of specialized EAP frequency lists to illustrate vocabulary dominant in academic areas. Such an approach requires analysis of this type of language in order to design appropriate materials to meet the lexical needs of international EAP students.

Coxhead and Nation (2001) note that learning academic vocabulary is of great importance to the EAP learner for three reasons. First, it occurs more frequently in a wide range of academic texts than in nonacademic

[1] The CD-ROM was produced with funding from the Centre for Teaching Enhancement (98 TL/142), and the technical support of software programmer Suzanne Wright and multimedia graphic designer Coleen McCants at the University of Nottingham.

texts. Second, there is evidence that nontechnical academic vocabulary is less known than technical vocabulary. Third, the nonspecialist EAP teacher is more likely to be able to help learners with this type of vocabulary than with technical vocabulary. There are also strong arguments for the teaching of technical vocabulary, especially when the learners' needs center around using the language in a particular discipline. In this case, it is advantageous to learn lists of specific words, which can enable the learner to communicate successfully within one discipline.

Jones and Haywood (2004) conducted an exploratory study of the acquisition of one type of vocabulary – formulaic sequences – in an academic setting. Through the use of awareness-raising reading and writing tasks, they guided EAP students to notice formulaic sequences used in academic discourse first and then produce specific sequences in their essays over a period of ten weeks. By the end of the study, although the students produced only a limited number of sequences accurately and appropriately in their essays, their ability to recognize formulaic sequences in academic texts had improved considerably.

Other studies focusing on specialized vocabulary and collocation in different academic disciplines have been conducted. Gledhill (2000: 115), for example, analyzed the discourse function of collocation in medical research article introductions. He claims that a representative and specialized corpus of research articles can be used to teach languages for specific purposes. Mudraya's study (2006: 235) focused on a combination of data-driven corpus-based methodology and the lexical approach in order to create awareness-raising materials "to help students acquire engineering formulaic multi-word units." Ward (2007) examined collocations that included complex noun-phrase formation found in engineering textbooks aimed at undergraduates. His corpus-based study revealed that there was a widespread use of technical collocations in the textbooks. He also observed that the collocations in these types of texts were highly specialized and rarely occurred in other subdisciplines. He argued that technicality involves specialization and difficulty, and for this reason, it should be tackled by EAP practitioners.

The above research shows that vocabulary is an important component of EAP proficiency, and just as importantly, it is amenable to training. However, it is clear that academic writing is still the focus of many corpus-based studies of specialized vocabulary and phraseology. Although writing will continue to be prominent in EAP pedagogy, it is also important to design materials that reveal the use of disciplinary-specific language in academic spoken genres. Thus, the rest of this report will focus on the spoken vocabulary used in academic contexts.

Analyzing student needs

Since the 1970s, the importance of needs analysis and target situation analysis has been emphasized in order to design EAP materials that address some of the problems students have with academic discourse in general, and with specific disciplinary language in particular (Geoghegan 1983; Hutchinson & Waters 1987; Hyland 2006; Kim 2006; Weir 1983; West 1994). Hyland (2006: 2) defines EAP as "specialized English-language teaching grounded in the social, cognitive and linguistic demands of academic target situations, providing focused instruction informed by an understanding of texts and the constraints of academic contexts." This definition makes clear that any informed EAP teaching must first start with a firm understanding of the EAP context itself, and what students need to know about it.

It is important to enable EAP students to become familiar with features found in both academic written and spoken discourse. Considerable emphasis has been placed on writing, as this is a widespread method for assessment in higher education. However, a growing number of academic departments now assess student participation in academic seminars where discipline-specific topics are discussed requiring the use of complex vocabulary and phraseology. International students may find this type of activity difficult for three main reasons. They may lack the confidence to take part in academic discussions where there is fast speech, rapid turn-taking, and complex elaboration of arguments. They may also be used to traditional academic contexts where they are only required to listen and take notes. Another reason could be their limited knowledge of disciplinary vocabulary and phrases to discuss issues relevant to their specific subject.

Jones (2003b) conducted a small-scale needs analysis survey to explore the relative importance of the above reasons. She questioned international students attending presessional and insessional EAP programs, and home students and staff in selected departments at the University of Nottingham to identify the perceived difficulties international students had with the language used in different spoken genres at college. She asked about three aspects of the discourse international students found themselves in: (1) the settings where they had to interact with native speakers of English; (2) the types of speakers they often had to speak to either on or off campus (e.g., tutors, other students, departmental administrative staff, shop assistants); and (3) the purposes of the interactions (e.g., asking questions, expressing opinions, complaining). According to the needs analysis survey, a lack of appropriate vocabulary and phrases related to specific disciplines, as well as

the use of fast speech and colloquial expressions, seemed to be responsible for the students' inability to communicate successfully with native speakers of English in seminars.

The survey results prompted the compilation of the Needs-Driven Spoken Corpus (NDSC) (Jones 2003b). Corpus analysis is considered to be an invaluable method for the development of EAP teaching materials, and the corpus does not have to be very large provided that it is well targeted, as is the NDSC (see below). This body of raw data can be adapted to raise awareness of grammatical, lexical, or pragmatic features, which otherwise might be overlooked. Flowerdew (2005: 9) comments on the value of small, specialized corpora of EAP / ESP texts, "where the analyst is probably also the compiler and does have familiarity with the wider socio-cultural context in which the text was created, or else has access to specialist informants in the area." She goes on to add that small corpora, which include features from the students' own sociocultural environment, are very useful for pedagogic purposes, as the teacher can act as a kind of "mediating specialist informant of the raw corpus data."

The NDSC consists of native and nonnative English for EAP use, which includes a combination of academic discourse, the language of service encounters, and casual conversation. This corpus consists of four subcorpora (Table 10.1): academic (seminar discussions and a conversation recorded on an open day in an academic department), community (conversations recorded in various settings on campus), commercial (conversations in various shops), and social (conversations in halls of residence). The spoken genres included in the corpus closely match those mentioned by students in the needs analysis survey.

The materials discussed in this chapter are based on the academic subcorpus only, specifically academic seminar data from three complete seminars recorded in the School of English Studies, the Business School, and the School of Law at the University of Nottingham between 1997 and 1998. These schools were chosen because of the large numbers of international students they attract. The number of words in this subset of the academic subcorpus was 24,119 words.

The participants in the Language and Gender seminar (English Studies) were approximately twelve undergraduate home students and the lecturer. The speakers taking part in the International Law seminar were the lecturer and about 20 postgraduate students (a combination of native and nonnative speakers). In the Business Entrepreneurship seminar, students worked in small groups and the speakers in the data were two native speakers and a nonnative speaker of English with a good command of English.

Table 10.1: *Composition of the Needs-Driven Spoken Corpus (NDSC) (Jones 2003)*

Spoken contexts in the NDSC	Number of words
Community subcorpus:	
Library, student life, community / voluntary work	13,948
Commercial subcorpus:	
Travel agents, restaurants, shops	11,835
Academic subcorpus:	
Seminar discussions on "Language and Gender," "International Law," "Business," and open day in the Politics Department	26,613
Social subcorpus:	
Multiparty conversations and a conversation between two students in halls of residence	13,728
Total number of words in the NDSC	66,124

Identifying key vocabulary in the academic seminar data

After the selection of the academic seminar data for the development of materials for vocabulary practice, it was necessary to identify the vocabulary and phrases to include in the materials on the CD-ROM which our groups of international students would use. The seminar data were POS (part of speech) and semantically tagged using *Wmatrix* (Rayson 2001) before producing POS and semantic domain frequency lists and concordances. The data were compared with the BNC Spoken Sampler, used as a normative native corpus, to identify key categories that were overrepresented in the academic seminar corpus. Although the academic subcorpus is very different from that of the BNC Spoken Sampler, a statistical comparison between the two corpora could be made by using log-likelihood ratios. Log-likelihood ratios are considered to be the most reliable and accurate statistical method for the comparison of two corpora of a different size. The higher the log figures, the greater the differences between the two corpora. By convention, log-likelihood ratios above 7 indicate real differences, as the threshold of significance equivalent to the chi-square value of $p < 0.01$ is 6.6 (Leech, Rayson, & Wilson 2001: 16–17).

Table 10.2 shows the results of the statistical comparison between the academic subcorpus and BNC Spoken Sampler. Column 1 gives information

Table 10.2: *Results of the statistical comparison between the academic seminar data in the NDSC and the BNC Spoken Sampler*

Item	NDSC	%	BNC Spoken Sampler	%	Log-likelihood values
Unit of measurement, e.g., "in," "cc"	94	0.39	159	0.02	375.40
Singular common noun	2961	12.28	84446	8.59	326.63
Plural common noun	1043	4.32	24407	2.48	261.01
Article	1492	6.19	38652	3.93	257.13
-s form of lexical verb	405	1.68	6603	0.67	247.40
General adjective	1292	5.36	36947	3.76	140.49
Possessive pronoun, prenominal	433	1.80	10384	1.06	99.76
that (as conjunction)	302	1.25	6644	0.68	91.54
of (as preposition)	495	2.05	12770	1.30	86.78
Plural proper noun	21	0.09	79	0.01	57.77
General preposition	1213	5.03	40428	4.11	44.66

on the grammatical categories that were overrepresented in the sample corpus, compared with the BNC Spoken Sampler. Columns 2 and 4 show the number of occurrences of such grammatical categories in the sample corpus and the BNC Spoken Sampler, respectively. The percentages of key categories in both corpora are shown in Columns 3 and 5. The POS list was sorted by log-likelihood value in descending order according to the level of overrepresentation of key categories in the sample corpus.

We find that a number of grammatical categories occur significantly more often in the academic seminar subcorpus than in general spoken English, as represented by the BNC Sampler. An example is the single common noun, which had a log-likelihood value of 326.63, indicating a very high level of overrepresentation. Each individual academic seminar was also compared with the BNC Spoken Sampler, and the level of overrepresentation of the single common noun was as follows:

Academic seminar on Language and Gender	32.43
Academic seminar on International Law	230.50
Academic seminar on Entrepreneurship	93.74

Research into variation between speech and writing has revealed that specific genres characterized by a high density of nouns followed by complex postmodifying prepositional phrases indicate a high informational content (Biber 1988; Biber et al. 1999). The relatively high density of single

common nouns in the Law and the Business seminars, and to a lesser extent, the Language and Gender seminar, suggest that the discourse in these seminars did have a relatively high informational content.

We can also learn something about academic collocation from Table 10.2. Among the grammatical categories overrepresented in the academic subcorpus, we find the following:

1. Single common nouns.
2. Plural common nouns.
3. Articles.
4. Adjectives.

All of these are important elements of phrases where the noun is the noun-phrase head. According to Biber et al. (1998: 5), "a corpus-based approach allows researchers to identify and analyze complex 'association patterns.'" Single and plural common nouns as well as articles and adjectives indicate a high level of "informational" content. Rayson et al. (2002) also found that nouns, adjectives, and prepositions, *of* in particular, were very common in informative writing and in the *context-governed* part of the Spoken BNC. In academic seminars, there are likely to be numerous subject-specific phrases where the noun is the noun-phrase head which may be useful for students to understand as chunks and eventually use.

Words and phrases in the academic seminar data where the noun was the noun-phrase head were selected on the basis of the frequency lists as the result of the corpus analysis and also according to their usefulness. The decision on usefulness was initially made intuitively by the lecturers who had delivered the lectures (Language and Gender; International Law). Students discussing the content of their PowerPoint presentation and assignment (Entrepreneurship) were subsequently consulted to ensure the choices made were appropriate. Schmitt (2000: 144) notes that "frequency is not the only criterion for choosing words to teach explicitly. Another is words particularly useful in a specific topic area (e.g., technical vocabulary)." In addition, colloquial words and phrases were chosen, as international students rarely encounter such language in the EAP classroom unless the EAP tutor makes a special effort to teach them. Contrary to what students might expect, colloquial language is also used in academic discussions, as well as general words and phrases.

Based on the corpus statistical analysis and the lecturer and student input, three types of vocabulary and phrases were selected to be included in our material. It is important to note that most, but not all, of these words and

Table 10.3: *List of words and phrases included in the Sound Understanding: Listening / Language Awareness Tasks CD-Rom (Jones 2003a)*

Technical (N=19)	General (N=22)	Colloquial (N=7)
Entrepreneurship	intuitive	gut feeling
Entrepreneurial audit	template	blokey
(to set up) a venture	accomplishment	stuff like that
economic sanction(s)	take the initiative	a hundred grand
coercion	violate	the thing is
acquiescence	recalcitrant	the only thing that's bothering me
economic might	agenda(s)	
a regional body	shift	
in the light of practice	leeway	
peer group	endorse	
single-sex group(s)	tacit consent	
mixed-sex group(s)	at the outset	
peer group(s)	the beauty of the idea	
age group(s)	good point	
hedging device(s)	at the end of the day	
backchanneling device(s)	the problem is	
take the floor	the only problem I've got	
regional arrangement(s)	sell the idea	
imposition of sanctions	run with an idea	
	take a breath	
	state of flux	
	carve out	

phrases have a noun as the head of the phrase. Examples of these are listed on the next page:

Developing materials for discipline-specific vocabulary derived from corpus data

The vocabulary selected from the above analysis was embedded into learning materials incorporated onto a CD-ROM learning program.

This program consists of a number of audio extracts from seminars on Language and Gender, International Law, and Business Studies, and a

series of interactive tasks based on those extracts. It can be used either in guided teacher sessions or independently, as there are clear explanations and feedback throughout the program. These activities are organized in two parts on the CD-ROM.

The features of academic spoken discourse that are portrayed in the program are as follows:

1. The context of academic seminars (Part 1).
2. Types of interaction (Part 1).
3. Speech acts commonly found in academic seminar discussions (Parts 1 and 2).
4. Structure of seminars and patterns of interaction (Part 2).
5. Disciplinary vocabulary (Part 2).
6. Summary of whole seminar (Part 2).

The section devoted to practice of disciplinary vocabulary and phrases includes three activities based on the language used by the lecturer and students in the three seminars: Find the Word, Crossword Puzzle, and Odd One Out. A total of 73 words and phrases were included in these activities, and of these, four were recycled.

The Find the Word activity requires the user of the program to find specific words and phrases horizontally, vertically, or diagonally in a grid. When the student finds and clicks on the correct word or phrase, it is automatically highlighted in light red.

The words and phrases in this task from the seminar on International Law include: *state of flux, deadlocked, agendas, endorse, at the outset, acquiescence, economic might, tacit, carve out, take the initiative, economic sanctions*, and *violate*. From the seminar on Language and Gender: *gut feeling, anecdotally*, and *stuff like that*, and from the seminar on Entrepreneurship: *attributes, template, drive, venture*, and *entrepreneurial audit*.

An example of the Odd One Out activity, based on the seminar on International Law, is provided below. The correct answers are indicated for the reader's convenience by **bold** print, although in the exercise, they appeared in normal font. This task focuses on collocations.

In this game, you are given five four words or phrases taken directly from the seminars on the CD-ROM. Next to each key word is a list of four words or phrases, only one of which does not fit in well with the context of the key word. This is called "the Odd One Out." See how many you can get right. You have two attempts at each word. NB the "Odd One Out" may be grammatically correct but you have to judge which word or phrase **best** matches the **context**.

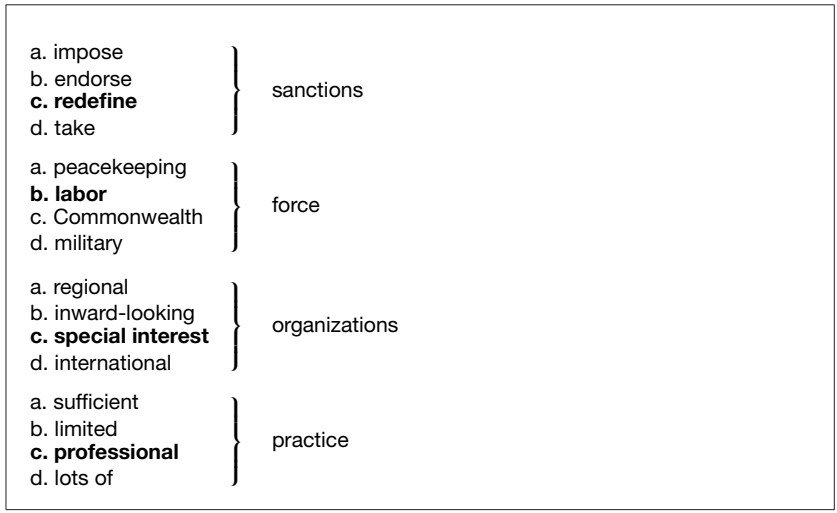

Figure 10.1: Example of 'Odd one out' Activity on the CD-ROM

A list of target collocations to be included in the Odd One Out task is provided below:

From the International Law seminar: (*impose* / *endorse* / *take*) *sanctions*; (*regional* / *peacekeeping* / *military* / *Commonwealth*) *force*; *economic* (*organization* / *might* / *measures*); *make* (*a difference* / *a statement* / *a decision*); *legal* (*argument* / *issue* / *authority*); and *seek* (*authority* / *approval* / *advice*).

From the seminar on Language and Gender: (*invite* / *get* / *give*) *a response*; (*hedging* / *backchanneling* / *softening*) *devices*; *female* (*speaker* / *conversation* / *chair*); *mixed sex* (*group* / *conversations* / *interaction*); (*create* / *show* / *express*) *uncertainty*; and *in a* (*personal* / *university* / *social*) *context*.

From the seminar on Entrepreneurship: *entrepreneurial* (*process* / *audit* / *idea*); *manufacturing* (*idea(s)* / *industry* / *experience*); *land* (*owner* / *company* / *titles*); (*set up* / *own* / *start*) *a company*; *specific* (*attributes* / *skills* / *competencies*); and (*gas* / *phone* / *nuclear*) *industry*.

The Crossword Puzzle task requires the user to identify the correct word or phrase based on any of the three seminars included on the CD-ROM. An example of a phrase included in this activity from the International Law seminar is *sanctions were imposed*. The clue for this phrase is "international organization took measures against (a state)." Another example of a phrase in this puzzle from the same seminar is *a regional body*, with the clue for this phrase being "an organization operating in a particular part of the world."

These vocabulary activities were deliberately included toward the end of Part 2 of the CD-ROM to ensure the user of the program had numerous opportunities to encounter the words and phrases in the three seminars before attempting to do such activities.

An exploratory study to check the effectiveness of the vocabulary materials on the CD-ROM

The empirical evaluation of materials effectiveness is something that is often neglected in materials development. In order to explore the effectiveness of our corpus-based CD-ROM materials, we carried out a small-scale study with EAP students.

Participants

Two guided sessions were conducted with students on two different specialized presessional programs: one group of students in Law and Business (August–September 2006), and one group of teachers of English from Japan, who were on an in-service teacher-training program (November 2006) at the Centre for English Language Education (CELE) at the University of Nottingham. The Law and Business group consisted of 29 students[2] (16M, 13F) in their mid-twenties / early thirties. Their proficiency was roughly equivalent to 6.5 / 7.0 IELTS, and they had offers from the University of Nottingham. All the students were highly motivated and showed great interest in the tasks on the CD-ROM, as the material was relevant to their discipline. The group of Japanese teachers of English consisted of nine students (7M, 2F) in their late thirties / early forties. They were on a teacher development program funded by the Japanese Ministry of Education, Culture, Sports, Science, and Technology, and their proficiency ranged from 5.5 to 6.5 in IELTS. The teachers were all experienced junior or senior high school teachers, and they were also highly motivated.

Instruments

At the end of Part 2 of the CD-ROM program, there is a summary of the main topics covered in each of the three seminars. The texts of these summaries were used to design an adapted C-test, designed to measure whether the participants were able to remember the target words and phrases from the

[2] Of the 17 students on the specialized Business Presessional Course who had participated in Session 1 using the CD-ROM, only three were able to take part in Session 2, when the C-test was administered.

various extracts they had listened to and read in the transcripts and in the glossary. A sample C-test is illustrated in Appendix 10.1. It can be seen that the emphasis was placed on phrases. The scoring scale below was used to assess the accurate use of those phrases:[3]

3 = correct phrase
2 = correct phrase but problems with morphology
1 = some idea of phraseology but could not get the correct phrase
0 = no idea of phraseology

Unfortunately, we were not able to run a pretest to determine the students' previous knowledge of the vocabulary and phrases. However, we were able to test a small group of four students at CELE who were similar to the participants in the study in that they were also studying EAP on a presessional program at CELE. Their results show that, although a few of the phrases in the test were partially produced, on the whole, most were unknown. This provides some evidence that any learning by the students in the study was due to using the program, and not previous knowledge.

Procedure

The learning treatment consisted of two teacher-led sessions with independent study in between. In the first teacher-led session, the students were asked to focus on the tasks in Part 1 that were relevant to their discipline (English, Law, or Business). A sample task worksheet for this session focusing on the language used in the Language and Gender seminar is found in Appendix 10.2. In this session's the students were asked to concentrate on the tasks that focused on the context of seminars, types of interaction, and speech acts, such as eliciting information and replying to questions. At the end of this first teacher-led session, there was a brief discussion on the students' performance and to what extent this type of material had helped them cope with fast speech and specialized vocabulary and phrases. After the initial exposure to this material, the students were advised to use the program independently to consolidate what they had learned. The program was available to them on the desktop of all computers in the IT room at CELE. After approximately two weeks, a second teacher-led session was conducted with each group. Appendix 10.3 includes the worksheet for this session, again based on the Language and Gender seminar. In this session, the students focused on the "Listening for detail" tasks and practiced understanding various speech acts. They also did the vocabulary activities, and

[3] This scoring approach is the same as was used by Jones and Haywood (2004) for formulaic sequences.

at the end of this second session, the discipline-relevant C-test was given to the students to determine to what extent they had learned the vocabulary and phrases to which they had been exposed in the teacher-led sessions and independently.

Results

INTERNATIONAL LAW GROUP

The results of the C-test given to the group studying on the specialized Law Presessional Course are presented in Table 10.4. The mean score for the whole group was 2.4 on a scale of 3, which reveals a high level of knowledge of the selected phrases. Students 2, 3, 5, 6, and 10 in particular, were able to achieve very high scores, followed by Students 4, 9, and 11. Even though Students 1, 7, 8, and 12 obtained the lowest scores, they showed a considerable ability to produce the chosen items in the test. As regards the phrases themselves, most were produced without difficulty, as shown by the high mean scores per phrase. *Tacit consent* and *in the light of* were the exceptions, but the mean score of 1.8 for these phrases does show that the students had some idea of phraseology, although they were not able to produce the whole phrase accurately.

BUSINESS GROUP

As can be seen in Table 10.5, of the 17 students who had participated in the first session, only three were able to take part in the second session due to a heavy workload on their specialized Business Presessional Course. The group mean score was 2.1, showing a reasonably good knowledge of the terms selected for the C-test on the subject of entrepreneurship. Student 1, who obtained a mean score of 2.8, demonstrated a high level of knowledge of specialized phraseology, while Students 2 and 3 both had a mean score of 1.7. It is interesting to note that both students produced what could be considered, strictly speaking, business-related phrases such as *venture*, *to set up*, *key management skills*, and *sell the idea*, but not phrases of a more general and colloquial nature such as *the only problem I've got*, *the thing is*, or *at the end of the day*.

GROUP OF TEACHERS OF ENGLISH FROM JAPAN

Table 10.6 shows the results of the C-test administered to the group of teachers of English from Japan. Like the Law group, the group of teachers obtained a mean score of 2.4. Student 3 was able to produce all phrases

Table 10.4: *Results of C-test (specialized Law presessional course students)*

Student	*economic sanctions*	*Security Council*	*regional arrangements*	*economic sanctions were imposed against*	*in the light of*	*regional body*	*carve out*	*impose sanctions against*	*impose sanctions*	*tacit consent*	*economic sanctions*	*a state of flux*	Mean
1	1	3	2	2	0	0	3	3	3	2	2	3	2
2	2	3	2	2	2	3	3	3	3	3	2	3	2.6
3	3	3	2	3	3	3	3	2	3	3	3	2	2.8
4	3	3	2	3	0	3	1	3	3	2	3	3	2.4
5	3	3	2	3	3	3	1	3	3	3	3	3	2.8
6	3	2	3	3	3	3	3	3	3	0	3	3	2.7
7	2	3	2	2	2	3	1	2	2	0	2	0	1.8
8	3	3	2	3	1	1	0	3	3	1	3	2	2.1
9	3	3	2	2	1	2	3	2	3	3	2	3	2.4
10	3	3	3	3	3	2	3	3	3	3	3	3	3
11	3	3	3	3	3	3	1	2	2	1	2	1	2.3
12	3	3	3	2	0	0	1	2	2	1	3	3	2
Mean	2.7	3	2.3	2.6	1.8	2.2	2	2.6	2.8	1.8	2.6	2.4	2.4

Table 10.5: *Results of C-test (specialized Business Presessional Course students)*

Student	*venture*	*to set up*	*entrepreneurial audit*	*key management skills*	*entrepreneurial audit*	*the only problem I've got*	*the thing is*	*sell the idea*	*run with the idea*	*the beauty of the idea*	*at the end of the day*	Mean
1	3	3	2	3	2	3	3	3	3	2	3	2.8
2	3	3	1	3	1	0	0	3	2	3	0	1.7
3	3	3	2	3	2	0	0	2	1	3	0	1.7
Mean	3	3	1.7	3	1.7	1	1	2.7	2	2.7	1	2.1

Table 10.6: *Results of C-test (teachers of English from Japan)*

Student	*gut feeling*	*single-sex groups*	*mixed-sex groups*	*peer groups*	*age groups*	*hedging devices*	*reduce the force*	*back-channeling devices*	*checking devices*	*take the floor*	*take a breath*	Mean
1	3	3	3	0	3	2	3	3	3	3	3	2.6
2	3	2	2	3	3	2	3	3	3	3	3	2.8
3	3	3	3	3	3	3	3	3	3	3	3	3
4	3	3	3	3	3	2	0	3	3	3	3	2.6
5	3	3	2	3	3	3	0	3	0	0	3	2
6	0	3	3	3	3	2	3	3	3	0	0	2
7	3	3	3	3	3	3	2	3	3	0	2	2.5
8	2	3	3	3	2	2	2	3	3	0	3	2.4
9	0	0	0	3	3	2	3	3	3	0	0	1.5
Mean	2.2	2.6	2.4	2.7	2.9	2.3	2.1	3	2.7	1.3	2.2	2.4

accurately and obtained a mean score of 3. This result also matched his performance on the in-service training course. His language level and teaching skills were of a high standard. Students 1, 2, 4, 7, and 8 obtained high scores whereas Student 9 had the lowest mean score (1.5). However, he was able to get a score of 3 for those phrases that were closely related to the subject of the type of language used by male and female speakers in the Language and Gender seminar such as *peer groups*, *age groups*, *reduce the force*, *backchanneling devices*, and *checking devices*.

Discussion

The main aim of the two teacher-led sessions focusing on the language used in the three seminars on the *Sound understanding* CD-ROM was to provide the students with numerous opportunities to understand, recognize, and eventually produce key vocabulary and collocations related to the main subject of each seminar in order to ascertain to what extent this kind of language had been acquired by the students. During both guided sessions, the students were able to listen to extracts of the seminars and perform interactive tasks to ensure comprehension of key information and language. Feedback was provided throughout the program and clear explanations were given in the Summary in Parts 1 and 2. In the case of tasks with audio recordings of extracts and transcripts, key vocabulary and collocations were highlighted in dark red and this language was also included in the Glossary, which students could refer to at any time. The tasks in the Vocabulary Check section in Part 2 of the program focused specifically on the target words and phrases that were considered to be most frequent or useful in the seminars and therefore worth learning. Saliency of this type of language in the program may have enabled the students to recall and produce the phrases in the C-test. Studies conducted on the effects of saliency of words and phrases on learning confirm that it is beneficial to draw learners' attention to specific language (Bishop 2004; de Ridder 1999, 2000, 2002). Also, exposure to both the spoken and written word could have helped the students to grasp meaning better rather than relying only on the spoken word.

As expected, not all groups were the same, but all of the three groups did show that they knew quite a bit about discipline-specific vocabulary and phraseology. Read (2000) suggests that there are different ways of interpreting C-test results. One of these is that, when the test is context-dependent, which was the case with the test used for the present study, "the test score is a measure of the test-takers' knowledge of the deleted words" (p. 114). The test results suggest that after exposure to the material on the CD-ROM, learning was involved, and the students were able to understand

text and formulaic language, and that they gained the ability to do the test by filling in the blanks correctly in many cases.

In sum, this exploratory study provides initial evidence that the CD-ROM program was effective to some extent in facilitating the acquisition of academic vocabulary and collocations. However, such a small exploratory study inevitably contains certain limitations.

The study was conducted with a small sample (24 students) and, although the C-test results were encouraging, it is not possible to claim that the same high mean scores can necessarily be replicated with a larger sample. Additionally, the fact that the students were able to produce key vocabulary and phrases in the C-test does not guarantee the same success in a freer situation, i.e., in real seminars in the students' academic departments, as the conditions in the study in question were highly controlled. Furthermore, in the text on International Law, the phrase *economic sanctions* appeared three times and variations of this phrase (*economic sanctions were imposed*, *impose sanctions*, and *impose sanctions against*) were also included in the C-test. This might have led to the students' high mean score of the C-test. Similarly, in the text on Entrepreneurship, the term *entrepreneurial audit* was included twice in the text. However, when the summaries were written during the production of the CD-ROM, these sequences were considered to be key expressions for the students to remember, as they were used extensively during the actual seminars. Therefore, it was important to draw the students' attention to their prominence in the context of the discussions that took place.

Conclusion

This chapter has described part of the needs analysis survey focusing on the problems encountered by international students when interacting with native speakers in formal and less formal academic settings. This survey informed the compilation of the Needs-Driven Spoken Corpus (NDSC), from which the academic seminar component was selected for comparison with the BNC Spoken Sampler in order to identify overrepresented key categories. Of these categories, only those which formed noun phrases were chosen for more in-depth analysis in this chapter: single common noun, adjective, and article. Specific discipline-specific words and phrases, as well as colloquial expressions, from the seminars were selected, on the basis of the quantitative analysis, and included in listening and awareness-raising tasks on the *Sound understanding* CD-ROM. Three groups of students at CELE used the program in two guided sessions and independently, as part of an exploratory study. A subsequent C-test suggests that they did

acquire some of the discipline-specific lexis taught in the material. However, although the students' mean scores were high on the whole, one has to be cautious and not claim that the high results were necessarily due to exposure to the material on the CD-ROM, as there was no test prior to the use of the program.

Discipline-specific vocabulary and phrases are important in academic spoken discourse, and it would have been hard to develop materials without a principled approach. A needs analysis and corpus analysis can provide good information about what a student needs to know to operate in this discourse environment. The materials developed, described in this chapter, were based on this principled approach and there is some evidence that they were effective. We hope this blueprint of methodology can be useful for EAP teachers to try in their own teaching situations.

Discussion questions and tasks

Reflection

1. Needs analysis surveys informed the compilation of the Needs-Driven Spoken Corpus (NDSC), covering three main spoken genres: academic seminars, the language of service encounters, and casual conversation. Why do you think these genres were chosen by the respondents?
2. After conducting the statistical comparison between the academic seminar data in the NDSC and the BNC Spoken Sampler, the authors chose key words and phrases where the noun was the nounphrase head. Look at Table 10.2. What other items on this list of overrepresented categories would have generated key words and phrases worth analyzing?
3. Look at the three types of vocabulary and phrases included in the vocabulary activities, listed on pages XX – XX. Is there another way in which they can be classified?
4. The authors conducted a small-scale study with EAP students to explore the effectiveness of the corpus-based CD-ROM materials. Do you think that, if a similar study were conducted with a larger student population studying the three same disciplines (Business, International Law, and English Studies) in your teaching environment, similar results would be obtained?

Evaluation

5. Look at the description of the three vocabulary activities included on the CD-ROM: Find the Word, Crossword Puzzle, and Odd One Out

on pages XX–XX. Evaluate the effectiveness of these activities by making a list of strengths and weaknesses.

Adaptation / Design

6. Look at the three types of vocabulary and phrases included in the vocabulary activities, listed on pages XX – XX again. Design a vocabulary activity where these words and phrases could be used in context.

Appendix 10.1

Sample C-test: Seminar on Language and Gender

In the Language and Gender seminar, the lecturer asked the students to look through some data in the form of transcripts of conversations and identify the language features used by male and female speakers. At the beginning of the seminar, the students gave their views, using their *g__ feel_____* on the type of language used by speakers within different groups: *sin _____ -se ___ groups, mix_____ -se___ groups, pe___ groups* and *ag ___ groups*.

Before the students started analyzing the language in the conversations, the lecturer reminded them of specific language features in a framework, that they were supposed to identify in the conversations, e.g., *hed _____* devices (devices used to *red _____ the for___* of an utterance), *ba ___ chan_____* devices (devices that show support for the speaker / interest in what they are saying / showing that the speaker is being followed and encouraged) and *chec_____* devices (devices used by the speakers to try to check that they are being listened to. The lecturer describes different strategies used in conversation with the word *take*: *take the flo _____* (to start speaking in a debate or discussion) and *take a bre _____* to pause to take air in before you continue to speak.

Appendix 10.2

Task worksheet for Guided Session 1: Seminar on Language and Gender

SESSION 1

In your Discussion and Seminar Skills class, you will be covering different aspects of academic discussion, e.g., the difference between a lecture and a seminar, roles of speakers, participation in seminars, interrupting, silence, and so on.

I have developed this program to help international students to observe how different speakers participate in academic seminars.

Aims of this session:

- To learn how to use the program.
- To identify different patterns of interaction in the School of English Studies.
- To become familiar with the topics of the discussion.
- To become familiar with the language used by the speakers.
- To become familiar with the vocabulary relevant to the topic of discussion.
- To become aware of what goes on in a seminar.

Procedure:
It is important that you look at the Introduction – do **not** skip it, as this will guide you so that you are able to use the program appropriately.

Read the instructions carefully.

Also, there is a logical progression of tasks, so please do **not** skip sections.

Reactions / Ideas:
Please spend some time making notes of what you learned through this program today.

Discussion:
Discuss with your partner(s) any aspects of the seminar that you listened to:

- How did you cope with fast speech?
- What subject-specific vocabulary did you learn?
- How would *you* have behaved if you'd been a participant in this seminar?
- How much lecturer intervention was there?

Appendix 10.3

Task worksheet for Guided Session 2: Seminar on Language and Gender

SESSION 2

In Session 1, you had the opportunity to listen to extracts from a seminar on Language and Gender, which took place in this university, and hopefully

you were able to become aware of what goes on in seminars as well as the roles that speakers play.

Aims of this session:

- To listen to the extracts from the seminar *in greater detail* (Listening for Specific Information) in Part 2.
- To identify the different expressions (vocabulary, phrases, and grammatical structures) that speakers use to perform different *functions*, e.g., agreeing, disagreeing, and so on.
- To follow the thread of the discussion.

Procedure:
Make sure you finish listening to all extracts in Part 1.

After reading the Introduction to Part 2, do the tasks in the "Listening for Specific Information" section.

Then do the tasks in the "Identifying Functions" section. If you get some answers wrong, listen to the extracts again until you think you have identified the correct function(s).

Do the tasks in the "Identifying Speakers' Purpose and Feelings."

Finally do the tasks in the "Vocabulary Check."

Reactions / Ideas:
Please spend some time making notes of what you learned through this program today.

Discussion:
Discuss the following with your partner:

- Were you able to understand the detailed information provided in the seminars?
- Were there any useful phrases / structures that you would like to note down so that you can use them yourself when you participate in seminars in your department? If so, which ones did you choose? Compare notes with your partner(s).
- Did you have any difficulties understanding either the content or the language? Why? What can you do *now* so that you are able to perform well in seminars in your department?

References

Aguilar, M. (2004). The peer seminar, a spoken research process genre, *Journal of English for Academic Purposes* 3: 55–72.

Basturkmen, H. (1999). Discourse in MBA seminars: Towards a description for pedagogical purposes. *English for Specific Purposes* 18(1): 63–80.

Basturkmen, H. (2002). Negotiating meaning in seminar-type discussion and EAP. *English for Specific Purposes* 21: 233–42.

Basturkmen, H. (2003). So what happens when the tutor walks in? Some observations on interaction in a university discussion group with and without the tutor. *Journal of English for Academic Purposes* 2: 21–33.

Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.

Biber, D., Johansson, S., Leech, G., Conrad, S., and Finegan, E. (1999). *Longman grammar of spoken and written English*. London: Longman.

Bishop, H. (2004). The effect of typographic salience on the look up and comprehension of unknown formulaic sequences. In Norbert Schmitt (ed.). *Formulaic sequences*. Amsterdam: John Benjamins, pp. 227–44.

Coxhead, A., & Nation, P. (2001). The specialised vocabulary of English for academic purposes. In J. Flowerdew & M. Peacock (eds.). *Perspectives on English for academic purposes*. Cambridge: Cambridge University Press, pp. 252–67.

de Ridder, I. (1999). Are we conditioned to follow links? In K. Cameron (ed.). *CALL and the learning community*. Exeter: ELM Bank Publications, 195–16.

de Ridder, I. (2000). Are we still reading or following the links? Highlights in CALL materials and their impact on the reading process. *Computer-Assisted Language Learning* 13: 183–95.

de Ridder, I. (2002). Visible or invisible links: Does the highlighting of hyperlinks affect incidental vocabulary learning, text comprehension and the reading process? *Language Learning and Technology* 6: 123–46.

Flowerdew, L. (2005). An integration of corpus-based and genre-based approaches to text analysis in EAP / ESP: Countering criticisms against corpus-based methodologies. *English for Specific Purposes* 24: 321–32.

Geoghegan, G. (1983). *Non-native speakers of English at Cambridge University*. Cambridge: The Bell Educational Trust.

Gledhill, C. (2000). The discourse function of collocation in research article introductions. *English for Specific Purposes* 19: 115–35.

Hutchinson, T., & Waters, A. (1987). *English for Specific Purposes*. Cambridge: Cambridge University Press.

Hyland, K. (2006). *English for Academic Purposes: An advanced resource book*. Abingdon, Oxon: Routledge.

Jones, M. (2003a). *Sound understanding: Listening and language awareness tasks*. CD-ROM focusing on the language of academic seminars, funded by the Centre for Teaching Enhancement, University of Nottingham.

Jones, M. (2003b). The development and exploitation of a Needs-Driven Spoken Corpus for students of English for Academic Purposes. PhD thesis, Lancaster University, UK.

Jones, M., & Haywood, S. (2004). Facilitating the acquisition of formulaic sequences. In Norbert Schmitt (ed.). *Formulaic Sequences*, pp. 269–300.

Kim, S. (2006). Academic oral communication needs of East Asian international graduate students in non-science and non-engineering fields. *English for Specific Purposes* 25(4): 479–89.

Leech, G., Rayson, P., & Wilson, A. (2001). *Word frequencies in written and spoken English*. London: Longman.

Lindemann, S., & Mauranen, A. (2001). "It's just real messy": The occurrence and function of "just" in a corpus of academic speech. *English for Specific Purposes* 20: 459–75.

Mauranen, A. (2004). "They're a little bit different": Variation in hedging in academic speech. In K. Aijmer & A.-B. Strenström (eds.). *Discourse patterns in spoken and written corpora*. Amsterdam: John Benjamins, pp. 173–97.

Mudraya, O. (2006). Engineering English: A lexical frequency instructional model. *English for Specific Purposes* 25: 235–56.

Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Poos, D., & Simpson, R. (2002). Cross-disciplinary comparisons of hedging: Some findings from the Michigan Corpus of Academic Spoken English. In R. Reppen, S. M. Fitzmaurice & D. Biber (eds.). *Using corpora to explore linguistic variation*. Philadelphia: John Benjamins, pp. 3–21.

Rayson, P. (2001). *Wmatrix: A web-based corpus processing environment*. Computing Department, Lancaster University, UK.

Rayson, P., Wilson, A., & Leech, G. (2002). Grammatical word class variation within the British National Corpus Sampler. In P. Peters, P. Collins & A. Smith (eds.). *New frontiers of corpus research*. Papers from the Twenty-First International Conference on English Language Research on Computerized Corpora, Sydney 2000, pp. 295–306.

Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.

Schmitt, N. (2004). *Formulaic sequences*. Amsterdam: John Benjamins.

Ward, J. (2007). Collocation and technicality in EAP engineering. *Journal of English for Academic Purposes* 6: 18–35.

Weir, C. (1983). *Identifying the problems of overseas students in tertiary education in the UK*. Unpublished PhD thesis, University of London.

West, R. (1994). Needs analysis in language teaching. *Language Teaching* 27: 1–19.