Walsh, S. and O'Keeffe, A**.** (2007) "Applying CA to a modes analysis of third-level spoken academic discourse". In H. Bowles, and P. Seedhouse (Eds) Conversation Analysis and Languages for Specific Purposes. Berlin: Peter Lang, pp. 101-139.

STEVE WALSH / ANNE O'KEEFFE

Applying CA to a Modes Analysis of Higher Education Spoken Academic Discourse

1. Introduction

Given the dominance of English as the main language of academia (and we acknowledge the political implications of this statement), an accurate pedagogical description of it is important for those of us who are attempting to prepare international students for the challenge of this Language for Specific Purposes. Major work is emerging in its description using corpus linguistics (for example, Biber *et al*. 2002; Mauranen 2002; Poos/Simpson 2002; Swales 2002; Biber 2003, 2006). Here we wish to add another dimension to this work by 1) illustrating how Conversation Analysis offers many synergies as a complementary tool to corpus linguistics, 2) approaching the data itself not in terms of discipline, but at the more localized level of mode of interaction (see below), and 3) investigating the pedagogical implications that our modes-based analysis provide. In this chapter, we offer an approach to characterizing the interaction of LSP classrooms, using an ESP context as an example.

The chapter is organized in seven sections. In section 2, an overview of the four classroom modes is offered. These are then analyzed and exemplified using a conversation analysis (CA) methodology in section 3. The fourth section provides an overview of the complementary corpus linguistics (CL) methodology, offering a more quantitative analysis of the data. In section 5, we take one of the four classroom

modes (classroom context) and analyze the classroom talk using both CA and CL approaches. Section 6 considers classroom applications: what is the relevance of the approach advocated in this chapter for teachers, teacher educators and materials designers? Finally, we offer our conclusions in section 7.

2. Higher education classroom modes

In this section, a framework for analyzing higher education LSP classroom talk is presented and exemplified. The framework, SETT (Self-Evaluation of Teacher Talk), emphasizes the fact that interaction and classroom activity are inextricably linked, and acknowledges that as the focus of a lesson changes, so interaction patterns and pedagogic goals change. Characterizing higher education teaching in this way is not intended to offer an all-encompassing description nor a means to 'code' interaction patterns. Rather, the intention is to offer a framework and a metalanguage which may be used to interpret interaction in the context of higher education LSP classrooms.

Central to the framework is the institutional discourse notion that much communication in the workplace is essentially goal-oriented – the classroom is no exception to this phenomenon. As a lesson progresses, teachers' pedagogic goals are constantly shifting in order to take account of their 'agenda' of the moment, to deal with unexpected problems, to vary the interaction and so on. More importantly, as goals shift, the language used to realise them must also vary. For example, an LSP teacher may, at various stages in the course of a lesson, have goals such as eliciting reactions from students, explaining a new vocabulary item or setting up a group-work task. Where pedagogic goal and language use are working together, it is more likely that the teacher will create opportunities for learning. If however, the language used to achieve a particular pedagogic goal is not appropriate, opportunities for learning may be obstructed (see Walsh 2002).

In this study, pedagogic goals were discussed with participating ESP teachers during playbacks of audio-recordings of their lessons which teacher had analysed using the SETT framework. In this way, it was possible to gain an *emic* (insider) understanding of what was 'really happening', of decisions taken and interactional adjustments made. Teachers commented on shifting pedagogic goals and corresponding language use in relation to the four modes described here. Thus, it was possible to ascertain (through a professional dialogue with the teacher) the extent to which pedagogic goals and language use were *mode convergent*: that is, complementary. The interview data from these discussions have not been included here owing to limitations of space.

According to Drew and Heritage (1992: 7), much of the research on L2 classroom interaction portrays the classroom context as something static, fixed and concrete. The reason for this is that most studies have had one of two central goals, attempting to account for either the nature of verbal exchanges, or the relationship between SLA and interaction (Wu 1998). Whatever their focus, most studies have referred to *the* L2 classroom context (singular), implying that there exists such an entity and that it has fixed and describable features which are common to all L2 contexts.

One explanation for this position is that many studies on classroom discourse have compared that 'genre' with 'real' communication without acknowledging that the classroom is as much a 'real' context as any other situation in which people come together and interact. Such invariant approaches to analyzing classroom discourse do not take adequate account of the relationship between language use and pedagogic purpose and do not recognise the goal-oriented nature of the discourse. When language use and pedagogic purpose are considered *together*, different contexts emerge, making it possible to analyze the ensuing discourse more fairly and more objectively (see, for example, van Lier 1988, Seedhouse 2004). Under this variable

view of contexts (plural), learner and teacher patterns of verbal behaviour
can be seen as more or less *appropriate*, depending on a
particular pedagogic aim.

A second possible explanation for the emphasis on 'the' single
L2 classroom context is that previous studies have tended to focus
heavily on IRF routines (Initiation-Response-Feedback). Following
the earlier work of Sinclair and Coulthard (1975), many studies of L2
classroom interaction have focused on the three-part exchange which
so often typifies the language of the classroom. While the contribution
these studies have made to our understanding of classroom discourse
cannot be denied, more recent studies employing longer stretches of
discourse suggest quite different interactional organizations (see, for
example, Jarvis/Robinson 1997, Kumaravadivelu 1999). By focusing
on longer stretches of discourse, more complex, complete relationships
emerge between interactions which are jointly constructed.
Moreover, such variable approaches acknowledge the fact that the L2
classroom environment is constantly shifting, that roles are constantly
changing in relation to the unfolding agenda, that participants play a
crucial role in *co*-constructing the interaction and that *different*
varieties of communication will prevail according to particular pedagogic
purposes. Several writers are now proposing that classroom
interaction should be investigated from a *multi-layered* perspective
which accounts for these different characteristics (see, for example,
Lantolf 2000, van Lier 2000, Seedhouse 2004).

Like other writers who adopt a variable view of classroom context,
the SETT framework, presented below, also adopts a variable
approach. Specifically, the design of the framework rests on four
assumptions. Firstly, all L2 classroom discourse is goal-oriented: the
prime responsibility for establishing and shaping the interaction lies
with the teacher; secondly, pedagogic purpose and language use are
inextricably linked – it is impossible to consider one without taking
account of the other; thirdly, any higher education L2 classroom

context is made up of a series of micro-contexts (termed *modes*) which are linked to the social, political, cultural and historical beliefs of the participants (cf. Kumaravadivelu 1999); fourthly, microcontexts are co-constructed by teachers and students through their participation, through face-to-face meaning-making and through a process of 'language socialization' (Pavlenko/Lantolf 2000).

2.1. L2 Classroom modes

In this section, the procedures used to establish the original SETT framework are described. The four modes of the framework are then exemplified in section 3. According to Seedhouse (2004), L2 classroom contexts should be seen as the *interface* between pedagogy and interaction. As such, they offer a means of understanding how the institutional business is completed. Here, the term *mode* is used to encompass the inter-relatedness of language use and teaching purpose. A mode is defined as "an L2 classroom microcontext which has a clearly defined pedagogic goal and distinctive interactional features determined largely by a teacher's use of language" (Walsh 2006: 111). A modes analysis recognizes that understanding and meaning are jointly constructed, but that the prime responsibility for their construction lies with the teacher.

The original SETT framework is based on a corpus of 14 higher education ESP lessons, totalling approximately 12 hours or 100,000 words[1]. The framework was constructed using the following procedures:

---

[1] The framework has since been applied to a much larger corpus of 1m words of academic spoken English recorded in two universities in Ireland. This corpus, LIBEL CASE (Limerick and Belfast Corpus of Academic Spoken English), comprises spoken academic data from the following contexts: lecture, seminar, small group tutorial, oral presentation and conference papers. 50% of the corpus was collected in each site and its design matrix spans subject areas and colleges within both institutions so as to achieve internal comparability and overall representativeness (see <www.mic.ul.ie/ivacs>).

1. The 100,000 words were transcribed using CA conventions to highlight turn transitions, topic shifts, overlapping speech, silence, etc.

2. The transcripts were then compared and discussed with the participating teachers through a series of awareness-raising workshops. At all times, the concern was to evaluate language use (teacher talk) in relation to the teachers' stated pedagogic goals.

3. From this initial analysis, and by looking at longer stretches of discourse, it was possible to identify several types of classroom talk, each with its own focus and characterised by specific interactional features such as display questions, direct repair, content feedback (see Table 1 below for a complete list).

4. The four modes emerged from our analysis as being the most representative of all types of interaction and as having the most clearly defined combinations of pedagogic goal and interactional features. Each mode can be characterised according to specific pedagogic goals and the language used to achieve them. Following these procedures, it was possible to identify four patterns, four micro-contexts, called *modes*: managerial mode, classroom context mode, skills and systems mode, materials mode. Each mode has distinctive interactional features and identifiable patterns of turntaking related to instructional goals. While other modes could almost certainly be identified (depending on the specific L2 context), these four are included as being representative of the interaction which takes place in the ESP classroom, because they provide clear-cut examples of different types of interactional patterning and because they can be used by teachers using samples of their own data as a means of raising awareness. By focusing on turn-taking mechanisms and topic management, and by looking beyond the IRF pattern at longer stretches of discourse, the aim was to provide a descriptive system which teachers can use to extend their understanding of the interactional processes

operating in their own classes. Below, we also illustrate the merits of a dual-approach by incorporating corpus-based analyses as a means of extracting core vocabulary.

Heritage and Greatbach's (1991) notion of 'fingerprints' is helpful to the present discussion. In that study, the researchers identify a number of socially constructed contexts in different institutional settings which they term 'fingerprints' to differentiate interactional organisations. For example, the 'fingerprint' of case-history taking in a doctor's surgery will have a different exchange and participation structure to that of making a diagnosis in the same context. Here, we are proposing that each L2 classroom mode has its own distinctive fingerprint, comprising pedagogic and linguistic features. Thus, the fingerprint of *classroom context* mode is markedly different to that of *managerial* mode; both are different again from *skills and systems* mode. The four modes, together with their interactional features and typical pedagogic goals, are summarized in table 1 below. Owing to the multi-layered, 'Russian doll' (Jarvis/Robinson 1997: 225) quality of classroom discourse, any classification is not without its problems and the present one is no exception. Tensions between and within modes do exist: rapid movements from one mode to another, termed *mode switching*; brief departures from one mode to another and back again, henceforth *mode side sequence*; the fact that some sequences do not 'fit' into any of the four modes identified. These have all posed problems for description. Moreover, the analysis is further complicated by the homogeneous and heterogeneous quality of classroom contexts (Seedhouse 2004); within a mode, every interaction is both similar to other interactions (homogeneous) and yet a unique encounter (heterogenous).

Table 1: L2 Classroom Modes (Walsh, 2006)

Mode	Pedagogic Goals	Interactional features
Managerial	- To transmit information - To organize the physical learning environment - To refer learners to material - To introduce or conclude an activity - To change from one mode of learning to another	- A single, extended teacher turn which use explanations and/or instructions - The use of transitional markers - The use of confirmation checks - An absence of learner contributions
Materials	- To provide language practice around a piece of material - To elicit responses in relation to material - To check and display answers - To clarify when necessary - To evaluate contributions	- Predominance of IRF pattern - Extensive use of display questions - Form-focused feedback - Corrective repair - The use of scaffolding
Skills and system	- To enable learners to produce correct forms - To enable learners to manipulate target language - To provide corrective feedback - To provide learners with practice sub-skills - To display correct answers	- The use of direct repair - The use of scaffolding - Extended teacher turns - Display questions - Teacher echo - Clarification requests - Form-focused feedback
Classroom context	To enable learners to express themselves clearly	- Extended learner turns. - Short teacher turns

	• To establish a context • To promote oral fluency	• Minimal repair • Content feedback • Referential questions • Scaffolding • Clarification requests

In the next section, extracts from the data are used to characterize each mode through a description of its pedagogic goals and interactional features.

3. CA analysis of modes

In this section, a CA approach is used to provide a description of each mode, together with examples from the data. This description is based on the original 12 hours' (100,000 words) recordings of ESP lessons. Table 1 should be used as a reminder of the interactional features and pedagogic goals of each of the four modes.

3.1. Managerial mode

Managerial mode occurs most often at the beginning of lessons, as illustrated in extracts 1 and 2 below, characterized in the first instance by an extended teacher turn of more than one clause and a complete absence of learner turns. In each extract, the focus is on the 'institutional business' of the moment, the core activity. Typically, there is a considerable amount of repetition and some kind of 'handing over' to the learners which occurs at the end of each sequence. At this point, there is a movement to another mode: in extract 1, for example, the pedagogic focus is re-aligned away from directing learning (managerial mode) to analysing errors (skills and systems mode).

Extract 1
```
1 T Ok we're going to look today at ways to
improve your writing and at ways which can
```

```
be more effective for you and if you look at
the writing which I gave you back you will
see that I've marked any little mistakes and
eh I've also marked places where I think the
writing is good and I haven't corrected your
mistakes because the best way in writing is
for you to correct your mistakes so what I
have done I have put little circles and
inside the circles there is something which
tells you what kind of mistake it is so
Miguel would you like to tell me one of the
mistakes that you made (3)
```

When managerial mode occurs at the beginning of a lesson, the
teacher's main concern is to 'locate' the learning temporally and pedagogically
or spatially. Once learning has been located, learners are
invited to participate: "so Miguel, would you like to tell me one of the
mistakes that you made". Locating learning is an important first step
in building a main context; consequently, in many respects, managerial
mode functions as a support to the other three modes. We can say
that it is an 'enabling' mode.

Although it is most commonly found at the beginning of a
lesson, managerial mode may occur post-activity or as a link between
two stages in a lesson, as indicated in extract 2, where the teacher's
aim is to conclude an activity and move the lesson on. As in the
previous extracts, turn-taking is wholly managed by the teacher,
learners have no interactional space and the agenda, the pedagogic
goal of the moment, is firmly in the hands of the teacher. Once the
activity is concluded, the learners are organised into three groups and
the lesson moves from one type of learning (pair-work practice) to
another (open class checking in groups). Throughout, the teacher's use
of language and pedagogic purpose are at one: the language used is
appropriate to the pedagogic goal of the moment.

Extract 2

```
5 T all right okay can you stop then please
where you are ... let's take a couple of ...
examples for these and ... put them in the
categories er ... so there are three groups
all right this one at the front Sylvia's
group is A just simply A B and you're C
(teacher indicates groups) all right so ...
then B can you give me a word for ways of
looking (3) so Suzanna ... yeah
```

The transition markers *all right, okay, so* signal the end of one part of
the lesson and alert learners to the fact that the lesson has moved on,
that pedagogic goals have been realigned with a shift in focus to a new
activity. These discourse markers are essential for learners to follow
the unravelling interaction and 'navigate their way' (Breen 1998)
through the classroom discourse. According to Armador Moreno *et al.*
(2006), discourse markers in classrooms may perform a range of funtions simultaneously and this is
certainly borne out in our data. For example, in extract 2, the discourse marker *all right* performs four
of the five functions identified by Armador Moreno *et al.* (2006: 91-93).
In its opening position, it simultaneously serves as an 'attention getter'
and a 'means of recapping', while at the end of the same extract, it
performs the dual function of 'introducing a new topic' and 'motivating
students'. There is no doubt that discourse markers in classrooms
represent a complex and complicated class: they function like punctuation
marks in a written text, or intonation patterns in a spoken text
and are crucial to understanding. In cases where discourse markers are
not used, the boundaries between modes are difficult to detect and
learners may become confused as to what they are expected to do.
Most teachers have encountered comments by learners such as: "What
are we supposed to be doing?", "Where are we?", "What's the task?".
These, and similar questions, testify to the need for signposting and
the use of language which is related to the pedagogic goals of the
moment. To summarise, managerial mode is characterised by one,
long teacher turn, the use of transition markers and an absence of
learner involvement. Its principal pedagogic purpose is the management

of learning, including setting up a task, summarizing or providing feedback on one particular stage of a lesson.

3.2. Materials mode

In this mode, pedagogic goals and language use centre on the materials being used. In extract 3, learners are completing a cloze exercise on sports vocabulary and the teacher directs their contributions; the interactional organization is almost entirely determined by the materials and managed by the teacher. Teacher and learner turns are mirrored by the material: the teacher elicits responses (81, 83, 85, 87, 89, 92, 94, 98, 100) and learners respond (82, 84, 86, 88 91, 93, 95, 101). The sequence is 'classic IRF', the most economical way to progress the interaction, with each teacher turn functioning as both an evaluation of a learner's contribution and initiation of another one. There is only one turn (99) which is not determined by pedagogic goals, though it is related; unusually, it is a learner's correction of the teacher's pronunciation. Very little interactional space or choice of topic are afforded since the interaction is organized exclusively around the material. Pedagogically, the focus can be interpreted as providing vocabulary practice around a specific piece of material. Key items of vocabulary are elicited, confirmed and displayed by the teacher through echoes of a previous contribution. In this mode, teacher echo serves a useful function, confirming a contribution and amplifying it for the other learners. In other modes, however, its function may be less useful, and there are clear instances where it can even hinder learner involvement (see Walsh 2002).

Extract 3
```
81 T (laughing) reDUCE yeah ok deuce deuce my
name's deuce ok ... now ... see if you can
find the words that are suitable in in these
phrases (reading) in the world cup final of
1994 Brazil Italy 2 3 2 and in a shoot-out
... what words would you put in there? ((1))
82 L7 [beat]
```

83 T [what] beat Italy 3 2 yeah in?

84 L7 in a penalty shoot-out

85 T a what?

86 L7 in a penalty shoot-out

87 T in a penalty shoot-out very good in a
penalty shoot-out ... *(reading)* after

90 minutes THE?

88 LL the goals goals goals *(mispronounced)*

89 T [the match] was ... what?

90 L [match]

91 LL nil nil

92 T nil nil *(reading)* and it remained the same
after 30 minutes OF (3)

93 L5 extra time

94 T extra time very good Emerson *(reading)* but
then Italy?

95 L5 lost (2)

96 T but then Italy ... what?

97 L5 lost=

98 T =lost ok 3-2 in the penalty shoot-out after
Venessi and Bagio *(mispronounced)* both
missed

99 L Bagio *(correcting teacher's pronunciation)*

100 T Bagio yes Spanish *(reading)* this was the
fourth time that Brazil had?

101 LL =won=

102 T =won ...

103 LL /won won/

104 T the World Cup very good (5) and ((2)) what's
that word? ((5))

In materials mode, then, patterns of interaction evolve from the material which largely determines who may speak, when and what they may say; the interaction may or may not be managed exclusively by the teacher. Though learners have varying degrees of interactional space, depending on the nature of the activity, their contributions are still bounded by the constraints imposed by the task in hand.

3.3. Skills and systems mode

In skills and systems mode, pedagogic goals are closely related to providing language practice in relation to a particular language system or language skill. Typically, as in materials mode, the interaction follows a lockstep organization and the IRF sequence frequently occurs. Turn-taking and topic selection are determined by the target language and responsibility for managing the turn-taking usually lies with the teacher. Pedagogic goals are normally oriented towards accuracy rather than fluency and the teacher's concern is to get learners to produce strings of accurate linguistic forms and manipulate the target language. Direct repair and scaffolding have an important role to play as illustrated in the next extract.

In extract 4, a group of intermediate level students is practising simple past forms. The teacher's pedagogic goal, as evidenced in the interaction, is to get the learners to produce patterns involving the use of irregular simple past forms *went* and *broke*. The slight pause in (218), (indicated [...]), provides the teacher with an opportunity to scaffold the learner's contribution in (219). Scaffolding involves the 'feeding in' of essential language as it is needed and plays an important part in assisting learners to express themselves and acquire new language. It is followed in (225) and (227) by direct repair, which is also used in (233) and (235). Direct repair, involving a short, quick correction, is a useful interactional strategy since it has minimal impact on the exchange structure. Patterns of interaction are only slightly disturbed and the 'flow' is maintained.

This combination of scaffolding and direct repair is found extensively in skills and systems mode, enabling learners to attend to specific features of their interlanguage while keeping the interaction 'on track', in line with the teacher's pedagogic goals, the agenda of the moment. Getting learners to 'notice' patterns (Schmidt 1990, 1995) and identify relationships is a central goal in skills and systems mode. Little attention is given to meaning, to communicative function; the prime objective is to enable learners to understand and produce target forms. It is widely acknowledged that form-focused instruction

plays a significant part in the SLA process (cf. Doughty/Williams 1998; Ellis 2001); clearly, the teacher's handling of learner contributions through scaffolded instruction and repair are central to that process. Teacher language which helps learners 'build' (Jarvis/Robinson 1997) a contribution (through scaffolding) and which lets learners know when there is a problem (through direct repair) are arguably more likely to contribute to SLA than the lengthier and supposedly more sensitive types of feedback which teachers often provide in an endeavour to avoid loss of face among learners (Seedhouse 1997).

Extract 4
```
218 L5 =the good news is he went to the went to
[...]
219 T he went to what do we call these things the
shoes with wheels=
220 L2 =ah skates=
221 L6 =roller skates=
222 T =ROLLer skates roller skates so [he went]
223 L5 [he went]to=
224 L =roller SKATing=
225 T =SKATing=
226 L5 =he went to=
227 T =not to just he went [roller skating he went
roller skating]
228 L5 [roller skating he went
roller skating]=
229 T =lets hear it he went the good news [is]
230 L5 [the] good news is he went … eh skating …
231 T good he went roller skating=
232 L5 =the bad news is he …was broken his leg=
233 T =he? (2)
234 L5 he he has …
235 T simple past …
236 L he broke=
237 L5 =he broke he broke his leg=
```

Unlike materials mode, where language practice evolves around a piece of material, in skills and systems mode, it evolves from teacher prompts and is managed by the teacher. Indeed, learner contributions typically go *through* the teacher for evaluation, confirmation or repair. To summarise, in skills and systems mode, the focus is a specific language system or sub-skill. Learning outcomes are typically achieved through tightly controlled turn-taking and topic selection, determined by the teacher. Learners respond to teacher prompts in an endeavour to produce linguistically accurate strings of utterances. The interaction is typically (though not exclusively) form-focused, characterized by extended teacher turns, display questions and direct repair.

3.4. Classroom context mode

In classroom context mode, the management of turns and topics is determined by the local context; opportunities for genuine communication are frequent and the teacher plays a less prominent role, allowing learners all the interactional space they need. The principal role of the teacher is to listen and support the interaction, which frequently takes on the appearance of a naturally occurring conversation. In extract 5, with a group of advanced learners, the teacher's stated aim is 'to generate discussion prior to a cloze exercise on poltergeists' and learners have been invited to share their experiences. The turn-taking is almost entirely managed by the learners, with evidence of competition for the floor and turn gaining, holding and passing which are typical features of natural conversation. In (258) for example, the two second pause at the end of learner 3's turn is perceived by other learners as an invitation to take up the discussion and two learners take a turn in (259) and (260), before the original speaker (L3) regains the floor in (261). Topic shifts are also managed by the learners (in 264, 270, 273), with the teacher responding more as an equal participant (265, 268, 276, 278, 280), allowing the discourse to develop within the topic frames selected by the learners. Note how in (270) the sub-topic of 'neuroses' is not developed and the original

speaker retakes the floor in (271), shifting to a new topic in (273). The only questions asked by the teacher are referential (268, 276, 278) and extended learner turns dominate the sequence (for example, 256, 261, 267). Errors go unrepaired, there are no evaluative comments and the only feedback given is content-based, normally in the shape of a personal reaction.

Extract 5
```
256 L3 =ahh nah the one thing that happens when a
person dies ((2)) my mother used to work
with old people and when they died … the
last thing that went out was the hearing
((4)) about this person =
257 T =aha (2)
258 L3 so I mean even if you are unconscious or on
drugs or something I mean it's probably
still perhaps can hear what's happened (2)
259 L2 but it gets ((2))=
260 LL /but it gets/there are ((2))/=
261 L3 =I mean you have seen so many operation
((3)) and so you can imagine and when you
are hearing the sounds of what happens I
think you can get a pretty clear picture of
what's really going on there=
262 L =yeah=
263 L =and and …
264 L1 but eh and eh I don't know about other
people but eh ((6)) I always have feeling
somebody watching watch watches me=
265 T =yes=
266 L4 =YEAH=
267 L1 =somebody just follow me either a man or a
woman I don't know if it's a man I feel
really exciting if it's a woman ((4)) I
don't know why like I'm trying to do things
better like I'm eh … look like this … you
FEEL it … I don't know=
268 T =you think it's a kind of spirit =
```

```
269 L1 =I think it's just yeah somebody who lives
inside us and ((3))… visible area …
270 L4 I would say it's just neurotic problems
(laughter)
271 L1 what what …
272 L4 nothing nothing nothing…
273 L1 but have you seen city of angels=
274 L4 =no I haven't =
275 L1 =with eh Meg Ryan and eh Nicholas Cage it's
a wonderful story and I think it's true
actually=
276 T =and does this bother you =
277 L1 =what?=
278 T =this feeling that you get does it bother
you?=
279 L1 =it's eh you know when I am alone I'm ok but
if I feel that somebody is near I would be
nervous=
280 T =I would be very nervous …
```

The predominant interactional feature of extract 5 is the local management
of the speech exchange system; learners have considerable freedom
as to what to say and when. This process of 'topicalisation'
(Slimani 1989), where learners select and develop a topic, is significant
in maximizing learning potential. The defining characteristic of
classroom context mode, then, is interactional space: extended learner
turns predominate as participants co-construct the discourse. Teacher
feedback shifts from form- to content-focused and error correction is
minimal. In short, the orientation is towards maintaining genuine
communication rather than displaying linguistic knowledge.

4. Corpus linguistics as a complementary means of exploring modes

Conversation analysis and discourse analysis have prevailed as
approaches to the study of classroom interactions. Here we attempt to
illustrate that corpus linguistics approaches also have something to

offer in tandem with existing approaches (in this case CA). Aijmer/
Altenberg (1991: 1) describe corpus linguistics (CL) as 'the study of
language on the basis of text corpora'. CL has developed rapidly since
the 1960s largely due to the advent of computers and especially their
capacity to store and process large amounts of data. This has facilitated
the systematic analysis of large amounts of language and in turn
this has meant that descriptions (and prescriptions) about the English
language have frequently been contradicted by corpus linguists who
work with representative samples of naturally-occurring language
(Holmes 1988, Baynham 1991, Boxer/Pickering 1995, Kettemann
1995, Baynham 1996, Carter 1998, Hughes/McCarthy 1998, and
McCarthy 1998). A corpus is best described as "a large and principled
collection of [computerized] texts" in spoken or written form (after
Biber *et al.* 1998: 4) which is available for analysis using corpus
software packages (for further definitions see Renouf 1997, Sinclair
1997, Tognini-Bonelli 2001, O'Keeffe *et al.* 2007). Some debate
exists as to whether CL is a theory or a method (see Tognini-Bonelli
2001) or indeed whether it is a new or separate branch of linguistics.
As Kennedy (1998) notes, corpus-based research derives evidence
from texts and so it differs from other approaches to language which
depend on introspection for evidence. Increasingly, CL is being
applied to contexts and domains outside of the study of language itself
where the use of language is the focus of empirical study in a given
context. Among the many fields where CL is being adopted to complement
other methodological tools such as discourse analysis and
conversation analysis are contexts such as courtrooms (including
forensic linguistics, Cotterill 2003), the workplace (Koester 2000,
2006, McCarthy/Handford 2004), pedagogic and academic contexts
(Farr 2002, 2003, Swales 2002, O'Keeffe/Farr 2003, O'Keeffe *et al.*
2007), political discourse, advertising and the media (Carter/
McCarthy 2002, Charteris-Black 2004, Chang 2002, O'Keeffe 2002,
2003, 2005, 2006). In all of these cases, CL offers a useful approach
to the study of language, allowing for the quantification of recurring
linguistic features, which in turn can substantiate qualitative insights.

Alternatively, this process can go in the opposite direction, from qualitative insight to quantitative finding. However, in order to interpret language results qualitatively, some framework is needed beyond corpus linguistics. CA offers a very viable complementary approach. A good example of this combined strategy is found in Carter/ McCarthy (2002), who look at one BBC radio interview with the British Prime Minister, Tony Blair, using a dual approach to its analysis. Firstly, they apply the framework of CA and subsequently they conduct a corpus-based analysis on the same data. They conclude that the CA analysis shows that the interviewer and interviewee both adhere to and exploit the generic conventions of the interview in terms of turn-taking, topic management and participant relationships. The interviewer presses an agenda of getting the interviewee to commit to action; the interviewee, in turn, responds cohesively and coherently and yet avoids direct commitment to action and maintains his topical agenda without losing face (and with useful soundbites delivered along the way, which are likely to be extracted and quoted in subsequent national news bulletins). The application of corpus techniques to the transcript reveals much about the lexical environment, especially about the high-frequency key words. Carter and McCarthy show how CA and corpus linguistics can complement each other and offer a more integrated way of understanding how conversational agendas are achieved when the two methods are used in combination.
Here we will use the following corpus linguistic techniques to look at one of the classroom modes identified above. This will function as an exemplar of the application of corpus linguistics to the study of classroom discourse and its pedagogical implications.

4.1. Concordancing

Concordancing is a core tool in corpus linguistics and it simply means using corpus software to find every occurrence of a particular word or phrase. The search word or phrase is often referred to as the *node* and concordance lines are usually presented with the node word/phrase in the centre of the line with seven or eight words presented at either

side. These are known as Key-Word-In-Context displays (or 'KWIC' concordances). Concordance lines challenge us to read in an entirely new way, vertically, or even from the centre outwards in both directions. They are usually scanned vertically at first glance, that is, looked at up or down the central pattern, along the line of the node word or phrase. Here are some sample lines from a concordance of the word *way* using the Limerick Corpus of Irish English (LCIE) taken from (O'Keeffe *et al.* 2007):

```
ether in northern Ireland is no different in a  way then em what they were desperately
      you see it?  Some of you anyhow?  Now in a  way 'What Dreams may come' it's not
      subject to study in college in fact it's a  way of life and you find this right
      and how could he present things in such a  way that he would persuade people.
ul and the purpose of life is to live in such a  way that when you die your soul is
t he was obviously he obviously lived a certain  way of live and they wanted to know
  lem that they had to deal with in a different  way they couldn't deal with it by
asically in football stadium that's the easiest  way to describe it.  There is a large
sking for you ok I find this the most effective  way.  Ok now today em you have as well
speculative because there is no evidence either  way.  You can't have evidence about
 e theologian starts from the top and works his  way down.  The theologian will have
 rts from the ground so it speaks and works its  way up.  The theologian starts from
```

Figure 1. Concordance lines for *way* from LCIE.

Concordance line analysis is very productive in terms of looking at language patterning (for a detailed illustration see O'Keeffe *et al.* 2007).

4.2. Word frequency counts or word lists

Another common corpus technique which software can perform is the extremely rapid calculation of word frequency lists (or word lists) for any batch of texts. By running a word frequency list on a corpus, you get a rank ordering of all the words in order of frequency. This function facilitates enquiry across different corpora, different language varieties and different contexts of use. Below, for example are the first ten words from three very different spoken corpora. Even from the first ten most frequent words, we can see differing patterns of use (any of these words can then be concordanced to look at how they are functioning in the texts):

1) Service encounters: a sub-corpus of the Limerick Corpus of Irish English (LCIE) comprising shop encounters (8,500 words)

2) Friends chatting: a sub-corpus of LCIE, consisting of female
friends chatting (40,000 words);
3) Academic English: The Limerick-Belfast Corpus of Academic
Spoken English (LIBEL CASE, one million words of Academic
English[2])

**Table 2: Comparison of word frequencies for the ten most frequent words across four
different datasets**

	1	2	3
Rank order	**Shop (L-CIE)**	**Friends (L-CIE)**	**Academic LI-BEL**
	Spoken	**Spoken**	**Spoken**
1	you	I	the
2	of	and	and
3	is	the	of
4	thanks	to	you
5	it	was	to
6	I	you	a
7	please	it	that
8	the	like	in
9	yeah	that	it
10	now	he	is

Even from just the first ten words of these corpora, tendencies emerge
in terms of genres and contexts of use. The shop (column 1) and
casual conversation (column 2) results show markers of interactivity
typical of spoken English such as *I, you, yeah* (as a response token),
like, please, and thanks (see Carter/McCarthy 2006). Though the
academic corpus LIBEL (column 3) is also naturally-occurring
speech, the first ten words lack the interactive markers found in first
two columns. The academic corpus results resemble more the written

---

[2] Hereafter LIBEL

data, that is to say the high frequency of:

- Articles *a* and *the*, indicating a high instance of noun phrases;

- The preposition *of,* suggesting post-modified noun phrases ;

- High frequency of *that*, especially in academic corpora, points to its multi-functionality, as a subordinator (particularly following

- report verbs or in *it* patterns) and as a relative pronoun in relative clauses;

- Prepositions *for, in* suggesting prepositional phrases.

Conversely, the LIBEL data show a lack of:

- Interactive pronouns *I* and *you,* the only pronoun that figures in

- the top ten words is *it,* which is referential as opposed to interactive:

- Response tokens or discourse markers such as *yeah, like, now*.

4.3. Key word analysis

This function allows us to identify the key words in one or more texts. Key words, as detailed by Scott (1999), are those whose frequency is unusually high in comparison with some norm. Key words are not usually the most frequent words in a text (or collection of texts), rather they are the more 'unusually frequent' (ibid). Software compares two pre-existing word lists and one of these is assumed to be a large word list which will act as a reference file or benchmark corpus. The other is the word list based on the text(s) which you want to study. The larger corpus will provide background data for reference comparison. For example, we saw above that *the* is the most frequent word in the LIBEL corpus of spoken academic English (table 2); if we select one economics lecture from this corpus and generate a word list, we can also see that *the* is again the most frequent word. However, if we compare this economics lecture word list with the larger one from the LIBEL corpus using keyword software (such as that found in *Wordsmith Tools*), it will tell us which words occur with unusual frequency, or 'keyness'. These words are then referred to as the key words. (see table 3). Scott (1999) notes the key word facility provides a useful way of characterising a text or a genre and has potential applications

in the areas of forensic linguistics, stylistics, content analysis and text retrieval. In the context of language teaching, it can be used by teachers and materials writers to create word lists, for example in Languages for Specific Purposes programmes (e.g. English for Pilots, French for Engineers), where the key specialised vocabulary can be automatically identified, either from a single text (e.g. an aeronautical training manual) or from a corpus of specialised texts.
These words are then referred to as the key words:

Table 3: Key words from an economics lecture relative to a general corpus of academic lectures

1	tax	15	higher
2	income	16	percent
3	system(s)	17	rates
4	average	18	ordinary
5	basic	19	sixty
6	rate	20	marginal
7	supply	21	scheme
8	poor	22	labour
9	thousand	23	terms
10	impact	24	cost(s)
11	equity	25	characterised
12	under	26	workers
13	both	27	systems
14	figures	28	negative

4.4. Cluster analysis

The analysis of how language systematically clusters into combinations of words or *chunks* (e.g. *I mean, this that and the other,* etc.) can give insights into how we describe the vocabulary of a language (Wray 2000, 2002, McCarthy/Carter 2002). As O'Keeffe *et al.* (2007) note, the way in which words cluster together into chunks has implications

for what we teach in our vocabulary lessons and how learners approach the task of acquiring vocabulary and developing fluency. As a corpus technique, the process of generating chunks or cluster lists is similar to making single word frequency lists. Instead of asking the computer to rank all of the single words in the corpus in order of frequency, we can ask it to look for word combinations, for example 2-, 3-, 4-, 5-, or 6-word combinations. For example, using *Wordsmith Tools,* here are the 12 most frequent 3-word combinations from the LIBEL corpus:

Table 4: Most frequent three-word chunks in the LIBEL corpus (with statistical significance)

	Chunk
1	a lot of
2	and so on
3	I don't know
4	at the moment
5	in other words
6	a number of
7	a couple of
8	a little bit
9	a bit of
10	as I said
11	first of all
12	as well as

5. Synergy of CA and corpus approaches

In this section, we consider how CA and CL approaches might be combined to offer a greater understanding of the interactional processes which characterise ESP lessons. The analysis is based on a smaller corpus totalling approximately 20,000 words taken from 14

lessons from the LIBEL corpus. These data were analysed using a CA methodology which centred on turn-taking mechanisms in relation to the perceived goal of the moment and the stated (written) lesson aims given by the teacher. Interaction patterns were found to vary according to instructional activity; for example, establishing procedures to complete an activity resulted in a very different pattern of interaction to open-class discussion. The different patterns manifested themselves in the turn-taking, sequence of turns and topic management. According to Heritage (1997: 162-3), interactants' talk is 'context-shaped' by a previous contribution, and 'context-renewing' by subsequent ones; understanding is indicated by the production of 'next' actions. In other words, participants both contribute to and demonstrate understanding of the interaction through the ways in which turns are managed. In this way, it is possible to characterise both the relationship between talk, and actions and assess the extent to which the 'talk-in-interaction' is appropriate to the shifting agenda and pedagogic goals of the moment. We hope to illustrate here that by taking one of the modes, as outlined above, and showing how CL and CA can work together, we can provide applications for pedagogy in an EAP context. For the purpose of this stage of the paper, we have isolated a sub-corpus of 20,000 words of classroom data from the LIBEL corpus. This comprises extracts from 14 lessons, all recorded in a University ESP context, involving small groups of multilingual adult-learners from a number of different educational and linguistic contexts. Their language ability levels can be regarded as 'mixed', though most are upper-intermediate level students with an IELTS score of around 6.0 or higher. All of the extracts fall into the category of classroom context mode. As outlined above, this mode involves local management of turns and topics. We frequently find examples of genuine communication and the teacher plays a less prominent supportive role, allowing learners all the interactional space they need.

5.1. Keywords

126 statistically significant key words were found when we compared

the classroom context sub-corpus with a one-million-word corpus (the Limerick Corpus of Irish English, a collection of casual conversations, see Farr *et al.* 2002). Many of these words related to the specific tasks/content within the lessons, for example *membership, join*, *mature*, *PGSU.* These were eliminated leaving the following list which represents the key words of the classroom context mode itself.[3]

Table 5: the top 20 key word

N	WORD	FREQ.
1.	is	554
2.	can	236
3.	this	292
4.	think	256
5.	yes	156
6.	ok	74
7.	agree	43
8.	choose	27
9.	maybe	75
10.	some	101
11.	firstly	21
12.	opinion	24
13.	explain	25
14.	meaning	20
15.	our	69
16.	first	71
17.	across	26
18.	what's	18
19.	third	23
20.	lastly	11
21.	also	34
22.	second	36
23.	use	37
24.	fourth	14
25.	describe	12
26.	very	112
27.	speak	18
28.	should	52
29.	must	37

---

[3] In the process of elimination, we ran concordance searches of all the keywords so as to establish whether they related to classroom context mode processes.

30.	discuss	10
31.	one	149
32.	next	39

Obviously there are too many to go through in individual detail, but there are patterns of use that we can observe. A number of words relate to students' attempts to collaboratively sequence what they are doing (as exemplified by the concordance lines below). This illustrates the need for ordinals and other sequencing lexis as core vocabulary items for classroom activity. In these data, the key words *firstly, first, second, third, fourth, next, last* appear frequently.

Figure 2: Extract from concordance lines for *first*

```
13      agree?    198. L1: =yes this is first=    199. [yeah/yeah]    200. L3:  =we
14    one=    121. L3: =I think THIS is first one this is second one=    122. L2: =how?
15    nt free of charge...maybe this is first sentence      32. L2: my link word is
16    e=    119. L3: =I think this is first one=    120. L1: =also first one=    121.
17    think=    103. L2: =ok this is first sentence this second one=    104. L1:
18    89. L1: =but eh perhaps this is first this is first and this second=    90. L4:
19    s begin    L2: ok    L1: who is first?    L: you can (2) go first    L2: we
21    --------------------    1. L1: first we say...our sentence?=    2. L2: =yeah=
22    --------------------    1. L1: first eh...marriage is a thing of the past I
23 1))    2. L1: first across=    3. L2: =first across eh... is eh...is eh noun the word
24      rd have six...six letter...the first three letter is eh (2) sorry=  3. L2: [two
25    g a town is better than living...first sentence=    25. L2: =and what do YOU
26    ...I think so    27. L1: ((4)) my first sentence is the you can use facility
27    =r.o.p.e=    42. L1: =eh...in my first sentence I think I forget eh...other
28    sports union=    13. L2: =but no first...firstly=    14. L3: =firstly is Jack's
29      continue speak    47. L: the first letter is f?    48. L4: =no I don't speak
31    me begin to talk talk about the first topic...first topic is living in a town is
32    agree this is considered as the first sentence?    109. L1:  =this
33      think my sentence may be the first sentence because...eh...every
```

We also find other words related to collaborative activity, especially deictic items such as *this, across.* Also, when we concordance the key word *one*, we find that its most frequent collocates are related to deixis and sequencing: *this one is, next one, this one, maybe this one, second one, the second one, I think this one, next one is, the last one, the next one, follow this one, last one, one by one, the first one, the first sentence, the third one.*

The type of quantitative analysis offered by a CL approach adds

further weight to the CA and modes analysis described earlier in the paper. For example, in classroom context mode (exemplified here), learners are tentatively 'feeling their way' in the discourse, coconstructing meaning and checking understanding, ensuring that peers are working collaboratively towards a common goal. Their actions and dialogue are evidenced in the high frequency of 'organisational language', of sequencing and deixis. One of the key pedagogic goals of classroom context mode is 'to allow learners to express themselves clearly' and again this goal manifests itself in the high frequency use of positional language (*maybe this one, I think this one, follow this one*) where learners take a particular stance and then adjust their position in relation to another's contribution. By combining qualitative (CA) and quantitative (CL) approaches, a much more detailed, incisive position on 'what's happening' can be adopted. We also find lexical items which suggest abundant attention to how collaboration is achieved through negotiation and careful sequencing of turns. Examples of these include: *yes, maybe, agree* and *opinion*. These concordance lines for the word *opinion* bear this out:

Figure 3: Sample concordance lines for *opinion*

```
1      =yes eh...in my opinion...ok?...there are three unions at the University of
4    n land but in my opinion living in the town or living in the country is
5   s word...eh in my opinion I think it describe something      very...stable     12.
6      gree with this opinion I think living in the country is better than living the
14 Frank what is your opinion?    2. L4: I think we can take the ((3)) and then take the
17     ea what's your opinion (3) marriage or co habit?     28. L4: =marriage    29. L2:
18      : =whats your opinion?    11. L4: do you agree on this do you want marry?    12.
19     L: what's your opinion?=  13. L1: =eh...no I'm not very familiar with these
20      1: what's your opinion?    25. L3: =I don't think so    26. L4: but I think eh..
21  heard what's your opinion?    45. L1: =can you explain this?    46. L3: yes must
22    (3) what's your opinion?    81. L: =what?=    82. L2:  =I think this sentence=
2   = now what's your opinion West?=  7. L5: =I prepare a traditional Chinese song...the
```

Again, when this analysis is considered in relation to classroom context mode and a CA analysis, there is convergence of findings. The combined CA and modes approach offers insights into the ways in which teachers and students co-construct meanings through their unfolding 'talk-in-interaction'. This qualitative approach considers ways in which turns and topics are managed in relation to predetermined

pedagogic goals. Turning then to a CL approach, we are able to consider the precise language which is being used in the interaction and evaluate the extent to which it exemplifies the interactional features of classroom context mode. For example, from the concordance lines above, it is immediately obvious that the language being used serves to help learners clarify meanings for one another and scaffold each others' responses; clarification requests and scaffolding are two of the key interactional features of classroom context mode. Essentially then, what we are gaining from this combined CA and CL approach are insights into the ways in which particular pedagogic goals are realised through specific interactional and linguistic features. Not only do these insights have relevance to L2 teacher education, they offer learners an opportunity to develop a more sophisticated interactional awareness which will contribute to enhanced learning (see section 6 below).

5.2. Wordlists and chunks (word clusters)

When we looked at the word cluster combinations, or chunks, generated from the classroom context mode sub-corpus, we firstly isolated the top 20 items which had semantic integrity and we then eliminated lesson content related items. The following is therefore a list of the chunks that relate to classroom discourse during classroom context mode.

Table 6: Summary of 2- to 5- word chunks from classroom context mode sub-corpus

N	2-word	3-word	4-word	5-word

1	I think	I don't know	I don't think so	do you agree with me
2	you can	I don't think	I agree with you	no I don't think so
3	this is	you are right	what do you think	what do you think about
4	do you	do you agree	but I don't know	yeah I agree with you
5	we can	do you think	do you think so	
6	I have	I think so	yeah you are right	
7	it is	a lot of	I catch your mean	
8	can you	no I don't		
9	there are	can you explain		
10	you mean	I want to		
11	we should	if you have		
12		if you are		
15		so I think		

Here again we find many indicators of interactive and collaborative language, for example, the use of phrases with interactive and inclusive pronouns *I/you, we*. Also, we find some deictic lexis within this batch of high frequency chunks: *this is, there are, it is.* Most striking though are the chunks relating to the exchanging and ordering of ideas and opinions (in italics in table 6). These include some items within the two-word chunks which may not appear transparently related to exchanging ideas until they have been checked with concordance lines. These are exemplified here:

You can

Figure 4: Sample concordance lines of *you can*

```
44: ok    L1: who is first?    L: you can (2) go first    L2: we can...we can
45t this...this first this last...you can read it=    35. L2: =read it we hear it=    36.
46   sorry to interrupt you maybe you can wait a moment I explain the fifteen    70.
47 t if you have not enough money you can buy a smaller one and a second hand car and eh
48      if you have enough money you can buy a big house or luxury car but if you have
```

I have

Figure 5: Sample concordance lines of *you have*

```
2 d now?=    33. L3: =(laughs) yeah...yeah I have a boyfriend=    34. L1: =do you bring
3       ou have childrens?    26. L1: =yes I have a little son    27. L3: =do you have a
4       e?=    27. L2: [I think]  28. L3: =I have a sentence with secondly=    1. L:
5       24. =seven seven across    25. L4: =I have I have across I have this is a
6       ...no...this is last one    29. L2: I have another sentence...ehm...their
7       xamination successfully    38. L3: I have confidence all of our language teachers
9       o have  126. [twenty-two]    1. L4: I have four letters (3) I think something like
```

Another interesting two-word chunk is *can you*. When we concordance
it, as exemplified below, it occurs very often in the context of
student to student requests.

Can you

Figure 6: Sample concordance lines of *can you*

```
20      e...modest=    35. L4: =can you eh...can you give me...opportunity to speak?
21      n't...eh...I don't catch what you say can you give a little more detail    17. L1:
22      disagree    4. L4: =why? (3) Sunshine can you give me some reason=    5. L1: =yeah
23          ...your handwriting is..can you can you mind my take this one...your
24      72. L2: the other sentence you have...can you show me?...eh no just now you tell
25   10. L2: =I like this song very much…    can you sing a few words for us…yeah ok=
26  13. L2: =whats your sentence?=    14. L4: =can you speak out your sentence?=    15.
27     or disagree    36. L2: ((5))    37. L4: can you speak...clearly  38. L3: [loudly]
```

It is interesting here to note the extent to which the language patterns
which have been identified using a CL approach exemplify the key
characteristics (both pedagogic and interactional) of classroom context
mode. Of note in the examples cited here is the evidence for student*Steve*
student *scaffolding*; in their co-construction of meaning, learners frequently
ask for and give support to aid the process of task completion.
Can you, shown above, for example, is clearly an important two-word
chunk for collaborative work where students are working independently
of the teacher. Simply making learners aware of examples of
classroom language which will facilitate task completion and enable
them to work well with minimal breakdown is a necessary first step in
enhancing learner classroom interactional awareness.

6. Classroom applications (based on classroom context modes)

From the brief analyses above, we can see that there are pointers for us as teachers and materials designers concerning the type of classroom vocabulary that is needed for this one classroom mode. These can be summarised functionally as:

1. *Vocabulary items that help to sequence.* Much of the discourse in classroom context mode involves learners working independently of the teacher, or occupying much of the available interactional space. For this type of classroom organization to work effectively, students need to be able to sequence and organize their talk for their listeners and pay attention to the relationship between the ideas they are trying to express. This is especially important in a higher education context where, even in language classroom, learners are endeavouring to acquire critical thinking skills and do more than simply 'regurgitate' the ideas and opinions of others. Examples would include adverbs such as *firstly, secondly, finally;* discourse markers which indicate degree such as *not only…but, even more so, furthermore;* adverbial phrases which show comparison and contrast: *on the other hand, similarly.* There is a sense in which these words, when used in classroom context mode, help listeners 'navigate the discourse' (Breen 1998) and focus attention on what McCarthy (2002, 2003) refers to as 'good listenership'. All too often in freer discussions, teachers comment on the fact that students do not listen to each other; arguably, by focusing attention on words which allow learners to sequence their ideas and signpost more carefully for their audience, some of these problems might be alleviated.

2. *Vocabulary items that help to point spatially, temporally and personally.* As identified by Carter and McCarthy (2006), deixis is a core feature of spoken grammar, whereby we point to people, places and time within our shared space as interactants. Deictic lexis such as *this, that, it, these, those, all of us, some of*

us, those of us, people like us, over there, at home, in my country etc. are all needed if we are to communicate our shared worlds in the classroom during classroom context mode. However, many of these are pedagogically challenging. For example, as McCarthy (1998: 112) notes, notorious problems arise between languages over the usage of items such as demonstratives. Deixis with *this* and *that* in English, he points out, does not always correspond with the cognate forms in other languages (e.g. Spanish, German, Danish). In addition, many temporal deictic phrases are culturally specific, for example phrases like *in the old days, in the 50s, during the 80s* are very relative. For one speaker they may have positive connotations and for others, they may represent the opposite. This is also the case for spatial references because deictic centres are very relative. For example terms like *the Mainland, the Far East, Overseas, the Continent, and Eastern Europe* can all represent just one deictic centre (that is, Great Britain in this case).

3. *Vocabulary which helps students negotiate topics and turns.* One of the key features of the discourse found in classroom context mode is longer turns and freedom to self-select topics. Control of the floor is essentially handed over to learners, with teachers playing a minimal role in shaping their contributions and helping learners to express themselves clearly. And yet breakdowns in this kind of discussion arise frequently because learners do not always understand each other – especially learners coming from markedly different L1 backgrounds, such as Spaniards speaking to Japanese. Breakdowns result in a loss in the 'flow' of the discourse, reduced motivation and a sense of failure. Two to five word chunks were identified in the classroom context sub-corpus as having particular importance owing to the extent to which they enable learners to adopt a particular stance and provide a personal perspective. For example, *I mean, I think so, I want to, but I don't know* are useful phrases in

classroom context mode and do much to minimize the risk of breakdown. Further, these chunks offer discussants an opportunity to provide 'shared space' where risks can be taken and students are not afraid of making errors.

4. *Lexico-grammatical items relating to the expression of opinion and agreeing/disagreeing*. Again, in classroom context mode, much of the activity centres on what is typically referred to as an information or opinion gap task. While learners may already be able to exchange opinions and agree/disagree, from our data, it is immediately obvious that there exists a whole range of language chunks for performing these important functions. In a higher education context, it is crucial for learners to develop criticality since they will be expected to demonstrate this, not only in their academic writing, but also when taking part in seminars and tutorials with native-speakers. Under these conditions, ESP students must be able not only to present and defend a coherent position, they must also be able to clarify other students' opinions. Again, from our analysis of this particular mode, the following chunks were found to occur frequently: *I catch your meaning, yeah you're right, do you agree with me, you mean…?, can you explain…?, do you think so?* It might be argued that these are less commonly taught to the more traditional exponents of the functions expressing agreement/ disagreement or giving opinions. Yet if students are to 'survive' in the spoken academic context of an English-speaking university, they are phrases which must be taught and mastered, probably in the ESP context.

5. *Lexico-grammatical items and discourse strategies relating to the expression of requests*. Of immense importance to students participating in any open discussion is their ability to 'keep channels open' and ensure that the dialogue does not break down. There is now some evidence to confirm that learners do 'scaffold' each other in-task (Donato 1994, Ohta 1995) and that

learners can greatly assist each other through their appropriate use of the language of requests. In our data, we were struck by the high frequency of the phrase *can you* (see above) which, depending on the precise context, might be used to ask for a contribution, check meaning, seek clarification, or 'shape' someone else's contribution. In the current educational climate of constructivism, where collaborative meaning making and peer-oriented dialogue are considered to be crucial to any kind of learning, ESP students need to master a range of expressions which will enable them to make requests as they interact in a *classroom* context. Typically, learners are given the language they need to make requests in so-called 'real world' service encounters, whereas what might be more important is to offer them strategies and language for making requests as they participate in class discussions.

7. Conclusion

Corpus linguistics is a powerful means of analysing very large amounts of data very quickly and this can give clarity to our intuitions about language and further our description of it. However, if CL is to have application for language teaching, language description for its own sake is not enough. We need to be able to look at what is happening within spoken language in its micro-context. Here we have looked at the synergy between CL and CA using data from a higher education context, within the SETT framework. In this exploration, just by looking at one of these modes (classroom context), we have been able to find many insights which have transferability for materials design for Languages for Academic Purposes. Moreover, it is apparent in our analysis that the approach advocated here could be hugely beneficial for LSP teacher education as a means of enhancing teachers' interactional awareness.

As we hope to have illustrated here, a combined CA and modes approach offers insights into the ways in which meaning is co-constructed through unfolding 'talk-in-interaction'. Within this qualitative

approach, we can examine ways in which turns and topics are managed in relation to pedagogic goals. We also looked at the CL approach to language description and by applying it to our data, we are able to consider the specifics and to evaluate the extent to which interactional features of classroom context mode were exemplified. By combining CA and CL approaches, we gained insights into the ways in which particular pedagogic goals are realised through specific interactional and linguistic features. Essentially by combining qualitative (CA) and quantitative (CL) approaches, a much more detailed picture emerges from the perspective of discourse as well as offering us information about specific language strategies and core lexis required for successful interaction in the classroom in this mode.

References

Aijmer, Karin / Altenberg, Bengt (eds) 1991. *English Corpus Linguistics: Studies in Honour of Jan Svartvik*, London: Longman.

Amador Moreno, Carolina P. / Chambers, Angela. / O'Riordan, Stéphanie 2006. Integrating a Corpus of Classroom Discourse in Language Teacher Education: The Case of Discourse Markers. *ReCALL* 18/1, 83-104.

Baynham, Michael, 1991. Speech Reporting as Discourse Strategy: Some Issues of Acquisition and Use. *Australian Review of Applied Linguistics* 14, 87-114.

Baynham, Michael, 1996. Direct Speech: What's it Doing in Non-narrative Discourse? *Journal of Pragmatics* 25, 61-81.

Biber, Douglas, 2003. Variation among University Spoken and Written Registers: A New Multi-dimensional Analysis. In Meyer, Charles / Leistyna Pepi (eds) *Corpus Analysis: Language Structure and Language Use*. Amsterdam: Rodopi, 47-70.

Biber, Douglas, 2006. *University Language: A Corpus-based Study of Spoken and Written Registers.* Amsterdam: Benjamins.

Biber, Douglas / Conrad, Susan / Reppen, Randi, 1998. *Corpus Linguistics Investigating Language Structure and Use.* Cambridge: Cambridge University Press.

Biber, Douglas / Conrad, Susan / Reppen, Randi / Byrd, Pat / Helt, Marie 2002. Speaking and Writing in the University: A Multidimensional Comparison. *TESOL Quarterly* 36, 9-48.

Boxer, Diana / Pickering, Lucy 1995. Problems in the Presentation of Speech Acts in ELT Materials: The Case of Complaints. *ELT Journal* 49, 99-158.

Breen, Michael, P. 1998. Navigating the Discourse: On What is Learned in the Language Classroom. In Renandya, Willy A. / Jacobs, George M. (eds) *Learners and Language Learning*. Singapore: SEAMO Regional Language Centre.

Carter, Ronald A. 1998. Orders of Reality: CANCODE, Communication

and Culture. *ELTJ* 52, 43-56.

Carter, Ronald A. / McCarthy, Michael J. 2002. From Conversation to Corpus: A Dual Analysis of a Broadcast Political Interview. In Sánchez-Macarro (ed.), 15-39.

Carter, Ronald A. / McCarthy, Michael J. 2006. *Cambridge Grammar of English*: *A Comprehensive Guide to Spoken and Written English Grammar and Usage.* Cambridge: Cambridge University Press.

Chang, Phoebe 2002. Who's Behind the Personal Pronouns in Talk Radio? Cartalk: A Case Study. In Sánchez Macarro (ed.), 115-152.

Charteris-Black, Jonathan 2004. Why 'an angel rides in the whirlwind and directs the storm': A Corpus-based Comparative Study of Metaphor in British and American Political Discourse. *Language and Computers* 49/1, 133-150

Cotterill, Janet 2003. *Language in Court: Power and Persuasion in the O J Simpson Trial*. Basingstoke: Palgrave.

Donato, Richard 1994. Collective Scaffolding in Second language Learning. In de Lantolf, James / Appel, Gabrielle (eds) *Vygotskian Approaches to Second Language Research.* New Jersey: Ablex.

Doughty, Catherine / Williams, Jessica 1998. *Focus on Form in Classroom Second Language Acquisition*. Cambridge: Cambridge University Press.

Drew, Paul / Heritage, John (eds) 1992. *Talk at Work: Interaction in Institutional Settings*. Cambridge: Cambridge University Press.

Ellis, Rod 2001. *SLA Research and Language Teaching*. Oxford, Oxford University Press

Farr, Fiona 2002. Classroom Interrogations – How Productive?. *Teacher Trainer* 16/1, 19-22.

Farr, Fiona, 2003. Engaged Listenership in Spoken Academic Discourse: The Case of Student-tutor Meetings. *Journal of English for Academic Purposes* 2/1, 67-85.

Farr, Fiona / Murphy, Brona / O'Keeffe, Anne 2002. The Limerick Corpus of Irish English: Design, Description and Application, *Teanga* 21, 5-29.

Heritage, John 1997. Conversational Analysis and Institutional Talk: Analysing Data'. In Silverman, David (ed.) *Qualitative Research:*

Theory, Method and Practice. London: Sage, 222-245.

Heritage, John / Greatbatch, David 1991. On the Institutional Character of
Institutional Talk: The Case of News Interviews. In Boden, Deirdre /
Zimmerman, Don H. (eds) *Talk and Social Structure: Studies in
Ethnomethodology and Conversation Analysis*. Berkeley:
University of California Press.

Holmes, Janet 1988. Doubt and Certainty in ESL Textbooks. *Applied
Linguistics* 9, 21-44.

Hughes, Rebecca / McCarthy, Michael J. 1998. From Sentence to Discourse:
Discourse Grammar and English Language Teaching. *TESOL
Quarterly* 32, 263-287.

Jarvis, Jennifer / Robinson, Mark 1997. Analysing Educational Discourse: An
Exploratory Study of Teacher Response and Support to Pupils'
Learning'. *Applied Linguistics* 18/2, 212-228.

Kennedy, Graeme 1998. *An Introduction to Corpus Linguistics*. London:
Longman.

Kettemann, Bernhard. 1995. "Concordancing in English language teaching". *TELL and CALL* 4: 4-15.

Koester, Almut, 2000. Getting Things Done and Getting along in the
Office. In Coulthard, Malcolm / Cotterill Janet / Rock, Frances (eds) *Dialogue
Analysis VII: Working with Dialogue.* Tübingen: Niemeyer,
197-207.

Koester, Almut, 2006 *Investigating Workplace Discourse*. London:
Routledge & Kegan Paul.

Kumaravadivelu, Bernard 1999. Critical Classroom Discourse Analysis.
TESOL Quarterly 33/3, 453-484.

Lantolf, James (ed.), 2000. *Sociocultural Theory and Second Language
Learning*. Oxford: Oxford University Press.

Mauranen, Anna 2002. 'A good question': Expressing Evaluation in
Academic Speech. In Cortese, Giuseppina / Riley, Philip (eds)
*Domain-specific English: Textual Practices across Communities
and Classrooms.* Bern: Peter Lang, 115-140.

McCarthy, Michael, J. 1998. *Spoken Language and Applied Linguistics*.
Cambridge: Cambridge University Press.

McCarthy, Michael, J. 2002. Good Listenership Made Plain: British and

American Non-minimal Response Tokens in Everyday
Conversation. In Reppen *et al.* (eds), 49-71.

McCarthy, Michael J. 2003. Talking Back: 'Small' Interactional Response
Tokens in Everyday Conversation. In Coupland, Justine (ed.) Special
issue of *Research on Language and Social Interaction* on 'Small
Talk' 36/1, 33-63.

McCarthy, Michael J. / Carter, Ronald A. 2002. 'This, that and the other': Multiword
Clusters in Spoken English as Visible Patterns of
Interaction. *Teanga* 21, 30-52.

McCarthy, Michael J. / Handford, Michael 2004. 'Invisible to us': A Preliminary
Corpus-based Study of Spoken Business English. In Connor, Ulla
/ Upton, Thomas (eds) *Discourse in the Professions. Perspectives
from Corpus Linguistics*. Amsterdam, Benjamins, 167-201.

Ohta, Amy 1995. Applying Sociocultural Theory to an Analysis of
Learner Discourse: Learner-Learner Collaborative Interaction in
the Zone of Proximal Development', *Issues in Applied
Linguistics* 6/2, 93-121.

O'Keeffe, Anne, 2002. Exploring Indices of National Identity in a Corpus
of Radio Phone-in Data from Irish Radio. In Sánchez-Macarro
(ed.), 91-113.

O'Keeffe, Anne, 2003. 'Like the wise virgins and all that jazz' – Using a
Corpus to Examine Vague Language and Shared Knowledge. In
Connor, Ulla / Upton,. Thomas A. (eds) *Applied Corpus Linguistics: A
Multidimensional Perspective*. Amsterdam: Rodopi, 1-20.

O'Keeffe, Anne 2005. 'You've a daughter yourself?': A Corpus-based
Look at Question Forms in an Irish Radio Phone-in. In
Schneider, Klaus P. / Barron, Anne (eds) *The Pragmatics of Irish
English*. Berlin: Mouton de Gruyter, 339-366.

O'Keeffe, Anne 2006. *Investigating Media Discourse.* London:
Routledge.

O'Keeffe, Anne / Farr, Fiona, 2003. Using Language Corpora in Language
Teacher Education: Pedagogic, Linguistic and Cultural Insights.
TESOL Quarterly 37/3, 389-418.

O'Keeffe, Anne / McCarthy, Michael J. / Carter, Ronald A. 2007. *From Corpus

to Classroom: Language Use and Language Teaching.
Cambridge: Cambridge University Press.

Pavlenko, Aneta / Lantolf, James P. 2000. Second Language Learning as
Participation and the (Re)construction of Selves. In Lantolf
(ed.), 155-178.

Poos, Deanna / Simpson, Rita 2002. Cross-disciplinary Comparisons of
Hedging: Some Findings from the Michigan Corpus of
Academic Spoken English. In Reppen *et al.* (eds), 3-23.

Renouf, Antoinette 1997. Teaching Corpus Linguistics to Teachers of English.
In Wichmann *et al.* (eds), 255-266.

Reppen, Randi / Fitzmaurice, Susan / Biber, Douglas (eds) 2002. *Using Corpora to
Explore Linguistics Variation.* Amsterdam: Benjamins.

Sánchez-Macarro, Antonia (ed.) 2002. *Windows on the World: Media
Discourse in English.* Valencia: University of Valencia Press.

Schmidt, Richard W. 1990. The Role of Consciousness in Second Language
Learning'. *Applied Linguistics* 11/2, 129-158.

Schmidt, Richard W. 1995. Consciousness and Foreign Language Learning:
A Tutorial on the Role of Attention and Awareness in Learning.
In Schmidt, Richard W. (ed.) *Attention and Awareness in Foreign
Language Learning.* Honolulu: University of Hawaii Press, 1-
65.

Scott, Mike, 1999. *Wordsmith Tools*. Software. Oxford: Oxford
University Press.

Seedhouse, Paul 1997. The Case of the Missing 'no': The
Relationship between Pedagogy and Interaction. *Language
Learning* 47/3, 547-583.

Seedhouse, Paul 2004. *The Interactional Architecture of the Second
Language Classroom: A Conversational Analysis Perspective.*
Oxford: Blackwell.

Sinclair, John McH. 1997. Corpus Evidence in Language Description.
In Wichmann *et al.* (eds), 27-39.

Sinclair, John McH. / Coulthard, Malcolm 1975. *Towards an Analysis of
Discourse. The English Used by Teachers and Pupils*. Oxford:
Oxford University Press.

Slimani, Assia 1989. The Role of Topicalisation in Classroom Language Learning. *System* 17, 223-234.

Swales, John M. 2002. Integrated and Fragmented Worlds: EAP Materials and Corpus Linguistics. In Flowerdew, John (ed.) *Academic Discourse*. London: Longman, 153-167.

Tognini-Bonelli, Elena 2001. *Corpus Linguistics at Work.* Amsterdam: Benjamins.

van Lier, Leo 1988. *The Classroom and the Language Learner*. London: Longman.

van Lier, Leo 2000. From Input to Affordance: Social-interactive Learning from an Ecological Perspective. In Lantolf (ed.), 245-260.

Walsh, Steve 2002. Construction or Obstruction: Teacher Talk and Learner Involvement in the EFL Classroom. *Language Teaching Research* 6/1, 3-23.

Walsh, Steve 2006. *Investigating Classroom Discourse*. London: Routledge.

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

Wray, Allison 2000. Formulaic Sequences in Second Language Teaching: Principle and Practice. *Applied Linguistics* 21/4, 463-489.

Wray, Allison 2002. *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.

Wu, Benhu 1998. Towards an Understanding of the Dynamic Process of L2 Classroom Interaction. *System* 26/4 525-540.