

London Metropolitan University
The Faculty of Social Sciences and Humanities

A Multimodal Perspective on Modality in the English Language Classroom

Submitted by:

Michael John Howard

93000146

PhD (Applied Linguistics and Teaching English to Speakers of Other Languages)

Thesis Submitted in Partial Fulfilment of the Requirements for the Degree of Doctor of
Philosophy in the Teaching of English as a Foreign Language and Applied Linguistics

to

The Faculty of Social Sciences and Humanities

at

London Metropolitan University

2015

Contents

Abstract.....	5
List of abbreviations.....	6
Chapter 1 Introduction	7
1.1 Becoming a teacher; becoming a researcher	7
1.2 An overview of the thesis	14
1.3 A note on Language Leader Intermediate	19
1.4 A note on examples and illustrations	19
1.5 A note on the written style of this thesis.....	19
1.6 A note on PPP and ESA.....	21
Chapter 2 Modality: course books and beyond.....	22
2.1 Lead in.....	22
2.2 A course book understanding of modality.....	24
2.2.1 A brief history of course books	24
2.2.2 Language Leader Intermediate	28
2.3 Making it interpersonal.....	31
2.3.1 Problematizing a lesson on modal verbs	31
2.3.2 Interpersonal meaning and modality: insights from systemic functional grammar.....	34
2.4 Modality in spoken language: insights from language corpora.....	39
2.4.1 Modal auxiliary verbs.....	41
2.4.2 Adverbs	44
2.4.3 Adjectives.....	45
2.4.4 Aspect.....	45
2.5 The lesson: moving from form to function, or not?	46
2.6 Rounding up	49
Chapter 3 Context: finding my way through the ‘black hole’	51
3.1 Lead in.....	51
3.2 Context of situation: foundations.....	53
3.3 SPEAKING and communicative competence: bringing in the social	56
3.4 Context of situation revisited: field, tenor and mode	60
3.5 The lesson as context: the dynamics of the classroom	62
3.6 A multimodal perspective on context: more than words.....	66
3.7 Rounding up.....	68
Chapter 4 Terminology and theoretical underpinnings.....	70
4.1 Epistemic modality.....	70
4.2 Mode	74

4.3 Multimodality.....	79
4.4 Systemic functional grammar	81
4.5 Social semiotics	85
4.6 Context.....	87
4.7 Activities.....	89
Chapter 5 Four extracts of data	92
5.1 Lead in.....	92
5.2 Methodology.....	94
5.2.1 Transcription	96
5.2.2 Classroom ethnography	98
5.3 Extract 1 Establishing social roles	102
5.3.1 Modality: linguistic markers of interpersonal relations.....	107
5.3.2 Sound: the meaning potential of sound	112
5.4 Extract 2 Modal ensembles: interpersonal meaning in multimodal communication	116
5.5 Extract 3 Authority and affinity: a multimodal perspective.....	121
5.5.1 Making meaning through posture	124
5.5.2 Making meaning through pointing and beats.....	126
5.5.3 Making meaning through gaze	127
5.6 Extract 4 Changing roles: contextual variables and communicative resources.....	130
5.7 Rounding up	138
Chapter 6 Teaching implications: applying my research	141
6.1 Lead in.....	141
6.2 Corpora and modality: using corpora to inform teaching materials	142
6.3 Interpersonal meaning: uncovering meaning through discourse.....	149
6.4 Learners and context: adapting to the social construct of the classroom.....	151
6.5 Multimodal communication: integrating modes, raising intercultural awareness	155
6.6 Rounding up.....	160
Chapter 7 Conclusion	161
7.1 Teaching strategies: research into practice	165
7.1.1 Data-driven learning	166
7.1.2 Task-based instruction	170
7.1.3 Three Is.....	172
7.2 Handling real spoken data for teaching purposes	174
7.3 Rounding up	175
References	179
Appendices.....	188

Appendix 1 <i>Language Leader Intermediate</i> audio scripts	188
Appendix 2 <i>Language Leader Intermediate</i> lesson.....	190
Appendix 3 Transcription conventions	191
Appendix 4 You must be joking! BYU-BNC corpus.....	192
Appendix 5 Cirio advertisement	193
Appendix 6 Classroom data (CD ROM)	194
Appendix 7 Student consent form	195
Appendix 8 Speech functions.....	196
Appendix 9 Sample lessons.....	198
9.1 Data-driven learning: Epistemic modality in academic writing	198
9.2 Task-based instruction: A business proposition.....	199
9.3 Three Is: Multimodal spoken interaction	201
Appendix 10 Speaker information BYU-BNC corpus.....	203
Appendix 11 Transcripts	204
Appendix 12 Multimodal moments	208

Abstract

This thesis is a study of an engage, study, activate (ESA) lesson of teaching modals of present deduction. The lesson has been taken from a published English language teaching course book and is typical of the way modal forms are presented to teach epistemic modality in many commercially produced English language teaching course books. I argue that for cognitive, social, linguistic and procedural reasons the linguistic forms and structures presented in the lesson are not straightforwardly transferred to the activate stage of the lesson.

Using insights from spoken language corpora I carry out a comparative analysis with the modal forms presented in the course book. I then explore the notion of 'context' and drawing on systemic functional grammar discuss how modal forms function in discourse to realise interpersonal relations. Moving my research to the English language classroom I collect ethnographic classroom data and using social semiotic multimodality as an analytical framework I explore learner interaction to uncover the communicative resources learners use to express epistemic modality in a discussion activity from the same lesson.

My analysis reveals that the modal structures in the course book differ to some extent from spoken language corpora. It shows that the course book offers no instruction on the interpersonal dimension of modality and thus how speakers use signals of modality to position themselves interpersonally vis-à-vis their interlocutors. The data collected from the English language class reveals that during the lesson learners communicate modality through modes of communication such as eye gaze, gesture and posture in addition to spoken language. Again drawing from systemic functional grammar I explain how these modes have the potential to express interpersonal meaning and thus highlight that meaning is communicated through modal ensembles.

Based on these findings I propose a number of teaching strategies to raise awareness of the interpersonal function of modality in multimodal discourse, and for the use of language corpora to better inform teaching materials on selections of modality.

List of abbreviations

BNC	British National Corpus
BYU-BNC	Brigham Young University – British National Corpus
CEFR	Common European Framework of Reference for Languages
CELTA	Certificate in Teaching English to Speakers of Other Languages
CLT	Communicative Language Teaching
Dip TESOL	Diploma in Teaching English to Speakers of Other Languages
DDL	Date-driven learning
EAP	English for Academic Purposes
EFL	English as a Foreign Language
ELT	English Language Teaching
ESA	Engage, Study, Activate
IRF	Initiation, Response, Feedback
L2	Second Language
PPP	Presentation, Practice, Produce
S1, S2, etc.	Speaker 1, Speaker 2, etc.
SFG	Systemic Functional Grammar
SFL	Systemic Functional Linguistics
TEFL	Teaching English as a Foreign Language
Three Is	Illustration, Interaction, Induction

Chapter 1 Introduction

1.1 Becoming a teacher; becoming a researcher

My interest in this area has developed out of 15 years as an English language teacher in its many guises. I have arrived at it through a combination of experiences from the day-to-day interaction with learners and teaching materials as well as interim periods of self-study, professional and academic teaching-related courses. It is as much a reflection on where I am (and have been) in terms of language teaching as well as a study, critique and discussion of the path that has led me here. As a way of positioning myself in relation to my research then, I should say that when I first started teaching in a language school I was a complete novice in terms of practical pedagogies (teaching techniques) and whilst I was a graduate in English Literature and Education Studies I could best be described as the kind of English language teacher that Thornbury claims ‘may have never formally studied the subject that they are teaching’ (1997: xiv). What I mean by practical pedagogies is the tool kit of teaching techniques that language teachers learn and pick up on teacher training courses, during staffroom conversations, browsing the bookshelf and nowadays the Internet, and ‘on the job’; what Thornbury means by teachers who may not have formally studied the subject they are teaching I understand to mean someone who, like myself as a native speaker of English, knows how to use the language, but in the absence of any formal instruction regarding its ‘rules’, would have difficulties explaining how it works to someone learning it. So in my first ever lesson, teaching the present simple and present continuous, I entered the classroom with a pile of cut up pictures of people doing various activities and the ‘rules’ for the two ‘tenses’ written on a prompt card courtesy of friend who had previously taught English as a Foreign Language (EFL). I do not want to dwell too much on the outcome of the lesson here, suffice to say it went well, but in hindsight perhaps I can say that I’d been handed a tried and tested template to present and then practice the grammar focus of the lesson and that was enough for me to leave the classroom feeling that teaching and learning had taken place. Accordingly for the next few years it seemed sensible and practical for me to defer and accept handed down knowledge from more experienced practitioners of Teaching English as a Foreign Language (TEFL).

Go in to any academic bookshop and there’s normally a wide selection of books on TEFL. These range from activity books, grammar reference, course books, lesson fillers, dictionaries and vocabulary books, resource packs as well as books on teaching methodologies and approaches. Go in to any language school and there’s normally a smaller yet similar collection in the staffroom. In short there are plenty of published ideas out there for the TEFL teacher. A lesson can be put together ad hoc five minutes before the start of a lesson using recycled ideas – ideas that must be sound for the simple reason that they have been published. Moreover, it is common practice in many language schools to

use a published course book as a means of providing structure to the 'course'. The course book is normally divided into units or modules each with a language focus, be it lexical or grammatical, a selection of exercises, activities or tasks and a focus on 'the four skills': reading, writing, speaking and listening. Indeed, the 'course' often refers to the starting and finishing of a course book and I would say it is very often the course book, and related progress tests, that dictate the level of the learner. That aside it was very comforting for me as an unqualified novice with no formal knowledge on the subject I was teaching to have this knowledge base from which to draw. I was learning the rules of the language whilst I was teaching it.

I feel I need to add here that what I have written above is not intended to make me look unprofessional. Fifteen years ago English language teaching was a very different profession from what it is today in many ways. Visa regulations, globalisation, shifts in the global economy and the 'marketization' of English language teaching (Howatt, 2004) have been major factors affecting student numbers, the kinds of English being sought and offered, the number of teaching institutions, and the means of gaining employment. When I first started teaching it was sufficient for me to be a recent graduate with enough enthusiasm and confidence to find employment at a small private language school. I was 'unqualified' in that I didn't have a recognised TEFL qualification but as I have indicated above, there was enough published material readily available to make up for any lack of formal training or formal instruction in the language I was teaching. In terms of my job security it was enough to have 'bums on seats'.

The reader will be pleased to know that I am not writing a year by year, lesson by lesson account of my teaching career. I do though need to mention two more significant events: firstly, the CELTA (Certificate in English Language Teaching to Adults) accredited by Cambridge, and subsequently the Diploma in Teaching English to Speakers of other Languages (Dip TESOL) accredited by Trinity College, London. After two years in the same language school I decided to move to Italy and teach. To do so meant gaining a recognised TEFL qualification. The CELTA I did was a four-week full-time course comprising input sessions, observed teaching, a short written assignment on a specific learner, and an exam on explicit rule knowledge of English grammar. There was one set reading: *The Practice of English Language Teaching* (Harmer, 1991). Cambridge, who accredit the course, say in their overview that those who take the CELTA will 'learn the principles of effective teaching, gain a range of practical skills for teaching English to adult learners, get valuable hands-on teaching experience' (Cambridge website). Whilst these claims are obviously very general and arguably lacking in substance, I will say that having done the course I did gain practical skills and valuable teaching experience for teaching English to adults. This was mainly through watching the trainers teach and picking up tips during the input sessions. The examinable component was a bit more nerve racking with the worries of having

no formal training in English grammar at the back of my mind. However, after two years of preparing lessons using grammar references I found that I was well equipped to take the test: grammar is grammar after all, and rules are rules, the only difficulty for me is remembering them. The course was enjoyable and useful and while I wouldn't say that I taught very differently before and after I did feel refreshed with a few more tools in my teaching kit.

After I returned from Italy I spent the next few years teaching at several different private language schools in London. Perhaps it was the variety of students, the range of course materials, the chats and ideas from colleagues, the principles and educational ethos behind the schools themselves, employment prospects, as well as my own experience over the years that made me think I needed to get back into the classroom myself and delve more deeply in to Teaching English as a Foreign Language. I had begun to notice that different learners in different settings were more or less responsive to particular teaching techniques than others; some materials could be used over and over again whilst others were fated from the outset; lessons more often than not deviated widely from the lesson plan with aims and objectives harder to realise than the plan had set out to achieve; and not least I began to question some of the descriptions of grammar I had previously taken as read. Perhaps most significantly I began to tune in to a kind of English that was different to the English I normally heard around me. By that I do not mean the English produced by the learners per se, but the English of EFL: the kind of English found in course books; the kind of English I was teaching.

On reflection it was somewhat different to the English I used and experienced outside of the classroom. I started to look closely at course book English and think 'that sounds odd', 'would we say that?' And frequently I discovered there were certain lexicogrammatical structures that learners really never seemed to produce orally even though they understood the rules of use and could manage them 'on paper'. For example, modal finites, or modal auxiliary verbs (modal auxiliaries hereafter) as they are more commonly referred to in EFL. These prove to be an area of grammar in which learners have difficulty. In my experience learners often learn the rules by rote and have little difficulty doing gap-fill exercises in which a number of sentences need to be completed using a given set of words. However, when it comes to spoken discourse they frequently do not feature. Curious, I thought, as modal auxiliaries turn up regularly in TEFL course books and materials. I decided, therefore, to enrol on a Dip TESOL and this is where my journey proper begins.

The course is accredited by Trinity College, London and followed a similar format to the CELTA with input sessions, observed teaching, written assignments and an exam, all of which needed to be passed in order gain the diploma. As a teaching qualification it is highly revered in certain areas of the profession with an unwritten belief, held by many professionals and employers, that a teacher is not

fully qualified until they have it. There was, in addition, a much more extensive reading list. As I have said above, this is where my journey really begins and so I need to mention a section of the course reading along with a few quotations from each. Firstly then, *Practical English Usage* (Swan, 1980) a book aimed at intermediate students onwards and teachers of English. The introduction claims the approach and style is 'as practical as possible' and Swan explains that '[e]ach entry contains an explanation of a problem, examples of correct usage, and (when this is useful) examples of typical mistakes' (p.xi); the examples he uses, he points out, are 'as realistic as I can make them' (p.xii). *Grammar for English Language Teachers* (Parrott, 2000), written for teachers of English with two primary aims: 'to help you develop your overall knowledge and understanding of English grammar' and 'to provide a quick source of reference in planning lessons and clarifying learners' problems' (p.1); spoken and written data is authentic and taken from a range of contexts (p.2). *A Communicative Grammar of English* (Leech and Svartvik, 1975) for advanced students and teachers 'devoted to the uses of grammar, rather than to grammatical structure' (p.10) using a 'communicative approach'. *About Language* (Thornbury, 1997) written to make up for the shortfall in English teachers' knowledge of the subject they are teaching, by providing 'explicit *knowledge* about language' (Thornbury, 1997, p.x; original emphasis). Thornbury writes that as part of his commitment to real usage 'the examples chosen to illustrate features of the language systems have been collected from authentic sources' (p.xv). *Discourse Analysis for Language Teachers* (McCarthy, 1991) a 'study of the relationship between language and the contexts in which it is used' (p.5), of how real people use real language in use (p.1).

The above is only a selection of the course reading but I have cited these few texts because of their significance to how baffled I was. At first I thought I was just a slow learner but in retrospect it is more likely that it was due to the overall tension between the various approaches to grammar taken up by the writers and that left me confused - I should point out that we, as students, were not made aware that there are different approaches to the description of grammar. Whilst I am not critiquing any book in particular I found it hard to see the alignment between their different stances. For example, the crossover between the terms use and usage of language;¹ the importance/absence of importance that some writers put on using 'authentic' data; the presentations of language in context and study of discourse, contrasting with examples of isolated sentences written to sound 'realistic'; and discourse analytical, structural and communicative approaches all seemed to pull in different directions. On reflection the course reading was a hotchpotch and lacked consistency, but as I was familiar with

¹ I expand on these terms further on in my thesis (see 6.3 *Interpersonal meaning: uncovering meaning through discourse*). For clarity at this point 'use' refers to actual language use and 'usage' to conventionalised patterns of/for use. Widdowson (1978) notes the distinction 'is related to de Saussure's distinction between *langue* and *parole* and Chomsky's similar distinction between competence and performance' (p.3)

pulling books off the shelf to reference a particular language point for a given lesson I saw it as part of the general eclecticism of TEFL.

My overall experience of the Dip TESOL was negative. I will not dwell on that here but, in brief, 'the Dip', for me, was not much different to the CELTA. The assignments and exam were longer and the lessons needed to be more rigorously planned and timed; and there was an aversion to using course books with a preference for 'authentic' material, and a big emphasis on contextualisation: two elusive terms that have preoccupied me since. But apart from picking up a few useful tips along the way I wasn't much different to the teacher I had previously been. It is true to say that we looked historically at different approaches to language teaching but the assessed lessons differed very little from the CELTA in the way we presented and then practised discrete language items: trainees were expected to structure their lessons using a common template: engage, study, activate (ESA). Accordingly I finished the Dip TESOL with more questions than answers and this is why I have included the experience as a relevant part of my journey to my current research.

Here I want to pull together the various strands of my preamble. Firstly, as I have said I was learning the rules of the language whilst I was teaching it. So, using grammar references at the back of course books I, like my intermediate and upper intermediate level students, learnt that the modal auxiliaries *may, might, must, can, could, will, would* are used to talk about degrees of certainty, probability and obligation. Swan, Parrot, Leech and Svartvik and Thornbury, their relative stance on English grammar notwithstanding, confirmed this and described modal auxiliaries in similar ways: they divide into two groups, one to do with degrees of certainty (extrinsic meaning), the other to do with obligation and freedom to act (intrinsic meaning); they have rules regarding inversion and third person -s, rules about not being preceded by auxiliaries and having 'past tenses', i.e. rules of form and function. As I have said though, the rules weren't the problem, the problem was getting learners to use modal auxiliaries in spoken discourse.

Secondly, the word context. From the blurb on language course books, through Harmer to the authors mentioned above on the Dip reading, 'context' was a word that kept recurring. Context, I have discovered since, is a big word, but in earlier days my encounters with context left me with the impression that it was a clever and convenient way of packaging language items for presentation; I will give two examples.

During the CELTA the trainer showed us a technique for teaching 'modals of certainty'. He drew a picture of a house and labelled it 'John's house'. In one of the windows he drew a light shining (he was good at drawing). Next to the house he wrote the words *may* and *might*. He drew a line down the centre of the board and on the other side of the line he drew a house, labelled it 'John's house'

but in this house there was no light from the window. He wrote the word *can't* next to the second house. Using the pictures and the linguistic prompts, we were able to make speculations using degrees of certainty such as 'John must/may be home' and 'John can't be home'. This is a typical TEFL technique and would normally be at the presentation stage of a presentation, practice, produce (PPP) lesson model. In this model the language focus is presented by the teacher in 'clear situational contexts' (Harmer, 1998: 64), practised by the learners in a controlled way 'using accurate reproduction techniques' (ibid) such as oral drills, written gap-fill exercises, etc., and then produced in a less controlled way using a spoken or written fluency activity. Whilst this model has been critiqued (Willis, 1996; Skehan, 1998; Ellis, 2003 among others), it can be quite an effective way of teaching 'a small sample of language with the focus on a particular form' (Willis, 1996: 134). It is a tightly controlled progression of classroom activity (Samuda, 2001) with the teacher organising, controlling, monitoring, prompting and correcting at the practice and production stages. Ellis (2003) points to practical problems at the production stage as:

it is not easy to design tasks that require learners to use the target structure, as learners can always fall back on their strategic competence to circumvent it. One way out of this problem is to make it clear to the learners that they must use the targeted structure when they perform the task. (p.29)

He adds that this would result in a primary focus on form. My current research has developed out of an interest in the relationship between the presentation context and the production context, i.e. a presentation of language form(s) and function(s) using a 'clear situational context' not always leading to the production of the same form(s) at a later stage in the lesson.

PPP has been questioned in relation to the learning process (see for example, Lewis, 1993; Scrivener, 1994; Ur, 1996; Willis, 1996; Harmer, 1998; Ellis, 2003). However, as many language course books adopt this way of staging a lesson (Nitta and Gardner, 2005; Masuhara et al., 2008) there is a tension between some areas of theory and what is widely available in the form of published material. My second example uses another technique, one which was favoured and prompted on the Dip that I undertook and also used by course book writers: engage, study, activate (ESA). Engage: the students are 'emotionally engaged with what is going on' (Harmer, 1998: 66) using a discussion, listening, reading, etc. as a means of contextualising the language focus of the lesson; study: 'any teaching and learning element where the focus is on how something is constructed' (ibid); activate: the stage 'at which students are encouraged to use all and/or any of the language they know' (ibid: 67). It should be apparent then that PPP and ESA are not dissimilar. Whilst the engage stage allows for the learners to be more 'emotionally engaged' (a rather opaque term) than the presentation stage of PPP, and the

study stage has the potential for a more inductive approach to language analysis, they both share a linear progression from presentation to production. As I have said the technique is one used by course book writers and my next example comes from an intermediate level course book I was expected to use at a language teaching institution.

The learners are shown three pictures. They listen to three pairs of students at a museum of design talk about the pictures and order them in the sequence in which they are described, this stage 'engages' the learners. They listen again and decide whether six written sentences are true or false. The study stage focuses on modals of present deduction and progresses from identifying modal verbs in the transcript of the recording, through an analysis of meaning to rewriting eight sentences using the modal verbs *must*, *can't*, *could* and *might*. For the activate stage the learners are given three pictures of different designs and asked to make deductions in pairs about what the designs might be using the modals they have studied.

So the two techniques open with a contextualization of language. This language then becomes the language focus of the lesson and is later expected to be activated or become productive output by the learners through a speaking or writing activity. There has been much research in to formal instruction and second language acquisition (see for example, Ellis, 1994). However, my research does not investigate this and I present no evidence on the relative merits or otherwise of PPP or ESA on the learning process or language acquisition. What I am interested in looking at, therefore, is not the likelihood or otherwise of linguistic items being processed from one stage of the lesson to the next but how context affects linguistic output, and in particular modal auxiliaries, e.g. *may*, *must*, *can*, *could*. The two examples of teaching models I have given above seem to rely on an assumption that 'context' is something that remains constant across all stages of the lesson; and if context remains constant, it seems to follow that the contextualised language should remain constant too. Accordingly language used in one stage is assumed to be carried over in to a subsequent stage of the lesson. So in the ESA model the modals learners are presented with in the context of 'three pairs of students at a museum of design' are expected to be used by the learners in the context of them working in class in pairs making deductions about a set of pictures. My observations to date, however, reveal that this is seldom the case. Of course, as I have indicated above, teacher intervention is built in to the design of PPP and ESA and so with the teacher monitoring, learners are prompted to use the language focused on in the lesson and in this respect language can be transferred from one context to another. But, to reiterate, my research is not an attempt to investigate language processing during the lesson or to go over well-trodden ground on the merits and restrictions of PPP and ESA. My interest is on modality and context, how context is far from being constant during the course of a lesson, and the relationship between context and the communicative resources learners have to make meaning.

To do this I need to explore modality in greater depth. I need to go beyond typical TEFL descriptions and look at how it functions in discourse. I need to broaden my understanding of context and investigate how context affects language choice and what effect this might have on activities in the language classroom.

A final point I want to add here relates to two further qualifications: an MA Teaching English to Speakers of Other Languages and Applied Language Studies at London Metropolitan University, and an Advanced Diploma in Applied Linguistics with the Open University. I undertook these qualifications out of my own interest in the field and not as necessary requirements for my teaching career. I separate these two courses out from the CELTA and Dip TESOL as I see them as being more research-oriented than teaching-practice qualifications. Over my teaching career and during the teacher training courses I mention above numerous questions have arisen. Some of these questions I have discussed in depth on the MA and the Advanced Diploma. I see my PhD thesis as an extension of this process and an attempt to shed some light on a few more of those questions and by doing so to better inform my teaching practice.

1.2 An overview of the thesis

In *Chapter 2 Modality: course books and beyond* I give an overview on how the grammatical content of published ELT course books such as *Language Leader Intermediate* have been informed in the main by structural approaches to grammar. With specific reference to *Language Leader Intermediate* I show that the choice of modal structures follows a course book tradition of putting a primary focus on modal auxiliaries. I present arguments that contend the choice of grammatical content in course books is largely based on introspection and not empirical evidence. Using a corpora-based grammar reference, I show that the selection of modal forms and structures in *Language Leader Intermediate* does not reflect actual language use; there are discrepancies between structures favoured in spoken and written English, there are modal forms which are very rare in spoken English, i.e. modals with continuous aspect, and even the overall choice of modals is questionable if corpus findings are considered. Using insights from corpora and descriptions of modality from systemic functional grammar, I discuss the interpersonal function of modality. I claim that the notion of modality in TEFL is limited to teaching about degrees of certainty and possibility with no mention of its interpersonal function. Such descriptions do little to explain how modality is used interpersonally, what Van Leeuwen calls 'the essence of modality' (2005: 176). My classroom observations have revealed to me that whilst learners can use modal auxiliaries in gap-fill exercises, they often do not produce them in spoken discourse. It might be that explanations of modality that focus on rules of usage, for example degrees of certainty, do little to foster their production when learners 'engage interpersonally and express points of view' (Coffin et al., 2009: 226). It might be that classroom situations giving rise to

certain interpersonal relationships negate the use of modal auxiliaries altogether, for example pairings of learners where the relationship is asymmetrical. It might also be that modals finites are not frequently occurring linguistic items in spoken discourse; recourse to spoken language corpora can provide insights here. Modality may be signalled using other linguistic structures and communicative resources, and it is these that this thesis, in part, sets out to investigate.

‘Context’ has been a recurring and problematic term for me throughout my teaching career and is central to my understanding of modality. In the third chapter (*Context: finding my way through the black hole*) I discuss what the term ‘context’ might mean in TEFL; again with specific reference to *Language Leader Intermediate*. My starting point is comments such as the following from the Teacher’s Book of the course book already referred to:

[t]his lesson looks at famous designers and some of the things they designed. Students read a text about a famous American designer. This is followed by a listening activity on different designs and designers which contextualises the grammar focus of the lesson: modals of present deduction (*can’t, must, might/could*). Students then do a series of activities to explore the meaning and form of these modals before using them in a speaking activity. (Albery, 2008: 98)²

I am uncertain as to what is meant by the grammar is contextualised. If the writer means modal forms are presented in stretches of text, then perhaps co-textualised is a better term; if he means the grammar is contextualised through the relationship between the situation, the speakers and the choices of language they have available to them, then it hasn’t been explained in that way to the students. The teacher’s notes explain that the modals of present deduction mean ‘I personally believe this because of something I know or can see’ (ibid: 100) and gives no mention to the interpersonal dimension that gives rise to modality – surely an integral aspect of context. This is of significance as contextualisation is intended to be key to the students understanding of grammar:

[t]he course covers all the key grammar points. These points are all contextualised and students are generally encouraged to analyse and understand grammar through an inductive approach with reference to examples in texts. (ibid: 3)

Modality in context is central to my work and I explore not just how context is used to present language but also how it activates language. I do not believe it follows that language presented in one context, e.g. the listening activity outlined above to present modals, will give rise to the same modals in another context, e.g. students working in pairs. I argue that context is not something that can simply be ‘handed’ to students as a kind of constant, static communicative space into which linguistic features

² For the three audio scripts for this lesson see *Appendix 1*.

fit nicely, but dynamic in the way that it both shapes and is shaped by participants and their communicative resources. I discuss how context has been addressed by Malinowski, Hymes, Firth, and Halliday and how aspects of their work have informed the way context is treated in teaching English as a foreign language. I discuss too how the language lesson is a context and thus shapes and is shaped by the communication that goes on there. At times this leads to 'mismatches' between the teacher's lesson objectives and the ways learners go about engaging with the lesson content. Finally I discuss context from a multimodal perspective. From my observations of classroom interaction it has become apparent that participants use a range of communicative resources to make meaning; at times modes such as eye gaze and gesture are more prominent than the verbal mode.

The first two chapters are a means by which I can gain a better understanding of modality and context. From my discussions it has become clear that the lesson on modals of present deduction from *Language Leader Intermediate*³ referred to above is problematic. I have suggested that for cognitive, social and procedural reasons the linguistic forms and structures presented in the lesson are not likely to be uttered productively; at least not in the way the course books writers had intended, i.e. through a discussion replicating the listening dialogues which 'contextualise' the forms and structures. To that end what I want to investigate is what *is* happening when the learners engage in the activity, what communicative resources they use and in particular how modality is articulated. In *Chapter 4 Terminology and theoretical underpinnings* I lay out the theoretical approaches that I use in my data collection and analysis. I revisit modality as a means of refining the kind of modality I am researching. I introduce communicative modes and multimodality and discuss how communicative events comprise ensembles of modes to express meaning. I draw on the notions of field, tenor and mode as well as the ideational, interpersonal and textual metafunctions from systemic functional grammar to show how the lexicogrammatical system is a system of choice and that meaning is made through the meaning potential of the system. As my data is multimodal I adopt a social semiotic view of communication and show how the ideational, interpersonal and textual metafunctions can be used to discuss the meaning potential of modes of communication other than the verbal. I revisit context and introduce the concept of higher- and lower-level actions (Norris, 2004) as a means of honing in my research lens to avoid being drawn in to the 'black hole' of context.

The data for analysis has been collected ethnographically using video recordings of learners in English language classes. As ELT teaching materials have been a major factor in my arrival at this research topic, the learners were recorded doing activities taken from a published ELT course book. Ethnographic data collection has allowed me keep the camera rolling for the duration of the class

³ For a pdf of the complete lesson from *Language Leader Intermediate* see *Appendix 2*.

without it becoming too intrusive. The idea is that the learners will get used to the camera being there and not feel they need to 'perform' for the sake of my research. However, as classrooms can be very noisy places at times it has been necessary for me in practical terms to focus my recording equipment on small groups of learners rather than having a 'global' perspective on the classroom. That said recording small groups of learners within the classroom has become integral to my research. I understand the classroom to comprise a multiplicity of micro-contexts each being a dialectical relationship between the socio demographic of those who are part of it and their communicative resources. Accordingly the language used in one context might be significantly different to the language used in another. This has implications for teaching models that try to shoe-horn 'one size fits all' linguistic features in to speaking activities such as the one I have used for my data collection.

Having taught many hours of English language classes I have been both a passive and active observer of numerous communicative activities. And I have often been surprised how communication differs not only linguistically from the way that lesson plans and teaching materials lay down but also in the modes of communication that learners use to make meaning. Language, whether spoken or written, is typically seen as the primary means of communication in the language classroom. However, I see language as part of a 'multimodal ensemble' (Jewitt, 2009: 14) and at times one mode may or may not take a more central role in a given communicative event. In *Chapter 5 Four extracts of data* I explore how and why language learners use these other modes of communication, such as posture, voice and gesture, and default from using the specific linguistic items that have been the focus of the lesson. My data analysis takes a multimodal approach. Multimodal analysis is 'concerned with the socially and culturally situated construction of meaning' (Bezemer and Jewitt, 2010: 180) making it suitable for my research in to modality in context. It should be noted that multimodal analysis is not a theory but a 'field of application' (ibid) and the qualitative descriptions I put forward are not intended to be taken as generalisations. My analysis is a study and discussion of what I have observed ethnographically and how it relates to my understanding of communication. I use a lesson from *Language Leader Intermediate* on modals of present deduction. My data comes from video recordings of learners at the 'activate' stage of the lesson in which they engage in a discussion intended to practise the modals forms they have been learning. I use four extracts of data and explore various communicative modes, such as sound, gesture, proxemics and eye gaze.

As much of what I have said relates to teaching materials, training courses and teaching techniques, the last chapter, *Chapter 6 Teaching implications: applying my research* is a discussion on the teaching implications of my research. I will suggest how corpora can be used to inform teaching materials based on the argument that invented examples are misleading. As much of my discussion looks at the interpersonal function of modality, I propose that the teaching of modality should not be limited to

descriptions of certainty and possibility but include how speakers use modal structures to express social relations during spoken interaction. Another thread of my research is context and how context affects language choice. I discuss how the ESA teaching model for this particular lesson provides the learners with the opportunity of oral production of a very specific kind. ESA is designed for a 'one size fits all' approach to language teaching with language presented as a 'product' (see Skehan, 1998; Ellis, 2003 for a critique on language as product) to a homogenous group of learners. Based on the notion that context is dynamic I argue that a discussion activity during the activate stage is problematic. Drawing comparisons between my classroom-based research and understandings of multimodal communication, I argue for the inclusion of multimodal texts in the language classroom to expand the communicative competence to multicompetence and raise intercultural awareness.

In *Chapter 7 Conclusion* I pull together the ideas discussed in the previous chapters and following my indication that this research is 'practice-derived' I put forward teaching strategies that include the use of corpora, analysis of multimodal discourse and tasks to promote language use. This is followed by a discussion and example of handling real spoken data for teaching purposes.

1.3 A note on Language Leader Intermediate

Before continuing I feel I should point out that my research is not intended to discredit *Language Leader Intermediate*. In fact, I have used this book as a case-in-point because it provided the syllabus for the university EAP course I was teaching at the time of my data collection. It typifies other popular commercial courses (see 2.4 *Modality in spoken language: insights from language corpora*), as I mention later, and, as I discovered in relation to modality, was not particularly beneficial to students' learning. As I have said whilst they were able to use the modal structures, more or less proficiently, in isolated examples, they were typically absent in extended discourse both spoken and written. As my job as a teacher of EAP was to prepare students for university life in which modality figures prominently in spoken and written interaction, at both a personal and academic level. My thesis is an argument for more effective ways of teaching modality, 'fuller' descriptions of modality to include its interpersonal function, and selections of modality based on actual language use.

1.4 A note on examples and illustrations

A number of the spoken dialogues, course book examples and illustrations have been used more than once throughout my thesis. This has been done for two reasons. Firstly, for the convenience of the reader to relieve them of the cumbersome task of referencing back pages; and secondly, I revisit the same material as a means of adding to my analysis by approaching it from different analytical perspectives. This second point will become apparent over the coming pages.

1.5 A note on the written style of this thesis

Before moving on I want to provide an explanation for the writing style of my thesis. I have pointed out above that my research focus has arisen out of my own practice as an EFL teacher and for this reason I have opted for a reflexive, narrative approach to my writing. However, it would be wrong to suggest that I am simply providing a descriptive account of my classroom observations; my thesis sets out to answer a number of research questions and accordingly my reflexive, narrative approach has been informed by theoretical perspectives and theory-generated analysis. My research is practice-derived (English, 2012) and intended to provide valuable insights into multimodal communication and English language teaching, and to develop new ideas for language teaching pedagogy; it is written with both language teachers and researchers in the field in mind.

The narrative approach I adopt incorporates the use of my personal voice. I sit very much at a crossroads between what can be seen as two communities of practice: teachers of English as a foreign language and researchers in the field of language teaching (see Tavakoli and Howard, 2012; Tavakoli,

2015), and whilst I am an experienced English language teacher, I am somewhat of a novice in terms of writing academic research. That said I am keen to write as a 'knower', an expert in my field, and not someone who is struggling with the discursive identity and practice of a field to which I am a 'peripheral insider' (Lave and Wenger, 1991), i.e. the research community. To that end I have opted to write using my own 'writer's voice' (English, 2012: 103). The process of communicating disciplinary information in this way has allowed me to develop my own discursive identity and bring about a shift in agency (see English, 2012 for a detailed discussion on this area).

Reflexive writing of the kind I produce in my thesis, along with the use of personal voice, typifies much work emerging from the field of academic literacies (Scott, 1999; Lillis, 2001; English, 2012). I hope this less traditional or conventional style and approach does not make the reader feel 'uncomfortable'.

1.6 A note on PPP and ESA

Throughout this thesis there are frequent references to the teaching models of PPP and ESA (see *1.1 Becoming a teacher; becoming a researcher* and *2.2.1 A brief history of coursebooks* in this thesis for descriptions of these teaching models). My research has developed out of a lesson on modals of deduction structured around the ESA model. ESA, along with PPP, is a common way for lessons to be structured in many commercially produced ELT course books (Nitta and Gardner, 2005; Thornbury, 2006; Masuhara et al., 2008).

I acknowledge that PPP and ESA have been critiqued in the literature (Lewis, 1993; Scrivener, 1994; Skehan, 1998; Ur, 1998; Harmer, 1996, 1998; Ellis, 2003; Criado, 2013), but it is not my intention or the purpose of this research to investigate the relative merits or otherwise of either PPP or ESA as a way of structuring and delivering a lesson in general, or on the affect they have on language acquisition. To that end I make no claim that PPP or ESA are unfit for the purpose of language teaching. Indeed I agree with Willis who contends that PPP can be an effective way to teach ‘a small sample of language with the focus on a particular form’ (Willis, 1996: 134) (see also Criado, 2013 for positive criticisms of PPP at the psychological, psycholinguistic and pedagogical levels), and with Harmer (1996) who proposes that ESA contains three key elements to language teaching and learning, i.e. engagement, study and activation. It is for these reasons that the teaching strategies I propose towards the end of this thesis can be seen as a means of complementing the ESA lesson under discussion rather than replacing it.

My research is a small scale investigation of one type of ESA lesson used to teach a specific area of language, i.e. modals of present deduction. The data was collected at one teaching institution with a specific group of learners. The observations I make, therefore, are with this in mind.

Chapter 2 Modality: course books and beyond

- T There's picture 2
- S1 Oh=
- S2 =Octopus= ((laughs with S1))
(2 secs)
- S3 This is a:::=
- S2 =Yes=
- S3 =Spider (.) f(h)ork
(2.8 secs)
- S1 It's like a (1.5) a:: (2.1) extra::
(1.5 secs)
- S2 Fork (.) spaceship ((laughs))
(2.3 secs)
- S2 I dunno
- S3 It's like fork
((laughter))
(15 secs)
- S2 ((inaudible))

(For original data see *Appendix 6: Classroom data CD ROM* - Richard (2) 00.10.
Appendix 11, Transcript E.)

2.1 Lead in

Above is a transcription of three learners of English engaged in an activity of making deductions about a picture (*Figure 1* see overleaf).⁴ The activity forms the activate stage of an ESA lesson that has focused on modals of present deduction. It comes from a published language teaching course book and the learners are at intermediate level. The modals of present deduction the learners have just studied are *can't*, *might*, *must* and *could*. The rubric for the activity encourages the learners to use the following words and phrases: 'must', 'might', 'could', 'can't', 'I'm sure/certain ...', 'It's possible that ...', 'It's not possible that ...', 'maybe/perhaps'.

⁴ The object is a lemon squeezer.

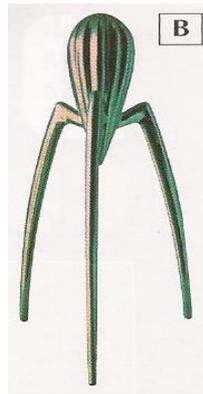


Figure 1

What is immediately apparent from the transcript is the complete absence of the suggested forms. My concern as the teacher of this group of learners was why they did not use the suggested forms or draw on the language focus of the lesson to complete the activity. The lesson objectives in the Teacher's Book state that by the end of the lesson the learners should have 'learned more about and practised using modals of present deduction (*can't, must, might/could*)' (Albery, 2008: 98); if this were an observed lesson on a teaching course such as the Diploma, I would certainly have failed. However, I argue that the activity might well be fated from the outset for a number of reasons. Firstly, the rather narrow description offered to the learners on modality and its use, secondly a misconception by the course book writers of the kinds of modality that play a part in spoken communication, and, not least, the activity itself and the way it is set up.

In this chapter I will give an overview of English language teaching course books and discuss how they have been informed by different approaches to teaching and learning language. I will make specific reference to *Language Leader Intermediate* (Cotton et al., 2008) and show the choices of modality found therein are based on a course book tradition and introspective choices of grammar. I will compare the selection of modal structures in *Language Leader Intermediate* with a corpus of spoken English to illustrate how they differ from actual language use. I will argue that the engage, study, activate (ESA) lesson on modals of present deduction in *Language Leader Intermediate* only prepares learners to use the modal forms in 'sentence level' practice exercises and activities and not extended discourse, i.e. a discussion. I then move to a broader description of modality informed by systemic functional grammar before returning to the course book lesson to argue that a discussion is sensitive to the interpersonal dimension of the participants. Thus I propose that learners need to be informed of the interpersonal function of modality in addition to the more traditional focus on degrees of certainty, possibility, etc. I finish by arguing that moving from presentation to activation of language in a lesson is not straightforward and I show why for a number of reasons it is unrealistic to expect learners to immediately produce linguistic forms they have just been presented.

2.2 A course book understanding of modality

2.2.1 A brief history of course books

In 1976 Wilkins wrote that the structure of language courses 'reflect different ways of looking at the objectives of language teaching and learning' (1976: 1). It follows that the structure of language course books reflects the objectives of language teaching and learning. To get a better idea of why modality is presented the way it is in *Language Leader Intermediate* we need a brief historical look at areas of English language teaching that have had an influence on the design of language courses and consequently language course books. Firstly, though, an overview of course books.

It is common for language schools to offer packages of language classes. These vary in length, scope, content, methodology and design but are typically referred to as 'courses'. Normal practice in many language schools in the UK and overseas is for courses to be structured using a set of published materials in book format. These materials are commercially produced, cover a range of 'levels' from Beginner to Proficiency, are divided into units or modules, cover the 'four skills' of reading, writing, speaking and listening, as well as grammar and pronunciation. Many of the texts are written specifically to present a particular language feature, but some are 'authentic' or 'semi-authentic'. They are colourful, glossy, contain pictures and cover a range of topics, tasks, exercises, activities, and games. There is an accompanying Teacher's Book providing instructions, support and tips on how teachers can use the course book. Some provide detailed lesson staging, objectives and timings for each exercise, activity or task to give an overall plan to how a lesson could be managed. Whilst the units or modules are often sequenced numerically there is not necessarily a definitive way of progressing through each book, however, they tend to progress from 'simple' to 'complex' forms and recycle vocabulary from earlier units/modules. Many come with additional work books, DVDs and progress tests based on the course book content. Littlejohn contends that the principle role of such materials is to 'structure the teaching and learning of English' (Littlejohn, 1992: 87). Course books, then, are a means of providing a language 'course'.

Language Leader Intermediate is a good example of the kind of course book outlined above. In the Teacher's Book the writers refer specifically to the 'course': 'We are pleased to welcome you to this new course: *Language Leader*', '*Language Leader Intermediate* is an international course with a global focus' (Albery, 2008: 2), '[t]he course has a topic-based multi-strand syllabus which includes comprehensive work on grammar, vocabulary, pronunciation and integrated skills' (ibid: 4). *Language Leader* has personal significance for me in that it formed the complete package of teaching and learning materials at one particular institution where I taught English for Academic Purposes. Courses strictly followed the course book with students enrolling for a fixed period of weeks which

corresponded to the number of units in the book. At the end of the course, learners were tested using a set of tests written and published by the same writers and publishing house. Littlejohn (1992) refers to the course book as providing 'a curriculum package' and this was very much what I experienced with *Language Leader*: the course was the course book and the measure of learners' progress was through tests based on the same materials. The transcript at the beginning of this chapter has been transcribed from a recording of two English language learners engaged in an activity from *Language Leader Intermediate*.

So the course book sits very much at the centre of the language course and as Wilkins has put forward the course is bound up with the writers' views on the objectives of language learning and teaching. Littlejohn (1992) recognises that there are probably as many course books as there are approaches to language teaching and learning. Whilst this might be an exaggeration, it does suggest a wide range of approaches taken by course book writers. As we have seen, course books focus on different areas of language in a variety of ways. My focus is on modality and so what follows is a brief outline of aspects of language learning and teaching in the last hundred years that have given rise to the kind of linguistic description of modality that we find in *Language Leader Intermediate*.

In the first half of the twentieth century English language teaching followed a structural syllabus. The teaching of language structures followed much the same pattern as academic grammar books. Learners were introduced to simple forms before more complex forms, for example, 'He writes' before the compound 'He is writing'; the structures 'were identified with sentences or their components in isolation rather than with utterances in context' (Howatt, 2004: 330). It was held that a structural syllabus provided 'a methodology for all occasions' (ibid: 301) which, once in place, remained unchanged from course to course. At this time 'the learning of a language is most commonly identified with acquiring mastery of its grammatical system' (Wilkins, 1976: 1). Mastery came gradually through a step by step approach to building up the whole structure of the language.

In the middle half of the century Hornby systemised the notion of 'sentence patterns' first put forward by Palmer in the 1930s. Sentence patterns 'emphasised the relationship *between* linguistic units' (Howatt, 2004: 217 *emphases in original*), an alternative to isolated grammatical 'parts of speech'. Whilst Palmer's sentence patterns worked at the level of the sentence, Hornby developed a classroom methodology which took language structures such as 'Can I help you?' and embedded them in simple situations. This became known as the situational approach. Situations were often acted out by the teacher in the classroom or with structures that could not easily be acted out the teacher would make use of pictures or invented situations. The shift in emphasis to classroom methodology determined 'both the way in which new patterns were taught and the order in which they were introduced' (ibid:

298) and gave rise to the tenet 'Courses of instruction should be built round a graded syllabus of structural patterns to ensure systematic step-by-step progress' (ibid: 300). Accordingly, these systematised patterns began to 'underpin the design of course books' (ibid: 298).

By the end of the 1960s the situational approach began to fall out of favour. The situational approach was based on 'the situations in which the learner is likely to need the language' (Wilkins, 1976: 16) and teaching 'the language that is necessary to perform linguistically in those situations' (ibid); the choice of situations being more or less subjective. The notion of situations giving rise to predictable language patterns was being overtaken by studies in how speakers use language to *do* things. Two major influences were Austin's *How to do Things with Words* (1962) and Hymes' notion of 'communicative competence'. Austin's 'speech acts', or language functions as they were more commonly known outside academic circles, described the functions that language is put to in communication, for example, promising, offering, deducing, etc. Speech functions were taken up widely in Communicative Language Teaching (CLT) with its emphasis on using 'real' language and the goal of 'communicative competence' i.e. knowledge of how to use language appropriately and effectively. Speech or language/communicative functions were seen as a means of giving learners access to language as it is used to *do* things rather than the formal study of abstracted language structures. Whilst this was a significant departure from the teaching of isolated language structures and situated patterns many believed that 'functions should be seen as an enrichment of old methods rather than as an alternative to them' (Wilkins, 1976: 339).

The inclusion of functional categories in English language course books was perhaps most influenced by Wilkins *Notional Syllabus* (1976). Using the starting point of what it is that we *do* through language Wilkins proposed organising language teaching 'in terms of content rather than the form' (1976: 18). Decisions of what to teach rested with the 'consideration of what learners should most usefully be able to communicate' (ibid: 19) and from there teach 'the most appropriate forms for each type of communication' (ibid). Wilkins notes, however, that these 'linguistic realisations of the communicative functions' (ibid: 41), i.e. the most appropriate forms, are introspective and not based on observed, objective research.

Two further events that had a significant impact on ELT and course book design are the 'Threshold Level' (T-Level) Project by the Council of Europe in 1971 and more recently in the early 2000s the Common European Framework of Reference for Languages (CEFR). The first was the precursor of the latter. The T-Level Project detailed situations, roles and activities in which a foreign language learner might be expected participate: the 'Threshold Level' referring to a 'common core' of communicative needs. Wilkins *Syllabus* was very much part of the project with an emphasis on acquiring basic

competence in 'meaningful functional objectives like "telling the time" and "buying food"' (Howatt, 2004: 340). Building on the importance of communicative needs, the CEFR 'describes in a comprehensive way what language learners have to learn to do in order to use a language for communication' (CEFR, 2001: 1) and 'provides a common basis for the elaboration of language syllabuses, curriculum guidelines, examinations, textbooks, etc.' (ibid). Howatt observes that the CEFR follows the same 'pedagogic principles' (2004: 368) as the T-Level and indeed developed out of it. Whilst he questions the validity of the CEFR in terms of pedagogic and educational perspectives, he notes its influence is apparent in curriculum design.

The framework uses a series of scaled 'Can Do' descriptors for aspects of communicative language competence, e.g. 'Can produce simple mainly isolated phrases about people and places' (Howatt, 2004: 56). Communicative competence is subdivided into three competences: linguistic, sociolinguistic and pragmatic. Each category is further subdivided but for the discussion here we need only a definition of two linguistic and one pragmatic category. Linguistic competence includes grammatical and semantic competence. Grammatical competence is described as 'knowledge of, and ability to use, the grammatical resources of a language' (CEFR, 2001: 112) and codifies linguistic units for form and meaning using traditional terminology; semantic knowledge 'deals with the learner's awareness and control of the organisation of meaning'. Pragmatic competence concerns the user/learner's 'knowledge of the principles to which messages are used to perform' and of principle concern here is performing communicative functions: functional competence (ibid: 123). Functional competence is defined as 'the use of spoken discourse and written texts in communication for particular functional purposes' (ibid: 125) and includes expressing and finding out attitudes using modality. The functional approach with its emphasis on communicative functions is believed to offer an alternative and complement to linguistic competence with its focus on form and meaning. In this regard it deals with the 'double articulation' (ibid: 116) of language. These competences are said to be part of the various means individuals have to perform language activities in specific social domains, i.e. personal, public, occupational and educational.

In terms of staging it has been observed that many language teaching course books use presentation, practice, produce (PPP) (Nitta and Gardener, 2005). PPP has been a mainstream EFL procedure for over 40 years and is a characteristic of communicative styles of language teaching (Cook, 2001: 227). Firstly the teacher presents the learners with the language focus, whether lexical, grammatical, or functional, in 'clear situational contexts' (Harmer, 1998: 64), it is then practised by the learners in a controlled way 'using accurate reproduction techniques' (ibid) such as oral drills, written gap-fill exercises, etc., and then (re)produced by the learners in a less controlled or freer way using a spoken or written activity. An adaptation of this model is engage, study, activate (ESA) and works as follows.

Engage: the students are 'emotionally engaged with what is going on' (ibid: 66) using a discussion, listening, reading, etc. as a means of contextualising the language focus of the lesson; study: 'any teaching and learning element where the focus is on how something is constructed' (ibid); activate: the stage 'at which students are encouraged to use all and/or any of the language they know' (ibid: 67). The sequence of activities from *Language Leader Intermediate* that are under discussion here adhere to ESA.

The account above is not exhaustive and is not intended to be a potted history of language teaching and learning or TEFL. I refer to these events for their relevance to the design of some English language courses and course books, specifically *Language Leader Intermediate*. In *Language Leader Intermediate* we can see close ties with structural, situational and functional approaches not only in terms of grammatical description but also in the selection of grammar and the topics used to present it.

2.2.2 Language Leader Intermediate

In the Teacher's Book Introduction the author writes that '*Language Leader* is not based on one particular teaching 'philosophy' or methodology' (Albery, 2008: 3). He does, however, elaborate by claiming 'we use a broadly communicative methodology' (ibid). As we have seen a communicative methodology emphasises what speakers *do* with language and so in *Language Leader Intermediate* modality is introduced in functional terms: modals of present deduction. The authors present what these particular modals can be used to do, that they can be used to 'make guesses (deductions) about the present based on evidence' (Cotton et al., 2008: 146). More description is added though by informing the learners of what the modals *report*, i.e. they are able to express degrees of possibility or certainty:

4b Match each modal with one of these meanings.

- | | |
|---------------------------|---------------------------------------|
| 1 It can't be true. | a) I think this is possible |
| 2 It might/could be true. | b) I'm certain that this is true. |
| 3 It must be true. | c) I'm certain that this is not true. |

(Cotton et al., ibid: 75)

Modality, then, is described using communicative functions supported by a traditional focus on form and meaning. In this respect it sits in line with the CEFR emphasising grammatical, semantic and functional competence. Whilst the exercise above is clearly an example of isolating sentences to study form and meaning, the course also employs a 'text-based approach' (Albery, 2008: 3). The texts which

form the listening exercise at the beginning of the lesson 'contextualises' the language in its functional terms (see overleaf).

A: Wow, look at that. What is it?

B: It's a drawing. I think it's a sort of flying machine.

A: Yes, it could be that. Who do you think it's by?

B: Mmm, it might be by da Vinci, I believe he did that sort of thing. Have a look at the sign. What does it say?

A: Erm ... yes, you're right, it is by da Vinci.

B: Goodness! It's in very good condition ... it says here he was born in 1452, so it must be over 500 years old.

A: Yes, and it's an amazing drawing.

B: Yeah, I read somewhere he was fascinated by birds and flying. Perhaps that's where he got his ideas for the drawing.

A: Yeah, you're probably right.

(Albery, 2008: 175)

Whilst *Language Leader* aligns itself with the CEFR (Albery, 2008: 5) and a communicative methodology (ibid: 3), I would argue, however, that the unit on modals of present deduction only really gives a nod to functional competence. The modals in question are introduced using functional terminology and despite one of the lesson objectives being the use of 'modals for present deduction' (ibid: 98) the practice exercises are really only practising form. From the perspective of functional competence the CEFR requires 'the use of spoken discourse and written texts in communication for particular functional purposes' (CEFR, 2001: 125) and as already pointed out, in terms of modality this means 'expressing and finding out attitudes'. Arguably this does not happen by matching form to meaning or rewriting sentences using a given set of modal verbs; although the last activity of the unit, in which the learners are supposed to 'activate' the language through pair work discussion by making deductions about a number of designs in coloured pictures, is perhaps where this might more realistically take place.

With structural, situational and functional approaches as well as the CEFR having a significant impact on course book design *Language Leader Intermediate* follows very much in an established and recognised tradition. Whilst some links are more tenuous than others, the influence of the various approaches to grammar can be seen both in the presentation of modals and in the methodologies engaged for teaching and learning.

Now I want to turn my attention to the writers' choice of modality, i.e. the modals auxiliaries that have been chosen for study purposes: *must*, *can't*, *might* and *could*. I believe this can be traced back to both the structural approach to grammar with its progression from simple to complex forms, still prevalent in course book design, and perceived 'usefulness' of language. This will also help to explain why modal auxiliaries are introduced at intermediate level in *Language Leader* and similar course books.

In reference to the grammar and syllabus design in *Language Leader* Albery writes '*Language Leader* follows an established syllabus progression' (Albery, 2008: 5). As I have already indicated, in course books of the kind that I have described here and of which *Language Leader Intermediate* is a typical example modality, in the form of modal auxiliaries, is typically introduced to learners at intermediate and upper intermediate levels. Indeed, in *Language Leader Intermediate* the unit on modals is midway through the 'course': unit seven of twelve. In his analysis of teaching materials Littlejohn concludes that the sequence of language content is to move from 'simple to complex in terms of surface structure' (Littlejohn, 1992: 98). I have already pointed out that in part the description of grammar in *Language Leader Intermediate* is based on a structural approach. The structural approach, as Howatt indicates, moves from simple to complex forms as a means of building up the whole structure of the language through the 'parts of speech'. In doing so learners acquire 'mastery of the grammatical system' (Wilkins, 1976: 1) through the 'gradual accumulation of linguistic items' (Littlejohn, 1992: 90). Arguably in terms of structure, modal verbs fall somewhere between simple and complex forms. In language courses and course books, then, modals, in the form of modal auxiliaries, are often introduced midway.

Littlejohn's research into teaching materials argues that linguistic form is 'presumably selected on the basis of usefulness to the learner' (Littlejohn, 1992: 88). However, it is not clear where the 'basis of usefulness' comes from. We have already seen the claims of Howatt and Wilkins on the subjective choice of useful language in the situational approach and traditional grammar; Wilkins notes the same when he says the linguistic realisations of his function categories are based on 'introspection'. 'Usefulness' in this respect is somewhat a matter of personal choice. The question of usefulness, and indeed what to teach language learners, is a pivotal and central question in language teaching and learning and not an easy one to answer. In its intention to develop communicative competence the CEFR puts an emphasis on the 'needs, motivations, characteristics and resources of learners' (CEFR, 2001: 3). Arguably expressing and finding out attitudes using modality and functional language is useful for learners towards acquiring this competence, but it does not make clear what is useful in terms of the linguistic items that realise it. I would suggest that following the introspective tradition the first choice arrived at is modal auxiliaries: functional categories grafted on to the tradition of

structural forms. Although on the surface this might be sensible enough it does not appear to be supported by linguistic data. What I intend to look at next, then, is what other modal resources are available to speakers and perhaps more typically used for making deductions.

What I have tried to show here is that the modal auxiliaries in *Language Leader Intermediate* are introduced at intermediate level following a course book tradition of progressing from simple to complex forms. The choice of modal auxiliaries also follows a tradition in course books, this time a tradition of perceived usefulness. The ‘established tradition’ which *Language Leader Intermediate* claims to follow is a course book template which has developed out of different approaches to grammar, teaching and learning that still have relevance to courses, course book design and organisations such as the CEFR.

2.3 Making it interpersonal

2.3.1 Problematizing a lesson on modal verbs

Just to reiterate my interest is on why learners can often manage the rules of form and meaning of modal auxiliaries at sentence level, i.e. the structural stuff but do not use them in spoken communication, i.e. functionally. The following exercise from *Language Leader Intermediate* provides the ‘meanings’ of the modals for the lesson (the exercise has previously been introduced in 2.2.2 *Language Leader Intermediate*):

4b Match each modal with one of these meanings.

- | | |
|---------------------------|---------------------------------------|
| 1 It can't be true. | a) I think this is possible |
| 2 It might/could be true. | b) I'm certain that this is true. |
| 3 It must be true. | c) I'm certain that this is not true. |

(Cotton et al., 2008: 75)

It is a relatively straightforward exercise and in my experience learners sail through it. Following this the learners are armed with the knowledge of how a selection of modal forms report meanings of certainty and possibility. The next exercise is intended to consolidate this knowledge by asking the learners to rewrite a number of sentences using the modals *must*, *can't*, *could* or *might*. For example:

‘I'm sure this design is by Armani’ (ibid)

By using the rules from exercise 4b and possibly a dictionary or recourse to the teacher to establish the approximate synonymy of ‘certain’ and ‘sure’, this sentence would most likely be rewritten: ‘This design must be by Armani’. So what happens between these exercises and the paired discussion

activity which produced the dialogue transcribed above? The learners have the forms, know the meanings but simply do not use them. The activity is set out as follows (see over):

6 Work with a partner to discuss what you think the designs below are. Try to use these words and phrases.

Must might could can't

I'm sure/certain ... maybe/perhaps

It's possible that ...

It's not possible that ...'

(Cotton et al., 2008: 75)



Figure 2

Firstly, the pictures are a rather random selection, which I argue has an impact on the way the learners approach the activity. As one learner said 'It's difficult because they are out of context'.⁵ The second, going back to the activity instructions is the additional modal adverbs and phrases containing modal adjectives the learners have now been given to use. Whilst this might throw a spanner in the works for the learners these added phrases are not of immediate concern to me. Not because I do not see a problem with throwing in odd words and phrases, but simply because, in the main, the learners do not even get as far as using the ones they have been introduced to. The reason for this I believe is threefold.

⁵ Information as to what the 'designs' are is not provided in the course book or teacher's book.

Firstly, I have shown above that the explanation of the modal auxiliaries in *Language Leader Intermediate* relates form to meaning: exercise 4b shows that *must*, *can't*, *might* and *could* report possibility and certainty. Arguably this is sufficient for rewriting sentences as in exercise 5. I also showed that the course book gives a mention to the functional use modals for making 'deduction based on present evidence'. This too might be sufficient for learners to make simple 'one line' guesses about the pictures. But the activity asks the learners to discuss in pairs and to do this the description of modality offered to the learners might not be sufficient.

Secondly, I have pointed out that the selection of linguistic items in the form of modal auxiliaries to realise the speech function of deduction is not based on researched linguistic data. I argue that it might not be the case that modal auxiliaries are the preferred modal resource used to perform this function. This then questions the notion of usefulness – a primary consideration when designing content for language courses and course books.

The third reason draws on the first two in terms of language content but more fundamentally the overall conception of the activity. The learners are asked to discuss a set of pictures. A discussion might manifest itself in a number of ways and will likely contain a number of different speech functions; it may even contain a speech function in which one, the other or both interactants make a deduction using a modal form. What a discussion is not is an exchange of the same speech function over and over again as in this invented example:

A: It's a coffee machine.

B: Perhaps it is a coffee machine.

A: I'm sure it's a coffee machine.

B: It can't be a computer.

A: It must be a coffee machine.

Whilst the above dialogue is not entirely improbable it is most likely not what the course book writers had intended when they came up with the activity and arguably not a discussion.

I will address each point below.

2.3.2 Interpersonal meaning and modality: insights from systemic functional grammar

The discussion activity asks the learners to work in pairs and make deductions about a series of pictures. At the level of the isolated sentence this shouldn't be too problematic. In my introduction I described a teaching technique in which a teacher trainer drew simple pictures on the board, provided linguistic prompts and asked us to make sentences. This kind of exercise is common in ELT for example

grammar practice books such as *How English Works* (Swan and Walter, 1997) which aim to practise 'short clear explanations of the rules of English grammar' (ibid: 1). For example the following sample taken from an exercise to practise 'how certain?' (ibid):

'Look at the pictures and write sentences with *must/may/might/can't*. Example:

1 *She must be ill. OR She can't be well.*



Figure 3 (Swan and Walter, 1997: 109)

But a discussion typically extends beyond the level of the sentence and involves more than one participant. The lack of distinction between an exercise to practise written sentences and an activity to generate a discussion, even if the intention is to practise the same language point, is an oversight by the authors and a point I will elaborate on later. For the time being I want to focus on an aspect of language that affects the kinds of linguistic resources of speakers, including modality, in a communicative activity such as a discussion.

I have said above that *Language Leader Intermediate* offers a rather narrow description of modality. Modality is introduced using modal auxiliaries and a couple of modal adverbs and adjectives with an emphasis on what the modals report and do. This is achieved through two exercises, the first matching verbs to meanings and the second rewriting sentences so that they carry the same meaning using the modal auxiliaries from the previous exercise.⁶ The exercises practise form and function. Neither exercise, however, takes into account the speakers of the example utterances. For example, rewriting the following sentence using *must*, *can't*, *could* or *might* does little to raise awareness or inform the learner of who the speaker is, who that speaker is addressing and the implications of the choice of language for that particular interaction.

1. Raymond Loewy is definitely one of the most influential designers of all time.

(Cotton et al., 2008: 75)

The point I want to make is that neither exercise prepares the learners for a discursive activity such as a discussion in which choices of language are sensitive to the participants and the context. What I

⁶ See *Appendix 2* for the full lesson from *Language Leader Intermediate*.

want to look at here is a theory of language which takes in to account the speakers, their social roles and their purposes and goals in communication. To do so I need to draw on terminology from systemic functional grammar.

Systemic functional grammar (SFG) builds on the approaches to grammar that we have seen so far and for which *Language Leader Intermediate* relies for describing modals. However, it develops these different approaches by relating grammar to meaning, function and context. Language is seen as a system of choices. Choices relate to the social goals and purposes of the speakers within a given context. Meaning is thereby derived from these choices of language in relation to context.

So we can see that in SFG an emphasis is put on who the speakers are, who they are speaking to and the context in which the speaking takes place. And these appear to have not been taken in to account for the discussion activity in *Language Leader Intermediate*.⁷ Neglecting the speakers and the context fails to recognise the interpersonal function of language. In SFG the interpersonal metafunction refers to the way speakers 'engage interpersonally and exchange points of view' (Coffin et al., 2009: 226). All communication has an interpersonal dimension but in some communicative events it is more evident. I would argue that a discussion is a communicative event in which this is the case. As the participants of a discussion position themselves in relation to each other their choices of linguistic/communicative resources reflect and convey interpersonal relations. The strength of these communicative resources will vary as participants seek to put forward arguments for and against the proposition under discussion; there will be degrees of authority and affinity sensitive to who the speakers are and what they are discussing.

There is no one-to-one relationship between the forms and structures of language and interpersonal meaning. We cannot say that interpersonal meaning is realised by x, y and z although x, y and z may well function as indicators of interpersonal meaning in an utterance within a given context. For example, the choice of personal pronoun, copula verb in the declarative 'I'm sure this design is by Armani'⁸ are not specifically indicators of interpersonal meaning, but they have the potential to realise interpersonal meaning from the relationship between the choices made and those that were not made.⁹ The speaker could have asked 'Do you think this design is by Armani?' If this utterance were

⁷ In terms of the activity in *Language Leader Intermediate* (see Appendix 2 exercise 6) the term 'discussion' is arguably inaccurate if we take a discussion genre as being a debate of arguments for and against a proposition (see Coffin et al., 2009). I use the term in relation to the activity as it is referred to as a 'group discussion' in the Teachers' Book for *Language Leader Intermediate* (Albery, 2008: 98).

⁸ Example reproduced from 2.3.1 *Problematizing a lesson on modal verbs*.

⁹ The copular *be* identifies states of existence. Adjective phrases as subject attributes express stance (Biber et al., 2008: 142), e.g. 'I'm sure this design is by Armani'. Taken together then this utterance signals high modality.

part of a discussion, the participants would be engaged in a different interpersonal relationship with a bearing on how points of view might be exchanged. From an interpersonal perspective a discussion as a communicative situation 'is one of roles and relationships, identity, perspectives, and power' (Coffin et al., 2009: 354). We can see here why the speakers are of importance and relevance to the discussion activity.

Lexicogrammatical resources which help to realise roles and relationships, identity, perspectives, and power can be found in the modal system; modality being important to interpersonal meaning. The modal system 'enables a speaker to grade the strength of their commitment to a proposition or proposal' (Coffin et al, 2009: 365) when intruding into a text. The modal resources express commitment to what they are saying as well as expressing relations among participants (Halliday, 1978: 46). Modality can be grouped under two headings: epistemic and deontic modality (see 4.1 *Epistemic modality* where I draw attention to different terminology in SFG for referring to epistemic and deontic modality). Deontic modality relates to the speaker's view of the desirability of a situation and includes meanings such as obligation and permission. Epistemic modality relates to 'the speakers' assessment of the validity of what they are saying' (Coffin et al., 2009: 169). It includes meanings such as certainty, probability, and possibility. The lesson in *Language Leader Intermediate* focuses on epistemic modality.

In language course books modal auxiliary verbs are a typical way of introducing learners to modality. However, the range of modal resources extends beyond the modal auxiliaries, to lexicogrammatical forms 'that are not traditionally seen as modal at all' (ibid: 366) (see below and continued overleaf; based on Coffin et al., 2009: ibid). What we have here is a lot more than the ability to express possibility, certainty, obligation, etc.

1. Verbs:
 - a) modal auxiliaries *can, could, will, would, etc.*
 - b) semi-modals *need to, dare to, etc.*
 - c) lexico-modal auxiliaries *have got to, be bound to, etc.*
2. Modal adverbs *probably, possibly, surely, perhaps, etc.*
3. Modal adjectives *possible, probable, etc.*
4. Modal nouns *the possibility, likelihood, etc.*
5. Mental process verbs *think, believe, etc.*
6. Verbal process verbs *suggest, propose, etc.*
7. Material process verbs *allow, guarantee, etc.*

8. Vague language¹⁰ *a bit, any, rather, etc.*
9. Prepositional phrases *in all likelihood, in fact etc.*

As we have seen, modal resources enable the speaker to grade their commitment to a proposition. This can be done using a range of modal resources from high to low (Halliday, 1985/1994). So 'I'm sure this design is by Armani', 'I think it's by Armani', and 'It looks rather like Armani to me' express high, median and low modality respectively. Of course, at either end of the scale of speaker commitment is the possibility of asserting the belief that something either IS or ISN'T (for a similar discussion see Coffin et al., 2009), for example: 'It's an Armani', or 'It isn't Armani'. Modality operates along a commitment cline from high to low modality and this forms the basis of their conventionalized meaning (see Martin and Rose, 2003/2011 for a discussion on modality for setting up a semantic space between 'yes' and 'no'; Halliday, 1985/1994 for a discussion on polarity and modality). However, conventionalized meaning is 'tweaked' when the forms are uttered in context. For example, the proposition 'I think it's by Armani' is modalized using the mental process verb 'think' which suggests median commitment and modality. However, in the context of a discussion where this statement contradicts a previous statement by another participant, it reflects a different kind of commitment to the proposition and as such implies higher modality. High and low commitment corresponds to whether the speaker is taking a closed or an open stance (Coffin et al., 2009) (or somewhere in between) in regard to the validity of what they are saying. For example, the utterance 'It's by Armani' presents a closed stance with the speaker showing high commitment. If the stance is closed and the modal commitment high as in the previous example, then the speaker is making a strong proposition through high modality. If such a proposition is expressed during a discussion it will have implications to how the other interlocutors respond; this being true of all propositions. So choices such as using *may* or *might* are not simply made on the speaker's belief of certainty, but also how they feel they want to position themselves regarding the validity of what they are saying in relation to their interlocutors. Modality becomes part of the dialogue as a means of acknowledging, agreeing with, opposing or simply 'closing out' the points of view of other speakers (ibid).

When I say that the description of modality is somewhat narrow in *Language Leader Intermediate I* do not mean in terms of the range of modal forms: the interlocutors are keyed with four modal auxiliaries and a couple of modal adverbs and modal adjectives. I would argue that practising the form and meaning of the modals at sentence level in the exercises we have seen, without emphasising their

¹⁰ The term vague language is a functional category rather than a word class. Vague language, like hedging devices, convey imprecision, uncertainty and allow the speaker to lessen the force of a proposition. It is a rather 'catch all' term and the examples above come from different word classes. As an example *in a way* is a prepositional phrase that signifies vagueness and so could belong to either group.

interpersonal function, may not be sufficient for learners to be aware of how they function in the context of a discussion. Modality is central to the interpersonal function of language as it enables speakers to express attitudes towards propositions, and to express a speaker's 'perceived relation' with an interlocutor' (Fowler and Kress, 1979: 188). Coffin et al. contend that 'The communicative functions of the modal system are highly sensitive to context and social purpose. Developing skills to use modalising lexicogrammatical forms effectively can really only be achieved within a particular communicative context' (Coffin et al., 2009: 382). What I propose then is the additional study of modal forms in relation to the communicative event, i.e. the discussion in conjunction with their form and meaning. By doing so the interpersonal aspect can also be taken into account and thus drawing together form, meaning, function and context.

2.4 Modality in spoken language: insights from language corpora

In this section I want to discuss the modalising forms selected by the writers of *Language Leader Intermediate*.

All utterances 'bear the signs of modality' (Hodge and Kress, 1988: 124). With the interpersonal function at play in all communication it follows that modality will in one form or another be present. As interlocutors position themselves in a dialogue modal expressions become a means through which they are able to commit themselves to a greater or lesser degree to what they believe is the truth value of what they or their interlocutors are saying. In the previous section I detailed some linguistic modal resources available to speakers from the perspective of systemic functional grammar. There is quite a range. What we find though in *Language Leader Intermediate* are the usual modal auxiliaries that turn up regularly in ELT materials (see, for example, Gairns and Redman, 2002; Soars and Soars, 2005; Kerr and Jones, 2007; Redstone and Cunningham, 2007; Wilson, 2007; Soars and Soars, 2009; Philpot and Curnick, 2011; Tilbury and Hendra, 2011).¹¹ I have suggested above that the selection of grammar in *Language Leader Intermediate* is not based on linguistic data and therefore the kinds of modality that might be present in real spoken communication. Part of the reason for this, I would suggest, relates back to the development of teaching materials outlined above. Many ELT course books follow a similar pattern in terms of the grammar they introduce and at what 'level'. Course books are 'frequently looked upon as carriers of grammatical structures or vocabulary items that *have* to be introduced to the learners' (Kumaravadivelu, 2003: 46 my emphasis). The structures that 'have

¹¹ I point out that the presentation of modals in these course books is based on PPP/ESA. These teaching models being common to many published ELT coursebooks (Harmer, 1998; Nitta and Gardner, 2005; Thornbury, 2006; Masuhara et al., 2008).

to be introduced' are referred to as 'all the key grammar points' (Albery, 2008: 3) in *Language Leader Intermediate*.

In the Teacher's Book Introduction it claims that *Language Leader Intermediate* 'follows an established syllabus progression' (Albery, 2008: 5) a progression of simple to complex forms. We have seen that this step by step progression is intended to lead to the acquisition of the 'whole' structure of the language. One of the steps along the way therefore is to introduce learners to degrees of certainty and possibility as ways of refining statements and commands. Modal auxiliaries are perhaps an obvious way of doing this. We have also seen that structures are chosen for their perceived usefulness to learners. In the goal of communicative competence the CEFR includes expressing and finding out attitudes using modality. Again the modal auxiliaries are perhaps the most obvious way of doing this. However, as pointed out choices of grammar have been largely subjective and more recently these choices have been criticized for being based on 'native speaker intuition'.

Intuition plays a large part in the writing of language teaching materials with many course books containing concocted texts (Carter, 1998), or semi-authentic texts as they are referred to in *Language Leader Intermediate* (Albery, 2008: 3). These texts are written specifically for teaching purposes with the design of putting specific features of language 'on display' (Widdowson, 1978). Texts of this kind highlight the language focus of the lesson. It is important also to note the word 'written' when referring to examples of language and teaching texts. Brown and Yule (1983) observe that for most of its history language teaching has been concerned with written language; written language forming the basis of 'correct usage'. When it comes to spoken language this is normally written as well-formed sentences 'translated into speech' (ibid) or 'spoken prose' as it has been called (Abercrombie 1963 in Howatt, 2004: 300). The reason for this is in part due to written language being 'accessible to conscious attention and systematic study' (Halliday, 1989: 96) as well as notions of correct usage being based on grammar as represented by written language (Tomlinson, 1998: 67). There is nothing new in highlighting the artificiality of spoken language in exchanges such as the following taken from *Language Leader Intermediate* to present modals of present deduction:

E: ... Look at that one over there. It must be a Ferrari, surely.

F: Yes, it can't be anything else. It's so red and stylish.

(Cotton et al., 2008: 176)

But what is of interest to me is whether modal auxiliary verbs are the preferred choice of modalising expression in a discussion of the kind set as a speaking activity in *Language Leader Intermediate* and to what extent speakers of English use modal auxiliary verbs to make deductions. Perhaps there are

more 'useful' modalising forms that learners of English need exposure to and opportunity to study and use to effectively communicate in discussions and when making deductions.

The modal forms in *Language Leader Intermediate* relevant to my discussion are termed modals of present deduction and are as follows:

Must, might, could, can't

Maybe/perhaps

I'm sure/certain ...

It's possible that ...

It's not possible that ...

She might be feeling ill.

You must be joking!

As I am interested to discover the extent to which these forms occur in spoken communication, I will discuss each one in turn in relation to findings from a corpus of English: the Longman Spoken and Written English Corpus. To do this I will not reference the corpus directly but refer to a corpus-based grammar based on the findings of the corpus. The reference grammar I am using for comparison is the *Longman Grammar of Spoken and Written English* (Biber et al., 2002). The corpus contains approximately 40 million words of text providing, the authors claim, 'a sound basis for the analysis of grammatical examples' (ibid: 3). The texts used to compile the corpus are taken from four 'registers': academic prose, conversation, fiction and news. A minor limitation, therefore, is the rather general term 'conversation' referring to spoken 'personal communication' (ibid: 4). The authors note that each register contains a number of sub-registers, however, the compilation does not identify sub-registers or list the topics of conversation. The corpus conversations are 'private (often domestic) talk' (ibid: 8) but include registers such as service encounters and one side of telephone calls, although the corpus examples do not specifically refer to these. Reference, therefore, to the 'sub-register' of discussion and the speech function of deduction would need to be sought through specific examples taken from the corpus. For the purpose of research I will refer only to the main registers. It is of relevance to note that planned speech such as lectures are not included in the register of conversation.

2.4.1 Modal auxiliary verbs

Must, might, could, can't

Of the nine modal auxiliary verbs¹² *Language Leader Intermediate* introduces four to express present deduction: *must, might, could* and *can't*. Each modal can have two types of meaning, what Biber et

¹² The nine modal auxiliary verbs are: *will, would, can, could, may, should, must, might, and shall*.

al., call personal or logical meaning, corresponding to the terms deontic and epistemic modality introduced earlier. *Language Leader Intermediate* draws on logical/epistemic meaning and so for the purpose of this research I will limit my discussion to logical/epistemic meaning.¹³ Biber et al. point out that modal verbs with epistemic meaning ‘usually occur with non-human subjects and/or with verbs that express states’¹⁴ (2002: 177) such as mental verbs and verbs of existence and verbs of relationship, e.g. *think, be, seem*. Modal verbs are more common in conversation than other registers. Whilst conversation has lower lexical density and fewer noun phrases than other registers it contains a higher frequency of verb phrases and clauses. This is characterised by primary verbs¹⁵ and modals. Epistemic meaning is more common than deontic meaning. The frequency of modal verbs in conversation is perhaps surprising as they have traditionally been seen as characteristic of writing (ibid). The higher frequency of modal verbs in conversation is perhaps to be expected due to the interpersonal nature of conversation and thus the use of modal verbs as a resource to express stance.

The overall frequency of the nine modal auxiliary verbs in conversation according to the corpus is approximately 21,000 per million words. Of course, frequencies do not tell us much on their own so it is worth noting that modal verbs are extremely common in conversation as a device to express stance (Biber et al., 2008: 178). Approximating 21,000 per million words, modal verbs are roughly 5,000 – 8,000 counts more frequent in conversation than in the other registers of fiction, news and academic prose with which the corpus has been compiled (see above). The findings in Biber et al. show that *will, can* and *would* in that order are by far the most common modal verbs in conversation.

Turning to the selection of modal verbs presented in *Language Leader Intermediate* Table 1 below outlines approximate frequencies per million words of the modals *must, might, could* and *can*.¹⁶ The three columns from left to right show their total frequency in the corpus, their frequency in the register of conversation, and their frequency in conversation with epistemic meaning.

Modal auxiliary	Total frequency in corpus (per million words)	Frequency in conversation register (per million words)	Frequency in conversation with epistemic meaning (per million words)
Must	800	750	450
Might	600	800	700
Could	1,800	2,000	1,400
Can	2,500	4,200	1,700

¹³ I will be using the terms epistemic and deontic rather than logical and personal in my thesis.

¹⁴ Lexical verbs in Biber et al. (2002) are classified slightly differently to SFG. Biber et al. identify seven semantic categories of lexical verbs including mental verbs and verbs of existence or relationship. In SFG they are referred to as processes, i.e. a mental process, an existential process and a relational process.

¹⁵ Primary verbs refer to ‘one of the verbs *be, have* and *do*, which can function as either auxiliary or main verbs’ (Biber et al., 2002: 459).

¹⁶ Biber et al. do not count the frequencies of *can’t* as a separate entry. Referencing the LGSWE my analysis therefore focuses on *can* and is followed by a commentary on *can* and *can’t*.

Table 1

Clearly *can* is by far the most common form. Again frequencies do not tell us much by themselves but as we know that modal verbs are extremely common in conversation as a mean of expressing stance, it is a fair to say that at 1,700 per million words *can* is a frequently occurring modal verb in conversation.

Biber et al. show that *can* is used mostly for ability but use is often ambiguous with epistemic meaning, e.g. 'Well you can get some cigarettes from there, can't you?' (2002: 179). Additionally it is also used with permission meanings, e.g. 'Can I use the bathroom?' 'Yes, you can.'¹⁷ These different meanings may help explain its overall frequency. Even so it is by far the most frequent of the four modal verbs under discussion here.

Biber et al. do not refer to the distribution of *can* or *can't* and use both in their examples. In an attempt to gain a clearer idea of their frequencies I accessed the Brigham Young University British National Corpus online (<http://corpus.byu.edu/bnc/> accessed 15/03/12), a corpus of 100 million words. A search for *can* in the corpus of spoken language revealed 37,105 hits. A search for *can't* in the same corpus revealed 12,723 hits. Whilst this shows a significant difference between the two, it does not tell us about the respective meanings of the modals and so closer reference to individual corpus entries is needed to establish whether *can't* is favoured with ability, epistemic or permission meanings.

Can't be is a modal form with epistemic meaning. It is a form used as a modal operator in *Language Leader Intermediate* i.e. *it can't be true, it can't be anything else, and it can't be that old*. Biber et al. do not make reference specifically to the form *can't be*. My own reference to the Brigham Young University British National Corpus for *can't be* reveals 599 hits. Comparison between the *Longman Spoken and Written English Corpus* and the BYU-BNC shows that *can't be* is significantly lower in frequency than the forms *can* and *can't*. In its negative form *can* would seem to relate more closely with lack of ability rather than having a modal meaning. However, closer reference to individual corpus examples can shed more light on this.

Findings from the corpus conclude that *could* and *might* are mostly used with epistemic meaning to express doubt 'with *could* showing the greatest degree of uncertainty and tentativeness' (ibid). Knowing what we do about modality and the use of 'past forms' to express indirectness this is not surprising. *Could* also has a high frequency and along with *can* there may also be some ambiguity with its use to express ability, although Biber et al. show that this use is not common at around 400 per

¹⁷ I feel it is worth pointing out that *can* is much more frequent than the rarely used *may* for permission meanings (Biber et al., 2002).

million words. Perhaps its use in questions such as ‘What could it be?’ adds to its tally. Again recourse to the corpus can verify this.

Must is the lowest of the group and this may be due to a mismatch between the interpersonal nature of conversation and the strength of *must* as a marker of stance.

There is no claim by the authors of *Language Leader Intermediate* as to which modal verbs are more frequent. It is perhaps interesting to note that they have been sequenced in a particular order but whether this is of significance to their perceived importance is uncertain. As I have already mentioned the language focus of the lesson is contextualised in three recorded dialogues. In these dialogues *could* is the most frequent modal verb with four counts, the other three having two counts each. The relatively even distribution of the forms is not surprising as the texts are there to present the forms.

2.4.2 Adverbs

Maybe/perhaps

With its high frequency of verb phrases it follows that conversation has a correspondingly high frequency of adverbials. In fact adverbials are only slightly less common than lexical verbs in conversation as ‘they often contain information central to the message’ (Biber et al., 2002: 357). There are many different types of adverbials including adverbs, prepositional phrases, clauses and noun phrases (Biber et al., 2002; Leech and Svartvik, 1975).¹⁸ They also serve a variety of functions and semantic roles. As markers of stance they approximate 7,000/million words in conversation, less than modal verbs and complement clauses at 26,000 and 15,000/million words respectively. Epistemic adverbials are by far the most common adverbials of stance in conversation. They comment on ‘the speaker’s judgement about the information in a proposition’ (ibid: 382). At roughly 5,800/million words they account for approximately 85% of all adverbials.

From this large group I now focus on the two epistemic stance adverbs presented in *Language Leader Intermediate*: *maybe* and *perhaps*. As epistemic stance adverbs they are used to comment on the speaker’s level of certainty or doubt of a proposition. Both adverbs occur approximately 200/million words in conversation; *perhaps* being more common in writing. Interestingly the most common epistemic stance adverb in spoken English is *probably* occurring significantly more frequently than *maybe* and *perhaps* at around 600/million words. However, this is appreciably less than the occurrence of the modal auxiliaries *can*, *could* and *might* (see Table 1 above).

¹⁸ Adverbials have a syntactic function defined by their broad semantics of referring to place, time and manner, for example. They can be identified using their ‘exchange classes’, i.e. the phrases that can occur in the same slot, e.g. Peter was playing *well* (adverb), Peter was playing *with great skill* (prepositional phrase), Peter was playing *to win* (clause).

2.4.3 Adjectives

I'm certain/sure

Attributive and predicative adjectives with a modal function are quite rare in conversation. One place they can be found is controlling post-predicate *that*-clauses. Post-predicate *that*-clauses are very common in conversation where they 'typically report the speech and thoughts of humans' (Biber et al., 2002: 312) at over 80% of all *that*-clauses. However, around 90% of these are controlled by verbs reporting speech or thought, e.g. 'Did you know that Kathy Jones had a brother here?' (ibid). The most frequent adjective and, indeed, as Biber et al. point out the only one that is especially common (ibid: 318) is *sure*.

It's possible that ...

It's not possible that ...

Extraposed *that*-clauses make up only a fraction of *that*-clause types in conversation. Again they are used to express degrees of certainty relating to the idea in the *that*-clause. The use of the dummy *it* reports attitude or stance without making it clear who it is attributed to and makes the clause impersonal. The interpersonal nature of conversation largely negates the use of this impersonal style which is much more common in news and academic prose. Biber et al. note that the most frequent adjectives controlling extraposed *that*-clauses are *clear*, *(un)likely*, *(im)possible* and *true*.

2.4.4 Aspect

She might be feeling ill

You must be joking!

As an extra nugget of information there is a 'Grammar Tip' in *Language Leader Intermediate* referring to the use of the continuous aspect with modal verbs:

'We can also use modal verbs with a continuous form:

She might be feeling ill. (= I think she's feeling ill.)

You must be joking! (= I'm certain you're joking.)

(Cotton et al., 2008: 75)

It has been observed in the LGSWE that relatively few modals occur with progressive aspect (Biber et al., 2002: 184); *must* being one of the occasional exceptions.

Using the findings of the LGSWE we can see that some of the modal forms presented in *Language Leader Intermediate* are more common in conversation than others. For example, the modal auxiliary *could* is a lot more common than *might*. We can see too the distribution of these modal forms across registers, for example, extraposed *that*-clauses are relatively rare in conversation compared to

academic prose. We can also see that some of the more common forms in spoken English have not been included for example, the adverb *probably* while rare structures such as modal verb + progressive aspect have, for example *She might be feeling ill*. Findings such as these can help to give insights into the usefulness of teaching certain forms and their distribution in relation to particular communicative events and registers. It might be more appropriate to leave extraposed *that*-clauses to writing activities, for example, than introduce them for a speaking activity. The notion of usefulness and what to teach is a theme I will return to later in my thesis when I explore the implications of my research for materials design and teaching (see *Chapter 6 Teaching implications: applying my research*).

2.5 The lesson: moving from form to function, or not?

The culmination of the lesson is a discussion¹⁹ in which learners work in pairs and draw on the language focuses of the lesson to make deductions about three pictures of different designs. The activity is presented as follows (reproduced from above 2.3.1 *Problematizing a lesson on modal verbs*):

6 Work with a partner to discuss what you think the designs below are. Try to use these words and phrases.

Must might could can't

I'm sure/certain ... maybe/perhaps

It's possible that ...

It's not possible that ...

(Cotton et al., 2008: 75)

One of the unit objectives claims that by the end of the lesson the learners will have 'learned more about and practised using modals for present deduction (*can't, must, might/could*) (Albery, 2008: 98). The discussion is the stage of the lesson in which the learners 'activate' the language they have studied. Earlier in the lesson the language focus was 'contextualised' using three dialogues of 'three pairs of students at a museum of design' discussing three objects. The assumption is that the learners will incorporate the modals in to their discussions and resemble something like the three discussions from the listening. I would argue though that the transfer of the modals from form exercises to functional use in a discussion is not so straightforward firstly, in terms of a discussion as a communicative event, and secondly, in terms of the participants.

To explain my first point I want to draw on the notion of genre. Genre can be described as 'a staged, goal-oriented social process' (Martin and Rose, 2007 in Coffin et al., 2009: 249); its structural

¹⁹ Again I point out that 'a discussion' is perhaps not the best term for the activity. I use the term as the activity is referred to as a discussion in *Language Leader Intermediate*. See footnote 7 for more information.

organisation, along with the grammatical features, it is likely to share with other texts with the same social purpose (Coffin et al., 2009). Based on this description a discussion can be seen as a genre. Whilst a discussion can occur in a variety of professional and social settings and is quite an umbrella term it is characterised as a genre by the similarities it has with other texts and distinctions it has with others. The genre of a discussion is, for example, different to the genre of asking for and giving directions, although the latter may well include the former if at some point the participants engage in a discussion as a sub-genre: they may discuss the best way whether it is quicker to walk or take a bus. Coffin et al. indicate that the social purpose of a discussion is to 'consider different perspectives on an issue' (ibid: 261) includes generic stages such as 'issue', 'arguments for/against' and 'position' and contains lexicogrammatical features such as present tense, declarative mood, interrogatives (in spoken mode), modality and generic nominal groups.

Above I have given an example of a 'discussion' that whilst not altogether improbable is slightly absurd. The reason for its absurdity is the two speakers engage using the same (and only) speech function of deduction five times one after another. Such a model is unlikely as a discussion because it does not follow the typical structural organisation that we might expect from the genre. If we take one of the, albeit concocted texts, from *Language Leader Intermediate* used to present the modals (see below, reproduced from above 2.2.2 *Language Leader Intermediate*) we can identify something more like the genre of a discussion. This has been done in part by generic staging giving the text a beginning, middle and end as well as through certain lexicogrammatical features as a way of achieving a specific social purpose.

A: Wow, look at that. What is it?

B: It's a drawing. I think it's a sort of flying machine.

A: Yes, it could be that. Who do you think it's by?

B: Mmm, it might be by da Vinci, I believe he did that sort of thing. Have a look at the sign. What does it say?

A: Erm ... yes, you're right, it is by da Vinci.

B: Goodness! It's in very good condition ... it says here he was born in 1452, so it must be over 500 years old.

A: Yes, and it's an amazing drawing.

B: Yeah, I read somewhere he was fascinated by birds and flying. Perhaps that's where he got his ideas for the drawing.

A: Yeah, you're probably right.

(Albery, 2008: 175)

It opens with an orientation to the object under discussion through an interrogative. Speaker B replies with an 'argument for' to which speaker A gives an evaluation. Speaker A asks for elaboration and speaker B provides this by identifying and classifying and so it goes on until speaker A takes up his final position in agreement with speaker B. All of this is characterised by interrogatives, declaratives, modality, third person pronouns, etc. typical lexicogrammatical features found in discussions.

Part of the reason I believe why the speaking activity is never realised as a discussion by the learners is that they are not primed to engage in a discussion. The exercises leading up to this activity have only prepared them to make deductions or guesses using simple sentences, at best. In this respect it is no different to the practice exercise from *How English Works* (Swan and Walter, 1997) given above in 2.3.2 *Interpersonal meaning: and modality: insights from systemic functional grammar*. What my data reveals is that at times the activity triggers off guesses and so the activity becomes more like a guessing game than a discussion (see responses to *Figure 7* in 5.1 *Four extracts of data* and the *Figure 12 'Lady Gaga' dialogue* in my *Conclusion*). Perhaps this is not surprising as the learners have been informed that the modal forms are used for present deduction 'to make guesses'²⁰ (Cotton et al., 2008: 146). What occurs, then, is a procedural mismatch (Kumaravadivelu, 2003: 87) between how the course book writers planned the activity and how the learners actually engage with it.

Now I want to turn to the participants of the discussion themselves.

Kumaravadivelu points out that teaching materials are the result of careful planning by the writers but not 'the result of any interactive process of classroom events' (2003: 46). Part of this careful planning defines 'not only what the learner is to talk/write about but also how they are to do this' (Littlejohn, 1992: 75) i.e. through the reproduction of the given language. This is the process adopted in *Language Leader Intermediate* with specific modals being introduced for reproduction in the speaking activity. The idea is that learners are presented with language, practise it in a controlled way and then reproduce it. It needs to be remembered that course books such as *Language Leader Intermediate* are '[c]ommercially produced for mass consumption' (Kumaravadivelu, 2003: 46) and designed to structure seemingly identical lessons (Littlejohn, 1992: 76). Whilst this is all very well on paper it does not take into account the fact that language learners are 'complex and nuanced social beings' (Frith and Wagner, 1997 in Block 2003: 4) in a social situation which 'exerts an impact on the communication that takes place in it' (Dornyei and Murphy, 2003: 74). The interactive process of classroom events, therefore, may well not give rise to the kind of language the course book writers had carefully planned; the teacher monitoring and intervening with the 'appropriate forms' notwithstanding.²¹

²⁰ The terms 'deduction' and 'guess' are used to mean the same thing in *Language Leader Intermediate*.

²¹ See Ellis (2003) for a discussion on 'practical problems' of this nature with PPP.

Models of language teaching such as PPP (presentation, practice, produce) and ESA (engage, study, activate) can be an effective way of teaching 'a small sample of language with the focus on a particular form' (Willis, 1996: 134). In both the produce and activate stages learners can be drilled with oral repetition or asked to produce written sentences to replicate the language focus of the lesson.

However, when learners engage in a discussion they bring with them their own personalities, beliefs, and expectations; they are not simply clean slates. These factors may have a bearing on how they perceive and interact with other members of the group. In addition as individuals, and as part of the wider social group of the classroom, their identities are in a constant state of flux as they readjust themselves to the group dynamic (Norton, 2000 in Kumaravadivelu, 2003: 48). Consequently there may well be a tension with how they see themselves as individuals and the roles they start to assume, or are given, as a member of a new group (Dornyei and Murphy, 2003: 111). With some members of the group they may feel they have much in common. With others there may be conflict as a result of personality clashes, disagreement over how activities should be carried out or from communication breakdown (ibid: 138). Conflicts may be voiced or result in individuals preferring not to work together. Individuals may come to the group as 'outsiders' entering an established social grouping or they may be part of an established group unaccepting of change. Individuals may simply not be on speaking terms with members of the group thus affecting who they believe is worthy of having a voice in the group (Bourdieu cited in Norton, 2000). They may feel uncomfortable carrying out activities the teacher asks of them; they may be uncomfortable using certain kinds of language or discourse that are 'highly sensitive to context and social purpose' (Coffin et al., 2009: 382), even if this is the language focus of the lesson. Alternatively they may focus on achieving their own objectives without conflicting with peer pressure. All of these factors may, or indeed may not, play a part in determining the interpersonal relations of the pairing or group and the communicative resources of the participants: Kress and Fowler (1979) points out that all language is addressed to someone and that relationships are generally asymmetrical. I return to this in *5.6 Extract 4 Changing roles: contextual variables and communicative resources*. (See Nunan, 2000 for a discussion on the experiential view of learning.)

2.6 Rounding up

In this section I have argued that the speaking activity intended to activate the language focus of the lesson is problematic. The exercises leading up to the activity present and practice the form and meaning of four modal auxiliary verbs. Whilst such exercises can prepare learners to integrate the language into simple sentences or utterances, they do not raise awareness of the interpersonal function of modality and thus how participants in a discussion use modals to negotiate interpersonal relations; I have tried to show that this is as much a part of modality as it is a means of expressing

degrees of certainty or possibility. Whilst this might sound overcomplicated for intermediate learners of English it need not be. If the purpose of the lesson is to integrate modality in to a discussion, looking at how and why speakers use modality in such a genre might be a sensible place to start. Language learners do not need to become linguists just made aware of how language is used in different communicative settings to make meaning. I will return to this later in *Chapter 6 Teaching implications: applying my research*.

I have also tried to show that the course book *Language Leader Intermediate* has emerged out of approaches to language teaching and learning over the last hundred years. It incorporates structural and functional approaches to grammar, uses situations to contextualise language, employs a communicative methodology, and uses PPP and ESA as a means of sequencing lessons. Whilst all of these have their relative merits none of them seem to have been developed to any degree in the unit on modals. I have also argued that the selection of modals is typical of what we find in ELT course books but not based on researched linguistic data. Therefore it might be more suitable and useful to teach lexicogrammatical features of modality that turn up in spoken discussions if the objective of the lesson is for learners to engage in a discussion. Again I will return to this in my teaching implications when I look at usefulness, awareness raising activities and alternative ways of teaching modality.

Finally, I have argued that the speaking activity does not account for the interpersonal relationships and dynamics of the language classroom. This is perhaps inevitable in a commercially produced 'one size fits all' course book. However, what this means in terms of the lesson is that certain activities may not be realised in the way the writers had intended. I have suggested that the dialogues in the speaking activity are supposed to resemble the listening extracts which contextualise the language focus of the lesson. However, it is important to understand that the given example (albeit contrived)²² represents 'two friends discussing in a museum', whereas for the students it is two learners paired up by the teacher in a *classroom*. In other words, these are two entirely different contexts. In the next section I will explore this further when I discuss the dialogic of modality and context.

²² I discuss language on display below in 3.2 *Context of situation: foundations*.

Chapter 3 Context: finding my way through the 'black hole'

[L]anguage is said to function in 'contexts of situation' and any account of language which fails to build in the situation as an essential ingredient is likely to be artificial and unrewarding.

(Halliday, 1978: 28-29)

3.1 Lead in

Situating language in 'context' is axiomatic in much of English language teaching these days. We can find numerous references to it in the literature (for example, Ellis, 1994; McCarthy, 1991; Young, 2009), in handbooks for language teachers (e.g. Kumaravadivelu, 2003, Thornbury, 1997; Woodward, 2001), and in teaching materials (e.g. Gammidge, 2004; Albery, 2008). However, 'context' is a 'big' word and by that I mean one that is used across disciplines such as anthropology, ethnography, linguistics, discourse analysis, and language teaching with considerable breadth of meaning and application; with each one arguably having left their mark on the use of context in English language teaching. We have 'local' and 'global' contexts (Brown and Yule, 1983), 'static' and 'dynamic' contexts (Widdowson, 2004), linguistic, extra linguistic and behavioural contexts (Duranti and Goodwin, (1992), psycholinguistic contexts (Ellis, 1994), 'forgotten contexts' (Blommaert, 2005) as well as 'serious' 'thought provoking' and 'light-hearted games' as contexts (Gammidge, 2004). In addition the term also seems to be used interchangeably with situation (Woodward, 2001), setting, (Skehan, 1998) and co-text (Ellis, 1994). Context is 'big' in another sense in that it spans the infinitely small, i.e. a single word utterance to the potentially infinite, i.e. a text of many words inter woven with cultural, social and political discourses pertaining to the who, when, where and why of its making (Blommaert, 2005; McCarthy, 1998). For this reason it has been described as 'the black hole' (McCarthy, 1998). It has been claimed in fact that context eludes definition (Finch, 2000; Widdowson, 2004) and that attempts to unravel it make it such fertile ground for discussion (Duranti and Goodwin, 1992:2).

In this chapter, I want to explore context and discuss it in relation to English language teaching and learning. My starting point is the lesson on modals of deduction from *Language Leader Intermediate* that I discussed in the last chapter. Whilst I point out that this particular lesson is a typical course book lesson in terms of the way the modal structures are presented (see 2.4 *Modality in spoken language: insights from language corpora*) and is modelled on a commonly used EFL teaching sequence found in language course books (Harmer, 1998; Thornbury, 2006), I am not critiquing the ESA model per se. What I am attempting to unravel is what is meant when the course book writers

talk of contextualized language (e.g. Albery, 2008: 99) and the relationship of this to spoken language produced by learners in production or activation type activities.

To recap I refer again to the ‘Lesson topic and staging’ in the Teachers’ Book:

[t]his lesson looks at famous designers and some of the things they designed. Students read a text about a famous American designer. This is followed by a listening on different designs and designers which contextualises the grammar focus of the lesson: modals for present deduction (*can’t, must, might/could*). Students then do a series of activities to explore the meaning and form of these modals before using them in a speaking activity.

(Albery, 2008: 98)

As I have previously pointed out, the staging follows an ESA (engage, study, activate) sequence: a sequence closely aligned to PPP (presentation, practice, produce) (Harmer, 1998). Woodward describes PPP in the following way:

[t]his teaching sequence involves setting up a situation, eliciting or modelling some language that fits the situation, having students practise the new language in a controlled way and then encouraging students to use the new language in a freer way for their own purposes and meanings or in differing, artificially constructed contexts.

(Woodward, 2001: 126)

In the previous chapter I argued that the descriptions and choices of modality offered by the course book writers did not lend themselves well to the paired discussion activity intended to activate the language. Here I want to turn my attention to the problem of ‘context’ in this regard. The extracts from Albery and Woodward above appear to make the assumption that language which is contextualised in a situation at the presentation/engage stage of a lesson will be ‘produced’ or ‘activated’ by the learners at a later stage of the lesson. Indeed, this is a desirable outcome of PPP and ESA. The writers of *Language Leader Intermediate* prepare for this by firstly contextualising the language focus through a listening exercise, i.e. pairs of students discussing designs at a design museum (engage), then through exercises on meaning and form (study), and round off with an ‘artificially constructed’ speaking context in which the language focus of the lesson simply fits into or gives rise to, i.e. a paired discussion to make deductions about a selection of pictures/designs (activate). The implication is that the context of ‘three pairs of students at a design museum’ gives rise to the same lexicogrammatical features as the context as pairs of students working together in the classroom. However, I argue that this is not the case. The context, artificial or otherwise, which presents the language focus of the lesson, is not the same as the classroom context for a number of

reasons which I will discuss below. Accordingly the language generated by the pairs during their discussions may differ.

The literature on context is vast and whilst I acknowledge the numerous contributions made my aim here is to provide a selective background to aspects of context that have had a significant impact on English language teaching. Reference to these aspects of context frequently occur in the literature and on teacher training courses; I will provide examples as illustration. I will follow this with a discussion on further aspects of context and argue that contexts which provide linguistic input in teaching materials are a different thing to language activation contexts in the language classroom; and for this reason there is no simple transfer of language from one context to another.

3.2 Context of situation: foundations

As we can see from the Halliday quote at the beginning of this chapter context is often used alongside 'situation' and it is a phrase that puts the two words together where I start. Malinowski put forward the phrase 'context of situation' in 1923 whilst studying indigenous culture in the Trobriand Islands. He explains it as:

an expression which indicates on the one hand that the conception of context has to be broadened and on the other that the situation in which words are uttered can never be passed over as irrelevant to the linguistic expression. (1923: 306)

The broadening of the 'conception of context' refers to need to look beyond the immediate linguistic context, i.e. the co-text if we are to fully make sense of an utterance. Contemporary linguistics of the time favoured written down, often 'dead', languages for the purpose of linguistic analysis. Consideration of 'external factors' beyond the isolated, monologic utterance was outside of the field of linguistics (Volosinov, 1929/1986). As an ethnographer Malinowski argued that language only becomes 'intelligible when placed in contexts of situation' (1923: 306) and believed the study of 'an object alive' would yield 'concrete conclusions' (ibid: 308) in respect of meaning. Language does not happen in isolation and thus the 'context of situation' in which an utterance is made is integral to its meaning.

For example, knowing what we do about the modal finite 'must' for expressing obligation and necessity it seems fair enough on the surface to say that the isolated utterance *You must be joking!* is used to make deductions (Cotton et al., 2008). However, if we also know that this utterance came as a response to the utterance *You can go up and get a bit of reading done in your bed at eight o'clock*²³

²³ This example has been taken from the Brigham Young University British National Corpus online (accessed September 2014). For original extract and information on participants see *Appendix 4*.

we can say that whilst the modal form still comments on the truth value of the proposition, it is not making a deduction in the sense of 'I'm certain that you're joking' (ibid: 75). What is needed to fully understand the meaning is to know something about the context of situation in which it was uttered. I will return to this example and develop this discussion below and in 4.4 *Systemic functional grammar*.

Malinowski's context of situation was coined while he was studying the language of a people with a very different culture and environment from his own, originating in the fields of anthropology and ethnography. However, the concept came to be widely accepted in the study of language and linguistic analysis: the term being elaborated by Firth in his 1950 paper 'Personality and language in society'. Firth came into contact with Malinowski at the School of Oriental and African Studies (SOAS) where he was appointed Chair of General Linguistics in Britain in 1944. He was much influenced by Malinowski's work and applied the context of situation to his own linguistic analysis. And, significantly for language teaching, he integrated the notion of situation into a course of Japanese to RAF pilots in World War II (Howatt, 2004: 309) unifying language and social activity.

Throughout the 1950s and 60s the use of 'situations' to teach English became increasingly common. The term situation was taken up by A.S. Hornby, who developed the Situational Approach. He believed that '[l]anguage is needed for situations and should be taught with situations as the starting-point' (Hornby, 1950b: 156) and saw the importance 'of enabling the learner of a foreign language to form links between new words and constructions and real situations' (1950a: 98). Hornby's methodology embedded grammatical patterns in simple situations illustrated by picture sequences or classroom actions to illustrate meaning. For example, the approach would take a 'situation' such as walking towards the window of the classroom to illustrate 'I am walking to the window' or use a series of illustrations to enable the telling of a story which might be extended to real situations in real life outside the classroom (1950b: 155). By the 1970s, however, as different theories on language use and language learning developed the Situational Approach was starting to run its course.

Wilkins argued that a situational syllabus was insufficient for more than 'narrowly definable contexts of learning' (Wilkins, 1979: 84). Based on an assumption that 'our choice of linguistic forms is restricted by the nature of the situation in which we are using language' (Wilkins, 1979: 83) a situational syllabus centred around 'predictions of the situations in which the learner is likely to operate' (ibid). Wilkins saw this as problematic. Firstly in terms of describing what a 'situation' actually is and following on from that the difficulty of relating the language therein to language learners' needs. For example, just what the situation of walking towards the window is and how it or the structure meets learners' needs is questionable.

Widdowson too criticised the contextualization of language items in the Situational Approach. He argued that the approach only taught the 'signification' of language forms and not their communicative functions (Widdowson, 1979: 118). For example, the present continuous *I am walking to the window* only illustrates the form of the 'tense' and how it relates to other forms: *I walked to the window, I am going to walk to the window*; it does not teach its communicative function i.e. that of commentary.

However, whilst the approach itself was coming under criticism, the relevance of situations was not at question. Arguing for the inclusion of communicative functions in a Notional Syllabus Wilkins allows for 'situationally oriented units' (Wilkins, 1979: 86); a syllabus designed around functions rather than unique situations. Widdowson likewise saw the continued need for situational settings but that the teaching of English needed to be extended to include communicative acts. For the Threshold Level (see previous chapter) situations were very much at the centre of the learning objective and form the starting point of determining language activities and language functions. Van Ek describes situations as 'the complex of extra-linguistic conditions which determine the nature of a language-act' (Van Ek, 1979: 105) and whilst acknowledging them to be 'strictly personal and unique' (ibid: 105) he wavers their heterogeneity to concentrate on four components: the social rules which the learner will be able to play; the psychological roles which the learner will be able to play; the settings in which the learner will be able to use the foreign language; the topics which the learner will be able to deal with in the foreign language (ibid: 106). And thus seeing the communicative situation as important in understanding communication is quite different to the organisation of language teaching around different situations as in the Situational Approach.

For Malinowski the context of situation was intended to yield clearer understandings of language use through ethnographic accounts of sociocultural frameworks. Yet the broadening of context to include 'the situation' of a language event has filtered down into English language teaching in different ways. For Wilkins and Widdowson, among others, situations provide context for communicative language functions; for Hornby a situation is a means for illustrating language forms. In this second vein situations are created to display language for teaching purposes. In the absence of knowing anything of the who, why, when and what of the context of situation, we are in a less than certain position for understanding why certain choices of language were made over others; I discuss this in more detail when I introduce contextual variables in *3.4 Context of situation revisited: field, tenor and mode*. The following dialogue extract from *Language Leader Intermediate* to present modals of deduction is an example of language on display. The selection of modality here has no more bearing on the situation (and vice versa) of 'pairs of students at a design museum' (Cotton et al., 2008: 75) than let's say a husband and wife discussing pictures in a magazine at home on a Sunday afternoon.

G: What's that?

S: I've no idea. What could it be?

G: It might be a spaceship. Well, a toy spaceship. It's the right shape.

S: No, I'm certain it's not that.

(Cotton et al., 2008: 176)

Situations of this kind are common in English language teaching course books acting as static spaces for the presentation of language. Accordingly statements such as the following 'a listening activity which contextualises the main grammar focus' (Albery, 2008: 96) are somewhat fuzzy. The main grammar focus in the dialogue has been co-textualized in that the 'situation' created by the surrounding text provides an illustration of what the modal forms signify in relation to each other, i.e. degrees of certainty, but the modal forms themselves cannot really be said to be contextualised as they arguably do not reflect sensitivity to the social purpose of the participants, i.e. the context.

Perhaps it is understandable that context of situation has come to be used in this way in some EFL course books. Malinowski's work was not initiated for the teaching of language nor did he provide a more detailed framework of analysis that could be used to explore the relationship between the context of situation and language (Widdowson, 2004; Young, 2009). Moreover, as we have seen ELT course books are very frequently designed around a structural syllabus with language structures embedded in situations as a means of presenting language; theirs is not the function of analysing the relationship of language and context. Context or situation provides language input.

The 'context of situation' is a very large frame for the study of language and just how much of the situation is necessary to make sense of an utterance needs qualifying. Earlier I referred to context being potentially infinite and the bearing that numerous factors such as social and cultural histories can have on meaning. Whilst this is true Halliday points out that some features of the situation are more relevant than others (1978: 29), with background of persons, actions and events being necessary components to derive meaning. Returning to the exchange above:

A: You can go up and get a bit of reading done in your bed at eight o'clock

B: You must be joking!

it is necessary to know something about who the speakers are, the relationship between them, where and why the exchange is taking place. I return to this below.

3.3 SPEAKING and communicative competence: bringing in the social

Another important perspective on the relationship of context to the understanding of language came out of the work of the anthropologist Dell Hymes. Concerned with the functions of language in social

life Hymes broke away from structural linguistics and descriptions of abstracted rules of language to the study of contextualized language use. In the late 1960s he proposed the term 'ethnography of communication' as an approach to exploring language in its social contexts. His fieldwork revealed that utterances were shaped in socially-contextualized ways and he foregrounded two important factors to the use of language: what speakers can and do say and the specific contexts in which it occurs (Johnstone and Marcellino, 2010: 4). He labelled this spoken interaction a speech event. As a framework for interpreting and understanding speech events he developed the mnemonic SPEAKING (see overleaf).

Setting
Participants
Ends
Acts
Key
Instrumentalities
Norms
Genre

Two important points to note here are the use of a non-linguistic unit being used for the interpretation of speech (Duranti and Goodwin, 1992: 25) and the primacy of *la parole*, actual speech for the study of language. The importance of the non-linguistic unit sees a shift in the definition of context: context not being solely a linguistic phenomenon. The primacy of speech underscores Hymes' belief that 'appropriateness' and 'rightness' of utterances are derived from socially constructed contexts and not written down abstractions. SPEAKING can be elaborated as follows:

The Setting refers to the place and time in which the communicative event takes place; Participants to the speakers and hearers and their roles and relationships; Ends to the stated or unstated objectives the participants wish to accomplish; Acts to the form, content and sequence of utterances; Key to the manner and tone of the utterances; Instrumentalities to the channel (oral or written) and the code (formal or informal); Norms to the conventions of interaction and interpretation based on shared knowledge; Genre to the categories of communication such as lecture, report, essay, poem, etc. (adapted from Kumaravadivelu, 2003)

As an example if we again take the utterance *You must be joking!* from the exchange:

A: *You can go up and get a bit of reading done in your bed at eight o'clock*

B: *You must be joking!*

and know it is a mother and child at home, and that the mother is attempting to get the child to go to bed early, and that the child has questioned her decision and responded by telling her he will go to bed at his normal bedtime, and that the child is starting to become irritable, we are in a much better position to interpret the utterance. I return to this example and develop this discussion below in 4.4 *Systemic functional grammar*.

Hymes' SPEAKING framework has been widely used in English language teaching (e.g. Cook, 1989; Cook, 2001; Kumaravadivelu, 2003). McConachy draws on the mnemonic as a teaching device to language learners 'for making salient the myriad sociocultural factors that influence language use' (2009: 124). And a similar framework is used by Brown and Yule to give learners access to 'globally' determined features of context in spoken language: 'speaker, listener, place, time, genre, topic'

(Brown and Yule, 1983: 63). But perhaps the biggest influence of Hymes' work on English language teaching is his notion of communicative competence.

Out of his ethnographic observations of language use he redefines Chomsky's notions of linguistic competence and linguistic performance to that of 'communicative' competence. Hymes argued against linguistic theory concerned primarily with an ideal-listener in a homogenous speech community, and the lack of relevance given to 'the actual use of language in concrete situations' (cited in Brumfit and Johnson, 1979: 3). For Hymes actual language use supported by a theory 'within which sociocultural factors have an explicit and constitutive role' (Hymes, 1979:

6) was central. His theory, which integrated the linguistic with other forms of communication, asked four questions: whether something is 1) *possible* i.e. grammatical within the language; 2) *feasible*, i.e. able to be processed; 3) *appropriate* i.e. suitable to context; and 4) whether it is actually *performed* i.e. part of language usage. Clearly language use and context are inseparable.

With its shift from an emphasis on language form and the Situational Approach to language use in context and how to mean, proponents of the communicative approach to language teaching in the early 1970s welcomed the notion of communicative competence taking it as their motto (Howatt, 2004). Brumfit called for a rethinking of methodologies away from 'the accurate construction of the target language' (1979: 187) to more fluency-based needs starting with communication and focusing on the *process* of language learning. The means was communication, the goal communicative competence: the ability to use language appropriately.

However, whilst the term 'communicative' is common currency in mainstream English language teaching these days, it is arguably somewhat removed from Hymes' original usage and indeed the beginnings of the communicative approach to teaching. Cook points out that many teaching techniques in the communicative style promote active practice similar to audiolingualism (Cook, 2001: 216) and that whilst the teaching of communicative functions may affect syllabus design it does not always affect teaching methods (ibid: 212). In the Introduction to *Language Leader Intermediate* we are informed that the course uses a communicative methodology (Albery, 2008: 3). What is being referred to are the kinds of exercises, activities and tasks that are carried out by the learners in pairs or groups with the intention of learning language by doing - the teacher taking 'one step back' (Cook, 2001: 214). But, as already pointed out, the language input in *Language Leader Intermediate*, whether structural forms or communicative functions, are presented to the learners for production and

practice – ‘communication’ or ‘communicative competence’ being concerned with language as a *product*, quite different to the *process* referred to by Wilkins and Brumfit.

Howatt (2004: 253) refers to the oversimplification of the term ‘communicative competence’ in mainstream English language teaching. Perhaps this is understandable given the difficulties of applying linguistic theory to practice as Brumfit pointed out in 1979 with specific reference to a communicative methodology (1979: 206). So whilst course books writers may make claims such as ‘we use a broadly communicative methodology’ (Albery, 2008: 3) there are large parts of the course book that do not. Albery confirms this by saying ‘*Language Leader* is not based on one particular teaching “philosophy” or methodology’ and adds the course is informed by ‘what works in the classroom’ (ibid). The lesson on modals of deduction is a good example. Although the learners are engaged in ‘communication’ the outcome of the lesson, in terms of the uses of modality, cannot be said to have raised communicative competence in a Hymesian sense.

3.4 Context of situation revisited: field, tenor and mode

The notion of context of situation as put forward by Malinowski and then Firth was taken up and developed by Halliday. For Halliday the relationship between language and social life meant that given ‘certain facts’ a great deal of what speakers say in communicative interaction is predictable (1975: 129). The certain facts or ‘environmental determinants of text’ (ibid: 130) define a situation ‘type’ or ‘social context’ setting it apart from other situation types. So, for example, a discussion is a different situation type to a service encounter or commentary to a football match given their relative environmental determinants; these determinants Halliday refers to as field, tenor and mode.

Below is a framework in Coffin et al. (2009) based on aspects of field, tenor and mode used in text analysis. I have used this framework to provide a brief analysis of a transcript of a discussion from the lesson on modals of deduction in *Language Leader Intermediate* (the example is reproduced from 2.2.2 *Language Leader Intermediate* above). My analysis is in parenthesis, the transcript follows the framework. Coffin et al. refer to field, tenor and mode as contextual variables.

1. Field

- The social activity taking place (*a discussion between friends*)
- The topic (*a picture by da Vinci in a museum*)
- The degree of specialisation (*basic knowledge of da Vinci by speaker B*)
- The angle of representation (*exhibit as object of discussion, uncertainty of object, speakers as agents of mental processes – making deductions*)

2. Tenor

- The social roles and relative social status in terms of power (*friends of equal status, speaker B has a little more knowledge*)
- The social distance (*little social distance between speakers*)
- Speaker/writer persona, i.e. general stance and assumed degree of alignment/agreement between interlocutors (*Speaker A asks more questions, speaker B gives more information, speaker A accepts speaker B's information*)

3. Mode

- The degree of interactivity (*spoken and very interactive*)
- The degree of spontaneity (*unplanned spoken interaction, some previous knowledge*)
- The communicative distance in time and space from the events discussed (*language accompanies action in time and space*)
- The role of language, i.e. the degree to which it interacts with other meaning-making resources such as gesture (*spoken language the main resource, interaction with written information*)

A: Wow, look at that. What is it?

B: It's a drawing. I think it's a sort of flying machine.

A: Yes, it could be that. Who do you think it's by?

B: Mmm, it might be by da Vinci, I believe he did that sort of thing. Have a look at the sign. What does it say?

A: Erm ... yes, you're right, it is by da Vinci.

B: Goodness! It's in very good condition ... it says here he was born in 1452, so it must be over 500 years old.

A: Yes, and it's an amazing drawing.

B: Yeah, I read somewhere he was fascinated by birds and flying. Perhaps that's where he got his ideas for the drawing.

A: Yeah, you're probably right.

(Albery, 2008: 175)

Using these contextual variables the analyst is able to make predictions about the kind of language contained in a particular situation type or social context. Taking the example again of a discussion we would expect to find the present tense (and past where appropriate), declarative mood with some use of interrogatives in spoken mode, modality, and generic nominal groups (Coffin et al., 2009: 214-215). Certain situations, therefore, have a range of possibilities of what a speaker can do in order to

make meaning; possibilities realized through the linguistic system and what Halliday has termed the meaning potential of the language.

For Halliday, as with Hymes, language is seen as a communicative tool and so perhaps it is not surprising that field, tenor and mode with their relationship to extra-linguistic phenomena (Brumfit and Johnson, 1978: 44) share much in common with Hymes' SPEAKING categories (Halliday, 1975). In terms of language teaching Halliday makes clear that language needs to be taught in relation to the communicative functions for which it is used (Halliday, 1978). Indeed, his ideas relating meaning to grammar and language functions to structures were integrated into early communicative language teaching (Rivers, 1968/1981; Littlewood, 1981; Howatt, 2004). For language teaching materials it is possible, whilst acknowledging that it is arguably not what Halliday had meant, to see how artificially constructed texts can be built around particular social contexts as a means of presenting language functions and structures. The dialogue from *Language Leader Intermediate* above is an example. The context is two friends discussing exhibits at a design museum, discussions give rise to modality: the contextualized language focus of the lesson.

It is important to note at this point that Halliday's take on the context of situation was somewhat different to Malinowski's. Whereas Malinowski saw the context of situation as a set of descriptive spatio-temporal goings-on, a kind of back drop to the speech event, for Halliday it was constitutive of the speech event (Halliday, 1975): '[t]he context plays a part in what we say; and what we say plays a part in determining the context' (Halliday, 1978: 3). Context of situation, or context as I will refer to it from now on, is not simply a static set of factors which determine language but a dynamic constituent that shapes language.

3.5 The lesson as context: the dynamics of the classroom

What I have tried to show above is how language input in English language teaching has moved away from isolated abstractions to language in context or situations and how this move has been informed by disciplines such as anthropology and ethnography and the study of language in sociocultural situations. I have also shown that situations and contexts have been interpreted differently in ELT over the last 60 years leading to some very different approaches to teaching. These differences notwithstanding, language in context is a mainstay in the literature, in teaching handbooks and teaching materials. And whilst there has not been extensive research carried out on the effect of contextualised language input on language acquisition, Kumaravadelu points out the positives of research that has been done on the importance of context in areas such as comprehension and production as well as understanding pragmatic, semantic and discourse factors of utterances (2003: 214).

In this section I want to return to the topic of producing or activating language input which has been contextualised at earlier stages of the lesson. Just to recap lesson techniques such as PPP or ESA present or engage learners with 'contextualised' language which is to be used 'productively' in subsequent oral or written activities. As I have already pointed out this is a common teaching sequence found in published English language teaching course books. What I suggest is that whilst contextualised language input may well aid comprehension, the same language may not necessarily be carried over by the learners to the production or activation stage. *Language Leader Intermediate* presents a model of context which cannot easily be replicated in the language classroom. I will argue that the 'new contexts' of the activation stage which come about as a result of the interaction between the who, the where, the when, the why – the dynamics of the context - are potentials to lead learners away from the language focus of the lesson and illustrate this with examples from my data. (See also Nunan, 2000; Ellis, 2003 and Kumaravadivelu, 2003 for further discussions on how learners interpret classroom tasks.)

As I have already indicated the part of the lesson that I am concerned with here is the activation stage, see below (reproduced from 2.3.1 *Problematizing a lesson on modal verbs*, above). The instructions in the course book for this stage of the lesson read:

[w]ork with a partner to discuss what you think the designs below are. Try to use these words and phrases.

must might could can't
I'm sure/certain maybe/perhaps
It's possible that ...
It's not possible that ...

(Cotton et al., 2008: 75)

The stated objective in the Teacher's Book for this stage of the lesson is the use modals in a speaking activity.

From the example of Halliday's work on field, tenor and mode above we can see that the social activity the participants engage in is a significant factor in shaping context and language. The 'success' of the activation stage, with learners drawing of the suggested range of modal forms, relies on it being a particular type of social activity, i.e. a discussion. As I have pointed out, the lesson follows an ESA procedure. Let me summarize as follows:

Engage: learners listen to a text which contextualizes the language input.

Study: the form and meaning rules of the language input become the focus of a matching exercise.

Activate: the learners are given picture prompts and 'use' the language input in a discussion.

However, from the data I have collected there are a number of mismatches (Kumaravadivelu, 2003) between the desired manipulation of the activation stage and learner's interpretation of it that act as an obstacle to its realisation as a discussion. Consider the following (for original data see *Appendix 6: Classroom data CD ROM – Richard (2) 16.00*):

- S1 What she's doing?
- S2 He=
- S1 =He ((laughter))
- S1/S2 (3 secs) uhm/uh
- S2 So here is the mouse (.) so he's searching something (=
- S1 =yeah research=
- S2 =research
- (2secs)
- S2 ↓ maybe he's ↑ scientist and trying to find out something (2 secs) here's the diagram so it shows (.) maybe how fast runs ↓ the mouse and it affects that and that and there is the result of that=
- S1 =You have a very good imagination [S2 ((laughs))] () I don't know (.)

(See *Transcript A, Appendix 11* for full transcript.)

In this example S1 gives a detailed description of what she thinks is 'happening' in the picture and does so using the continuous aspect 'a scientist trying to find out something', 'She's searching something', 'doing research'. Indeed her opening question, which she answers for herself, asks exactly that 'What she's doing?' In answer to this question she makes highly modalized statements about the picture; the only suggested modal form comes towards the end of her extended turn when she says 'it shows maybe how fast runs the mouse'. What is happening here is what Kumaravadivelu has termed a procedural mismatch: the way the students go about completing the activity or seeking a resolution to a problem differing from the expectations of the teacher. The picture does not generate a discussion as intended but an attempt to make sense of what the person in the picture is doing by describing it.

This next example illustrates a different kind of mismatch, a pedagogic mismatch, and this refers to a discrepancy between the teacher's and learners' perceptions of the learning objectives of an activity (for original data see *Appendix 6: Classroom data CD ROM – Richard (1) 26.34*).

- S1 I think massage for headache
T ok
S1 (demonstrates) I can see before yeah? (looks at teacher)

(See *Transcript B, Appendix 11* for full transcript.)

Here S1 makes a guess at what she thinks the object is based on something similar she has seen before. Having made this guess she turns her attention to the teacher for clarification. What is happening here is a mismatch between the pedagogic aims of the activity. The teacher wants a discussion whereas S1 sees the activity as one of demonstrating 'correct' answers. There is no attempt at discussion and the other three students in the group wait for the teacher to accept or reject the answer.

In this short extract there is also what Kumaravadivelu calls strategic mismatch. A strategic mismatch refers to the means by which learners go about completing an activity differing from what the teacher had planned. The activity is intended to facilitate a discussion with learners drawing on the modal forms they have learnt to make deductions about the picture. Here though S1 takes the 'simplest route' by making guesses and then turning to the teacher for clarification. There is no procedure here for discussion with fellow students. Mismatches of this kind appear throughout my data (see, for example, *Appendix 11 Transcripts*).

In *Transcript D* there is a procedural mismatch in which S1 knows what the picture is and attempts to tell the teacher. On failing to get the teacher's attention he turns to the other two members of his group and identifies the object in the picture. This could be a possible reason why a discussion fails to develop.

Pedagogic mismatches are common across the data. In *Transcript B* and *Transcript C* different speakers attempt to identify the object in the picture and thus their perception of the activity appears to differ from the teacher's perception of the activity as a discussion. The 'fork' dialogue at the beginning of *Chapter 2 Modality: coursebooks and beyond* and the 'helicopter' dialogue in *5.6 Extract 4: Changing roles: contextual variables and communicative resources* are two further examples of a pedagogic mismatch. The attempts by learners to produce 'correct' answers may be a result of the initiation, response, feedback model common to these dialogues.

The above examples, in addition to *Transcript D*, are also indicative of strategic mismatches. The learners take the 'simplest route' by making guesses and in so doing circumvent the process of deduction by means of a discussion.

The 'success' of the lesson and meeting the objectives is measured on the extent to which the learners use the target language, and so to ensure that this happens, the teacher needs to monitor at the activation stage to prompt and intervene when necessary to ensure the correct forms are used. This is all very well if the intention is to simply practice the language input orally. However, there is a difference between this kind of exercise and a discussion. As I have pointed out in the previous chapter there is a tension between methodologies in *Language Leader Intermediate*. The authors claim to use a communicative methodology (Albery, 2008) and whilst PPP and ESA might be considered a communicative type of teaching technique (Cook, 2001) there is some ambiguity in just what is expected of the learners at the activation stage. This results in what Cook has referred to as the incompatibility between teaching models (ibid). Indeed, the rubric states '[w]ork with a partner to discuss what you think the designs below are. Try to use these words and expressions' (Cotton, 2008: 75 my emphasis). Are the writers asking the learners to practise the forms or to engage in a discussion? The answer(s), I argue, has implications for the kinds of context created and thus the relationships of the interactants and their communicative resources.

3.6 A multimodal perspective on context: more than words

So far my discussion has centred mainly on the linguistic features of language and their relationship to context. Indeed whilst the work of Malinowski, Firth, Halliday, Hymes, among others, point to the importance of the social as an integral part in the construction of context very little of this has filtered down into mainstream ELT. As language teaching is concerned with the job of teaching language, study of the social might on one level be deemed as unimportant. However, my argument is that 'the social' within the contexts of the classroom needs to be an essential consideration for not only the activation of language, and I have attempted to show above how this is the case, but also for an understanding of all communicative resources learners use, or do not use, in the communication of meaning.

In this section I want to turn my attention to social semiotic multimodal analysis. Social semiotic multimodal analysis developed largely out of Halliday's theories on social semiotics and systemic functional grammar. Halliday's notion of socially situated language led proponents of social semiotic multimodal analysis such as Kress and Van Leeuwen to emphasize the context of communication, the 'sign-maker', and the situated choice of communicative resources (Jewitt, 2009). In brief social

semiotics views participants in 'semiotic activity as connected and *interacting* in a variety of ways in concrete social contexts' (Hodge and Kress, 1988:1; emphasis in original). Within these social contexts participants have recourse to 'socially shaped and culturally given' modes (Kress, 2009: 54 in Jewitt, 2009) as resources for making meaning. Multimodality views all communication as consisting of sets of semiotic resources, for example a spoken interaction may consist of language, posture, gaze, sound, etc.; these resources being socially and culturally shaped. Within multimodal research no communication is viewed as monomodal.

Jewitt explains that the focus of social semiotic multimodal analysis is 'on mapping how modal resources are used by people in a given community/social context' (2009: 30) and thus the centrality of context and the social process in sign-making (ibid). She adds that multimodal research 'is strongly underpinned by social and cultural theories of representation and communications', and identifies 'the link between representation and communication and the social, cultural and historical conditions of societies' (ibid: 5). Social semiotic multimodal analysis explores the range of modal resources available to make meaning within specific social contexts; the choice of mode in any communicative event being a complex interaction of social, cultural, emotional, psychological and physiological conditions (ibid).

Arguably then a fuller and more detailed analysis of any communicative event would need to take into account a variety of modes; although, of course, this will depend on what the analysis or the research perspective is. Social semiotic multimodal analysis is able to reveal how modes interact in given social contexts. It might be the case that one mode is more prominent than others, for example as Van Leeuwen points out 'giving and demanding "goods and services" is more often realized non-verbally' (Van Leeuwen, 2005: 252) than by spoken language. Or we may find various modes being employed simultaneously to form modal ensembles (Kress, 2009: 64 in Jewitt, 2009) as in the case of asking for and giving directions in which verbal deixis (just go down there, these buildings) is accompanied by deictic gesture such as pointing and/or physical movement of the body towards the direction being indicated. In the extract overleaf it is hard to imagine deixis only being communicated verbally:

Um... if you go... actually because you're here it probably just go down there just walk along the road (<S1> Yeah) on the right you'll see a pub called the Victoria Park (<S1> Ok) park's behind it you're right on top of it the park park's basically behind all these buildings (<S1> Ah, ok) you'll probably get to quick quicker by walking through the estate I don't know but definitely you have to go down turn to your left when you see the pub on your right there's actually big gates that's the entrance

With social semiotic multimodal analysis the social context is intimately connected to mode and vice versa. Choice of mode depends on what is being communicated, when it is being communicated, where it is being communicated, how it is being communicated, and by whom. There is a dynamic here between context shaping mode and mode shaping context. So to bring my discussion back to the language classroom, learners have a range of modes at their disposal to make meaning as they gauge and shape the contextual variable of field and tenor, which in turn partly shapes the context for other participants who make their own corresponding choices of mode within the interaction. We can see this in this brief exchange between two students making deductions.

S1: It's a shell.

S2: (pause) Hmmm (tilts head to one side and avoids eye contact with S1)

The object of the activity was to use modals of deduction in a discussion about a variety of objects. Here S1 makes a closed stance statement with high modal commitment making the possibility of a challenge less likely. If we know something of the contextual variables of tenor, i.e. the social roles and status of the participants, for example, that S1 was a Turkish male student who contributed regularly and often to oral discussions and that S2 was a younger Japanese female student who was often quiet during spoken activities, we can broaden our understanding of the interaction. That S2 does not use spoken language to comment on S1's statement is perhaps understandable. However, she does make a contribution to the interaction using the modes of silence (a pause), sound 'Hmmm', movement (tilting the head) and eye gaze. A fuller analysis of these modes will be given in the next chapter, but it is worth a comment here on Van Leeuwen's discussion of nasality and nasal sounds. Citing Crystal (1970) Van Leeuwen argues that nasality occurs when people of lower social status address those of higher social status. On nasals he claims their semiotic potential 'lies in the way they "keep the sound inside"' (Van Leeuwen, 1999: 208). From this perspective both nasality and nasal sounds have the potential to convey modality. So what we see with S2's 'Hmmm' is a way of conveying meaning and modalising her commitment to the interaction as she feels fitting to the context that S1 has in part shaped. S1 does not in turn respond to S2. Again this has meaning potential. S1 has made a statement, S2 expresses doubt in the truth value of that statement and the 'discussion' comes to an end.

3.7 Rounding up

What I have tried to show in this chapter is that the term context has a long and complex history in English language teaching. However, whilst fields such as anthropology, ethnography, linguistics, and social semiotics have considerably broadened the concept of context in the way that it is understood

and in what it encompasses, many ELT course books still interpret context as co-text, a means of presenting language for study, and as static spaces giving rise to predictable linguistic features, for language practice. My understanding is that the term context needs to refer to more than the immediate linguistic co-text towards something that encompasses the social and cultural relationships of the participants, the activity and the communicative resources available. I would argue the interpretation of communication in the context of the language classroom, for example, is considerably enhanced using the contextual variables of field, tenor and mode. These variables allow us to see context as being something more than just 'there' as an ornamental backdrop to language: the variables interact to shape a dynamic of social interaction with communication and meaning being realised through a choice of modes.

Using an activity from *Language Leader Intermediate* I have tried to show how the rather limited concept of context taken by the authors inhibits a full understanding of the choices of modality speakers make and creates problems during the activation stage of the lesson, in which the learners invariably do not use the suggested modal forms. However, even if the learners are not using the target language, they are not simply sitting doing nothing; and even if they were, meaning would still be communicated. I have also suggested that the contextual variables of field, tenor and mode can have a bearing on the dynamic unfolding of language and context. In the *Chapter 5 Four extracts of data* I will explore the kind of communication taking place during the activation stage of a lesson on modals of deduction and will use social semiotic multimodality for my analysis.

A last word on 'the black hole' of context. I have referred the potentially infinite when describing context. Whilst I have extended my understanding of context to include the sign-maker and multimodality there are still numerous details that can arguably be included in a fuller description of context. Lighting, seating arrangements, the clothes learners are wearing, for example, whilst perhaps working on a less conscious level, are all signifiers of meaning, and thus contributors to context. Three points on this. Firstly, my analysis does not claim to be exhaustive. What I offer is simply an insight in to a broader understanding of interpreting social interaction within the language classroom. Secondly, I will make a distinction between setting and context: context being described in the terms that I have outlined above. Setting can be seen as the physical environment and whilst this does arguably impact on the participants, the activity and communicative modes, it will not feature as prominently in my research as other communicative modes. Thirdly, I will be using Norris' (2004) notion of higher- and lower-level actions as a means of 'framing' my research lens. This will allow me to focus my research without getting drawn into the black hole of context.

Chapter 4 Terminology and theoretical underpinnings

In this chapter I pull together the various themes so far discussed. I will define the terms relating to the main themes and theoretical underpinnings of my thesis. And I will explain, with examples, how these form the basis of my methodological approach to my data analysis in the next chapter: *Chapter 5 Four extracts of data*.

4.1 Epistemic modality

Firstly I will recap on my understanding of modality and how it is referred to in my research. 'Modality means the speaker's judgement of the probabilities, or the obligations, involved in what he is saying' (Halliday, 1985/1994: 75). At this point I need to restate that modality can be grouped under two main headings: deontic and epistemic.²⁴ Deontic modality relates to the speaker's view of the desirability of a situation and includes meanings such as obligation and permission. Epistemic modality relates to 'the speakers' assessment of the validity of what they are saying' (Coffin et al., 2009: 169) and includes meaning such as certainty and possibility. My research focuses on epistemic modality. Epistemic modality does two things: it comments on the truth value of a proposition by allowing the speaker to grade the strength of their commitment to an utterance (Coffin et al., 2009: 365), and along with deontic modality it forms part of the interpersonal metafunction of language and relates to the way speakers 'engage interpersonally and exchange points of view' (Coffin et al., 2009: 226).

In *Chapter 2 Modality: coursebooks and beyond* I critiqued the choices and descriptions of modals in mainstream ELT course books, in particular *Language Leader Intermediate* (2008). My argument was that modality is often restricted to modal finite verbs (modal auxiliary verbs) under the titles modals of certainty or deduction, for example *may, might, must* with no mention of their interpersonal meaning. What is offered to language learners is limited to descriptions of form, meaning and function, e.g. *must* means the same as 'I am certain that this is true' and functions as a modal of deduction (ibid). I argued that whilst this is not incorrect such descriptions limit the use to which the modals can be used. I gave the example of an activity from *Language Leader Intermediate* in which students were

²⁴ Halliday (1985/1994) uses different terminology to make this distinction and indicates that the terms 'epistemic' and 'deontic' modality are used in philosophical semantics to refer to probability and obligation respectively (see, for example, Kreidler, 1998 and Saeed, 2009). In SFG, degrees of probability and usuality are referred to as 'modalization' and degrees of obligation and inclination are referred to as 'modulation' (see also Eggins, 1994). Whilst acknowledging these terms, I use the terms 'epistemic' and 'deontic' throughout my thesis as I am more familiar with these terms and feel they can be more easily distinguished than the terms 'modalization' and 'modulation'.

asked to enter into a discussion to deduce the purpose of a selection of designs as a way of illustrating the difficulties of using the forms in connected discourse with multiple participants. My data reveals utterances were often limited to simple clauses containing a modal finite making a discussion difficult to get started, i.e. engaging interpersonally and exchanging points of view.

So for the purpose of my research I am focusing on both the truth value and interpersonal function of epistemic modality. As a lynch pin for bringing these two aspects together I refer to Hodge and Kress who claim all utterances 'bear the signs of modality' (Hodge and Kress, 1988: 124). This allows me to start with the interpersonal function of language and broaden the concept of modality beyond the modal auxiliaries. Modality can thus be realised in terms of lexicogrammatical structures such as modal nouns, adjectives and adverbs, mental, verbal and process verbs, vague language, mood, tense, deixis and hesitation phenomena. The following conversation (reproduced from above 3.5 *The lesson as context: the dynamics of the language classroom*) between two language learners discussing a design from *Language Leader Intermediate* illustrates how some of these features help realise interpersonal relationships.

- S1 What she's doing?
S2 He=
S1 =He ((laughter))
S1/S2 (3 secs) uhm/uh
S2 So here is the mouse (.) so he's searching something ()=
S1 =yeah research=
S2 =research
 (2secs)
S2 ↓ maybe he's ↑ scientist and trying to find out something (2 secs) here's the diagram so it shows (.) maybe how fast runs ↓ the mouse and it affects that and that and there is the result of that=
S1 =You have a very good imagination [S2 ((laughs))] () I don't know (.)

(See *Transcript A, Appendix 11* for full transcript.)

Inaudible speech, hesitation and a question open the discussion. S1 might have produced something inaudible as a means of offering comment whilst lacking confidence to actually 'voice' it. If we take the utterances as *sotto voce*, then it has arguably been modalized. S2s first contribution is a hesitation, although it is not clear whether this is in direct response to S1 or an indicator of uncertainty as to what the design is. Taken together with S1s following question 'What she's doing?' we see not only uncertainty but also sensitivity to each other's relative position by not wanting to assert themselves

too forcefully in to the discussion – and both modalized utterances. After more hesitation S2 starts to describe the picture. We see use of the present tense as one would perhaps expect in a discussion but the speaker's use of 'is' is significant. The purpose of the activity is to enter into a discussion and make deductions about the design. What we see, however, is S2 describing what she sees in front of her. The highly modalized utterances with 'is' allow her to express what she sees as 'real' over what she is uncertain about. Describing what 'is' happening allows her to take some amount of control in the discussion. It also allows her to take a closed stance; high modality and a closed stance indicate strong commitment. S1's follows S2's description by commenting 'You have a very good imagination' and thus seemingly content with what has been described, at least not to challenge it.

Hodge and Kress go further than verbal language in their discussion of modality and argue that every semiotic act is modalized (1988: 123). I will elaborate on this when I discuss social semiotics but as a working definition social semiotics views all semiotic acts and processes as social and is concerned with the participants, relations and structures in social acts and processes; semiotic acts and processes realise the social acts and processes of participants, relations and structures (1988: 122). Social semioticians view all semiotic acts and processes as having meaning potential and, thus, modality is present in all semiotic acts and processes (ibid: 123).

Kress and Van Leeuwen claim that '[w]e routinely attach more credibility to some kinds of message than others' (2006: 154). Some messages appear more real or lifelike and these are said to convey higher modality and thus represent a more 'truthful' image of reality. A stick man sketched with a pencil is seen to be less real than a high-definition colour photograph. Van Leeuwen explains how visual modality is conveyed by degrees of detail, saturation, depth, light, shade and tone (1999: 159). In a similar way verbal language is modalized by degrees: present/past tenses, closeness and distance markers, as well as using modal verbs 'must, might, could be'. The degree of modality conveys the truth value of the message. It needs to be pointed out here that notions of 'real' and 'truth' pertain not to actual objective truth, but truth as seen by the speaker, writer, illustrator, etc.; again in the same way that the truth values of verbal utterances are modalized by the speakers and writers (Hodge and Kress, 1988; Kress and Van Leeuwen, 2006).

In their work on visual communication they show, for example, how colour in pictures acts as an indicator of modality: 'the more that is taken away, abstracted from the colours of the representation, the more colour is *reduced*, the lower the modality' (Kress and Van Leeuwen, 2006: 159; emphasis in original). The practice of using soft focus and soft colours in advertising tends towards a promise of the product rather than a sharp focus and highly saturated colours representing 'reality'. An example is their discussion of a coffee advertisement. The advert uses soft colours and soft focus to lower

modality and thus represent ‘a fantasy or promise’ (ibid) of what using the product ‘might be like’ (ibid). However, they add that highest modality does not coincide with maximum saturation of colour as there is a point where maximum colour saturation lowers modality. For example, I have in front of me an advert for Cirio chopped tomatoes from Tesco Food Family Living magazine (April 2014: 110).²⁵ The image of a glossy tin of tomatoes stands proudly on a shining white ‘catwalk’ while either side of the catwalk sit rows of plump, ripe tomatoes; the caption in an azure ‘sky’ reads ‘Cirio, the true taste of Italy’. The advert is highly saturated in colour. It is not a photograph. The image the advertisers have created pushes the boundary of what we will accept as real and unreal by a careful manipulation of colour as a marker of modality. As a reader of the image I know it is not real but close enough to ‘reality’ for me to fantasize that the tomatoes in the tin are really as ripe and tasty as they look in the picture.

The above is an example of how modality has been described in a mode other than the verbal. What I hope to show is that modality is present in other modes of communication such as sound, posture and gaze. Van Leeuwen has written extensively on the modality of sound (1999, 2009, for example). Using the same reference to degrees of truth that we find in visual and verbal communication he proposes a list of articulatory parameters including pitch, variation, range, loudness, depth, voice quality and timbre (1999). He applies his ideas to music, television and the individual voice. Below is an extract from a transcript of an oral discussion (for original data see *Appendix 6: Classroom data CD ROM – Richard (2) 24.34*). I will discuss this in greater detail in *Chapter 5 Four extracts of data*, for now I just want to provide an example of how Van Leeuwen’s ideas can be used to explore how degrees of modality are articulated by the participants.

- 1 (3)
2 T: was this?=
3 N: =yeah was this?
4 N: (3)looks like a diagram.
5 T: (.)uh is this u:h >correct direction or rotation
6 N: (((laughter))
7 T: (.)o(h)r something<
8 T: (.)<you↑think?>
9 N: (1)I have no idea what that is.

The two speakers are discussing a picture from *Language Leader Intermediate* (2008) and making deductions. I have used transcription conventions²⁶ from Conversation Analysis (Ten Have, 1999) to indicate intonation and sound. What I want to focus on are the deictic markers *this* and *that*. Deictic markers such as *this/that*, *here/there* are a means signalling modality (Hodge and Kress, 1988) by

²⁵ See *Appendix 5* for the Cirio advertisement.

²⁶ See *Appendix 3* for transcription conventions.

referring to closeness and distance. Here I discuss the meaning potential of sound in *this* and *that* as markers of modality. Referring to frontality Van Leeuwen makes a distinction between ‘the dimension of front (for example [i] as in *heed*) vs back (for example [a] as in *hard*)’. He adds this ‘can give rise to metaphorical extensions related to the idea of distance’ (Van Leeuwen, 1999: 146) and gives examples from Dutch, German, English and French. If we look at the extract above, the two participants initially engage with the picture and ask ‘was this?’ (Lines 2 & 3). After a short discussion, however, containing pauses, hesitations, vague language, questions and uncertainty N says in Line 9 ‘I have no idea what that is’. I would argue that initially the participants engage with ‘this’ picture but on realising the apparent difficulty of deciphering it, the speakers closeness and engagement with the picture diminishes and it gets referred to as ‘that’. This is only a brief example but it shows how verbal language can combine with sound to express modality. I return to the meaning potential of sound in *Chapter 5 Four extracts of data*. It is worth adding that closeness and distance manifest themselves literally during this discussion with N and T physically positioning themselves closer to the picture in the initial stage of the discussion and moving away from it in Line 9. Indeed, N picks up the picture in Line 3 when she asks ‘was this?’ and pushes it away from her in Line 9.

To round up my research looks at epistemic modality i.e. the way speakers position themselves in relation to the truth values of what they are saying. Epistemic modality forms part of the interpersonal function of language and expresses social relations among participants. In my analysis of modality I will include modal auxiliaries, modal nouns, modal adjectives and adverbs, as well as modalising forms such as vague language, deixis, tense and questions. I have shown above how modality can also be conveyed in modes other than the verbal and I will draw on a range of modes in *Chapter 5 Four extracts of data*. In the next section I will discuss mode.

4.2 Mode

- S1 Excuse me, do you know how to get to Victoria Park from here?
S2 That’s it straight across there straight..across there see the gates?
S3 ^LThat’s Victoria Park over there dear look (<S1> Ah) ^Lsee that
(<S1> Yeap) see that car there (<S1> Yeap) look there’s a car go up
down there see the gates (<S1> Yeap) that’s the park
S2 ^LBy the car
S1 Ok thank you very much

I recorded the above dialogue for my MA dissertation (Howard, 2008: 67). At the time I was researching spoken and written dialogues as teaching resources. One of the threads of my research

looked at how shared situational information was conveyed using verbal language. For example, in the extract above the mental process verbs *look* and *see* along with the deictic markers *that* and *there*. After collecting, transcribing and analysing quite a number of similar dialogues, interesting questions were raised regarding the actual function of verbal language in these kinds of exchanges. It was beyond the scope of my MA dissertation to answer these, but on revisiting the dialogue above it is apparent that the verbal language in part has a phatic function. The park that S1 was looking for is indicated by S2 in the second line of the dialogue and by S3 in her first utterance; S3's longer contribution is arguably adding 'the frills', reducing social distance and signalling friendship.

More significant for my current research is how S2 and S3 draw on language, proximity, body position and gesture as a means of conveying the information S1 is looking for as well as having a social or interpersonal function. As I have already mentioned, at the time I was researching spoken language and so I only made voice recordings. I did, however, make field notes to complement the recordings and so as to illustrate the interaction I will detail these. S2 and S3 were an elderly couple walking away from the park that S1 was looking for. S2 was a man; S3 a woman. All three were walking along a narrow path. S1 and S2 would have passed each other shoulder to shoulder. On asking for directions all three came to a stop and faced each other. When S2 responded to S1 he turned in the direction of the park and pointed. At the same time S3 positioned herself closer to S1 so that the three participants made a small semi-circle facing in the direction of the park. During the rest of the exchange the park, the gates and the cars were all indicated using pointing.

Looking back over the dialogue now it is difficult to imagine the interaction without pointing being used alongside the imperatives 'look see that car', 'go up down there' and the deixis 'that car', 'that's the park'. Above I refer to the amount of phatic language particularly from S3. Arguably the shift in proxemics by S3 to reduce physical distance mirrors the way she reduces social distance through language. What this tells me is that in this dialogue language forms part of an ensemble of communicative resources for making meaning, whether this meaning is carried in the form of information giving or through interpersonal relations.

The range of communicative resources available for making meaning I will refer to as Modes. Modes can be understood as 'systems of representation' (Norris, 2004: 11). In the extract above I have identified the modes of spoken language, proxemics, gesture and body position; to these we could add sound, as in voice quality and timbre that I refer to in 4.1 *Epistemic modality*. These, as do all modes, display rules and regularities (ibid) that are culturally shaped by the people who use them (Jewitt, 2009: 21). The spoken mode has a system of rules and regularities such as syntax, grammar and intonation which is culturally shaped and thus intelligible to those who use or understand the

system. In the extract above I draw attention to the mode of proxemics. Proxemics is 'the ways in which individuals arrange and utilize their space' (Norris, 2004: 19). Again this system will have rules and regularities which are culturally shaped. Whilst individuals might not be aware of this system being available to them (Norris: 2004: 20) the participants utilize space as a means of signalling social relations.

So spoken language is one mode of communication and in an interaction such as the example of asking for and giving directions used above there are a number of modes functioning simultaneously to make meaning. Norris notes that when we observe an interaction the task of noticing all the communicative modes participants are using can be overwhelming (2004: 12); to the modes of spoken language, proxemics, gesture, body positioning, and sound already indicated, we could add facial expressions, head movements, and gaze, for example. In short there is a lot going on. In the light of such complexity we might be uncertain as what can realistically be classed as a mode. I have indicated above that spoken language and sound are modes. Both systems follow rules and regularities which are culturally shaped. But what about proxemics? Is proxemics a system of representation with rules and regularities? Kress (2009: 59) proposes asking three questions as a means of identifying a mode:

1. Can it represent what goes on in the real world – actions, states and events?
2. Can it represent social relations to those engaged in communication?
3. Can it represent these meanings which are internally coherent and which cohere with their environment?

He concludes that if the answer to all three questions is 'yes', then it is a mode. Addressing the three questions to spoken language and sound we can quite safely conclude they are indeed modes. If we think about proxemics then and how the three participants approached each other, how S1 might have asked his question at a distance he felt appropriate, how the physical positioning of the participants changed as the individuals took turns to speak and finally how the leave-taking might have taken place, we can see how in this kind of interaction meaning is made, how relations are shaped and how the actions cohere. Proxemics is clearly a mode.

As I have indicated above, in this interaction there is a complexity of modes happening simultaneously. And I have suggested that the function of language has as much a social role as one of information giving: arguably S2 and S3 could simply have pointed in the direction of the park in response to S1's question. Spoken language is one mode among many; indeed, it might not be the central mode of communication in this instance. It has been argued that language has been seen as the central channel of communication with other 'nonverbal' modes being subordinate to it (Kress et al., 2001; Norris, 2004). However, recent studies in multimodality claim that this is not always the case (e.g. Kress et

al., 2001; Norris, 2004; Van Leeuwen, 2005; Jewitt (ed.), 2009). Van Leeuwen, for example, notes that 'giving and demanding "goods and services" is more often realized non-verbally than giving and demanding information' (2005, 252). Modes then might be subordinate to one another with one or more mode taking a superior role. Or modes might play an integral part with each other. If we think about the similarities and differences in three types of interaction for asking for and giving directions, i.e. face-to-face, over the phone or face-to-face with the use of a map, we can think about how modes might be superior, subordinate or integrated with each other. As an example I would argue that the interaction I have presented above sees spoken language and gesture integrated to provide information, whilst spoken language, proxemics, body positioning, amongst others, are integral to the social function.

The relevance of this to my research is my understanding that classroom interaction and the making of meaning comes about through a variety of modes. Using an ELT activity from *Language Leader Intermediate* I have observed that learners do not often draw on the language focus of the lesson to carry out their discussions. However, far from doing nothing I argue that they are engaged in communication. What I intend to investigate is how that communication takes place. More specifically, as the activity is designed around making deductions, i.e. expressing certainty/uncertainty I am interested to know how this manifests itself. Hodge and Kress borrow a term from linguistics and apply it to all semiotic systems to claim that all 'utterances' bears signs of modality (1988). So I will be looking at how modality is expressed using different modes. The extract and pictures below provided an example (for original data see *Appendix 6: Classroom data CD ROM – Richard (2) 18.44*).

T: So what do you think about the first picture?

S: Signal transmission / something like that

(See *Transcript 3 Appendix 11* for full transcription.)

In this extract where "/" represents the boundary of a tone group, it appears that S is fairly certain what the design in the picture is. There is vague language in the second tone group indicating modality but 'signal transmission' is a clear attempt to make a guess at what she thinks it is. The utterance may well be an elided affirmative statement such as 'It's a signal transmission'. It is worth noting that this exchange is between a teacher and a student as feedback following the student discussing the picture with another student. S had already discussed and decided it was a 'signal transmission'. However, on speaking to the teacher, who S presumes has the 'correct answer', she appears to become less certain. Her intonation rises at the end of the second tone unit forming a question and thus having meaning potential of modality as we have already seen. The pictures overleaf in *Plate 1* further bear out her uncertainty. (S1 is on the left of the pictures. The sequence is from left to right.)

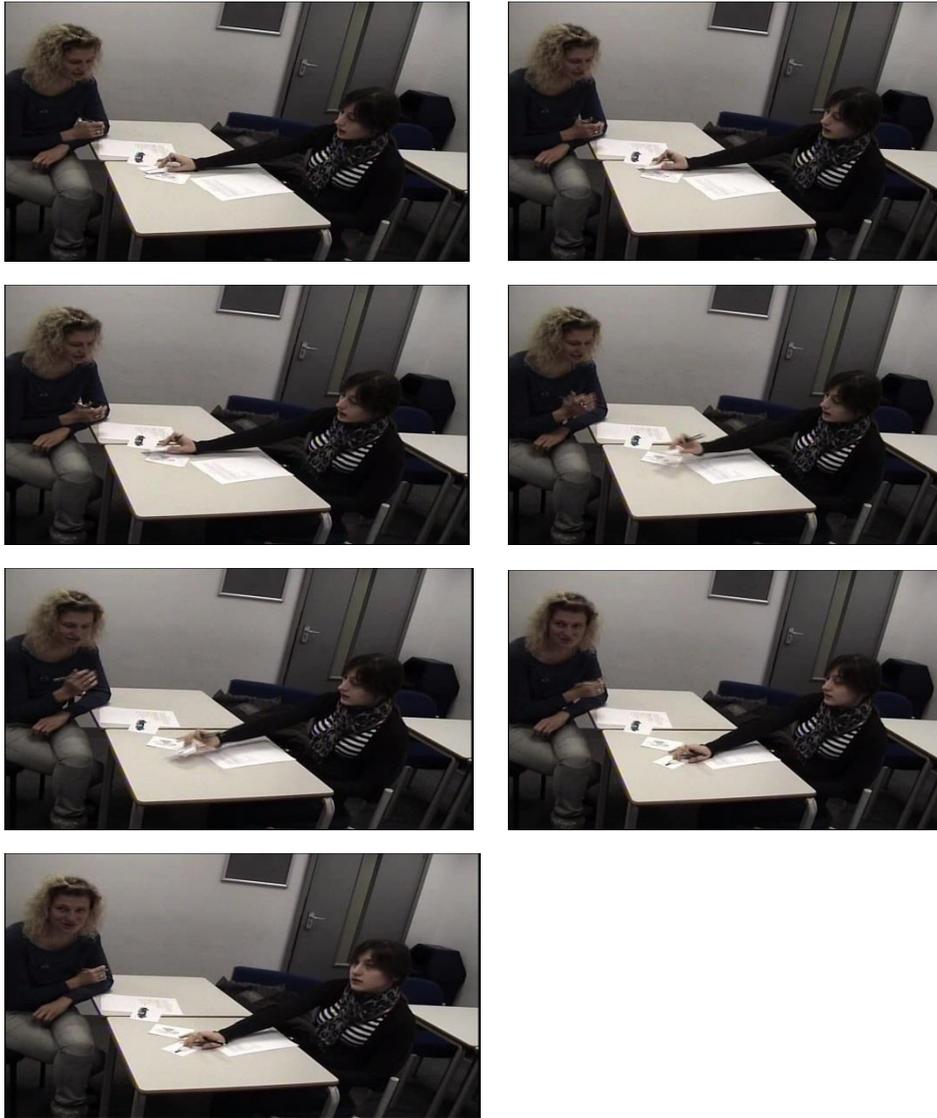


Plate 1

Before she attempts to speak she shakes her head from left to right. Whilst saying 'transmission signal' she slowly raises her right hand in a circular motion before letting it come to rest back on her left arm although now forming a somewhat more 'closed' position between her arms and body. All of these elements are 'synthesized into a single smooth action' (McNeill, 1985; 260). She engages T with eye contact when she says 'something like that' but at the same time shakes her head again. I will explore this in more detail in *Chapter 5 Four extracts of data*, but I argue that her head movements and gesture, along with her choice of language and intonation pattern express uncertainty. Her choice of modes and the way they integrate signal low modality. 'Signal transmission' could well then be an elided form of 'It could be a signal transmission'.

Before moving on it is important to make two final points. Firstly, as has already been pointed out, mode 'is a socially shaped and culturally given resource for making meaning' (Kress, 2009: 54). It cannot be assumed therefore that modes have shared meaning potential across cultures, sub-cultures

or even between individuals if they belong to different sub-cultures (Norris, 2004: 20). As systems of representation the modes of, say, pointing and gesture may well have similarities as well as differences in their rules and regularities in the same way as verbal language; meaning may not be shared. McNeill indicates that pointing has a standardized form 'within a given culture' (2005: 12) e.g. with one or two fingers, the full hand, or the face, lips or nose. In terms of the language classroom, in which learners from a variety of sociocultural backgrounds converge, the misunderstanding of modes is likely to occur.

Secondly, it is not my intention to make an inventory of signs within modes which carry modal functions. It is not the purpose of my thesis to ascribe precise meanings to certain phonemes, voice quality, proxemics etc. that can be coded in dictionaries or grammars: a head shaken from left to right, a hand raised in a circular movement, an elongated phoneme /m:::/ do not correspond to the grammatical modals *may*, *might*, *perhaps*, etc.; as Scollon and Scollon point out, '[t]he question of modality must necessarily be developed within each mode' (2009: 180).²⁷ However, I do argue for the potential of semiotic resources such as sound, posture, gesture and head movement to convey modality and I discuss this in more detail in *Chapter 5 Four extracts of data*. Nor am I suggesting correlations between modes in the way that has been done with kinesics in which paragraphs, sentences, words and phrases are said to be equivalent to postural positions, head and arm position, hand movements, facial expressions and gaze, (see Hinde, 1972). My concern is only with how meaning is conveyed in specific types of social context.

4.3 Multimodality

I have indicated above that in social interactions and communication meaning is made through a complexity of modes. And I have shown how the modes of gesture, pointing, proxemics, sound and language have meaning potential to express epistemic modality in spoken interaction. These interactions then can best be described as multimodal. An underlying assumption of multimodality is that 'representation and communication always draw on a multiplicity of modes, all of which have the potential to contribute equally to meaning' (Jewitt, 2009: 14). Traditionally, however, language has held a central role in the analysis of communication (Norris, 2004; Scollon and Scollon, 2009). Whilst multimodality does not set out to 'side-line language' (Jewitt, 2009: 2) proponents argue that viewing language as always central 'limits our understanding of complexity of interaction' (Norris, 2004: 2) and

²⁷ I am of course aware that certain signs carry conventionalised meanings within cultures such as moving the head horizontally back and forth to mean 'no' in many western cultures. My research focuses on semiotic resources in context and thus how participants 'tweak' these resources to convey meaning through interaction. For example, the vigour with which the horizontal movement of the head back and forth is carried out arguably grades the strength of commitment to that 'utterance'.

thus language, from a multimodal perspective, is seen as 'part of a multimodal ensemble' (Jewitt, 2009: 14).

In recent years there has been increase in interest in the role of other modes of communication in making meaning. Disciplines such as ethnography, anthropology, education, linguistics, and musicology to name a few have drawn on multimodal perspectives to research areas as diverse as children's drawing, gesture and movement in tourist spaces, visual design, IKEA tables and classroom interaction – the starting point for these being a social interpretation of mode to create meaning. Jewitt argues that whilst multimodal communication is not in itself 'new' (2009: 1), in that people have always used image and non-verbal modes to communicate (ibid), the technological means by which communication can be produced, accessed, circulated, and recorded is. This has practical implications in for example the way in which social interactions such as classroom activities can be recorded, transcribed and analysed. In addition, as Jewitt notes, 'different modes have differential effects for learning' (2009: 15) shaping knowledge and learner identities 'What can be done and thought with image, writing, or through actions differs in ways that are significant for learning' (ibid).

The title of my thesis is '*A Multimodal Perspective on Modality in the English Language Classroom*'. Using a multimodal framework I am able to explore how modality is communicated through a variety of modes. Video and sound recording allows me to revisit temporal communicative sequences that form multimodal ensembles. In the example above I am able to see how participants co-construct meaning using sound, gaze, gesture and language, for example. Furthermore I am able to analyse the modal configuration of interactions, i.e. the hierarchical relationships among modes in an interaction. So for example, whether language or gesture takes a more superior or subordinate role. My argument is that much modalising of commitment is carried out non-verbally through gestures, gaze, body positioning, etc. than through, say, modals verbs. A multimodal framework will allow me clearer insight into this. My findings I intend to feedback in to materials design for teaching epistemic modality. I suggest that by analysing multimodal data, learners are arguably able to 'see' how interactants shape meaning and the various communicative modes participants use to modalize commitment to their propositions.

Multimodality is said to refer to a 'field of application rather than a theory' (Bezemer and Jewitt, 2010: 180); multimodality offers a way of looking at the world. I have pointed out above that the starting point of multimodality is social interpretations of situated meanings. To varying degrees research traditions concerned with these aspects such as discourse analysis, interactional analysis, linguistic ethnography and social semiotics have taken an interest in multimodality. A multimodal framework then has given me a means through which to see the world and collate data. To make sense of that

world and data I draw on theories from systemic functional grammar and social semiotics. I will discuss these below.

4.4 Systemic functional grammar

Systemic functional grammar (SFG) is a strand of linguistics outlined by Halliday (e.g. 1985/1994). It is functional because it looks at how language functions; systemic because it offers a paradigmatic perspective of the options available to make meaning. SFG is concerned with relating language to meaning, function and context by investigating 'how people use language in real-world situations to achieve their social goals and purposes' (Cotton et al., 2008: 201). Take for example the utterance referred to above *You must be joking!* Cotton et al., (2008) tell us this structure is used to make deductions and has the same meaning as 'I am certain that you are joking' (ibid: 75). The description they offer tells us what the structure does. However, it does not tell us how it functions to make meaning in a specific context. It also gives us no understanding as to why this particular structure might be chosen over others e.g. 'I am certain that you are joking', 'You are joking!', 'You're not serious', etc. SFG explores the choices of language participants have to make meaning i.e. the meaning potential of language.

In *Chapter 3 Context: finding my way through the 'black hole'* I refer to the contextual variables of field, tenor and mode. Just to recap and summarize the three variables are field (the activity type), tenor (the social roles and relationships of the participants), and mode (how the text is produced and transmitted). So if we take the utterance *You must be joking!* and know that it forms part of a spoken conversation between a mother and child, we are in a better position to be able to make comments on its meaning and function and why this particular structure was chosen over alternatives. To illustrate further consider the following:

- A: Oh yeah! So you're gonna get rid of me early today.
- B: Yippee! (pause) No I just want
- A: Why mum?
- B: I just think you should have a good sleep.
- B: I'll go to bed the usual time.
- A: Eight o'clock.
- B: I don't wanna (pause) I'm not!
- A: You are!
- B: I said, I'm not!
- A: You can go up and get a bit of reading done in your bed at eight o'clock.
- B: Must be da-- **you must be joking!**
- A: I'm not joking, I'm serious!
- B: Tt! (banging)
- A: Wish I was going to my bed at eight o'clock. (banging)
- B: But mum
- A: Do that again
- B: yo-- will you leave us alone!
- A: You go to bed eight o'clock.
- B: (sigh)
- A: You need the rest.
- B: (screaming) I couldn't give a blooming damn!
- A: I don't care. You want yo-- (pause) the rest.
- B: (crying) But I'm not!
- A: You are. Now do²⁸

As already mentioned above A and B are a mother and child. A is attempting to get her child to go to bed early. B does not want to do this and appeals. The mother attempts to reason with the child but the child loses his temper. The child's utterance *You must be joking!* can thus be interpreted as a rejection or rebuttal of his mother's wishes. In this context the function of *You must be joking!* is not really one of making a deduction so the child's range of choices would not really include 'I'm certain

²⁸ For the original dialogue taken from the BYU-BNC corpus see *Appendix 4*.

that you are joking'.²⁹ The meaning of B's utterance in this context is more likely to be something similar to 'No, I am not going to bed at eight o'clock!'

In SFG then, language is seen as a resource to make meaning. With meaning taking place on three simultaneous levels: the ideational, the interpersonal and the textual. These three levels of meaning are referred to as metafunctions. The metafunctions map on to the three contextual variables of field, tenor and mode that I have already outlined in 3.4 *Context of situation revisited: field, tenor and mode*. This can be seen in the diagram below (*Figure 4*) from Eggins and Slade (1997: 51).

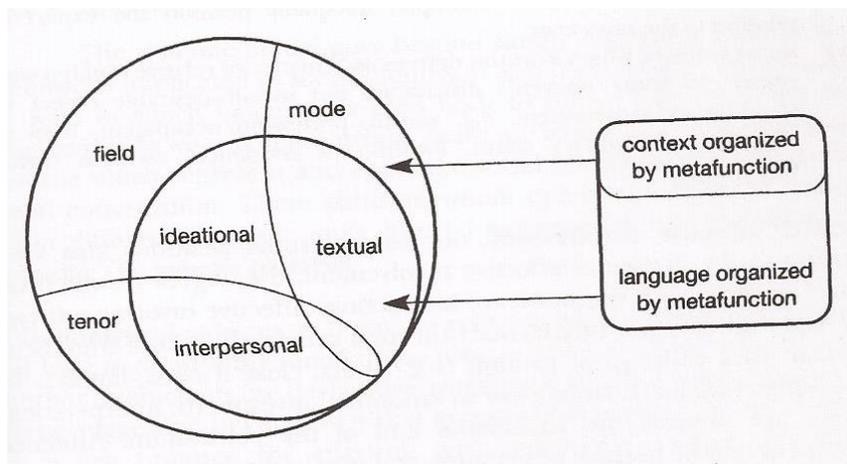


Figure 4

In this diagram we can see within the context organized metafunctions of field, tenor and mode there are the three language oriented metafunctions: the ideational, the interpersonal and the textual. To draw this together we can say that **field** 'refers to what is happening, to the nature of the social action that is taking place' (Martin and Rose, 2003/2011: 297) and is realised through the **ideational** metafunction i.e. 'language as expressing the speaker's experience of the external world, and his own internal world, that of his own consciousness' (Halliday, 1978: 45); **tenor** 'refers to who is taking part, to the nature of the participants' (Martin and Rose, 2003/2011: 297) and is realised through the **interpersonal** metafunction, i.e. language for 'expressing relations among participants in the situation, and the speaker's own intrusion into it' (Halliday, 1978: 46); and **mode** referring to 'what part language is playing, what it is that the participants are expecting language to do for them in the situation' (Martin and Rose, 2003/2011: 297) realised through the **textual** metafunction, i.e. 'its structure as a message in relation to the total communication process' (Halliday, 1978: 46).

²⁹ It is interesting to note that the child makes a false start 'Must be da – you must be joking!' leading to speculation as to what alternative the child might have been considering.

In summary the three language oriented metafunctions can be explained as follows:

the ideational metafunction used to talk about and represent the world;

the interpersonal metafunction used to engage interpersonally;

the textual metafunction used to create cohesive text.

In SFG these metafunctions are realised through different features of the lexicogrammatical system. The ideational through verbs (referring to 'processes' encompassing actions, states and processes proper), complements (referring to participants in the verbal process), adjuncts (referring to the circumstances of the verbal process); the interpersonal through mood, modality, lexis, formality of language; the textual through cohesive markers, textual reference, clauses, lexical density, nominalisation, interruptions, repetitions, etc.

What this enables the analyst to do is to take an utterance, for example, *You must be joking!* and by looking at its lexicogrammatical structures explore how meaning operates on these three levels simultaneously. For example, the subject pronoun 'you', the auxiliary verb 'be' and present participle 'joking' form the ideational meaning and tell us something about how the speaker 'sees the world' i.e. how he talks about and makes reference to a past event. The declarative mood, high modality and informal language of the interpersonal function tell us something about the status, intimacy and relations of the participants. And finally the low lexical density, simple clause structure, nominalisation and anaphoric reference through the subject pronoun 'you' of the textual function tell us that it is quite likely to be a spoken response to a question. If we add to this what we know about the contextual variables of field, tenor and mode e.g. the kind of activity and the topic, social roles and distance, interactivity and degree of spontaneity, then we are in a much better position to make comments on the meaning, function and the choice of utterance.

All interaction contains an interpersonal dimension, although in some kinds of interaction the interpersonal dimension is particularly evident (Coffin et al., 2009; Eggins and Slade, 1997). My data has been collected from a speaking activity in a language classroom in which learners are engaged in a discussion. Arguably in such an activity interpersonal relations are very much of importance. In a discussion participants are aware of social identities and relations and this manifests itself through choice of language and other communicative modes. My focus is on modality, part of the interpersonal metafunction. By using systemic functional grammar I will be able to investigate the choices of modality used in the data in relation to the participants and the social context.

4.5 Social semiotics

Systemic functional grammar was developed by Halliday through his theories of language as social semiotic (1978). As we have seen in SFG the social is very much at the centre of linguistic analysis. In this section I want to show how the dimensions of the ideational, interpersonal and textual metafunctions in SFG have been adopted by the field of social semiotics and applied to other semiotic resources (modes).

Social semiotics is concerned with 'the social interpretation of language and its meanings to the whole range of modes of representation and communication employed in a culture' (Bezemer and Jewitt, 2010: 183). Halliday's ideas of socially situated language have been seen as the starting point for social semiotic theory (Jewitt, 2009). Social semioticians such as Hodge and Kress (1988) and Kress and Van Leeuwen (2006), for example, have drawn extensively from Halliday's ideas and social semiotics has developed around a theory that all communicative modes are built around a system of choices in the same way that SFG views language as a system of choices. Social semiotics also holds that the meaning of signs within modes are socially and culturally determined as well as sensitive to the context in which it is used. Just as language has meaning potential, so too do other communicative modes. Above I have shown how the modes of gesture, sound, proxemics, language, and visual communication are able to convey meaning. I have suggested that in some contexts they have the potential to express modality; although they do not necessarily equate to modal expressions per se.

Social semiotic multimodality takes up Halliday's metafunctions and applies them to other modes of communication (semiotic resources). Kress and Van Leeuwen detail how the metafunctions can be used in visual design (2006). They note '[l]ike linguistic structures, visual structures point to particular interpretations of experience and forms of social interaction' (ibid: 2) and explain how colour can convey the three metafunctions (ibid: 229). The ideational 'can be used to denote people, places and things as well as classes of people, places and things and more general ideas' (ibid) and they discuss as examples the colours of flags and colours on maps. Colour can also convey the interpersonal 'and is used to *do* things to or for each other' (ibid, emphasis in original) i.e. to impress, to intimidate, to subdue. Colour can also function textually in the way that colour schemes create cohesion in interior design or in text books. Stenglin has used the metafunctions in a discussion on space and communication in exhibitions (2009). She notes the ideational function of space enables classification according to function, i.e. a kitchen, a bathroom, a bedroom. The interpersonal is able to convey relationships between space and its occupants and lists two main categories 'security and insecurity' (ibid: 278) and the textual is seen in the way series of spaces are organised into meaningful wholes.

Above I have shown how the ideational, interpersonal and textual metafunctions are realised through language; here I will give my own example to show how the metafunctions can apply to gesture. Norris describes gesture as a 'deliberately expressive movement' (2004: 28) and sets out four categories: iconic, metaphoric, deictic and beat gestures. For the purpose of the discussion here I will refer to metaphoric gestures. Norris defines metaphoric gestures as depictions of pictorial content as well as abstract ideas and categories: 'abstract notions are given form and shape in the imagery portrayed in the motion and the space of the gesture' (ibid: 29). Earlier using the screen shots of two students engaged in a discussion to deduce what the design in a picture might be I made reference to a hand gesture: a circular movement of the hand. I will elaborate on this in *Chapter 5 Four extracts of data*, here I want to provide some indication as to how this gesture might convey meaning across the three metafunctions.

If we take the gesture as a conduit metaphor (McNeill, 2005) meaning something like 'This is my idea', the open upturned hand presents the idea to the discussion group and conveys ideational meaning. The gesture is loose and relaxed, it does not cut across anyone's space and does not appear threatening to the other participants. The way the gesture is carried out in respect to the contextual variables of tenor conveys its interpersonal function. Norris (2004) describes gestures as consisting of one or more phase i.e. preparation, pre-stroke hold, stroke, post-stroke hold, and retraction (p.30), thus the way the gesture links internally through the preparation, the stroke and the retraction and the way it links 'externally' with S1's other modes of communication e.g. her head movement and verbal language represents its textual function; it also links textually in a metadiscursive sense in that it marks the boundaries of her utterances.

However, a word of caution is called for when analysing modes in this way. Jewitt points out that '[w]hen several modes are involved in a communicative event all of the modes combine to represent a message's meaning. The meaning of any message is, however, distributed across all of these modes and not necessarily evenly' (Jewitt, 2009: 25). I have shown earlier how the deictic gesture of pointing is integrated with spoken language when giving directions. It is possible that the ideational and textual meanings are conveyed in this way, whereas the spoken language conveys ideational, textual as well as interpersonal meaning. Furthermore we need to be cautious when assigning meanings to modes (Norris, 2004). Norris and McNeill both assert that speech and gesture co-occur and 'combine into a system of their own in which each modality performs its own functions' (McNeill, 2005: 9). The gesture in the example above may well have more of an interpersonal function, arguably modalising the commitment of her spoken utterance.

We have seen in SFG how meaning and context are bound and thus the meaning of any mode is really only deducible from its context. The meanings I have assigned to gesture in this example have come about through an analysis of the interaction as a whole.

4.6 Context

So far in my thesis I have already written extensively on context. In this section I want to detail how I am using context in my data analysis. Firstly I want to recap on context as dynamic and then provide an explanation as to how I will focus my research lens to avoid the 'black hole' of context.

I have stepped away from the idea of context as static, as an ornamental backdrop for the presentation of language to a view of language as interactive with context; context as dynamic. Central to this idea is Halliday's view that 'context plays a part in what we say; and what we say plays a part in determining the context' (1978: 3), which I will expand on here. The research traditions that I am drawing from and have referred to above see language and meaning as being integral with sociocultural context. As Halliday notes, language is 'the exchange of meanings in interpersonal contexts of one kind or another' (Halliday, 1978: 2). In systemic functional grammar sociocultural contexts are viewed in terms of the contextual variables of field, tenor and mode outlined above. These variables take into account the kind of activity participants are engaged in, their social roles and relationships and the means through which this activity is carried out. For example, a classroom discussion involving two or three English language learners using spoken language. Given these contextual variables meaning is realised through three language metafunctions referred to as the ideational, interpersonal and textual and it is these metafunctions that shape and are shaped by the language choices participants engaged in an interaction make. So for example a classroom discussion might form a spontaneous exchange of ideas in which questions are asked, statements made and notions challenged, the choice of which is determined by the intimacy or otherwise of those engaged. However, at any point the discussion might descend into an argument or a conversation on an unrelated subject and bring about a shift in the context and thus the field of activity, the roles and relationships of the participants and their choice of language. It is through the dynamic interaction of these variables and metafunctions that language functions and meaning is made.

Earlier I made reference to context being described as a 'black hole' (McCarthy, 1998) and as being 'potentially infinite' (Blommaert, 2005). Whilst these descriptions are helpful in providing some idea of the complexity of context in terms of what can be included in it, for the purpose of my research I need to focus my research lens a little more. I am not attempting to define context and my thesis is not an exhaustive exploration of the elements of context that affect meaning. What I am attempting

to do is investigate how a selection of modes have the potential to make meaning within a specific social activity. And to do that I need some way of framing that activity as well as a means of looking at how communicative modes function individually and collectively.

In my data the learners are carrying out a speaking activity from *Language Leader Intermediate* referred to in the rubric as a discussion. A discussion is a social activity. This particular discussion forms part of an ESA (engage, study, activate) lesson on modals of deduction. A discussion could be analysed as a speech genre or speech event i.e. relatively stable types of communication governed by rules or norms (Eggins and Slade, 1997) as you might see, for example, in conversation analysis. This would allow me to analyse generic stages, language functions, turn taking, etc. as well as lexicogrammatical features, for example. However, because of their primary focus on verbal language as the unit of analysis this is not a suitable description for my research. To allow for other communicative modes in my analysis I will adopt an integrative multimodal approach in which action is the central unit of analysis (Norris, 2004 & 2009).

Norris subdivides actions into higher-level and lower-level actions. A discussion is an example of a higher-level action. This higher-level action is made up of lower-level actions such as utterances, gestures and gaze. However, it is more complex than it first appears. Norris defines lower-level actions as 'the smallest meaning unit in which a social actor draws upon a communicative mode such as gesture, posture, spoken language or layout, and constructs meaning' (2009: 82) and so a discussion is constructed via chains of lower-level actions such as sounds, intonation units, clauses, utterances and turns to form spoken language, and finger, hand and arm movements to form gestures, etc. Added to this higher-level actions are in turn embedded within higher-level actions. For example, a discussion within a higher-level action of a lesson within a higher-level action of a course syllabus. And, of course, lower- and higher-level actions co-occur with other lower- and higher-level actions: speech, gesture and proxemics co-occur in a discussion, for example.

As complex as this seems it does allow modes to be analysed and discussed in ways that are manageable. Earlier I mentioned modal configurations. This refers to the hierarchical relationships between modes that construct higher-level actions (Jewitt, 2009). A face-to-face discussion is composed of numerous chains of lower-level actions again making the task of analysis overwhelming. And whilst these chains of actions comprise modes that make meaning my research is an investigation into those modes that arguably carry modality within a specific higher-level action/social activity.

4.7 Activities

My research focuses on a speaking activity from a published ELT course book. The activity is the last stage of an ESA (engage, study, activate) lesson on modals of present deduction. I have provided an overview on how an ESA lesson can be structured in *1.1 Becoming a teacher; becoming a researcher* and in *2.2.2 Language Leader Intermediate* above, shown how it is similar to PPP (presentation, practice, produce) and pointed out that it is a lesson staging model favoured in many ELT course books. Here I want to focus on the term activity. I will explain what is meant by an activity in ELT, compare this with how it is used in *Language Leader Intermediate* and contrast activity with task.

Richards and Rodgers indicate that '[p]ractice activities should involve meaningful learning and language use' (2001: 66). Whilst this might seem rather an obvious statement these days in the western approach to ELT, it has not always been the case. It wasn't until the early days of communicative language teaching (CLT) in the 1960s and 70s where the idea really began to take shape. CLT was in part influenced by the work of Austin *How Language Works* (1962), Halliday *Language as Social Semiotic*, (1978), and Hymes paper *On Communicative Competence* (1972). These theorists described how language is used and this had an impact on language teaching. With CLT the focus was on using language communicatively, i.e. to communicate effectively in everyday situations and there was a movement away from two-part lesson structures in which language was firstly presented then followed by tightly structured pattern practice (Howatt, 2004).

CLT also favoured a process approach to language learning (Brumfit, 1979). Using the language became part of the learning process. The use of role plays, information gaps, games, jigsaws and pair work were seen as meaningful activities in which using the language facilitated learning. Activities promoting interaction and co-operation became a growing trend in CLT in the 90s (Jacobs and Ball, 1996) during which time the work of educationalists such as Piaget and Vygotsky became influential (ibid). Learning, communication and meaning developing through the process of interaction and co-operation.

Activities now abound in ELT. A Google search for ELT activities, for example, shows 3,770,000 hits (25/06/14). There are mix and match activities, quizzes, word wheels, crosswords, games, reviews, reading, writing, speaking, listening, critical thinking activities and so on. *Language Leader Intermediate* too is full of activities such as discussions, presentations, interviews, debates, using a dictionary, planning, negotiating, deduction and speculation activities; in terms of speaking they claim there is 'at least one explicit speaking activity per lesson' (Albery, 2008: 5). The use of such activities matches with the writers' claims that the book uses a communicative approach and offers practice in 'real-life situations' (ibid: 3). For the purpose of my research and discussion then I am referring to the

following discussion as an activity (see *Figure 5* below reproduced from 2.3.1 *Problematizing a lesson on modal verbs*). It is communicative, it is interactive, and it focuses on language use.

6 Work with a partner to discuss what you think the designs below are. Try to use these words and phrases.

must might could can't
I'm sure/certain ... maybe/perhaps
It's possible that ...
It's not possible that ...



The image contains three distinct visual elements. The top-left is a photograph of a person's hands working on an open computer tower case, with a monitor displaying a green waveform. The top-right is a black and white photograph of a woman in a black dress and gloves, looking upwards. The bottom-center is a digital illustration of a person in a red shirt interacting with a large, glowing, multi-colored mesh structure that resembles a complex network or a futuristic interface.

Figure 5 (Cotton et al., 2008: 75)

Why this is of importance to my research I will now explain. Activities have been described as the forerunners of tasks (Howatt, 2004). Tasks and activities though are significantly different when it comes to teaching and learning. A task in relation to task-based instruction is designed to 'encourage naturalistic acquisitional mechanisms, cause the underlying interlanguage system to be stretched, and drive development forward' (Skehan, 1998: 95). As a definition of tasks in task-based instruction, Skehan proposes the following:

- Meaning is primary;
- There is some communication problem to solve;
- There is some sort of relationship to comparable real-world activities;
- Task completion has some priority;
- The assessment of the task is in terms of (non-linguistic) outcomes.

He goes on to point out that there are major differences between tasks and conventional language activities 'in which there is a much greater focus on structure' (Skehan, 1998: 96). He points out tasks:

- Do not give learners other people's meanings to regurgitate;
- Are not concerned with language display;
- Are not conformity-oriented;
- Are not practice-oriented;
- Do not embed language into materials so that specific language structures can be focused on.³⁰

As well as using activities *Language Leader Intermediate* claims to use a task-based approach (Albery, 2008: 3). Arguably though the 'activity' from *Language Leader Intermediate* above conforms more to what a task is not than to what a task is. It does give learners other people's meanings to regurgitate, it is concerned with language display, it is conformity and practice-oriented and it does have embedded language specifically to focus on structures.

Now this is not a problem in itself. There is nothing wrong with activities and I am not criticising the ESA model per se. Where there is a problem, however, is in the discrepancy between task and activity. A discussion is arguably a task; yet the writers of the course book have designed the activity to practise linguistic structures. I have already indicated the potential pitfalls of this when it comes to putting the activity into practice so I will not reiterate that here. I make reference only to define my terminology. The activity presented above is an activity designed, in the main, to practise language structures.

³⁰ This is not to suggest that task-based instruction is based on an absence of focus on form. See Willis, 1996; Littlewood, 2004 and Willis and Willis, 2010 for discussions on focus on form in task-based learning and 7.1.2 *Task-based instruction* below in which I propose a task-based model which integrates a language focus.

Chapter 5 Four extracts of data

5.1 Lead in

In this chapter I will discuss four extracts of students carrying out a speaking activity. During the activity the students were discussing what the design in Picture A is (*Figure 6* below). Picture A is taken from *Language Leader Intermediate* (Cotton et al., 2008) from the lesson on modals of present deduction. I will provide more information on the participants on introducing each extract.

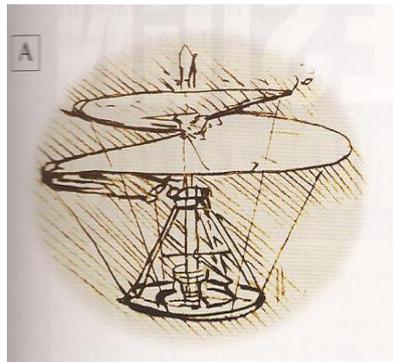


Figure 6 Picture A (Cotton et al., 2008: 75)

In the extracts Picture A is given as a means of activating modal structures that have been taught to the students. However, I have taught extensively with these materials and from my observations the suggested structures do not often occur. I have discussed this above and argue that for cognitive, social and procedural reasons the linguistic structures are unlikely to be activated. What I am interested to discover, therefore, is what modal resources speakers use when talking about Picture A.

I have limited my research to discussions about Picture A. This has been done for methodological as well as practical reasons. Using social semiotic multimodality I am taking an inductive approach to discuss how patterns and structures of modality occur in a small sample of data. I do not claim that my observations are exhaustive and the semiotic modes I have chosen to discuss do not in any way encapsulate the full complexity of multimodal interaction. What I hope to show is that modality markers are realised through a number of linguistic structures and occur in modes other than the verbal. Multimodal analysis posits that meaning is conveyed through the integration of modes in interaction. However, whilst acknowledging this, I have, for the purpose of my focus on interpersonal meaning and modality, been selective of the modes within each interaction I draw on for discussion. Further it is beyond the scope and indeed the purpose of this thesis to provide a detailed inventory of the full range of modal resources available to participants in spoken interaction.

Picture A affords continuity to my data. It is one of six images from the course book (see *Appendix 2: Language Leader Intermediate Lesson*). It is a sketch of Leonardo da Vinci's flying machine. I argue that each of the pictures presents learners with different cognitive and communicative challenges. Skehan (1998) uses a three-way distinction to discuss the challenges of tasks. These are code complexity, cognitive complexity and communicative stress. I apply these to the discussion activity. For example, my data and field notes tell me that the following picture (*Figure 7*) presents a different challenge to learners than Picture A.



Figure 7 (Cotton et al., 2008: 75)

As I have previously mentioned the activity often turns into a guessing game rather than a discussion about what the design might be; also the images tend to promote description rather than discussion. Responses to the above picture are often 'It's a car', 'It's a fast car', 'It's an old car', 'I think it's expensive' (see, for example, *Appendix 11, Transcript G*). Arguably what the design is in this picture is more opaque than Picture A (see 4.1 *Epistemic modality* and *Extract 4 Changing roles: contextual variables and communicative resources* for commentary on 'truthful' images of reality).

The following pictures (*Figure 8* overleaf reproduced from 2.3.1 *Problematizing a lesson on modal verbs*) too present their own challenges. They activated some weird and wonderful answers requiring 'good imagination', as one student put it, but little discussion using modals of deduction. During feedback a frequent and interesting comment was that there is no context for the images, thus making them more difficult to discuss. Picture A yielded some interesting data and for the reasons I have given here seemed to me a sensible choice for analysis.



Figure 8 (Cotton et al., 2008: 75)

5.2 Methodology

My research questions arose out of my classroom observations as an English language teacher. I was interested to observe the communicative resources learners used to convey epistemic modality during a speaking activity. With this in mind collected ethnographic data during lessons using video-recording equipment (see 5.2.2 *Classroom ethnography*). Groups of learners were recorded carrying out a speaking activity from *Language Leader Intermediate* intended to activate a number of modal structures (see *Appendix 2 Exercise 6*). All the participants in my research were aware that they were being video recorded and that the recordings would be used for my research purposes only. Consent forms (see *Appendix 7*) were given out and signed by all participants. The data was stored on the recording equipment before being transferred to a CD ROM and then uploaded onto my computer using the software package ELAN, which I downloaded from the Max Planck Institute for Psycholinguistics (the CD ROM containing all of the data included in my thesis is included in *Appendix 6*). As with all (multimodal) communication, there are numerous modes in play at all times through which meaning is made. Before transcribing the data I therefore needed to make a decision as to which modes I would focus on; this was done not only for the practical reasons of making the transcriptions readable but also for methodological reasons. Whilst viewing the data it became apparent that certain modes came to the fore and within the particular interactions had a more dominant role in conveying meaning. I was, therefore, selective on which modes to observe and transcribe. The modes that have become the focus of my discussion are spoken language, sound, gesture (including pointing and beats) as well as posture (including head movement). As already mentioned, these modes seemed to carry the most meaning in the interactions and this can be

attributed to a number of reasons. Firstly, spoken English is the main focus of the language lessons I observed. The students had come to the class to learn and use English: the activity they were engaged in was a speaking activity. However, there was more to this choice than it being simply what the learners were supposed to be doing. Ong (1982) writes on the primacy of speech and whilst acknowledging the countless ways humans communicate using all their senses and the richness of some modes of 'non oral communication' such as gesture, he contends that 'in a deep sense language, articulated sound, is paramount' (p.7). In a similar vein Norris notes that spoken language has a 'high information value' (2004: 66). I am taking spoken language, then, not only as it was the 'main' communicative resource of the activity, but arguably a central mode of communication in social interaction. Furthermore in regard to other communicative modes, Norris notes that language does take a central role with modes such as gesture and gaze often being 'subordinate to spoken language' (2004: 53). As a point of departure, then, my analysis focuses on spoken language and then works outwards, so to speak, to the other communicative modes of my research.

In *Extract 1* I look at the meaning potential of sound. The quote from Ong above highlights the centrality of sound in spoken communication, indeed spoken language is 'the world of sound' (Ong, 1982: 6) and I was therefore interested to observe *how* sound conveys meaning. Gesture seemed to be a much used mode across the data. In extracts 2, 3 and 4 I bring in the modes of gesture, gaze and posture. Gesture is described as being 'exceedingly rich' (Ong, 1982: 6) and there exists a large amount of literature devoted to this mode of communication. This is perhaps not surprising as gesture is often interdependent with spoken language (Norris, 2004). Similarly gaze is a mode of communication occurring simultaneously with spoken language. Gaze has been described as 'a fundamental resource in the initiation, or inhibition, of interaction' (Kidwell, 2014: 1326). Body posture, including head movement, again seems to be intricately bound to spoken language and as Bohle points out 'researchers agree that bodily coordination is an indispensable prerequisite for any verbal exchange' (2014: 1302).

With these communicative resources in mind I began the process of transcribing the data. In the next section (*5.2.1 Transcription*) I describe in more detail how I went about this. I point out once again that my research focus is on epistemic modality and this has been a primary aim for interpreting my data: I am observing how communicative modes have the potential to communicate modality. I am aware that communicative modes integrate to form modal ensembles (Jewitt, 2009) (see *5.4 Extract 2 Modal ensembles: interpersonal meaning in multimodal communication*) as '[i]ndividuals in interaction usually do not utilize one mode completely separately' (Norris, 2004: 53), thus my analysis has been with this in mind. Meaning is derived from the integration of modes (Norris, 2004) and therefore I have been mindful not to assign meaning to modes in isolation from one another. However,

for the purpose of analysing interpersonal meaning and modality, I have limited my discussions to selected modes within each extract. There are, of course, practical reasons for this too as my analysis has, of necessity, had to stop somewhere for the reasons I have discussed above pertaining to the 'black hole' of context.

The multimodal approach taken in this thesis draws attention to the range of communicative resources or modes people use to make meaning. To handle the data analytically I take a social semiotic approach. This approach builds on Halliday's (1978) ideas of language as social semiotic. Within social semiotics communication and representation are viewed as taking place across a range of modes, each with its own cultural, social and historical use to realize particular social functions (see 4.5 *Social Semiotics* above for a fuller description). In this chapter I show how interpersonal meaning and modality are communicated using the modes of spoken language, sound, gesture and posture. I draw on Halliday's concept of speech functions, take a social semiotic approach to an analysis of sound and use the notion of the ideational, textual and in particular interpersonal metafunctions to show how the modes referred to above are used to communicate modality and thus the social roles, affinity and authority of the participants in four classroom interactions. This is explained in more detail over the following pages.

5.2.1 Transcription

The data I collected was video recorded. I would like to have incorporated these video recordings into my thesis using a multimedia format, but this has not been possible due to the nature of the this paper-bound thesis. A CD with the original data is included in *Appendix 6*.³¹ Transcribing data is not only necessary because of the technical restraints of including moving visual images in a paper-based thesis, but because the re-making of video into a transcript, multimodal or otherwise, has allowed me to observe, focus on and analyse communication in order to gain insights into my research aims. As with all transcription there are elements which are lost, but 'accuracy' of a transcript is not intended as a replica of reality; as Bezemer and Mavers (2011) point out, transcription is a transduced and edited analytical tool providing a particular 'professional vision' (Goodwin, 1994 in Bezemer and Mavers, 2011).

Social interaction comprises ensembles of modes (Jewitt, 2009). In this chapter I focus on specific modes within extracts of data and discuss how they integrate to make interpersonal meaning and express epistemic modality. The modes I have selected for discussion purposes I refer to as lower-

³¹ Included on the CD are all the data from which the transcriptions contained in this thesis were made. The CD also contains further samples of discourse in which the different groups of learners are engaged in the same *Language Leader Intermediate* activity that forms the basis of this research. There are approximately 125 minutes of data on the CD.

level actions, a term I have borrowed from Norris (2004). To do this I do not transcribe ‘everything’ that is going on during the interactions, only the necessary amount of detail to allow a particular analysis (see above 5.2 *Methodology*). This is, of course, an important consideration and just how much detail to include reflects the ‘research aims and theoretical underpinnings’ (Flewitt et al., 2010: 52) of my research. I am primarily interested to observe how given modes have the potential to signal modality, but I bear in mind that during these interactions other communicative modes are present and I am aware of their potential to convey meaning. My analysis focuses on configurations of spoken language, sound, gesture and posture and this will be evidenced below across the data extracts. What I offer is a sample and I add that ‘[t]he task of a multimodal transcript is not to analyse the images that are incorporated, but rather to use the images to describe the dynamic unfolding of specific moments in time’ (Norris, 2004: 65).

The recordings were transferred to CD ROM and then uploaded to ELAN, a software package downloaded from the Max Planck Institute for Psycholinguistics. This software allows ‘the creation of complex annotations on video and audio resources’ (<http://tla.mpi.nl/tools/tla-tools/elan/> accessed July 2014). Annotations can range from a single word to multiple tiers of text. Again, as I point out above, the inclusion of video recordings has not been possible and so such technical annotation is not included in this thesis. What ELAN has allowed me to do is slow down and speed up visual as well as audio and capture actions in 100ths of seconds the ‘fleeting details of interaction’ (Kidwell, 2014: 1325) and thus provide a detailed focus and to approach my data using a ‘micro-analytic time frame’ (Flewitt et al., 2010: 53). This has proved particularly helpful when analysing modes such as gesture and eye gaze; it is also an invaluable tool for analysing synchronized actions.

After uploading the data onto the software and viewing it, I made numerous notes for each interaction and then transcribed each one ‘by hand’. As there are no strict guidelines for multimodal transcription and a diversity of transcribing conventions exists amongst researchers (Flewitt et al., 2009; Litosseliti, 2010; Bezemer and Mavers, 2011) I have used different formats. This has proved useful as different formats have allowed me to focus on different aspects of interaction. For example, I have used conventions from conversation analysis (see *Appendix 3*) in *Extract 1* and *Extract 4* to analyse speech functions, linguistic structures of modality and the meaning potential of sound. Although it would have been possible (if rather inaccessible in terms of readability) to transcribe all the data using language (see, for example, Flewitt et al., 2009), I have, in the subsequent extracts, relied significantly on the use of images to provide richer detail and thus allow the visualisation of the participants’ lower-level actions during complex interactions. The tabular format in *Extract 2* allows for a detailed verbal description of gesture, speech, head movement and gaze illustrated with photo captures plotted

against timings in seconds; and the sequences of photo captures in *Extract 3* and *Extract 4* focus on the integration of the modes of gesture, gaze and posture with spoken language.

I have tried to include as much annotation as I feel is needed for my particular analysis without making the transcripts overcomplicated and difficult to 'read'. The process of transcription was carried out with my research aims in mind and this is reflected in how I have transcribed the data. I point out that 'image does not have words' (Bezemer and Mavers, 2011: 196), and describing, for example, a gesture is not only complicated but may take many words; thus describing actions has been less than straightforward. Again I note that transcription involves *transduction* from one mode to another. Whilst this changes 'the material structure and dynamism of how meanings are expressed' (Flewitt et al., 2009: 52), it opens up 'transmodal moments' (ibid) allowing the analyst to see data differently. Finally I draw attention to the representational choices (Bezemer and Mavers, 2011) involved in the framing, selecting, highlighting and analysis of transcription. Transcription is 'semiotic work' (Kress, 2010 in Bezemer and Mavers, 2011) and is carried out with the 'purpose of gaining analytical insights and developing theoretical arguments' (Bezemer and Mavers, 2011).

Finally, unless otherwise stated the transcriptions in this thesis form separate communicative episodes centred around one of the pictures from the *language Leader Intermediate* activity (see *Appendix 2* for the complete set of pictures). Transcribed snippets of discussions are indicated in the thesis and the reader directed to *Appendix 11* for the complete transcription.

5.2.2 Classroom ethnography

I became interested in the subject of my thesis through my work as an English language teacher. I noticed that modal verbs were problematic for learners to use in extended spoken discourse. In my experience whilst the rules of form and meaning seemed to be readily grasped, and this was evident in practice exercises such as gap-fills, when it came to freer practice activities such as speaking, modal verbs were typically absent. This raised my curiosity and I was interested to investigate why learners did not use them in spoken discourse despite 'knowing' their form and meaning.

I did not want to approach my research as a 'researcher' or a classroom observer, i.e. as an 'outsider'. I had done systematic classroom observation as part of the Dip TESOL using observation schemata to record quantitative data such as error correction of pronunciation, but I was aware of what has been termed 'the observer's effect' (Blommaert and Jie, 2010: 27), that is teachers and learners adjusting themselves to the presence of the observer in the classroom. At the time I felt my findings were not as 'authentic' as they might have been considering all involved knew that observation was taking place: my presence in the classroom was short term, intrusive and however quiet and diminutive I tried to be, my being there proved disruptive. Accordingly I was sensitive to this in my current research and

wanted to stay very much as the teacher and observe lessons that I was directly involved in. I wanted to see what was happening during the lessons by getting as close as I could to an authentic, everyday teaching and learning experience: I wanted direct observation through being immersed in the 'field' (Gordon et al., 2001). And I wanted to take a 'global' view on how learners were communicating during a specific activity with the aim of collecting qualitative rather than quantitative data. I knew the learners were not using modal verbs but they were, more often than not, engaged in the activity, and so my aim was to observe, through 'long-term engagement in the situation', (Walford, 2008: 9) what was unfolding and the communicative resources emerging from the learners' interactions. I point out again that my research emerged out of my practice as a teacher, as someone directly involved in the teaching and learning experience, an 'insider'. It was with the intention of 'participating' (Walford, 2008) in the classroom environment that I set out to collect, observe and comment on behaviour that was typically characteristic. Of course, there is always a danger when a teacher is observing in the familiar setting of the classroom of focusing on what one 'expects', 'knows', or is familiar with (Delamont, 2008), and for this reason 'reflexive observation' is called for, an awareness of subjectivity and of converting it into an 'objective' account (Blommaert and Jie, 2010). These observations would then, in an inductive sense, lead to the development of theory, based on empirical evidence (ibid).

Further, as the preceding chapters of my thesis show, my epistemological position is the situatedness of language, on how language, or more generally communication, and context form a dynamic. My research sets out to investigate the communicative resources participants draw on in a specific social interaction, i.e. a speaking activity from an ELT course book. To that end the methodology I would adopt would need to view language as a social tool, as part of social behaviour, structure and relations (Blommaert and Jie, 2010) and thus a part of the context in which it functions; as Blommaert and Jie point out 'language is context' (ibid: 7). With the above in mind I set about collecting data using video-enabled ethnography.

With its beginnings in social and cultural anthropology (Walford, 2008; Blommaert, 2010), ethnography orientates towards the study of social life by focusing 'attention on the complexity of separate social units' (Blommaert and Jie, 2010: 5). These beginnings are significant as they point to the function of language as a social tool (ibid: 7), a theme I pursue throughout my thesis (see, for example *Chapter 2 Modality: course books and beyond* and *Chapter 3 Context: finding my way through the 'black hole'*). The classroom is a social setting and teaching and learning a social activity. By using video recordings and field notes I would be able to collect rich, detailed accounts of the lessons and study the social processes involved in communication whilst the learners were carrying out a specific social activity, i.e. a discussion. But the use of video recording is more than just practical in this sense

as it enables the researcher to observe how members of social groups 'discursively construct events, identities, and academic content' (Baker et al., 2008).

Ethnography has, somewhat erroneously, at times been seen as synonymous with description (Blommaert and Jie, 2010; Dicks et al., 2011) and criticised as being 'insufficiently rigorous or systematic' (ibid). Visual phenomena, sound, places and objects are treated as 'detail, background or context' rather than seen as giving meaning to social settings and situations (Dicks et al., 2001: 230). Blommaert and Jie, however, depart from this view and argue that ethnography is much richer than mere description and 'involves a *perspective* on language and communication, including ontology and epistemology, both of which are of significance for the study of language and society, or better, language *as well as* of society' (2010: 6). The significance of language as well as society is what ethnography brings to my research. As I have indicated, I was interested to explore how language and social activity interrelate. Following Halliday, language can be seen as a social tool enabling humans to act as social beings. In ethnography, as Blommaert and Jie note, 'questions about language take the shape of questions of how language works and operates for, with and by humans-as-social-beings' (2010: 8); social groups are constituted in and through discourse-in-use (Baker et al., 2008).

Ethnography then, accords with social semiotics and multimodal theories, which centre on socially shared communicative systems of representation. In both approaches the making of meaning is contextually determined for carrying out social functions, and not restricted to one mode of communication. In terms of ethnography, Blommaert and Jie point out that 'the distinction between linguistic and nonlinguistic is an artificial one since every act of language needs to be situated in wider patterns of social behaviour' (2010: 10).

However, whilst I am attempting to combine these methodological approaches I am aware that their 'epistemological compatibility' has been called in question. Dicks et al. highlight that due to their respective histories and associations (Dicks et al., 2011: 227) there are theorists who see them as complementary and those who do not. Kress argues for a merger as multimodality and ethnography achieve different objectives. He explains 'that although multimodality both invokes and relies on the social, it does not itself provide a base for the social'; ethnography can offer insights that multimodality cannot. Pink sees ethnography and multimodality as two distinct theoretical and methodological approaches (ibid: 231). Arguing from the perspective of sensory ethnography her view is at odds with 'observing' and 'looking at' meaning for analysis. She proposes that the researcher needs to share experience and share knowledge with participants. As a way of showing parity Dicks et al. point out that ethnographers and social semioticians 'are interested in examining the diversity of resources that people use in their everyday worlds, and both do so from a perspective that favours social over

cognitive explanations' (ibid: 228) and this accords with my perspective of the social in language teaching and learning.

Of course learning is a very personal activity and I acknowledge there is always a risk of a subjective viewpoint being imposed onto another person's unique account of the world by the observer. Indeed my participatory role will in part have helped shape the classroom interactions³² and my observations led to subjective ways of constructing the 'objects' of my research: "what becomes 'objective' as a scientific *result* is subjective as a scientific *process*" (Blommaert and Jie, 2010: 66 emphasis in original). Subjective accounts cannot be avoided (ibid) and thus reflexivity is called for when making claims about my research foci. For example, due to the interpersonal function of modality it is perhaps to be expected that my presence as the teacher, as a contextual variable, will have had a significant bearing on the modal resources the learners draw on (see discussion in *5.6 Extract 4 Changing roles: contextual variables and communicative resources*). I have pointed out that my thesis is based on a small sample of data and I am careful not to overgeneralise in my analysis.

³² See *Plate B, Appendix 12* for a brief illustration on how the teacher's (and camera's) presence affected the posture of two students in a paired discussion. See also *3.5 The lesson as context: the dynamics of the classroom* for how the teacher/researcher's presence may, in part, bring about pedagogic mismatches during an activity.

5.3 Extract 1 Establishing social roles

T: What's this?

N: yeah what's this? Looks like a diagram

T: is this correct direction or rotation (N: laughs) or something you think?

N: I have no idea what that is some equipment

T: some equipment sss maybe that way's good uh yeah I have no idea as well

N: uh this looks like it rotates some kind of fan or I don't know

T: uh no lamp? no?

N: (inaudible)

T: (inaudible) I have no idea of this I don't know what is it what it is

N: no

The two participants in the dialogue above were MA TESOL students at London Metropolitan University (see *Appendix 6: Classroom data CD ROM – Richard (2) 24.34* for original data). N is a native speaker and T can best be described as an 'expert user' of English (Rampton, 1990). As modality is part of the interpersonal function of language, I will firstly discuss how interpersonal relations are established between the participants. I will do this by analysing speech functions as defined by Halliday (Halliday, 1985/1994; Eggins and Slade, 1997; Coffin et al., 2009). Following that I will discuss the modality markers used in the dialogue as a means of showing affinity with propositions and establishing and maintaining social relations. I will then turn my attention to the way the speakers use sound as a semiotic resource and comment on how certain features have meaning potential to express modality. I focus on sound as it became apparent during observation that meaning was being communicated through not only *what* was said but also *how* it was said. The higher level action in this extract is an oral discussion. The lower level actions I focus on are spoken language and sound.

I provide the original course book dialogue overleaf (*Figure 9*), which 'contextualises the grammar focus of the lesson: modals of present deduction (*can't, must, might/could*) (Albery, 2008: 98), for reference. (I have referred to this dialogue above.)

A: Wow, look at that. What is it?
 B: It's a drawing. I think it's a sort of flying machine.
 A: Yes, it could be that. Who do you think it's by?
 B: Mmm, it might be by da Vinci, I believe he did that sort of thing. Have a look at the sign. What does it say?
 A: Erm ... yes, you're right, it is da Vinci.
 B: Goodness! It's in very good condition ... it says here he was born in 1452, so it must be over 500 years old.
 A: Yes, and it's an amazing drawing.
 B: Yeah, I read somewhere he was fascinated by birds and flying. Perhaps that's where he got his ideas for the drawing.
 A: Yeah, you're probably right.

Figure 9 (Albery, 2008: 175)

It is quite apparent that the two dialogues are very different. This is probably not surprising as the first one is naturally occurring speech whereas the second is scripted to display certain features of language. I have already discussed this in detail above so will not reiterate that here.

In *Language and Control* Kress and Fowler point out that 'All language is addressed to someone, and involves an addressee: it is relational...relationships are generally asymmetrical' (1979: 63). Using this as a starting point what we can expect to see in the interaction above is unequal distribution of power between the participants realised through different uses of language. Eggins and Slade use Halliday's concept of speech functions to provide information about the participants in a situation and to show how power is distributed among them (1997: 190). There are four basic speech functions because the **items** exchanged in an interaction can be **information** or **goods and services** and the **roles** of the speakers in the interaction can be **giving** or **demanding**. The four speech functions are statement, offer, question and command. In this section I will detail this using the discussion between the two speakers T and N introduced at the beginning of this section. Firstly, the four speech functions are illustrated in Table 2 below.

Speech Role	Item (Information)	Item (goods and services)
Giving	<i>Looks like a diagram.</i> (Statement)	<i>Would you like to look at the diagram?</i> (Offer)
Demanding	<i>Do you think it looks like a diagram?</i> (Question)	<i>Look!</i> (Command)

Table 2

These four speech functions can be extended if the speaker's role is one of **initiation** or **response**; and if responding whether **supporting** or **confronting**. This gives us a total of twelve speech functions. If we look at Table 3 overleaf, we can see how this maps out. So along with the four speech functions

already listed above of statement, offer, question and command, we have acknowledgement, contradiction, answer, disclaimer, acceptance, rejection, supply, and refusal to supply.

	Initiation	Response	
		Supporting	Confronting
Giving information	<i>Looks like a diagram.</i> (Statement)	<i>Yes it does.</i> (Acknowledgement)	<i>No it doesn't.</i> (Contradiction)
Demanding information	<i>Do you think it looks like a diagram?</i> (Question)	<i>Yes, I think it does.</i> (Answer)	<i>No, I don't think it does.</i> (Disclaimer)
Giving goods and services	<i>Would you like to look at the diagram?</i> (Offer)	<i>Yes, I would, thanks.</i> (Acceptance)	<i>No, I wouldn't.</i> (Rejection)
Demanding goods and services	<i>Look!</i> (Command)	<i>Ok</i> (Supply)	<i>No</i> (Refusal to supply)

Table 3

(Based on Coffin et al., 2009)

So using these twelve speech functions I am able to look at the discussion between T and N and by analysing who is demanding/giving information, initiating/responding, etc. make inferences regarding how social roles and status is distributed. Firstly I need to set out the dialogue in a more detailed way and to do this I will use the notion of 'moves' from conversation analysis.

Egins and Slade explain '[t]he move is closely related to the turn-taking organisation of conversation...A move is a unit after which speaker change could occur without turn transfer being seen as an interruption' (Egins and Slade, 1997: 186). A move though is a unit of discourse rather than a unit of grammar and so whilst 'most clauses are moves, and most moves are clauses' (ibid) subordinate clauses do not function as separate moves as they 'do not select independently for mood' (ibid: 187). For example, 'I have no idea what that is' is treated as one move.

At the head of this section the dialogue was laid out in terms of speaker 'turns' i.e. all the talk produced by one speaker before another speaker gets in' (ibid: 184). However, for the analysis of speech functions turns are not suitable as one turn can consist of more than one speech function, for example 'yeah what's this, looks like a diagram'. Here we have one turn, two clauses and two speech functions. The dialogue is reproduced overleaf with the moves indicated in Arabic numerals.

T: (1) What's this?
 N: (2) yeah what's this? (3) Looks like a diagram
 T: (4) is this correct direction or rotation (N: laughs) (5) or something (6) you think?
 N: (7) I have no idea what that is (8) some equipment
 T: (9) Some equipment sss (10) maybe that way's good (11) uh (12) yeah I have no idea as well
 N: (13) Uh this looks like it rotates (14) some kind of fan or (15) I don't know
 T: (16) uh no (17) lamp? (18) no?
 N: (19) (inaudible)
 T: (inaudible) (20) I have no idea of this (21) yeah (22) I don't know what is it (23) what it is
 N: (24) no

There are 24 moves in total realising 22 speech functions. Move 19 is inaudible and moves 11 is a filler. The moves can be grouped as follows.

Speaker	Initiation	Response	
		Supporting	Confronting
T	12	2	1
N	6	1	

Table 4

Table 4 above shows 18 initiations in total and 4 responses; 3 supporting responses and 1 confronting response. Of the 18 initiations 12 were made by T and 6 by N. Of the supporting responses 2 were made by T and 1 by N. The confronting response was made by T. Initially then we can see that T took a more vocal position in the discussion with twice as many moves as N. This may suggest that T is taking a more dominant position in the discussion with N remaining largely uncommitted; additionally T makes the one confronting response. However, if we take a closer look at the class of speech functions and see how they are distributed a clearer picture emerges. Table 5 below shows the choices of speech functions with the corresponding number of moves.

Speaker	Initiation	Response
T	Statement 7 Question 5	Disclaimer 1 Contradiction 1
N	Statement 4 Question 2	Disclaimer 1 Acknowledge 1

Table 5

To find so many statements and questions for giving and demanding information is perhaps not surprising considering the interaction was a discussion. Correspondingly responses come in the form of disclaimers (move 7 'I have no idea what that is'; move 12 'yeah I have no idea as well') and acknowledgements (the elided 'no' in move 24). I pointed out above that T made twice as many moves

as N. What is interesting at this stage is the number of statements and questions. Whilst T initiates more, she asks nearly as many questions as she make statements. And of her five questions only one receives a response from N. If we look at the kind of statements being made, we can see that N is making more 'guesses' using relational verbs about what she thinks the design is: move 3 'Looks like a diagram'; move 13 'this looks like it rotates'; move 14 'some kind of fan'. Whereas T's statements are more concerned with uncertainty as to what the design is. Indeed the string of moves 20, 21, 22 and 23 are to this effect. In contrast what N appears to be doing through her statements is attempting to identify the design.

It is evident then that whilst T talks more her moves come in the form of questions and statements of doubt. N talks less but arguably more confidently; at least she is prepared to make deductions. It is also evident that there is not a reciprocal relationship between questions and responses. Questions in the main go answered. And thus whilst the participants are positioned in such a way that they have shared access to the same communicative space³³ and the picture under discussion (*Plate 2* see below) they do not seem to be engaged verbally with each other, rather they are 'thinking aloud'. However, in terms of the relative status of the participants, N appears to be more confident and assertive.



Plate 2

As has been pointed out the move is a unit of discourse. To find out how moves function over extended discourse Eggins and Slade use a more 'subtle' distinction of the 'types' of initiations and responses (1997: 191). For example, in move 4 T initiates with a question demanding information. She does so by asking a closed question using a polar interrogative 'is this correct rotation?' This type of question differs in its discourse function from say an open question using a wh-interrogative, e.g. Which direction do you think it is? Again this distinction can provide insights into the roles of the participants: asking a polar interrogative question puts the interlocutor into a different position as it

³³ I explore the relationship between participants, posture and communicative space when I discuss the F-formation in 6.5.1 *Making meaning through posture*.

somewhat limits the response to as 'yes' 'no' answer rather than seeking an opinion through an open question.

Similarly T's moves 20, 21, 22 and 23 all function as statements giving information. In Eggins and Slade's classification these moves would be classed by their discourse function, i.e. as 'elaborations' a means of 'restating' (1997: 202) what has been said. If we look at these four moves alongside the initial count of seven statements for T, we can see that a significant number of her statements express restating uncertainty. It is worth noting that these moves span 11 seconds. During which time N sits in silence.

It is beyond the scope of my thesis to go any further into a discussion of speech functions³⁴. For the purpose of my current research I have discussed speech functions to illustrate inequality between participant roles in an interaction. An analysis of speech functions reveals much about the relationship between the participants. In the dialogue above T makes more moves, asks more initiation questions and makes more initiation statements and whilst this gives the initial impression of a person taking a dominant role in the discussion the types of questions and statements reveal that she is more often seeking support from her interlocutor than giving opinions. On the other hand N makes fewer moves and remains silent for longer periods but engages with the exercise more in terms of making the deductions. I want to turn my attention now to how the participants use modality.

5.3.1 Modality: linguistic markers of interpersonal relations

Working on the premise that all utterances contain modality (Hodge and Kress, 1988) and that modality is part of the interpersonal function of language, it is safe to say that unequal social roles in an interaction will be, in part, signalled through modality. To that end I will now return to the discussion between T and N and discuss what features of modality are present. This is not intended to be a comparison with the dialogue from *Language Leader Intermediate* although what is immediately noticeable is the absence of modal verbs.

I have laid out the dialogue again, this time using conventions from conversation analysis (CA) and have opted for a vertical format in which the utterances of different speakers are written one below the other (Ten Have, 1999: 89). I am using a model from CA for two reasons. Firstly, whilst my data comes from two students engaged in a language teaching exercise in which the social import is arguably not conversation, they are engaged in 'talk-in-interaction' which 'whatever its character or setting' (ibid: 4) is suitable for conversation-analytic study. Secondly, for the purpose of analysis CA is

³⁴ For a fuller description of the speech functions and their discourse functions for this dialogue see *Appendix 8*.

concerned with both *what* is said and *how* it is said; as I have pointed out above, in this section I am as interested in the *how* as the *what* and the analytical tools of CA allow me to do this.

I should also add here a couple of comments on transcription. I have indicated that I am following a vertical format in which one utterance follows directly below another. This to me seems the most straightforward way of laying out the dialogue in terms of readability. Utterances follow the speaker moves which I identified in the discussion above on speech functions in 5.3 *Extract 1 Establishing social roles*. Again this seems to be a suitable way of formatting as it aids clarity by allowing each move to be transcribed separately on a different line. Whilst I acknowledge that strings of sound do not necessarily correspond to written words, I have, in the main, used standard orthography to represent sounds as words. I have done this because it aids accessibility and to show that the participants are using conventional linguistic norms (ibid: 80). That said T is a non-native speaker of English and N's linguistic background at times gives rise to marked variations. Non-standard orthography, therefore, has only been used where it seems relevant without attempting to make the speakers sound too informal or colloquial.

Perhaps most importantly I should add that this, as indeed all, transcription is not a substitute for the actual recording. It is inevitably selective and theory laden 'produced with a particular purpose in mind' (ibid: 77). (Dialogue continued overleaf.)

- 1 (3)
2 T: wos this?=
3 N: =yeah wos this?
4 N: (3)looks like a diagram.
5 T: (.)uh is this u:h >correct direction or rotation
6 N: (((laughter))
7 T: (.)o(h)r something<
8 T: (.)<you↑think?>
9 N: (1)I have no idea what that is.
10 N: (.)some equipment
11 T: (3)s:ome equipment s::
12 T: <maybe that way is: good.>
13 T: (.)uh<yeah I have no idea as well.>
14 N: (.)uh this looks like it rotates
15 N: some kind of(.)f::an or
16 N: (.)I don't know.
17 T: (.)uh°no°
18 T: (2)la:mp?
19 T: (2)no↓
20 N: [°()°
21 T: (.)°()°(1.5)°I have no idea of this.°
22 T: (7)yeah I don't know what is it what it is.=

23 N: =no.

(Example above reproduced from above: 4.1 *Epistemic modality*. See Appendix 6: *Classroom data CD ROM Richard (2) 24.34* for original data.)

As well as the more 'obvious' modalising forms such as vague language in move 15 'some kind of', the relational verb 'look' in moves 4 'looks like' and 14 'this looks like' and 'a rejection 'uh no' in move 16, there are a number of other markers of modality. I will discuss these below. Perhaps at this time we need to remember that during interaction participants seek to impose their idea of the truth of things. This truth is not objective, but the truth as seen by the speaker. The degree of truth is signalled through modality and thus modality and interpersonal relations are closely bound. A speaker's choice of high or low modality not only commits that speaker to what he/she sees as the truth, but also positions him/herself interpersonally with his/her interlocutor as he/she tries to assert his/her version of the truth. So modal forms in the dialogue include question forms, present tense, pauses and hesitation, deixis and laughter.

Firstly, the opening question 'wos this?' [what is this]. T could have launched straight in to the discussion with a statement, perhaps indicating what she thought the design was. As it is the question she asks sets up a particular relation with the other participant: one participant is seeking information from (a more knowing) other. In this respect the choice of question is significant if we compare it with alternatives such as 'What do you think this is?' or 'What would you say this is?' The use of the relational structure *x is y* asserts 'that there is a classification and that it has the status of fact' (Hodge and Kress, 1988: 126); 'this' is an identifiable object.³⁵ Interestingly this is the same question used to initiate the discussion in the course book dialogue where the two speakers do in fact 'discover' what the design is.

The verb 'is' is a marker of high modality. The question 'wos this?' therefore signals high modality. This contrasts with the verbal and mental process verbs in the rather more complex 'What would you say this is?'³⁶ and 'What do you think this is?' With these forms the speaker would be rather tentatively asking her interlocutor in the manner of 'What would you say this is if I were to ask you (and even if I did, I don't expect you to know)'. Both 'think' and 'say' signal lower modality as does the use of 'would' by signalling distance. Arguably then T is expecting or hoping that N can say what the design is, which sets up a particular kind of power relationship: she is deferring to N.

³⁵ Other relational processes are 'seem' and 'appear'. These processes carry lower modality than 'is' as rather than making the judgement *x is y* they make a judgement that *x seems/appears to be y*. I discuss 'seem' in 5.6 *Extract 4 Changing roles: contextual variables and communicative resources*.

³⁶ The formulaic utterance 'What would you say this is?' is related to both politeness and commitment. It is polite through the use of a certain degree of commitment.

However, the initiation question is immediately repeated by N in move 3 'yeah was this?' Lexical repetition can signal affinity by providing supportive encouragement through a registering move (Egins and Slade: 1997) and a reciprocal relationship between the participants becomes evident. But there is more happening here. N does not take up the position of knower conferred on her by T; instead she asks the question back in the manner 'I don't know what it is. You tell me what it is'.

Repetition occurs again in moves 10 and 11 'some equipment' and in moves 7 and 12 'I have no idea'. By repeating what one participant has said, the other participant shows affinity with that speaker. Again we can see a reciprocal relationship being maintained. Both of these instances then signal high modality (see 2.3.2 *Interpersonal meaning and modality: insights from systemic functional grammar* for a discussion of low, median and high modality). Repetition occurs elsewhere in the dialogue. The utterances 'I have no idea what that is' and 'I don't know' occur three and two times respectively, with an elided form in move 24 'no (I don't know what that is either)'. Both utterances are used by both speakers with N initiating both times. In the case of 'I have no idea what that is', whilst the speaker moves do not follow on immediately from one another, T's 'yeah I have no idea as well' in move 12 is clearly referring back to N's move 7.

T and N use 'I have no idea' before 'I don't know' and both speakers end the dialogue with this utterance. Coffin et al. explain that mental process verbs are concerned with 'the world of consciousness rather than the external, material world' (2009: 303) i.e. sensing phenomena. And Kress and Hodge note that mental process verbs can be used modally to 'indicate the authority of an utterance or the relation of the speaker to the utterance' (1979: 126). I would argue that the string of utterances 'I have no idea of this, yeah, I don't know what it is' moves from the relational realm of fact to one of sensing and can be interpreted as something like 'I have no idea what this is as fact and cannot even think/guess what it is'. According to Biber et al., the verb 'know' is more often used to report what a speaker does not know (2002: 324) and the structure 'I don't know' is high frequency at 1000/million words. *No*-negation is very rare in conversation although 'I have no idea' does occur as a collocation (ibid: 245). T and N's use of both structures could indicate a marked choice for emphasis. And, of course, the modal value is also retrievable from the context as both speakers use 'I don't know' as a way of finalizing their spoken turns.

I want to turn my attention now to the use of present tense. The use of the present throughout this dialogue is perhaps understandable as discussions tend to use the present tense. However, what the present tense also allows is for the speakers to signify 'proximity in time, and hence verifiability' (Hodge and Kress, 1988: 126). Present tense can be used to signal things that are closer to the 'truth': something that 'is' is more verifiable than something that 'was'; it has higher modality. To this effect

the present stands in contrast to the past tense and hypothetical markers which can signify temporal distance and thus low affinity. In this particular dialogue this is evident if we again consider the choice of question 'was this' over the removed, hypothetical 'What would you say this is?'

Proximity is also signalled using deixis. The dyads here/there, now/then, this/that can all indicate temporal and spatial closeness/distance. As with the use of present tense deictic markers of closeness and distance can signal high/low modality. At the beginning of the discussion N says 'was this' (move 3). As discussed the relational *x is y* asserts that 'this' picture can be classified, what the design is can be stated as fact. 'This' signals high modality. However, the design proves difficult to pin down and after N remains silent for approximately ten seconds and five moves she says in move 9 'I have no idea what that is'. The object under discussion becomes a removed 'that' and her affinity with it lowered.

Looking back over the transcript the number and length of pauses or hesitations stands out. This interaction was just short of one minute. So for most of the time there was no speaking. Indeed there was a pause of three second before either participant made a bid to speak. Pauses can show a need to think carefully before speaking. They can also act as a face-saving mechanism: when uncertain, do not speak. Pauses and silence can signal non-collaboration or a dispreferred response (Liddicoat, 2011) and thus have the potential to signal modality (Kress, and Hodge, 1979; Hodge and Kress, 1988). I have pointed out about that the number of responses to questions is nonreciprocal. So just as repetitions can signal high affinity, the lack of responses can signal low affinity between participants. The inaudible comments in moves 19 and 20 might be due to recording difficulties; however, comments made *sotto voce* can also suggest a lack of commitment or uncertainty.

The tag 'no' in move 18 seeks agreement and expresses modality. T's only statement about what the design might be comes in the move before when she says 'lamp'. From the analysis so far it should be apparent that T plays a more supportive and less assertive role in the discussion. And so when she makes her comment 'lamp' with rising intonation and no response comes, she seeks confirmation again from N and in doing so adds modality to her comment 'lamp'. Whilst she makes a guess the rising intonation on 'lamp' and the tag 'no' suggest low modality.

Finally N's laugh is of relevance for this discussion. In move 4 T asks 'is this correct rotation?' referring to which way up the picture should be looked at. Following N's laugh T hedges 'or something' and then seeks support from N 'you think?' N's laugh could be a means of passing 'comment' on T's question. T appears to perceive it in this way from her follow up moves. The act of laughing in this instance is interpreted as conveying high modality in relation to the previous move.

I have shown above that the dialogue between T and N is full of modality markers. However, the degrees of modality fluctuate throughout the discussion as the participants make propositions and respond and react to each other. Through this it is possible to see how the participants align themselves to each other through degrees of affinity and what kind of social roles they take on. Social roles can also be inferred from the analysis of how speech functions are distributed. Whilst this analysis has thus far been primarily linguistic, I have made reference above to laughter and silence; suggesting that they carry meaning. In the next section I will discuss the function of sound as a semiotic resource for T and N.

5.3.2 Sound: the meaning potential of sound

Whilst I was collecting and listening back to my data I was interested to discover that modality is not only signalled by *what* is said but also *how* it is said. For example, by elongating the initial consonant sound in the modal finite 'might' a speaker can add a degree of tentativeness to a proposition. However, the impressions that sound gives come through combinations of features. Elongation by itself may not give the same impression as elongation combined with nasality. Van Leeuwen points out that with nasal sounds such as [m] there is no mouth opening and thus 'The sound is "kept" inside...it remains a *hum*, a *murmur*, a *mumble*' (1999, 149). Perhaps, then, the combination of being 'kept inside' and elongated adds something to the strength of the modal: m:::ight. What the speaker produces in word initial position is an *Mmmm*, 'the most non-committal of reactions' (ibid). I am not arguing here a case for sound symbolism or suggesting that elongated sounds in general are an indicator of modality but from my observations it appears that speakers use their vocal apparatus to produce distinct sound qualities and it is these in combination (ibid: 150) which convey meaning. In this section I want to discuss the semiotic value of some of these sounds. To do so I am using a social semiotic approach to sound. This differs from phonetics and phonology which are concerned with the technical dimension and structural properties (Finch, 2000) of speech giving meaning to the lexicogrammatical system. It also does not cover prosodic features such as the attitudinal and accentual function of intonation (see Roach, 1983). A social semiotic approach looks the 'materiality' of sound (Van Leeuwen, 1999) and the meaning potential of dimensions such as voice quality and timbre i.e. pitch, loudness, nasality and speech sounds such as plosives, fricatives, semi-vowels, etc.

To start with I want to focus on the deictic markers *this* and *that*. These two words are interesting because in EFL they are not normally connected with modality. A typical description for use in an EFL context has *this* to talk about a person or thing which is literally close to the speaker, and *that* to talk about a person or thing which is 'more distant from the speaker, or not present' (Swan, 1980: 565). However, I have indicated above (4.1 *Epistemic modality*) that deixis is a marker of modality (Hodge

and Kress, 1988: 125) in that it allows speakers to position themselves metaphorically to the truth value of propositions although interestingly they do so by being markers of 'closeness' and 'distance'. In this section I want to discuss the sounds of these words as a semiotic resource and an indicator of modality. Firstly, Van Leeuwen writes that the dimension of front and back vowels 'can give rise to metaphorical extensions related to the idea of distance' (Van Leeuwen, 1999: 146). So perhaps it is not surprising that the frontal vowel sound [i] is found in the deictic marker *this*: an indicator of both literal and metaphorical closeness and the back vowel [a] is found in *that*: an indicator of distance. In move 9 of my data, N says 'I have no idea what that is' whilst moving away from the picture on the table in front of her; earlier in move 2 she had picked up the picture and asked 'yeah was this?' Initially she engages with the activity and refers to 'this' picture, but on realising the apparent difficulty of deciphering it she moves away from it and refers to it as 'that'. Arguably this physical positioning of the body and vocal apparatus coupled with the literal and metaphorical meanings of the deictic markers signal her affinity with the design. In addition 'that' has a fricative in word initial position. Van Leeuwen states that fricatives 'are particularly good at expressing anything to do with friction, whether literally or figuratively' (ibid: 148). It is likely that some amount of friction is being caused to N by not being able to realise the picture she is attempting to deduce. What I am suggesting is that the composite parts 'form larger semiotic gestures whose meaning and effect derive from the way they are combined' (ibid: 150).

Throughout the dialogue the hesitator 'uh' is used four times. 'Hesitation phenomena' is a marker of modality (Hodge and Kress, 1988: 125) indicating careful consideration of what the speaker has to say (Van Leeuwen, 1999: 44). Again as I am interested in *how* something is said I think it is interesting that the hesitator has a back vowel which as I have pointed out above can signal distance. Whilst claiming a metaphorical extension to the notion of 'backness' and distance, it is perhaps obvious that hesitation is signalled by a back vowel: a sound produced far back in the mouth. To this, however, could be added the dimension of aperture. It has not been possible to convey this in my transcription but the mouth and oral cavity are fairly closed when producing this hesitating sound. As a larger semiotic unit both the *what* and *how* of 'uh' can be seen as an indicator of modality.

I think it is worth restating that the picture the two students were attempting to deduce was a drawing of Leonardo da Vinci's flying machine; it is also worth pointing out that they never really came close to making that deduction. The whole dialogue, therefore, is one of speculation and guesswork. Accordingly a high amount of modality is evident. Only two suggestions are made on what the drawing might be of, once in a highly modalized utterance by N in move 15 and then by T in move 18. N's utterance 'some kind of(.)f::an or' contains vague language 'some kind of', a pause (.), an elongation of word initial sound 'f::' and a contrasting conjunction 'or'; moreover, in her next

utterance, move 16, she says after a pause 'I don't know'. All of these elements are modal operators but what I find of particular interest here is the elongation of the [f] in 'fan' when she puts herself 'out there' and makes a guess. It is of interest because T does the same thing with the vowel sound in 'lamp' when she makes the only other 'concrete' guess of the dialogue. Again as I have pointed out above I am not suggesting that the elongation of a phoneme is an indicator of modality per se, but taken a part of a semiotic unit, i.e. N's complete utterance and T's utterance spoken with rising intonation, the elongation of the phoneme in question adds to the modality of the utterance.

'[S]ound can be used to represent our environment and to represent the actions and interactions of people' (ibid: 93), two sound qualities I want to briefly look at now are breathiness and softness. In move 7 T uses the contrasting conjunction 'or'. Prior to this move, in move 5, she queries the direction or rotation of the picture being discussed '(.)uh is this u:h >correct direction or rotation' which gives rise to laughter by N in move 6. As a result the conjunction in move 7 is rather breathy 'o(h)r'. Breathiness is often associated with intimacy (ibid: 140) but it can also indicate emotions such as laughter, crying, etc (Ten Have, 1999: 214). T is obviously confused by the picture and trying to make sense of it and perhaps feels a bit embarrassed by N's laughter. The full utterance of a 'breathy' contrasting conjunction and vague language follows a short pause and suggests she is modalising her previous utterance in move 5: she is retracting into vagueness and uncertainty.

Softness and loudness are associated with both social and physical distance (Van Leeuwen, 1999: 133). In move 17 T says in a comparatively soft voice 'no' to N's suggestion that the picture might be of a fan. Obviously 'no' carries with it high modality and could be quite a confronting response to N. To lessen her claim on territory (ibid) T modalizes her utterance by softening her voice and thus maintains the social distance between her and N.

Later in move 21 T uses a soft voice again when she says 'I have no idea of this'. An utterance of high modality it comes very near the end of the dialogue in which the speakers have not been able to make sense of the picture they are discussing. T seems to be stating defeat but the softness of her voice arguably tempers this. Softness of voice is again used by both speakers towards the end of the dialogue. Perhaps not surprising both times occur when speech was uncertain/inaudible on the recording: both speakers being tentative and doing so by speaking softly and inaudibly.

I want to finish with a final point on tempo. While I was listening to the recording it became apparent that the tempo of the utterances increased and decreased from time to time. It might well be expected that conversational speech and the kind of talk-in-interaction T and N were engaged in is less regular than other spoken registers (Van Leeuwen, 1999: 43). However, speeding up and slowing down appear to convey meaning for the participants in the dialogue. In move 5 T slows down

considerably while she tries to work out the correct way to look at the picture. Her utterance is preceded by hesitation phenomena and a pause and the reduced tempo adds to what is a highly modalized utterance. Conversely in move 12 she speeds up when she appears to decide on how the picture should be seen 'maybe that way is good'; her decision is further marked by falling intonation.

What I have presented above is not an exhaustive description of the semiotic value of sound in this piece of data. What I have attempted to show is that sound is a semiotic resource in the making of meaning, in this case modality. Some of my descriptions may seem obvious, for example, the hesitator 'uh' as a way of expressing modality. However, it is not only *what* is uttered but *how* it is uttered that combine to make meaning. In this way 'uh' becomes meaningful in any particular spoken context when the dimensions of frontality, aperture, length of the vowel sound, etc. integrate with other semiotic modes. Both speakers use sound as a modal resource helping to establish their respective social roles in the interaction.

5.4 Extract 2 Modal ensembles: interpersonal meaning in multimodal communication

- T: what about the second picture?
S1: this one yes?
S2: signal transmission [s]omething like that
S1: [yes]
T: so uh an aerial you mean?

(See *Transcript C, Appendix 11* for full transcript.)

This next piece of data comes from two pre-sessional students discussing the same picture. The dialogue has already been discussed in part above in 4.2 *Mode* and the sequence of stills has been reproduced from the same section (see *Plate 1*).

The two students had just discussed Picture A together and the dialogue above comes as a response to the teacher's question 'What about the second picture?' I have chosen this extract because of the brevity of S2's spoken response and from what we can add to its meaning through an analysis of how it integrates with her other communicative modes. S2 says 'signal transmission/something like that' (where / refers to a move). These two moves correspond to two clauses and are possibly elided forms of 'It's a signal transmission or it's something like that'. Kress and Hodge indicate 'The most straightforward model for presenting a comment or judgement is the relational' (1979: 113) i.e. *x is y* and this is what S2 has done using the relational verb 'is' to present a comment on how *x* is classified using the classification *y*. Further the intonation on both of these moves is falling, indicating that S2 is making a statement and so we might draw the conclusion from this that she is quite certain about what the design in the picture is. However, if we look at how speech integrates with other communicative modes, we can see that this first impression is not the full picture (see Table 6 overleaf).

The higher-level action is a response to the teacher's initiation. The exchange follows an initiation, response, feedback exchange (IRF). The lower-level actions are speech, gesture, head movement and gaze.³⁷

³⁷ A reading of left to right is of no relevance here. Gesture should not be taken as an 'anchor mode' (Flewitt et al., 2010: 53).

Time	Gesture	Speech	Head movement	Eye gaze	Photo capture
19:23:859	Arms crossed (right on top of left). Holding pen in right hand.		Shakes head	Looking at picture.	
19:24					
19:25	Right thumb raises up. Hand opens. Arm and hand move upwards. At level of right shoulder stops. Hand closes. Arm moves towards body and lowers. Comes to rest on left arm. Palm downwards.	Signal transmission	Looks up towards teacher	Eye contact with teacher.	
19:26			Head position directed towards teacher		
19:27	Right arm at rest on left arm.	Something like that			
19:27:433					

Table 6

G's response is a little over three seconds long. After the teacher asks the question 'What about the second picture?' S2 sits with her arms folded looking at the picture and shaking her head. It is two seconds before any other actions commence. She then simultaneously adjusts her head position towards the teacher, gestures with her right arm and says 'signal transmission'. As her arm comes to rest on her left arm she engages eye contact with the teacher and says 'something like that'.

I am viewing the four modes of gesture, speech, gaze and head movement as a modal ensemble, i.e. an integrated whole (Jewitt, 2009) and inferring meaning through their interrelationship. I will briefly say a few words about each mode before commenting on how they function together to make meaning.

G's first action is to shake her head. Horizontal rotation of this kind has a conventional meaning 'no' in many western cultures. Norris (2004) points out that such a movement comes as a second part to an adjacency pair and is often performed in place of spoken language. In this case the first turn in the adjacency pair being the teacher's 'What about the second picture?'

When S2 speaks she integrates it with gesture. It has been observed that gesture and speech are very closely bound and often co-occur (Norris, 2004; McNeill, 2005). This gesture follows the typical movement phases of preparation, stroke and retraction (Norris, 2004). These three phases are seen as 'obligatory' phases in most gestures. During the preparation phase the limb moves away from the rest position; this can be seen when S2 moves her thumb and opens her hand. The stroke is the full movement of the gesture to its 'peak'; S2's upward movement of the arm. The retraction is the limb returning to its rest position; the return of S2's arm to rest on her left arm again.

At the beginning of the interaction S2's gaze is focused on the picture under discussion. This shifts as she reaches the end of her first spoken move and comes to rest on the teacher. When it does so she utters 'something like that'. Gaze typically signals 'focused interest' (Forgas, 1985: 154). Norris informs us that the hearer gazes more at speaker than the speaker at the hearer, and that speakers typically look towards the hearers at the end of phrases (2004: 37). This accords with S2's gaze in this extract. Norris adds that gaze is to some extent sequenced and subordinate to spoken language (ibid). Again this tallies with S2's actions as her gaze clearly 'follows' her spoken language.

I want to suggest that the configuration of the four modes expresses ideational, interpersonal and textual meaning. However, as I have already indicated, when several modes are integrated, meaning is not always distributed evenly (Jewitt, 2009) as modes can be dominant or subordinate in relation to one another.

The propositional content is expressed ideationally using, for example, relational clauses in the spoken mode: the elided 'signal transmission'. I have explained above how the relational *x is y* allows S2 to classify and identify what the picture is. Ideational meaning is also conveyed through S2's gesture. If we take the gesture as a conduit metaphor (McNeill, 2005), the open, upturned hand 'presents' S2's idea and has a meaning something similar to 'This is my idea'. The shaking of the head at the beginning of the extract conveys not knowing or uncertainty.

Interpersonal meaning is conveyed in part through speech, gaze and gesture. S2 answers the teacher's question with 'signal transmission' and then adds 'something like that'. These moves establish relations between the speaker and the hearer. The first move provides her answer, the second modalizes it. To understand this we need to consider the relations of the speaker and hearer in this

instance. S2 is a student addressing her teacher. Fowler and Kress contend '[m]odality in general establishes the degree of authority of an utterance' (1979: 196). And not only the authority of the utterance but also the authority of the speaker making an utterance. To that effect S2 'softens' her first move with her second.

Interpersonal meaning is also established here using gaze. Gaze has an important function for communicating interpersonal relationships (Argyle, 1972; Kidwell, 2014; Ricci Bitti, 2014). Gaze can suggest degrees of engagement and express attitudes such as cordiality, hostility, dominance and submission (Kidwell, 2014). Gaze can also 'emphasize, underline, and modulate the content and meaning of the concomitant verbal language' (Ricci Bitti, 2014: 1345). I suggest that in this example S2's gaze is expressive of co-participation as she moves it from the picture to the teacher. The repositioning of her gaze engages the teacher. Interestingly full eye contact is only established when she says 'something like that', i.e. the clause with low modality. As with all modes they are culturally shaped and their meaning is derived from their integration with other modes and the context in which they occur.

I have mentioned above how S2's gesture expresses ideational content (see also Kendon, 2004 for a discussion on how speech-related gesture gives meaning to an utterance's referential or propositional content). As well as being co-ordinated with the referential meaning of spoken utterances, gestures can express emotional states (Argyle, 1972) or have a pragmatic function (Kendon, 2004; Graziano, 2014). Kendon identifies modal gestures as a sub category of gestures with a pragmatic function: 'gestures having a *modal* function show the speakers own mental attitude towards his discourse, indicating the interpretive key to the discourse' (Graziano, 2014 emphasis in original). S2's movement is loose and relaxed. It is not wild, uncontrolled or in her interlocutors face. Again, as gestures are culturally shaped and their meanings derive from context and their integration with other modes, it is only through the modal ensemble that we can arrive at meaning (see the penultimate paragraph of this section).

The modes of speech, gaze, gesture and head movement are all intricately co-ordinated. The way that each mode coheres within itself, i.e. the preparation, the stroke and the retraction phases of gesture and the syntactically and grammatically coherent utterances within and across moves; along with the way that each mode integrates with the other modes, i.e. the co-ordination of eye gaze with spoken language and the synchronisation of gesture to speech, are all indicative of the textual metafunction. Each mode works in synchronisation and forms an integrated whole. The four modes identified are lower-level actions that make up the higher-level action of S2's response. This higher-level action is embedded within the higher level action of the discussion, or perhaps more appropriately defined as

an initiation, response, feedback (IRF) sequence initiated by the teacher. And so we can see S2's response also coheres textually to the teacher's initiation question and to the feedback move.

The description above is not exhaustive. The modes I have selected for analysis are only a selection. I could have broadened it with an analysis of sound and intonation as well as facial expressions for example; facial expression as a mode is closely linked to that of gaze (Argyle, 1972). What I have shown is how modes integrate to make meaning and I argue that for a full understanding of S2's response, modes in addition to the verbal need to be taken into account. At the beginning of this section I suggested 'signal transmission/something like that' was an elided form of 'It's a signal transmission. It's something like that', and that S2 was quite sure about this. However, it seems more likely it is an elided form of 'It could be a signal transmission or something like that.' Taken together the lack of eye contact on the first move, the head shaking prior to speaking, the gesture, and the modalized second move indicate a significant degree of softening in relation to the authority of her utterance.

It is worth noting again that IRF is a common sequence in teaching (Cook, 2001; Ellis, 1994).³⁸ What perhaps is unfolding in this exchange is S2 responding to the teacher's initiation question and expecting evaluative feedback. This would help to explain S2's tentative response: she is giving her answer to the teacher who has the answer. Modality in this exchange is as much about interpersonal relations as it is about deduction.

³⁸ The data collected for this research typically follows an IRF model in line with the structure of the activity on which this research is based. Recourse to the CD in *Appendix 6* and the transcripts in *Appendix 11* provide further examples. It is worth noting that *Transcript A* and *Transcript D* in *Appendix 11* do not open with a verbal initiation from the teacher but with the teacher presenting Picture A. This can be seen as a non-verbal initiation by the teacher.

5.5 Extract 3 Authority and affinity: a multimodal perspective



Plate 3

(See *Transcript D, Appendix 11* for full transcript.)

Using this next piece of data (*Plate 3*) I will discuss speech, posture, pointing, beats and gaze. I refer the reader to the discussion on gaze provided in 5.4 *Extract 2 Modal ensembles: interpersonal meaning in multimodal communication*. Firstly then I will define posture and pointing and beats as gestures. I will then provide a gloss for the interaction above (see *Appendix 6: Classroom data CD ROM – Richard (1) 27.35* for original data). This will be followed by a commentary on the three modes: posture, pointing and gaze and how the three participants attend to them. I will finish with an explanation as

to how these modes are used to signal interpersonal meaning and modality. The higher-level action is a discussion. Before I begin, a few word on the participants.

The interaction is between three pre-sessional learners discussing the same picture: the drawing of Leonardo Da Vinci's flying machine. The interaction is approximately 32 seconds long and begins just after the teacher has handed the picture to S1. S1 sits on the left hand side of the frame, S3 on the right-hand side of the frame and S2 in the middle. S1 is a Japanese male. S2 a female from Congo. S3 a female from Turkey. The red lettering in the stills is S1's speech and the blue lettering the speech of S3. The blue arrows indicate the direction of gaze. The sequence of frames is from left to right. The time of the frame in seconds is in the top right-hand corner.

Posture has been identified for its social function (Argyle, 1972; Norris, 2004; Jaworski and Thurlow, 2009). Argyle tells us that that it can 'convey interpersonal attitudes' (Argyle, 1972: 248) as well as emotional states and social status. Norris notes two important aspects to posture: the form of the body and the direction of the body. The form of the body refers to the limbs, the head and torso and analysis needs to consider the whole body as open or closed positions are measured in degrees. For example the arms and legs of a person may be crossed but the torso straight and the head uplifted.

Directional position refers to the way bodies orientate towards others; and again this is measured in degrees. For example while I was collecting data of asking for and giving directions for my MA dissertation, I found people oriented the direction of their bodies to different extents. Some turned their bodies squarely to me whilst they spoke, others would just rotate their torso. It is important to note that body form and body position are culturally and individually shaped although overlaps and universals have been identified (Argyle, 1972; Norris, 2004). Moreover, like all semiotic acts, meaning is derived from the context in which they occur and from how they integrate with other modes of communication. As a general rule though directional positioning of the body towards another person indicates engagement and positioning away from another person indicates disengagement (Norris, 2004).

Pointing is sometimes subsumed under the category of deictic gesture (McNeill, 2005). McNeill defines deictic gestures as a means of 'locating entities and actions in space vis-à-vis a reference point'. Jaworski and Thurlow add that indexical pointing gestures are as much about 'the actor as they are about the object' (2009: 257). They give examples of tourists pointing out landmarks and suggest that spatial relations pin-point self in relation to a particular place being referred to. We can see a similar thing in the examples of asking for and giving directions I referred to earlier. The reference point, in this case Victoria Park, was pointed out in relation to where the actors were discoursing. The pointing

gestures locates self in relation to the reference point. Its verbal equivalent is something like 'Look, we are here. In relation to us the place you are looking for is there'.

Beat gestures take the form of beating time. They are typically up and down beats or back and forth movements (Norris, 2004; McNeill, 2005). Beats are normally in rhythm with speech and usually mark emphasis signalling 'something the speaker feels important with respect to the larger discourse' (McNeill, 2005: 40). They usually accompany speech and act as a means of indicating or highlighting something. For example, they can be used to indicate features of intonation such as prominence. Beats can be audible as well as visual and act as a means of orienting focus to self in a similar way to pointing.

Whilst the terms deictic gestures and beat gestures are useful, McNeill observes that 'semiotic properties are dimensional and not categorical' (2005: 41); so we need to refer to dimensions of gestures rather than kinds. So, for example, we may find someone giving directions by pointing and using beats to accompany and empathize speech: 'that's the park'. In this instance deictic and beat gestures integrate to assist in making meaning.

In their discussion of movement tied to speech Ekman and Friesen propose that gestures, or illustrators as they call them, can 'repeat, substitute for, contradict, or augment the information provided verbally' (1969: 69). They observe that changes in rate are associated with mood and articulation problems: 'With excitement and enthusiasm about the topic or process of communication, people increase their rate of illustrator activity' (ibid). Activity also increases with difficulty finding words (Hinde, 1972) and at these moments illustrators or beats can help the speaker past the difficulty in speech and can act as a way of introducing or priming the 'speakers' verbal utterances. The use of beats can assist in synchronising utterances (ibid) and are 'intimately interrelated with the concomitant verbal behaviour on a moment-to-moment basis' (Ekman and Friesen, 1969: 69). They add that illustrators receive external feedback from observers who may not give verbal comments but usually pay visual attention (ibid).

The following is a gloss for the interaction above (see *Plate 3*). In the sequence S1 takes the picture from the teacher and lays it on the table for all the participants to see. He seems to immediately recognise it and starts to point at it with his right hand. While he is pointing his hand 'beats' with an up and down movement. At this stage his body is still half turned to the teacher and he looks back at the teacher still pointing at the picture. At this point S2, who had been looking through her bag, turns her gaze to S1. While this is happening S3 slides the picture over closer so that she can get a better look at it. No one has spoken up to this point. The teacher does not engage with S1. S1 then turns his attention away from the teacher but is still pointing at the picture and beating with his hand. S2

returns to looking in her bag and S3 focuses more closely on the picture bending her head and torso to do so. In frame 8 S1 focuses his gaze on S2 and in the next frame he says 'This is' and then pauses. Over the next three frames S1 moves his body position more squarely to the other participants and continues pointing at the picture. S2 finishes with her bag and turns her focus on the picture. S3 is still looking closely at the picture. After a pause S1 continues speaking 'first aeroplane it makes' then pauses again. S1 and S2 sit slightly more upright and still pointing S1 says 'It is made by Leonardo da Vinci'. At the same time S2 turns her gaze onto S1 and S3 straightens herself up and runs her left hand through her hair moving it away from her face. Just as S1 finishes speaking S3 looks at the teacher and says 'This is looking satellite'. S2 turns her gaze to S3, S1 turns his head and torso and looks towards the teacher. In the last frame S1 is no longer pointing but holding his hands together. S2 is looking again at the picture her shoulders turned away from S3 and her left arm across her body with her hand on her chin. S3 is holding the picture and focusing attention on it.

I will now say a few words about each mode and how each student attends to them.

5.5.1 Making meaning through posture

The three students are working as a group and sit in a small cluster. For most of the discussion the picture they are making deductions about lays on the table at a focal point central to that cluster. The positioning of the participants was intended to generate interaction with each other. At the beginning of the sequence S2 is trying to retrieve something from her bag but in terms of directional positioning her head and body are 'engaged' with the picture and the other participants. S3's directional positioning suggests the same. S1's position on the other hand is interesting as it somewhat distorts the group cluster. To illustrate this I will draw on the concepts of F-formation and vectors.

An F-formation refers to a transactional space in which one or more persons come together to do something and arrange themselves in such a way that they have 'direct and exclusive access' to that space (Kendon, 1990: 210). A vector is a means of visually representing interaction between actors and objects. Vectors do not necessarily exist in physical terms but can be formed by spatial arrangements of objects, by 'bodies or limbs' (Kress and Van Leeuwen, 2006: 59), and by gaze, for example. At the beginning of the sequence S1's body is turned towards the teacher, and camera. Whilst his form is somewhat open, his directional position is turned away from the group, arguably affecting the group dynamic (see *Plate 4* p.125). Kendon points out that F-formations refer to spatial and orientational organisation in terms of 'lower bodies' (Kendon, 1990: 211). S1's posture jars with

the transactional space and the F-formation is not established.³⁹ Indeed, at this stage S1 is attempting to get the attention of the teacher and not interacting with the group.

³⁹ See *Plate B, Appendix 12* for a further illustration and commentary on F-formation.



Plate 4

After he fails to capture the attention of the teacher he turns his head and torso in towards the group; his right arm too extends into the centre of the cluster. His legs, however, still remain in the direction of the teacher. The position does not look comfortable or suggest full engagement with the group (see *Plate 5* below).



Plate 5

When he starts to speak S1 shifts his body position and extends his head towards the other participants and so the upper half of his body, at least, attends to the group. S2 remains for the most part square with the group and the picture at its centre that is the focus of the discussion. Her hands are at times touching making her posture relatively closed. She does, however, handle the picture and shifts her body position towards S3 when she speaks. Her form and the direction of her body suggest she is engaged with the group. At the beginning of the sequence S3 sits upright but as she draws the picture in towards her for closer scrutiny her arms, head and torso take on a closed position. She remains in this position until she speaks; at which point she sits up straight and says 'This is looking satellite'. Argyle points out 'someone who is going to take charge sits in an upright position' (1972: 248)⁴⁰ and this seems to be what S3 is doing. As S3 speaks S1 withdraws his body, his right arm and his hands touch.

5.5.2 Making meaning through pointing and beats

S1 uses a hand gesture throughout most of the extract. It is a half pointing, half beating gesture with his right hand and index finger. At times he points to reference the picture as a deictic gesture, at

⁴⁰ See *Plate A, Appendix 12* for further illustration speakers' sitting upright posture.

other times he beats for emphasis, e.g. when he corrects his utterance 'it makes' to 'it is made by Leonardo da Vinci'. He beats to mark his correction 'made' and to augment the intonational prominence of his spoken utterance.⁴¹ His beats signal that he intends to speak and it is possible that S2 attends to this as she looks at him in frame 4. Interestingly at times his hand is rotated outwards so that the underside of his gesturing hand is facing upwards. This action is sustained until S3 cuts in with her spoken utterance 'This is looking satellite'. As she says this she also points. She uses the index finger on her right hand to point at the picture on which her hand rests. Her gesture is deictic and she does not 'beat time'.⁴²

5.5.3 Making meaning through gaze

In frame 4 S1 turns his gaze towards the teacher. As I have shown using the picture above the effect of this is to disengage himself with the group cluster. It is not certain whether he attempts to speak but does not get the attention of the teacher. S2 had been looking in her bag but turns her gaze to S1 as he turns towards the teacher. The combination of S1's priming moves to speak, i.e. his directional turn, his uplifted torso and head and his beats might have been the cause of this; her shift in gaze signals her interest in S1. However, as he does not speak S2 returns her attention to her bag. When S1 makes a bid to speak again in frame 6 he turns his gaze towards S2. Here the shift in gaze is used to synchronize speech (Argyle, 1972); S1 signals that he is about to speak. As S1 starts to speak both he and S2 focus their gaze on the picture. S2 returns her gaze to S1 moments later while he is still speaking as a way of giving 'feedback' (ibid). Throughout this interaction between S1 and S2 S3 has remained with her gaze on the picture. As soon as S1 finishes speaking S3 straightens up, fixes her gaze on the teacher and speaks. Her shift in gaze disengages her from the group indicating that her spoken utterance is for the teacher and not her fellow students. At the same time S1 directs his gaze to the teacher waiting for feedback and S2 turns her gaze onto S3 indicating she is listening. These shifts in gaze are interesting. S1 knows what the picture is and thus waits for the teacher to confirm this. S2 does not know what the picture is and her gaze to S3 mirrors her 'I'm listening' gaze to S1 earlier.

Whilst this extract is relatively short and spoken language is minimal, there is an interesting power struggle going on and we can see this through the modes of posture, gesture and gaze as well as speech. S1 knows what the picture is and makes an initial bid to speak. As I have suggested this comes across in his posture, gaze and gestures; all of which act as a means of priming himself to speak.

⁴¹ See *Plate C, Appendix 12* for a further illustration of a speaker using pointing and beats to mark a correction.

⁴² See *Plate D, Appendix 12* for a further illustration of a speaker pointing and beating in co-ordination with a spoken utterance.

However, his body position is directed more towards the teacher than with his group. The teacher though does not acknowledge his attempt to engage and so S1 is left out on a limb from the group. By the time he turns his attention to the group S2 is back looking through her bag and S3 is engrossed in the picture. S1 has lost the floor.

However, S1 is quite persistent though maintaining his beating and pointing and making a second bid to speak. This time he directs his gaze to S2 and manages to engage her. She returns his gaze moments later as a means of giving feedback and he gazes back at her as he reaches the end of his utterance. All of this, however, seems to go unnoticed by S3. She does not acknowledge S1's utterance and when she speaks, she looks directly towards the teacher announcing what she thinks the design is.

We can see now how these modes interact to convey interpersonal relations and modality. Pointing can fulfil an interpersonal function through the placing of self and others in spatial relations, and gaze and posture have the potential to fulfil an interpersonal function through the degrees of (dis)engagement they express with others. S1 uses pointing as a means of announcing he is about to speak, attracting attention, holding the floor and emphasising what he says. His gesture references both the object under discussion and himself in the way suggested by Jaworski and Thurlow (2009) by placing himself firmly in the centre of the frame 'This is me talking about this picture'. I have pointed out above that his hand is slightly rotated with the palm facing upwards. Here we can partly see a conduit metaphor (McNeill, 2005) with his idea shown and presented to the group. His body position indicates self-assertion. His form is open and upright and his shoulders square to those he addresses. He also uses gaze to attract the attention; he does this with the teacher and with S2. When he speaks he uses a relational structure to tell the group what the design is: 'This is first aeroplane'. He is sure what it is. In terms of modality S1 is signalling high modality claiming his right to speak based on the fact that he knows what the picture is. It is perhaps unfortunate for him that his degree of certainty and authority is not acknowledged by the others in the group. Whilst he is certain about what the design is, he takes some time to articulate it. In doing so he loses S2 and she returns her attention to her bag. When S1 starts to speak again S2 shows willingness and engages with the picture and then with S1 using gaze. S1 seems to be aware of this and returns her gaze at the end of his utterance, a common feature of gaze in discourse (Norris, 2004) and a sign that they are to some extent sharing the same perspective on the interaction and are synchronized in rhythmical coordination (Kendon, 1990).

S3, however, does not connect with the other participants at all. She does not acknowledge or engage with either student. Her body form closes them out as does her hair, acting like a curtain. In frame 14 she pushes back her hair, straightens her body, shifts in direction, gazes at the teacher and

announces 'This is looking satellite'. Her statement using the relational process 'looking [like]' makes a real world relationship. Making a judgement using 'looks like' lowers the modality if compared with 'This is [a] satellite'. However, taken together with her gaze, body position and pointing the statement carries some authority. At that point S1 stops gesturing and speaking and takes up a more closed body position by holding his hands.

Whilst acknowledging their ideational and textual functions, I have discussed the modes here for the way they signal interpersonal relations. Posture, gesture and gaze are used by speakers to prime and emphasise spoken utterances and to show engagement with objects and listeners. Gaze and posture are used to signal attention, engagement and feedback from listeners. As such they have the potential to express the authority of and affinity between speakers: a modal function.

5.6 Extract 4 Changing roles: contextual variables and communicative resources

- 1 T: S1 [laughs] what about picture A?
- 2 S1: (2) uh == (2) ==
- 3 S2: == flight ==
- 4 S1: == it seems uh an helicopter ==
- 5 T: == uh huh ==
- 6 S1: == or a something uh which uh fly ==
- 7 T: == ok ==
- 8 S1: which uh could fly (.) ==
- 9 T: == uh huh ==
- 10 S1: == and uh it seem it seems a thing a of a cartoon or ==
- 11 T: == right ==
- 12 S1: = something like this (1) ==
- 13 T: == hmm (2)
- 14 S1: == ↓ cartoon or ==
- 15 T: == uh yeah uh uh ==
- 16 S1: == it seems it doesn't seem a(hh) re real
- 17 T: right (1) hmm I I heard you saying something [indicates S4]
- 18 S4: ah it's made by < it is made by > (.) Leonardo da Vinci (.)
- 19 S4: isn't ==
- 20 S1: == °really°? ==
- 21 S4: == ↓ maybe (.)
- 22 S4: isn't it?
- 23 == [general laughter] ==
- 24 S2: == are you serious? ==
- 25 == [general laughter] ==
- 26 T: == OK right ==
- 27 S3: == (inaudible) ==
- 28 S1: == °I don't know°

In this section I will discuss how stages in a lesson affect student roles and thus the communicative resources they use. I will also show how the modes of posture, gaze and gesture can be co-expressive

with speech as a means of expressing affinity with propositions and can help to soften spoken language to reduce confrontation. The higher-level action is an initiation, response, feedback exchange; lower-levels actions speech, posture, gaze and gesture.

There were five students in the class: one Italian female, two South Korean females and two Japanese males. The students were divided into two groups: the Italian female (S1), one South Korean female (S2), and one Japanese male (S3); one Japanese male (S4) and one South Korean female (S5). The dialogue is taken from a whole class feedback session following group work (see *Appendix 6: Classroom data CD ROM – Richard (1) 03.09* for original data). The feedback takes the form of initiation, response, feedback. ‘Picture A’ refers to the sketch by Leonardo da Vinci’s flying machine.

S1 gives a lengthy response to the teacher’s prompt ‘S1 [laughs] what about Picture A?’ From the dialogue we can see that S1 is quite uncertain about what the design in the picture is and correspondingly her spoken reply contains a number of modal operators. She uses the hesitator ‘uh’ frequently; breathiness (Van Leeuwen, 1999) in line 16; vague language in line 6 and 12 ‘something’ and ‘something like this’; the modal verb ‘could’ in line 8; a contrasting conjunction ‘or’ in lines 10 and 14 ‘cartoon or’, which again acts as a modal structure. As she is not certain what the object is, the contrasting conjunction in these extension moves (Eggins and Slade, 1997) allows her to remain tentative, and so what she appears to be saying is ‘it could be a cartoon *or* it could be something else’.

Of particular interest are the structures ‘it seems’ and ‘it doesn’t seem’ in lines 10 and 16. The verb ‘seem’ is a relational process. As discussed previously in *5.3.1 Modality: linguistic markers of interpersonal relations* relational processes posit that *x is y*. With ‘seem’ the speaker’s judgment is modalized as rather than stating that *x is y* the statement takes the form *x seems to be y*. Thus with the verb ‘seem’ an effect is ‘produced in the speaker by external events of which he is merely an interpreter’ (ibid: 206). So S1 is only interpreting what the picture, i.e. ‘it’, ‘seems’, rather than what she herself, ‘thinks’ it ‘is’ or could be. In use of the verb ‘seem’ allows S1 to distance herself further and reduce her affinity with the proposition.

The use of negation with this verb adds to its modality. Kress and Hodge point out ‘Simple negation is the exemplary form of modal operation’ (1979: 144). To say what something ‘doesn’t seem’ is a highly modalized structure.

What is also of interest in regard to her comments ‘it seems a thing a of a cartoon’ and ‘it doesn’t seem a(hh) re real’ is the fact that the picture is a sketch. In their work on visual media Hodge and Kress (1988) and Kress and Van Leeuwen (2006) detail how some images are conceived as more real than others. The credibility given to a message or image and thus ‘what can be regarded as credible

and what should be treated with circumspection' (Kress and Van Leeuwen, 2006: 154) is made on the basis of 'modality markers'. I have briefly discussed this in reference to colour saturation in *Chapter 4 Terminology and theoretical underpinnings*. Here I make reference to the picture as a sketch. That the picture is a sketch does not necessarily mean that it conveys low modality; to some extent this will depend on what the image is trying to convey, the context and the purpose of its composition. But that the picture is a sketch and not a glossy photograph impacts on its sense of 'realism' and thus whether or not it conveys high or low modality. When we talk of 'real' it should be remembered that 'what is regarded as real depends on how reality is defined by a particular social group' (ibid: 158). In this instance the juxtaposition of the sketch with the photographic Pictures B and C perhaps decreases its sense of realism. Indeed for S1 the picture 'doesn't seem a(hh) re real' and so perhaps classifying it becomes more difficult. It is not surprising then that she uses expressions such as 'it seems', 'it doesn't seem real' and 'a cartoon'.

I now turn my attention to modes S1 uses co-expressively with the spoken mode.



Plate 6

In the picture (*Plate 6*) above S1 is making a gesture. Her right hand rotates in a circular movement suggesting it is iconic gesture: a referential symbol presenting the action of the helicopter blades. Iconic gestures 'present images of concrete entities and/or actions' (McNeill, 2005: 39); in so doing they embody aspects of semantic content (ibid).⁴³ S2 and S3 appear to visually attend to the semantic content of the gesture as they focus their gaze on it. The gesture and her spoken utterance are co-expressive, they refer to the same entity. However, as I have discussed above gestures are not strictly bound into categories and I want to suggest there is something in this gesture that is more than iconic. The peak of the gesture stroke occurs on 'helicopter' and remains in a post-stroke hold (Norris, 2004)

⁴³ See *Plate F* and *Plate G*, *Appendix 12* for further illustration of iconic gesture.

while she says 'or something'. At this point the iconic gesture takes on more of a metaphoric role. She is not sure if the design is a helicopter and so 'juggles' this idea with her open, outstretched hand. My interpretation of this gesture is that it is representative of 'helicopter' but also it is teasing out an idea; a gesture mirrored by the low modality of her verbal expression 'it seems an helicopter or a something'.



Plate 7

In this picture (*Plate 7*) S1 is using a conduit metaphor (McNeill, 2005) to 'hold her thoughts/ideas' and present them to the discussion.⁴⁴ However, this idea seems to be something she is wrestling with. She is using both hands and making small circular movements as if she is 'weighing up' her ideas. It is a movement reminiscent of her gesture in the previous frame and the low modality of her spoken utterance mirrors this 'it seems a thing a of a cartoon'.



Plate 8

These last three frames (*Plate 8*) come after her initial extended turn has come to an end and the teacher responds with 'uh yeah uh uh' in line 15. 'it doesn't seem a(hh) re real' is a prolonging move with logico-semantic relations (Egins and Slade; 1997) to her previous move. Such moves elaborate, extend or enhance by clarifying, restating, exemplifying, qualifying or modifying prior information. 'it doesn't seem a(hh) re real' clarifies/qualifies her comment 'it seems a thing a of a cartoon or something

⁴⁴ See also Kendon (1995: 275) on how fingers and hand can enact metaphoric gestures of 'drawing different things together'.

like this' in lines 10 and 12. During S1's turn the teachers had been reacting with registering moves providing 'supportive encouragement' (ibid: 204). The intonation on these is typically rising and thus the teachers convey that he is aware that something is to follow and is inviting S1 to continue. However, in line 15 the teacher's intonation becomes fall-rise suggesting 'response with reservations' (Roach, 1983: 139). (See below where \vee represents fall-rise intonation.)

15 T: == \vee uh \vee yeah \vee uh \vee uh ==

Eggs and Slade point out that prolonging moves 'seem to pre-empt possible challenges or queries' (ibid: 196) and can be used both defensively and assertively. Arguably, then, S1 picks up on the intonation pattern of the teacher and in so doing attaches a prolonging move to defend/assert/clarify her previous move. The use of negation lends her statement high modality and asserts her authority on it. The purpose of the activity was to discuss what the designs in the pictures were. But, for S1, this picture 'doesn't seem real' and therefore defies classification.

As she makes this statement her posture changes and again we can see this in the three frames above (*Plate 8*). Before this sequence she had been gesturing with her hands and her chin was raised in an open posture (*Plate 7*). Conversely now she is sat back, her left arm across her body, her chin lowered and her right hand resting on her lips and chin. It is a closed posture. If we take the utterance 'it doesn't seem a(hh) re real' to be a defensive prolonging move,⁴⁵ then the closed form of her body is co-expressive of this. Her gesture towards the picture at 3:41 in the middle of the three-frame sequence is a deictic movement of her head and right hand. This movement functions referentially (ideationally), interpersonally and textually in relation to herself, the picture and the surrounding discourse and could be glossed 'it doesn't seem real, this thing/picture'.

The next sequence of pictures (*Plate 9*) corresponds to lines 17 and 18 of the feedback session:

17 T: right (1) hmm I I heard you saying something [indicates S4]

18 S4: ah it's made by < it is made by > (.)

⁴⁵ Further examples of prolonging moves can be found in *Appendix 11*. For example, *Transcript B* 'S1 Yeah ((gestures to head)) I can see before (.) yeah? ((shrugs))' where S1's 'yeah?' addresses the teacher directly seeking clarity on her previous move. Also in *Transcript C* 'S2 Signal transmission [S1 yes] something like that' where 'something like that' can be understood as a modalized expansion on the previous move 'signal transmission'.

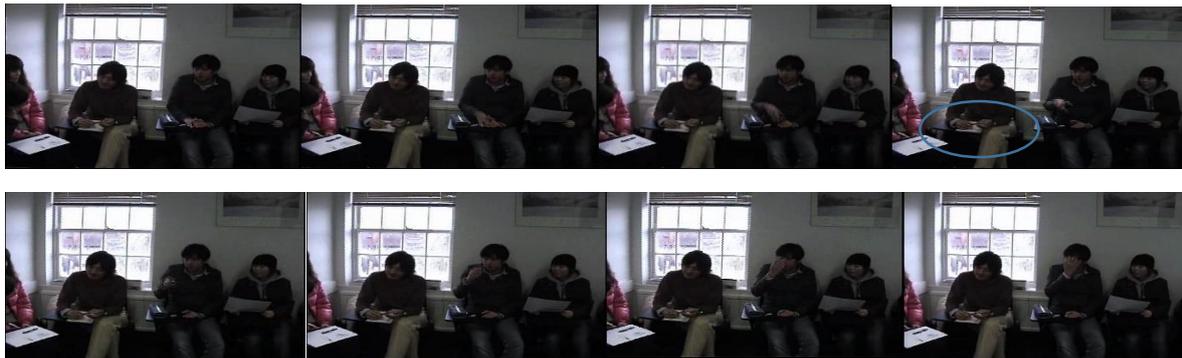


Plate 9

S4 had been discussing with S5 about the pictures. S5 had made an observation that there was a transport theme to two of the pictures and based on that they discussed what Picture A might be. S5 suggested that it is 'a kind of space ship'. S4 follows this up and suggests 'maybe alien will use it'; he adds 'I think this is one kind of UFO () but I'm not sure'. In doing so they entered into a discussion to co-construct what the design could be. However, to the teacher S4 takes a different approach. He knows what the design is and offers that information up. I have chosen this particular sequence for the way the gesture and gaze soften the spoken utterance. The pictures scan from left to right and are timed at twentieths of a second.

Again we see high modality in the relational clause identifying what the design is. If we take this feedback session to be IRF, which is arguably how the students perceive it, then S4 is in effect supplying the information that S1 could not. To soften the impact of his statement he does two things. Firstly, he uses a pointing gesture to reference the picture. It is interesting that he does not point to the picture that he and S5 had been discussing. His gesture references the picture that S1 had herself pointed to. In doing so he uses what McNeill refers to as 'the gesture space' (2005). And this sets up a relationship between himself, S1 and the picture. However, he does not extend his pointing gesture too far into this shared space (*Plate 10* see below). The full extent of his 'stroke' can be seen in frame four above in which the gesture space is indicated with a circle.⁴⁶ Arguably S4's gesture is less invasive than if he had extended his reach across S3 and pointed directly at the picture. His gesture has a referential, textual and interpersonal meaning and arguably modalizes his spoken utterance.

⁴⁶ See *Plate D* and *Plate E*, *Appendix 12* for further examples of participants gaze in conversation attending to the gesture space.



Plate 10

Secondly he avoids direct eye contact with S1 or any of the other participants including the teacher. Gaze has a number of functions such as registering attention, signalling turn-taking moves and emotions such as intimacy, involvement, dominance and even aggression (Forgas, 1985). As I have indicated, S4 is aware of his 'challenge' to S1 and so he softens the authority of his statement by avoiding eye contact. Like his gesture, his use of gaze has an interpersonal function and it acts as a means of modalising his spoken utterance.

In these next two frames (*Plate 11*) we can again see how gaze becomes part of interpersonal relations. S4 engages the teacher with his gaze when he says at the end of his move 'ah it's made by < it is made by > (.) Leonardo da Vinci (.)' in line 18 (see below). This is perhaps not surprising as S4 is responding to the teacher and his adjustment of gaze towards his interlocutor is a typical way of signalling an end of turn (Norris, 2004) as well as having a monitoring function (Kendon, 1990). The next frame, however, shows him lowering his head and gaze as he says in a lower pitch 'maybe (.) isn't it?' This response comes after S1 has reacted to his statement on Leonardo da Vinci by saying 'Really?' S4 has clearly acknowledged S1's spoken utterance and is responding to her with his head is slightly turned in her direction. His gaze and the modal adverb and question tag of his spoken utterance are co-expressive of his attempt to reduce the authority of his statement.



Plate 11

We can also see a smile emerging on his face and in the next two frames (*Plate 12*) there is general laughter from all of the five students in the group. Laughter here counts too in regard to interpersonal relations as it reduces friction. S4 initiates the laughter and defuses any possible friction caused by his 'challenge' to S1.



Plate 12

The interaction presented in this section comes in response to the teacher's question: 'S1 (laughs) what about Picture A?' In asking for feedback from the students the teacher has set up an initiation, response, feedback (IFR) exchange. Ellis (1994) points out that IFR exchanges tend to dominate classroom discourse, especially in teacher-controlled environments. The ESA teaching model is teacher-controlled (Samuda, 2001; Ellis, 2003). The question the teacher uses is a referential question requesting information (see Ellis, 1994). However, the kind of information he is requesting is restricted for two reasons. Firstly, an objective of the lesson is to practise using modals for present deduction (Albery, 2008: 98) and secondly, he knows what Picture A is. So arguably he is looking for one of two things: the use of modals for deduction or the 'correct' answer.

As I have pointed out ESA and IFR are common in the language classroom. Ellis points out that learners are often restricted to a responding role (1994). The students are familiar with ESA and IFR and following the teacher's initiation they take on their respective roles; firstly S1 provides feedback and, following her, S4.

Before moving on I want to draw attention once again to S4. I have discussed in *3.4 Context of situation revisited: field, tenor and mode* how '[t]he context plays a part in what we say; and what we say plays a part in determining the context' (Halliday, 1978: 3) and have shown how contextual variables play a part in this. S4 makes two contributions regarding Picture A. Firstly he discussed it with his partner S5, and then responds to the teacher's initiation. In discussion with his partner he says: 'maybe alien will use it () I think this is one kind of UFO () but I'm not sure'. When he responds to the teacher he says 'ah it's made by < it is made by > (.) Leonardo da Vinci (.)' The first thing that stands out is the content is clearly different. What is also clear is the way he commits himself to both propositions. With his partner he uses a number of modal operators: 'maybe', 'I think', 'one kind of', 'I'm not sure' to indicate what the picture might be; however, when he address the teacher he uses a relational clause to say what the picture is.

The point I want to make is that the contextual variables of field, tenor and mode differ in both contexts. He is no longer taking part in a discussion but in an IRF exchange initiated by the teacher.

He is no longer discussing with a fellow student but answering the teacher's question. Correspondingly the modes of interaction, i.e. spoken language, gesture, gaze and posture will reflect these different role relations and his angle of representation regarding Picture A. I will return to this below.

5.7 Rounding up

In this chapter I have discussed four interactions of learners discussing a picture of Leonardo da Vinci's flying machine from *Language Leader Intermediate*. Prior to collecting this data I had observed that during this activity learners tended not to use the suggested forms provided by the course book writers. I was aware that they knew the rules of form, meaning and function of modal verbs but when it came to extended spoken discourse they were typically absent. However, far from sitting in silence the learners always engaged with the activity to a greater or lesser extent. I was interested, therefore, to discover how the learners approached the activity, and keeping in mind the interpersonal nature of interaction, what modal resources participants used. Further, recent studies in multimodal communication and disciplines that view verbal language as not the only mode, or indeed central mode of interpersonal communication, led me to wonder what other modes were being used to convey meaning during the activity and in particular how modality was being articulated through configurations of modes.

In the first extract I showed how the social roles of participants are realised through speech functions. How the type and distribution of speech functions, for example, initiations, responses, questions and statements enables speakers to show affinity or lack of affinity with each other. I remarked on how different the dialogue was from the course book dialogue based on the same picture, the course book dialogue being used to present certain modal forms. And, with this in mind, I looked at how the two speakers were using modality markers to express interpersonal relations and modalize propositions. Using a social semiotic approach to sound I finished by discussing and providing examples of how certain sounds, in combination with other semiotic modes and viewed in context, had the potential to signal modality and were used by the speakers to do so.

The second extract was a short IRF exchange between a teacher and a pre-sessional student discussing the same picture. Using this dialogue I showed how a multimodal ensemble of speech, gesture, gaze and head movement provided a fuller understanding of the interaction than a verbal account alone could. Drawing on Halliday's notion of ideational, interpersonal and textual metafunctions I pointed out that whilst modes integrate, meaning is not always distributed evenly as

each modality has the potential to become dominant or subordinate and may at times act as a resource to communicate, to greater or lesser extent, its ideational, interpersonal or textual function.

In the third extract I focused on posture, gesture and gaze and showed how they were used by the students to prime and emphasise spoken utterances and to show engagement with their interlocutors and the object under discussion. I gave examples of how listeners used gaze and posture to signal attention, engagement and give feedback. With these functions they have the potential to express the authority of and affinity between speakers; a modal function.

With extract four I discussed how the modes of posture, gaze and gesture can be used co-expressively with speech as a means of expressing affinity with propositions and help to soften spoken language to reduce confrontation. Then focusing on one particular student I showed how changes in the contextual variables of field, tenor and mode during different stages of the lesson affect interpersonal relations and thus language choice.

Whilst these four extracts provide only a small sample they are typical of the interactions I observed and recorded. Inaudible speech, hesitations, pauses, silence, declaratives, questions, use of present tense, vague language, deixis, repetition, hedging, sounds such as elongation, softness/loudness and tempo, along with laughter, i.e. examples of the modal resources that I have discussed in this chapter and at other points in this thesis, can be found and cross-referenced on the CD in *Appendix 6* and in the transcripts contained in *Appendix 11*. What stands out most in terms of how the dialogues conform to the lesson objective, is the near total absence of modal verbs (see *Appendix 11* for examples). It could be argued that these dialogues come from learners of English. And while the learners may have been able to use the suggested forms in practice exercises, the forms have probably not entered into their interlanguage and so are unlikely to be activated in spoken interaction. For this reason I gave the same activity to native speakers of English and the outcomes were similar;⁴⁷ the participants in the first extract are a native speaker and an 'expert speaker' of English: in *Extract 1* there are no modal verbs. There are, of course, other markers of modality and I have discussed these above.

The interactions in all four extracts show interpersonal relations. They show that modality, articulated through a variety of verbal structures and other communicative modes, is used to convey interpersonal relations through the degrees of affinity speakers assign to both their own and their

⁴⁷ I point out that the purpose of this part of my research is not to make comparisons with a control group of native speakers. During the time that I have been engaged with this research topic I have replicated the discussion activity with students, friends and colleagues – with just about anyone who would afford me a few moments of their time. On most of these occasions with friends I did not make video or audio recordings.

interlocutors' utterances. Furthermore they show that interpersonal relations are not constant over the duration of an interaction, and that speakers are constantly (re)positioning themselves vis-à-vis their interlocutors. They show too how interactions are made up of modal ensembles and that these have meaning potential. Modes can be co-expressive or modes can have a dominant or subordinate role in relation to other modes.

Carrying out the above analyses revealed to me the difficulty of not only transcribing but also talking about communicative modes other than the verbal using verbal language. Transcribing spoken verbal interaction is time consuming and as the above examples show is often limited to certain features the analyst wants to focus on. However, there are established transcription conventions such as those I have borrowed from conversation analysis that have made the task of transcribing spoken language less of a foray into the unknown. There is also a metalanguage for discussing spoken verbal interaction, i.e. verbal language. In comparison it has not been easy to verbally describe, for example, hand gestures and to transcribe multimodal ensembles. I have found myself using different formats to do this to highlight and draw attention to certain modes for discussion but I have attempted to make them as reader friendly as I can. Just as significantly I am also aware of the possibility of criticism being levelled at descriptions of other communicative modes using terminology and concepts formulated for verbal language. The term 'utterance', for example, borrowed from linguistics to describe a co-occurrence of modes such a speech and gesture to form one communicative act as well as the notions of the ideational, interpersonal and textual metafunctions from systemic functional grammar applied to, for example, gestures and pointing (cf. Jewitt, 2009). These points raise important questions for how multimodal data is transcribed and discussed. However, I have tried to remain consistent with my methodological framework and argued why I feel it appropriate.

Finally I am aware that modes such as gaze, proxemics and gesture are cultural as well as individual and have therefore been cautious not to overgeneralize; indeed as I have needed to be cautious and avoided assigning meaning to modes in isolation. My analysis has been based on modal ensembles within specific contexts and it is from these that I have inferred meaning.

Chapter 6 Teaching implications: applying my research

6.1 Lead in

The three main areas of my thesis can be found in *Chapter 1 Modality: course books and beyond*, *Chapter 2 Context: finding my way through the black hole* and *Chapter 5 Four extracts of data*. I briefly revisit them here as they have specific relevance to *Chapter 6 Teaching implications: applying my research*.

In *Chapter 2 Modality: course books and beyond* I centred my discussion on how modality is often represented in ELT training and teaching materials and drew on one particular course book, *Language Leader Intermediate*, as a case in point. I argued that introspective selections of modals, such as those that appear in that course book, do not tally with the findings of corpora which suggest modal finite verbs are not the most frequent way of modalising propositions in spoken English. I also highlighted the absence of a focus on the interpersonal function of modality in approaches that rely on modality as form and limit it to functions of possibility and argued that neglect of this key aspect of modality limited learning. Practising form and function at the level of the sentence is not the same thing as using modals in extended discourse such as a discussion in which interpersonal relations are very much at the fore.

In *Chapter 3 Context: finding my way through the black hole* I provided a review of some of the major works on context that have filtered down to English language teaching. I noted, however, that for published ELT course books such as *Language Leader Intermediate* the term context is used to refer mainly only to co-text (see 2.4 *Modality in spoken language: insights from language corpora* for a list of similar course books). I argued that simply providing surrounding text, often purposely written, to present linguistic structures is not the same thing as providing a context in a Hymesian or Hallidayan sense in which variables such as the participants, their social roles, purpose of communication, etc. affect language choice. I provided the following example (above in 3.2 *Context of situation: foundations*) from *Language Leader Intermediate* to show that whilst the modal forms have been contextualised the arbitrary choice of modals cannot be said to tell us much about why a 'pair of students at a museum of design' (Cotton et al., 2008: 75) made these language choices any more than if they were made by a couple discussing pictures in a magazine at home over breakfast on a Sunday morning.

G: What's that?

S: I've no idea. What could it be?

G: It might be a spaceship. Well, a toy spaceship. It's the right shape.

S: No, I'm certain it's not that.

(Cotton et al., 2008: 176)

I also made reference to context being dynamic. And I drew on the notion of contextual variables to show that context is not just something that is 'there' as an ornamental backdrop to language but in part shapes and is shaped by language choice.

These two chapters then were a way of aiding my understanding of modality and context. *Chapter 5 Four extracts of data* was instigated by a curiosity to discover what communicative resources learners used to carry out a speaking activity. The activity was taken from the lesson of modals of present deduction in *Language Leader Intermediate* (Cotton et al., 2008). Having used the materials extensively I had noticed that the target language of the lesson was usually absent from the final speaking activity intended for its activation. Therefore I recorded EAP students carrying out the activity and identified a number of communicative modes the learners used to make meaning and articulate modality. The sample of data was small but revealed that modality is articulated through a variety of verbal structures other than modal auxiliaries in addition to other communicative modes; modality is used to convey interpersonal relations through the degrees of affinity speakers assign to both their own and their interlocutors' utterances.

In this chapter I pull together the discussions and findings from the three chapters referred to above to discuss teaching implications. In *1.5 A note on the written style of my thesis* I note that my research is 'practice-derived' with both English language teachers and researchers in the field in mind. The writing of my thesis has derived from practice and theory generated analysis and is inevitably oriented towards practice. I look at corpora and modality, interpersonal meaning, learners and context, multimodality and finally put forward three different teaching strategies which can incorporate the ideas discussed.

6.2 Corpora and modality: using corpora to inform teaching materials

I have noted above a common tradition found in many published ELT course books of following a 'form-focused grammatical syllabus' (Hughes, 2010: 406). And very often a form-focused syllabus follows a similar pattern. As Hughes comments 'the typical table of contents of a traditional intermediate grammar textbook will move through, say, rank-order topics from words and word classes, to phrase structure, to simple and complex sentences, and run through the tense system from 'simple present' to 'talking about the future: *will* versus *be going to*' (Hughes, 2010: 406).

Discussions surrounding the selection of grammatical content in ELT course books have been long running. For example, in the early days of CLT Wilkins made reference to the introspective nature of language choices which were not based on empirical research (Wilkins, 1976). More recently with the

establishment of large language corpora, arguments about the grammatical content in course books has taken on further impetus. It is widely acknowledged that course book writers rely on intuition and invented examples (McEnery and Wilson, 1996; Kennedy, 1998; Cheng, 2010) and frequently observed that course book language differs from actual language use (Carter, 1998; Chambers, 2010; Cheng, 2010; McCarten, 2010). Chambers notes that course books language is significantly different from actual language use particularly in relation to spoken language' (Chambers, 2010: 345) and both McCarten (2010) and Cheng refer to 'disparities' and 'discrepancies between the information found in grammar materials and the real-life language use that learners encounter' (Cheng, 2010: 323 see also McEnery and Wilson, 1996).

This is perhaps not surprising if we consider the history of English language teaching discussed earlier and in relation to spoken language the difficulties, until the last 40 years, of recording it for the purpose of analysis; along with the prestige afforded to written language as the preferred mode for linguistic analysis (Halliday, 1989; Kress, 1982; Tomlinson, 1998). Further, as Biber et al. indicate, language studies can be divided into two main areas: studies of structures and studies of use (Biber et al., 1998: 2). They argue that the traditional study of language has been that of structure with students being presented with 'concocted', 'unreal' dialogues (Carter, 1998: 47) in which language is 'put on display' (Widdowson, 1978: 53). And it is this study of structure that finds its way into many ELT course books in which the form-focused syllabus referred to above takes precedence.

The significance of this for language teaching is the argument that the study of structure alone leads to 'narrow notions of the linguistic competence of an ideal speaker-hearer in a homogenous society' (Paulston, 1992:38). The study of use, on the other hand, drawing on 'naturally occurring texts...created by users of the language for a communicative purpose' (Conrad, 2005:394) leads to communicative competence, including 'not only the linguistic forms of the language but also its social rules, the knowledge of when, how, and to whom it is appropriate to use these forms' (Paulston, 1992:98). Thus it has been suggested that the teaching of spoken grammar, drawn from spoken language corpora, can provide 'a missing link' (Cullen and Kuo, 2007: 361) between linguistic and communicative competence. Cheng points out that '[c]orpus data have helped researchers to identify patterning that differs from traditional models of the English language' (2010: 324); models which lack 'an empirical basis' (ibid). What is being argued is for corpus-based accounts of language use to be included in language teaching materials.

Biber et al. describe the characteristics of corpus-based analysis as follows:

- it is empirical, analyzing the actual patterns of use in natural text;

- it utilizes a large and principled collection of natural texts, known as a “corpus” for the basis of analysis;
- it makes extensive use of computers for analysis, using both automatic and interactive techniques;
- it depends on both quantitative and qualitative analytical techniques. (Biber et al., 1998:4)

The implications of this for teaching and learning are now widely discussed in the literature (e.g. McEnery and Wilson, 1996; Kennedy, 1998; Aijmer, 2009; O’Keefe and McCarthy, 2010). Here I will briefly gloss the above.

Empirical analysis of natural text is arguably beneficial for language teaching for a number of reasons; perhaps the most pressing being ‘[o]ne of the major benefits of the corpus-based approach to grammar teaching materials is that it can highlight the difference between assumptions about language structure in the abstract and what is found in real-world use’ (Hughes, 2010: 402). As I have already pointed out this is of importance to language learners if the goal is communicative competence i.e. knowledge of how to use language appropriately and effectively. To this end Biber et al. make clear that ‘[i]nvented examples that sound artificial – a familiar feature of many other grammars – are entirely absent’ from their substantial *Longman Grammar of Spoken and Written English* (LGSWE) (Biber et al., 2002: 3).

In terms of teaching approaches the use of ‘natural texts’ and ‘real-world use’ of language fits in with CLT’s advocacy of using ‘authentic material’. Kennedy (1998) supports the use of corpora to provide authentic examples of language in context to avoid presenting learners with ‘oddities’ and language items of ‘dubious use’ (p.282). Recourse to large amounts of naturally occurring language also sits well with a functional perspective, which emphasises language as a system of choice (Conrad, 2010: 227; Halliday, 1978); the analysis of a corpus is able to reveal the choices language users make in given contexts and registers. The use of corpora is also accessible to learners and facilitates inductive and exploratory approaches to learning, promoting learner autonomy and ‘encourages self-access’ (Kennedy, 1998). Data-driven Learning, which I will discuss below, uses the ‘tools and techniques of corpus linguistics for pedagogical purposes’ (Gilquin and Granger, 2010: 359).

The use of corpora is about more than simply establishing frequencies (Kennedy, 1998). Quantitative accounts of language allow us to see that, for example, the modal verb *must* is not a frequently occurring modal in conversation. Using concordance software we are able to access a word’s textual environment (Sinclair, 1991: 32) to look at its distribution across registers, explore contexts, discover meanings, and test out hypotheses. The modal verb *must* has both epistemic and deontic meaning.

Use of corpora reveals its use with deontic meaning is especially rare in conversation and this is likely due to the interpersonal nature of spoken communication and the likelihood of it leaving too strong an impression in face-to-face communication (Biber et al., 2002: 181). Further we are able to reveal information about how structures collocate and colligate: *must* with progressive aspect is very rare. Corpora then allows us to see structural characteristics as well as discourse patterns of use (Cheng, 2010: 323).

In 2.4 *Modality in spoken language: insights from language corpora* I discussed the selection of modal structures presented in *Language Leader Intermediate* for present deduction and compared them with the findings of the corpus-based *Longman Grammar of Spoken and Written English* (LGSWE) (Biber et al., 2002). To recap I will list the modal forms and briefly comment on their salience in the spoken register of the LGSWE.

Must might could can't – these four modal auxiliaries are distributed evenly in the three *Language Leader Intermediate* dialogues used to present the grammar focus of the lesson. The LGSWE tells us that *can*⁴⁸ is by far the most common modal verb at approximately 2500/million words; *can* is followed by *could* at around 1800/million words. *Might* is significantly less at about one third of the frequency of *can* and *must* is even lower at about one quarter of *can*.

I'm sure/certain that ... - post-predicate *that*-clauses are very common in conversation, however, around 90% of these are controlled by verbs e.g. 'Did you know [that] Kathy Jones had a brother here?' (Biber et al., 2002: 312). The LGSWE informs us that the most frequent adjective and the only one that is especially common is *sure* at 100

Maybe/perhaps – *probably* is more common than *maybe* or *perhaps* in conversation; *perhaps* is more frequent in writing. The LGSWE reveals a significant number of epistemic adverbials in conversation but these are less than modal auxiliaries and complements as a means of signalling modality. Epistemic modal adverbs make up only a small percentage of epistemic adverbials. Overall the modals *can*, *could* and *might* are more common than the modal adverbs listed here.

It's (not) possible that ... extraposed *that*-clauses make up only a fraction of *that*-clause types in conversation with the most frequent controlling adjectives being *clear*, *(un)likely*, *(im)possible* and *true*. Others pertaining to certainty include *certain*, *doubtful*, *evident*, *false*, *inevitable*, *obvious*, *plain*, *probable*, *right* and *well-known*.

⁴⁸ As indicated earlier some uses of *can* are ambiguous signalling either ability or possibility which might help to explain its frequency.

Modals with a continuous form e.g. 'She might be feeling ill' and 'You must be joking!' (Cotton et al., 2008: 75) are very rare.

From the corpus evidence in the LGSWE it is apparent that the selection of modal structures in the course book are not the most frequently occurring in spoken English. Indeed some of them are very rare. Whilst corpora evidence does show that modal auxiliaries are the most common means of signalling modality in spoken English, the modal verbs themselves are arguably not the best selection based on frequencies and register distribution. I pointed out that ELT course books such as *Language Leader Intermediate* follow a course book tradition, a tradition in which choices of language have often been based on intuition rather than evidence. The modal structures in *Language Leader Intermediate* have not been made on corpus-based evidence and accordingly reflect the disparity between real-life language use and linguistic structures presented to language learners mentioned above. Again perhaps this is not surprising if we consider the course book and the type of lesson sequence used to present and practise the modals. I have indicated that the lesson on modals of deduction follows an ESA sequence and so arguably the purpose of the lesson is to practise linguistic structures. Whilst this is not a problem in itself it goes against the claims of the course book writers who say *Language Leader Intermediate* encourages an inductive approach to grammar, and uses a task-based and communicative approach (Albery, 2008).

If the aims of a course book are to promote an inductive, task-based and communicative approach to language use, I argue for recourse to language corpora. Hughes (2010) outlines the following benefits of using corpora to inform teaching materials, which are arguably far from being idealistic:

- Corpus-based grammar materials and approaches to teaching grammar that are based on corpora can be adapted to suit several very different approaches and individual learning styles. For the analytical individual, whether teacher or student, corpora provide tangible quantitative data on which to test out ideas about language or from which to develop materials for the grammar classroom.
- At a more practical level and whatever their approach, corpora can provide teachers with an immense variety of samples of language which can be adapted for any form of teacher-developed materials.
- Corpus-based grammar materials lend themselves to task-based and communicative activities particularly well as they are readily adaptable to group and project work and autonomous learning contexts. The language learner can be encouraged to develop their own questions about grammatical points and investigate the answers for themselves via corpus examples.

- The teacher facing mixed ability group will also find corpus-based grammar materials beneficial. With some preparation the same task framework can be used for groups with the same instructions but applied to different language points.

(Hughes, 2010: 405)

Returning to grammatical content I argue for a rethinking of which modal structures, for example, are selected for study is called for. We have seen how the selection of modal forms presented in *Language Leader Intermediate* map onto data from spoken language corpora and discrepancies are startlingly evident. In addition there are irregularities between modal structures used in spoken and written English i.e. extraposed *that* clauses and the adverb *perhaps* are more common in writing. And so it needs to be pointed out that this is more than just about preference as what gets included, or left out, of language teaching and learning has implications for language learners. A study by Romer (2006) on English conditional *if* clauses found that ‘misrepresentation of conditionals coupled with inconsistencies across textbooks compounded difficulties faced by German learners of English’ (cited in Cheng, 2010: 325). Skehan (1998) cites Biber (1988) and Biber, Conrad, and Rippon (1994) who suggest that misrepresentations of language in descriptive and pedagogic grammars can be misleading. I argue that the modal structures common to written English which are presented to the learners to carry out a speaking activity are unsuitable and unhelpful. Biber et al. contend that conversation expresses stance and so recourse to spoken language corpora can provide evidence of how stance is actually signalled. For example, the mental process verb *think* is the fifth most frequent lexical verb in the LGSWE at approximately 2,400 counts per million words; more frequent than the modal verbs *might* and *must* and thus arguably useful for language learners as a means of expressing stance.

Reports on using corpora in language teaching are insightful. Flowerdew (2010), for example discusses the potential for enhancing L2 writing and I make reference to my own action research project which used corpora to assess EAP students’ use of epistemic modality compared with published academic writing as a means of providing corrective feedback. I compiled a corpus of student essays to discover how they used stance in their written arguments and my findings were in accordance with Hewings and Hewings who claim that ‘student writers make a much greater and more overt effort to persuade readers of the truth of their statements than do the published writers’ (2004: 114). The effect of this is in the students’ essays was the tendency to come across as being too forceful in their arguments and opinions. Examples from student essays and published academic writing were then compared, analysed and discussed in class in a lesson on writer stance. (See also Coxhead, 2010 on the influence of corpora on EAP pedagogy.) McEnery and Wilson (2009) draw attention to parallel corpora in language teaching which ‘focuses especially on the problems that speakers of a given language face when learning another’ (p.122) and give the example of German speakers learning about aspect in

English grammar. And they give further examples of how corpora has been used to look critically at language teaching materials. Kennedy (1998) discusses the use of corpora for 'studying lexical, grammatical and discoursal characteristics' across different varieties of English and English texts. On a more 'macro level' Granger (2009) and Johansson (2009) discuss corpora and second language acquisition.

Corpora have informed textbooks such as McCarthy and O'Dell's *Academic Vocabulary in Use* (2008) as well as the *Touchstone*, *Viewpoint* and *English Unlimited* series of course books published by Cambridge (Cambridge.org/corpus), *Real Grammar: A Corpus-based Approach to English* by Conrad and Biber published by Pearson, and informed the COBUILD (Collins Birmingham University International Language Database) dictionary. However, in the main corpora research has had little impact on language teaching not only in terms of materials but also in language teacher training. I have made numerous references above to the 'disconnect' between corpus evidence and language teaching course books and as Hughes indicates 'data-driven' materials based on corpora are rare (2010: 401). And whilst Aijmer (2009) points out that corpora 'can have an impact on syllabus design and on the preparation of textbooks, dictionaries, grammars and course-books' (p.7) and do inform dictionaries and grammars, 'with a few exceptions textbooks still shy away from corpora' (ibid). She adds that publishers of language teaching course books are 'loathe' to develop materials based on 'corpus-based insights' (Ibid: 408). In terms of language teacher training Cheng notes (2010: 328) that corpus linguistics and applications yet to become mainstream in language teacher education programmes and language teaching and McCarthy refers to the slow 'filtering down' of corpora evidence to teacher training programmes (McCarthy, 2008). Whilst Cheng and McCarthy refer specifically to language corpora in teacher training these concerns are endemic of a wider division between language research and language teaching; a divide discussed by Erlam (2008) and Tavakoli and Howard (2012). Likewise Hughes refers to a lack of communication between the 'teacher training community in applied linguistics and the research community' (2010: 402) and both Chambers and McCarten highlight the slow process of research finding its way into language teaching and learning (Chambers, 2010: 345; McCarten, 2010).

Despite the reluctance of publishers, course book writers and teacher training programmes to draw on the insights from language corpora, I argue that if communicative competence and language use are goals for language learners, reference to language corpora needs to inform language teaching course books as well as become an integral part of language learning. I will discuss this further in 7.1 *Teaching strategies: research into practice*.

6.3 Interpersonal meaning: uncovering meaning through discourse

Language teaching is not simply about presenting the structures which are most frequent in the language. By simply replacing one set of modal structures with another learners are not necessarily getting any closer to communicative competence if they are provided no instruction on their use. In this section, therefore, I want to expand on the distinction between teaching language structure and teaching language use. This is a significant distinction as it relates to what we teach, how we teach and which activities the learners are to perform. For example if we are teaching a grammatical structure we might first present it to the learners and then ask them to practise it in gap-fill exercises, sentence reformulations (as in *Figure 10* overleaf), or a speaking activity; the way a typical PPP/ESA type lesson is structured. Such an approach helps to ‘manifest our knowledge of the language system of English’ (Widdowson, 1978: 3) and promotes accuracy and correctness of usage: an aspect of language performance. However, as Widdowson points out ‘we are not commonly called upon simply to manifest our knowledge in this way in the normal circumstances of daily life’ (ibid). What we are called upon to produce are instances of language use: ‘the way the system is realized for normal communicative purposes’ (ibid: 18).

5 Rewrite the sentences below using *must, can't, could or might*.

- 1 Raymond Loewy is definitely one of the most influential designers of all time.
- 2 I'm sure this design is by Armani.
- 3 I'm sure this painting is not by Leonardo.
- 4 Maybe this painting is by Picasso, but I'm not sure.
- 5 Not many people are attending the exhibition. I'm sure they aren't promoting it very well.
- 6 This product looks dangerous to me.
- 7 People are not sure if it is a Loewy design.
- 8 I'm sure that designer is working very hard. I've seen a lot of her designs.

Figure 10 (Cotton et al., 2008: 75)

Widdowson is writing from the perspective of *Teaching Language as Communication* (1978) and whilst acknowledging that a knowledge of use must necessarily include a knowledge of usage, knowledge of usage alone is of little utility (1978: 18). Bringing the discussion back to *Language Leader Intermediate* I would argue that the lesson on modals of deduction only teaches about language usage. Although the final speaking activity is intended to engender use of the modal forms, realistically it is little more than an opportunity (or not) for the learners to produce modal structures verbally. The lesson has focused solely on rules of form and meaning ‘Use modal verbs to make guesses (deductions) about the present, based on evidence. The different modal verbs express different levels of certainty’

(Cotton et al, 2008: 146); no account of how the modal system in English is used ‘for normal communicative purposes’ has been provided.

And this is where my research is situated. As I have observed rules of usage seem to be readily grasped but problems occur when it comes to use; for example in extended discourse such as a speaking activity. I argue that the breakdown comes about, in part, from a lack of instruction on the communicative use of modality, i.e. as a means of expressing interpersonal stance. Informing a learner that ‘It must be true’ has the same meaning as ‘I’m certain that this is true’ is only part of the picture. As discussed previously *must* leaves a strong impression as it is bound up with speaker authority and power, it is, therefore, quite rare in face-to-face communication. So a knowledge of the use of structures is essential if the aim is for learners to communicate effectively and appropriately in any given context.

Biber et al. explain that in spoken language speakers ‘have a primary concern for their feelings, attitudes, evaluations and assessments of likelihood’ (Biber et al., 2010: 433) referred to as personal stance. To that end ‘[m]any of the most common grammatical features in conversation are used to express stance, including modal verbs, complement clause constructions, and stance adverbials’ (ibid). Such information provides a useful starting point for teaching the interpersonal function of modality. Egging points out that modality, along with mood, ‘are the keys to understanding the interpersonal relationships between interactants’ (1994: 196). And of course, understanding the interpersonal relationships between interactants, is key to understanding modality.

What I am suggesting in terms of teaching is tying modality in with expressing interpersonal stance in spoken communication. I have discussed elsewhere *A Discourse Approach to Teaching Modal Verbs of Deduction* (Howard, 2010) in which learners are presented with authentic or semi-authentic texts (see 7.2 *Handling real spoken data for teaching purposes*) and discuss the who, when, where and why, i.e. the context, of the modals choices. The argument I formulate is that by exploring aspects of the context in which utterances are made, learners are in a better position to understand language use. Of course, this will require a different teaching method to PPP or ESA and I discuss this below in 7.1 *Teaching strategies: research into practice*.

To illustrate my point about language use in context consider the following dialogue (continued overleaf):

Mother: Alright which cereal would you like for breakfast?

Tim: Oh, mum.

- Mother: Come and choose one then.
- Tim: I don't like it.
- Mother: You've got to have some cereal. You'll get hungry.
- Tim: I won't get hungry.
- Mother: You **must be hungry**, you hardly had any dinner last night. (Gives Tim the cereal) There we are.

<http://corpus.byu.edu/bnc/>

The dialogue has been taken from the BYU – BNC: British National Corpus, a corpus of 100 million words compiled between 1980s-1993. It is an extract from a conversation between a mother and her child, Tim. In the extract the mother is attempting to get her child to eat breakfast. I have tidied up the dialogue and will discuss the reasons for doing this in *7.2 Handling real spoken data for teaching purposes*.

The use of the modal structure 'You must be hungry' accords with descriptions of usage such as 'Use *must* to say that you are certain something is true' (Cotton et al., 2008: 146); it functions as a modal of deduction. We know that *must* is rare in conversation due to the strong impression it leaves. However, this is a mother speaking to her child. She is keen to impress on Tim that he must be hungry as he didn't have dinner and therefore he should eat now. *You must be hungry* is as much about a mother asserting herself as it is about degrees of certainty. Obviously there is much more going on in this dialogue in terms of power and modality and I will discuss this in *7.1 Teaching strategies: research into practice*, I provide this brief example here as a way of showing how the study of discourse relates language use to context and how the interpersonal function of modality can be made aware to language learners.

6.4 Learners and context: adapting to the social construct of the classroom

As well as context being integral to an understanding of language use, it plays an integral part in the realisation of language in the language classroom. I have referred to context as being dynamic, i.e. an interaction between the participants and the contextual variables of field, tenor and mode. Given its dynamic nature '[c]ontext should therefore be viewed not as a natural given, but as a social construct, the product of linguistic choices made by two or more individuals interacting through language' (Kramsch, 1993: 46). The emphasis is on 'social', a theme I have pursued throughout my thesis referring to the work of Halliday and Hymes, for example. I refer to Halliday again: '[t]he context plays a part in what we say; and what we say plays a part in determining the context' (1978: 3) and thus the

linguistic choices 'in turn hold together, control, manipulate, and maintain the social order, that is, the social organisation of classrooms, homes, and workplaces' (Kramsch, 1993: 46). In this section I will discuss context in relation to the activation stage of the lesson of modals of present deduction in *Language Leader Intermediate*. Arguing that context is created by the participants in interaction and thus it should come as no surprise if the linguistic forms they use to carry out the speaking activity differ from the predesignated structures presented by the course book writers. This is understandable if the participants are taken into account for as Kramsch points out '[i]n the classroom, the success of any communicative activity is heavily determined by the way the participants perceive the context of situation and shape it accordingly through their verbal and non-verbal behaviour' (ibid: 50).

I have discussed above that context is not a 'given', static space transferable from one stage of the lesson to another. I argued that language presented at one stage of the lesson may not necessarily be transferred to subsequent stage. Using the lesson from *Language Leader Intermediate* I explained that contextualizing modals structures in a discussion between pairs of students discussing designs at a museum of design, may not naturally give rise to the same modal forms, or even the same kind of discourse, during the activation stage of the lesson in which learners are discussing in pairs or groups in the language classroom. Part of the reason for this, as Kramsch observes, is the way participants shape context through their communicative resources; thus the same activity can give rise to different actualisations (see also Nunan, 2000; Kumaravadivelu, 2003 and 3.5 *The lesson as context: the dynamics of the language classroom* above). In *Chapter 5 Four extracts of data* I showed how the contextual variables of field, tenor and mode shape participant interaction and language choice. With changes in contextual variables, participants take on different roles opening up different language choices. Similarly in 3.5 *The lesson as context: the dynamics of the language classroom* I discussed different 'mismatches' between the ways the learners carried out the speaking activity from the way intended by the course book writers. For procedural, pedagogic and strategic reasons learners took different approaches to engage and to complete the activity. These mismatches gave rise to different contextual variables, different participant interaction and thus different language being activated.

My argument is that learners are not a homogenous group lacking personal characteristics, beliefs, and attributes; in the same way the language classroom is not a uniform construct. As Breen points out:

[a]lthough the language class may be one social situation, it is a different social *context* for all those who participate within it. The culture of the classroom is an amalgam and permutation of different social realities. This means that the content of the lesson (the language being taught) and the procedures of teaching and learning (the things being done) are both continually

interpreted differently as the life of that language class unfolds. The classroom is the meeting point of various subjective views of language, diverse learning purposes, and different preferences concerning how learning should be done. Such differentiation brings with it potential for disagreement, frustrated expectations and conflict.

(Breen, 2001: 129 emphasis in original)

I would add that this is to be expected; it is the norm. The language classroom is a place where teacher/learner, learner/learner mismatches frequently occur, teaching and learning takes different routes than planned, and 'context' is continually shaped and reshaped as the lesson unfolds. However, English language teaching has developed a culture of uniformity and sameness (Skehan, 1998) and this takes place across the board from publishing, syllabus design, teacher training and methodology; the way learners are perceived, teaching methods practised and syllabuses and teaching materials are designed takes little, or no account, of the socio-demographic of the learners, their educational histories, learner attributes, or the way they might be feeling on any given day and thus how they engage with the lesson, their teacher and classmates.

Skehan addresses the issue of uniformity and sameness in some detail and it is worth noting his main points. Publishers have commercial interests and are keen to 'develop a product that targets the widest purchaser group possible' (1998: 260). For this reason 'it is in the publisher's interests to treat all learners as the same, in order that a course book series will not lose appeal to any particular group of buyers' (ibid). Course books are big business and a lot of investment goes into their production (Gray, 2002). There is, therefore, a 'great pressure for it to be successful to return that investment' (McCarten, 2010: 413). For syllabus designers '[t]he units and sequences of syllabus design are regarded as being equally appropriate for all learners, and no account is taken of styles or preferences or abilities which might make some approaches to organising courses more appropriate for some learners than others' (Skehan, 1998: 260).

I have noted that *Language Leader* forms the syllabus at the institution where this research was carried out: the success of the learner is measured on their engagement with the syllabus, not vice versa. Classroom procedures do not explore 'how adapting a particular methodology for different learners types, or using different methodologies with different sorts of learner, might produce better results' (ibid); ESA, like PPP, puts learners in the passive role of receivers of language input wherever in the world it is implemented and with whatever group of learners. Teacher training concentrates 'on how entire classes can be organized; by training teachers how to implement official syllabuses and course books, and by testing in an approved manner. There is little emphasis, in most teacher training courses, on the development of techniques which serve to adapt material to the individual learner, or

in ways of fostering individuality in learning' (ibid). In my introduction I made reference to both the CELTA and DIP TESOL on which published ELT materials are used as a template to linguistic structures and language skills; they also encourage uniform lessons based on PPP and ESA, lessons which are then uniformly assessed by observers and examiners. And finally Skehan claims for reasons of accountability, administrators and educational authorities 'assume that all learners are similar' (ibid). At the most rudimentary level I refer to the 'level setting and testing' of learners into Elementary, Pre-Intermediate, Intermediate, etc.

Bringing this back to my research and the lesson on modals of present deduction, a brief look at the lesson staging and objectives in the Teacher's Book reveals much about the culture of sameness (reproduced from 3.1 *Lead in*):

[L]esson topic and staging

This lesson looks at famous designers and some of the things they designed. Students read a text about a famous American designer. This is followed by a listening on different designs and designers which contextualises the grammar focus of the lesson: modals for present deduction (*can't, must, might/could*). Students then do a series of activities to explore the meaning and form of these modals before using them in a speaking activity. Finally, students do a further speaking exercise to discuss what they would like to design.

Objectives

By the end of the lesson, students should have:

- Practised reading a text to extract specific information
- Practised listening to a text to extract specific information and language items
- Learned more about and practised using modals for present deduction (*can't, must, might/could*)
- Engaged in a group discussion.

(Albery, 2008: 98)

Here we can see how the course book, representing the syllabus and structured around a particular teaching method all presuppose a seamless lesson with a homogenous, 'faceless' group of learners all 'on task'; which on the surface, provides a nice package. However, as a teacher and researcher I know that this is far from the case and I have shown how this becomes evident throughout my thesis. In *Chapter 5 Four extracts of data* the activation stage of the lesson is clearly not what the course book writers had intended in terms of language output and engagement with the activity. However, it

would, be unfair to assess the learners, the success of the lesson, or indeed the teacher, based on these outcomes. Rather it is an oversimplification of the classroom as a social construct inherent in the lesson that leads to a different set of outcomes.

The point I want to make is that by giving learners a set of linguistic structures, pairing them up, and asking them to discuss a number of pictures does not create a 'given' context. In reference to the lesson under discussion here, the speaking activity intended to activate the language focus does not replicate the same context used to contextualise the language focus. The context(s) will be defined, in the main, by the way the learners go about the activity and thus the language they use to do so. There is perhaps some ambiguity as to just what the learners are being asked to do. Are they being asked to *discuss* the pictures, or orally 'activate' the modal structures i.e. using single clause utterances? The course book informs us that it adopts a communicative approach and so one would expect the former. I refer once again to Willis who argues that PPP is suited to teaching 'a small sample of language with the focus on a particular form' (Willis, 1996: 134), an argument which could perhaps be extended to ESA. Whether or not this is the case such teaching models have been critiqued (for example, Lewis, 1993; Scrivener, 1994; Ur, 1996; Harmer, 1998; Skehan, 1998; Ellis, 2003) and it is with this in mind that I propose complimentary approaches to explore and teach the interpersonal function of modality. These I will discuss below in *7.1 Teaching Strategies: research into practice*.

6.5 Multimodal communication: integrating modes, raising intercultural awareness

In *Chapter 5 Four extracts of data* I showed how in spoken face-to-face communication verbal language is one of many communicative resources speakers use to make meaning. Using a social semiotic multimodal framework I argued that modes such as eye gaze, head positioning, gesture, proxemics and sound all have meaning potential of an ideational, interpersonal and textual kind. Focusing on modality in a speaking activity I suggested that different modes combine together in modal ensembles but that meaning is not necessarily evenly distributed. For example, I showed how speakers were able to use verbal messages of low modality to make deductions and express ideational content, and combine these with modes such as eye gaze, gesture and laughter to convey interpersonal meaning and thus 'soften' utterances. I pointed out that my interest in this area came about as a result of watching language learners carry out speaking activities. The activity that forms the basis of my research comes from a lesson on modals of present deduction from *Language Leader Intermediate*. I had observed that whilst the learners were engaged with the activity and with each other, their recourse to verbal communication was considerably different to what the course book writers had laid out.

What my analysis revealed was that meaning was being conveyed across different communicative modes. Using terminology from SFG I was able to describe how modes other than the verbal had an ideational, an interpersonal and a textual function. I was able to study how these modes integrated and showed that at times certain modes come to the fore to communicate certain functions. It became evident whilst observing that participants were able to use different communicative modes in favour of others and that their interlocutors attended to the meanings being conveyed 'non-verbally' as much as they did to verbal messages. It too became evident that much of the interpersonal function of modes such as gaze, proxemics, and sound came to the fore; this is perhaps not surprising due to the interpersonal nature of face-to-face communication. Thus participants were able to signal social roles and relations as well as signal affinity with other participants.

At this point one may well ask what any of this has to do with teaching English as a foreign language. If we look again at the objectives of the lesson, we can see:

[b]y the end of the lesson, students should have:

- Practised reading a text to extract specific information
- Practised listening to a text to extract specific information and language items
- Learned more about and practised using modals for present deduction (*can't, must, might/could*)
- Engaged in a group discussion.

(Albery, 2008: 98)

There is no mention of pointing, eye gaze, proxemics or even sound; indeed one could argue the strangeness of an objective on a lesson plan which reads '[b]y the end of the lesson, students should have used pointing to indicate objects under discussion, eye gaze to signal intention to speak and elongating word initial phonemes to indicate they are being tentative'. But that is not the angle I am trying to pursue.

What I am arguing for is a broadening of the term communicative competence beyond the linguistic and functional orientated goals widely accepted in CLT (Sielloff Magnan, 2008: 373) towards a multicompetence and intercultural competence. Such competences open up a 'reflexive critical awareness of and engagement with communities that extend beyond their own' (ibid: 350) into what Kramersch has termed 'a third space' (Kramersch, 1993; Sielloff Magnan, 2008). The classroom activities that I have recorded show that whilst the learners may have 'deviated' from the course book objectives for the lesson, they are engaged in the process of making meaning. Of course there is communication breakdown, but in the main 'interactional synchrony' (Chick, 2001: 233) is evident,

clearly suggesting the learners are able to make sense of what they are doing together; as Chick argues 'synchrony contributes to the perception that purposeful activity and learning are taking place' (ibid). To that end I agree with Haught and McCafferty who argue '[g]estures, proxemics, facial expressions, posture and other mimetic features of interaction are a part of how meaning is encoded in an L2 and obviously deserve further attention' (2008: 158-159).

The notions of multicompetence and intercultural competence accord with the thinking behind the *Common European Framework of Reference* (CEFR) with which *Language Leader* aligns itself (Albery, 2008). In *Chapter 4* 'Language use and the language user/learner' we are informed:

[t]he language learner becomes *plurilingual* and develops *interculturality*. The linguistic and cultural competences in respect of each language are modified by knowledge of the other and contribute to intercultural awareness, skills and know-how. They enable the individual to develop an enriched, more complex personality and an enhanced capacity for further language learning and greater openness to new cultural experiences.

(CEFR, 2001: 43 emphasis in original)

Referring specifically to 'non-verbal communication' it lists the following which 'users of the framework may wish to consider' (p.88) for the learner to recognise/understand and/or use (ibid):

1. Practical actions accompanying language activities:

Pointing, e.g. by finger, hand, glance, nod. These actions are used with deictics for the identification of objects, persons, etc., such as, 'Can I have that one? No, not that one, that one';

Demonstration, accompanying deictics and simple present verbs and pro-verbs, such as, 'I take this and fix it here, like this. Now you do the same!';

Clearly observable actions, which can be assumed as known in narrative, comment, orders, etc., such as, 'Don't do that!', 'Well done there!', 'Oh no, he's dropped it!'. In all these cases, the utterance is uninterpretable unless the action is perceived' (CEFR, 2001: 88)

2. Paralinguistics

Body language. Paralinguistic body language differs from practical actions accompanied by language in that it carries conventionalised meanings, which may well differ from one culture to another. For example, the following are used in many European countries:

- gesture (e.g. shaken fist for 'protest');
- facial expression (e.g. smile or scowl);
- posture (e.g. slump for 'despair' or sitting forward for 'keen interest');
- eye contact (e.g. a conspiratorial wink or a disbelieving stare);
- body contact (e.g. kiss or handshake);
- proxemics (e.g. standing close or aloof).

Use of extra-linguistic speech-sounds. Such sounds (or syllables) are paralinguistic in that they carry conventionalised meanings but lie outside the regular phonological system of a language, for example, (in English):

'sh' requesting silence

's-s-s' expressing public disapproval

'ugh' expressing disgust

'humph' expressing disgruntlement

'tut, tut' expressing polite disapproval

Prosodic qualities. The use of these qualities is paralinguistic if they carry conventionalised meanings (e.g. related to attitudes and states of mind), but fall outside the regular phonological system in which prosodic features of length, tone, stress may play a part, for example:

voice quality (gruff, breathy, piercing, etc.)

pitch (growling, whining, screaming, etc.)

loudness (whispering, murmuring, shouting, etc.)

length (e.g. ve-e-e-ery good!).

(CEFR, 2001: 89)

It adds that 'many paralinguistic effects are produced by combinations of pitch, length, loudness and voice quality' (ibid). Whilst these might be slightly different conceptualisations of social semiotic

actions and multimodality as well as using different terminology, the acknowledgement of modes other than the verbal for making meaning is clear.

On communicative competence the CEFR says the following '[a]ll human competences contribute in one way or another to the language user's ability to communicate and may be regarded as aspects of communicative competence (2001: 101). And it lists linguistic competence, sociolinguistic competence and pragmatic competence, consisting in turn of discourse, functional and design competences. These competences are, to greater or lesser extents, included in language teaching and learning, and I argue the reference to 'all human competences' suggests there is scope here for a multimodal competence. I am not advocating lessons on how to point or how to say 'ugh', for example, but on raising intercultural awareness on how modes of communication configure to make meaning; in the same way that back-channel devices might be addressed in a lesson (see for example Bilbrough, 2007). It is important to note the reference to the terms multicompetence and plurilingual as I am not suggesting recourse only to the norms of the target culture, hence reference to Kramsch's 'third space'.

Study of communicative modes is becoming more accessible. Lee notes the emergence of corpora 'now fully multimedia in the sense of having transcripts that are aligned or synchronised with the original audio or video recordings' (2010: 114) (also Flewitt et al., 2009 and Ruhlemann, 2010). This opens up the possibility of using both audio and visual elements of interactional discourse for study and analysis.

The inclusion of other modes of communication in language study has been the subject of little research. However, in one such longitudinal study following McNeill's interconnection with thought and gesture, it was found that gesture played a part in the meaning-making process as well as evidence of gesture aiding 'self-regulation in both developing and conveying' thoughts in the L2 (Haught and McCafferty, 2008: 146). In a further study it was suggested that the ability to use metaphoric gesture with speech shows understanding of the 'illocutionary force' of idioms (ibid).

The inclusion of other modes of communication in language study would require alternatives to teaching sequences such as ESA. It would mean a more inductive, exploratory and analytical approach in which learners are more than passive receivers of knowledge. I will discuss this in the next section.

6.6 Rounding up

In this chapter I have shown there are discrepancies between structures presented to teach modality in a typical published ELT course book to those found in actual language use. I have argued for the need of language teaching materials to be informed by corpora rather than present learners with invented examples. My argument is based on the premise that for learners to attain communicative competence, recourse to actual language use is more useful than presenting them with invented examples that can be seen as misleading in their presentation of modal forms and structures.

I have also argued for the need to teach the interpersonal function of modality as modal structures do more than simply express degrees of certainty. Modality allows speakers to express stance and is thus bound up with issues of power. The literature and corpora tell us that there are a range of structures for this purpose in addition to the modal auxiliaries. Again for this purpose I have suggested corpora can inform teaching materials.

I argued that context is not a given but subject to contextual variables. Therefore in the language classroom activities can easily take different routes from those pre-planned in lesson objectives; this is to be expected and allowed for. Language learners are not a homogenous group and the way they perceive language activities will affect the way they respond and react and thus the communicative resources they use to engage with other learners, the activity and the teacher. This is part of the learning process.

Using a multimodal perspective I showed how meaning is communicated through different communicative resources. I argued for the importance of raising awareness of these resources in order to develop learners' multicompetence and intercultural awareness. I suggested that as meaning is bound up in modal ensembles and not just conveyed verbally, these should be studied in the language classroom to promote an understanding how verbal communication integrates with other communicative modes.

To meet these challenges I suggested a shift from teaching models which transmit language as product to teaching strategies which promote exploration of communication using inductive analysis and critical reflection. In the final chapter, I put forward strategies that include the use of corpora, analysis of multimodal discourse and tasks to promote language use and consider how these additional approaches could contribute to ELT practice more generally.

Chapter 7 Conclusion

My research has emerged from over fifteen years of teaching and observing English as a foreign language. As I have said at the beginning of my thesis, I learnt the rules of English grammar at the same time as I was teaching it. These rules came from grammar references such as Swan (1980), Swan and Walter (1997), Raymond Murphy (1994) and from the back of English language teaching course books such as the *Headway* series (Soars and Soars). In addition to receiving formal teacher training on a CELTA and Dip TESOL I learnt a lot about teaching on the job by picking up tips from colleagues, resources books and by trial and error. So I have been pretty much schooled in a TEFL tradition: one which uses standardised teaching models such as PPP and ESA and teaches pedagogic rules of language. I put my faith in the teaching methods I had been taught and the course books and grammar references I was using. For most of my early teaching career I taught in a similar fashion. When learners were not 'on task', or the lesson veered from the plan or the lesson objectives were not met, I felt something must have gone awry with the way I handled the lesson; I did not question the methods or the rules.

After several years of teaching I gained confidence and experience and questions began to formulate over the teaching methods I was using and the course books activities and descriptions of grammar. I observed discrepancies between the language that course books presented and the language I knew and heard around me; I noticed that teaching sequences needed to be rigorously controlled otherwise learners would deviate from the lesson objectives; and I discovered that activities designed for the use and production of linguistic items that had been the focus of the lesson often gave rise to different language and were engaged with in different ways by different groups of learners. This may all seem obvious to an experienced teacher or researcher. However, as a novice teacher, these observations rocked my little teaching world.

My thesis has been an attempt to hone in on a few of questions that have stood out. I have for many years been puzzled by modality. Firstly in terms of how pedagogic grammar matches actual language use, and secondly how it is taught and the affect this has on learners' awareness and use of modality. I was therefore interested to come to a broader understanding of modality in an attempt to discover why modal forms appear to be difficult for learners to use productively. Context has been for me an elusive term. The word crops up all over the place in TEFL, from language input and modelling language to creating context for language output. To me context always seemed to be talked about as if it were a handy little frame that could be set up to present language and then used as a means of controlling language output. In practice context proved to be a frame of uncertain dimensions and

less than easy to manipulate language with. I therefore wanted to gain a better understanding of what context is and how it affects language choice, modality in particular.

Observing modality in context brought me to my next question. As I have pointed out above, classroom activities intended to activate certain linguistic structures can not only give rise to different language than expected but also to a different kind of interaction than might have been expected. Being interested in multimodal communication and taking the statement that all utterances 'bear the signs of modality' (Hodge and Kress, 1988: 124) as my point of departure, I wanted to observe what communicative resources learners used in an activity to activate modals structures, especially as modal auxiliaries were typically absent.

To make this manageable I centred my research around a lesson on modals of present deduction in a published ELT course book. The choice of book was made on the basis that it had formed the teaching syllabus on an EAP course on which I had been teaching for some time. It provided me with a means of focusing my research and gathering video recorded data from the classes I was teaching. Moreover, I argue that the course book is typical of many such course books found in TEFL. The descriptions of modality and the lesson staging follow very much in a course book tradition and it shares similarities with descriptions of grammar found in pedagogic grammar references and lesson sequences used and promoted on teacher training courses.

My discussion of modality showed how epistemic modality is typically referred to in TEFL as a way of talking about degrees of certainty and possibility. And done so through the modal auxiliaries. The standard way of teaching modality is to match rules of form and meaning before giving learners a chance to practise them in written exercises and oral production activities – a PPP or ESA model. A lesson of this kind is based on a structural approach to grammar and focuses on language usage. I have pointed out the necessity of teaching usage but that learners also need instruction on language use. Using ideas from systemic functional grammar and corpus-based grammar I argued that the function of modality in discourse is to express stance in regard to a speaker's propositions and to express degrees of affinity to their interlocutor's propositions. So whilst modality is used to express degrees of certainty, it does so through the interpersonal function of language. The interpersonal function is overlooked in course books such as *Language Leader Intermediate*. The result of this, I argue, is that learners are not made aware of how modality affects social relations during discourse. Thus when engaged in a speaking activity such as a discussion, learners are uncertain as to how and when to use modal forms; or when they do use them, they are often uttered 'mechanically' simply to practise the form as in the following example of a learner making deductions about *Figure 12* overleaf:

- S2 Oh it frog
(1 sec)
- S2 She wanna (.) like frog (.) so (.) [S1 ah] ya=
S1 =Lady Gaga ((laughs))
S2 =((laughs)) ha Lady Gaga ah ((inaudible))=
S1 =She could be::: >a mother of °Lady Ga°< ((laughs))
S2 ((laughs)) (1.5) but she looks (.) very old
S1 ((inaudible)) or maybe (.) she (.) she might be::: (2) she might be cleaning (1.5) her house (.)

(Original data *Transcript 6: Classroom data CD ROM*. See *Appendix 11, Transcript F* for fuller transcription.)



Figure 12 (Cotton et al., 2008: 75 reproduced from *Figure 2* above)

In the short exchange above S2 initially makes a deduction and whilst a discussion of a limited kind develops the dialogue leaves the impression that the speakers are making guesses in an attempt to squeeze in the forms *maybe* and *might be* (see *Appendix 6: Classroom data CD ROM* – Richard (1) 11.55 for original data).

From teacher training courses I had been instructed in using teaching sequences such as PPP and ESA and guided on how to write lesson plans. For observation purposes a lesson would need to be constructed around a recognised teaching sequence. So with the ESA model learners would firstly engage with the language focus through, for example, a listening or reading activity, followed by study of the selected language focus and then activate the language in a speaking or writing activity; from ‘receptive’ to ‘productive’ skills. A simplified example would be to play a recording of two peoples having a discussion making deductions about a number of pictures (engage), match the rules of form

and meaning and complete gap-fill exercises using a selection of modal (study), put learners in pairs and have them discuss a number of pictures with the same modals as linguistic prompts (activate). Written into the lesson plan are the list of objectives, for example '[t]he objective of the lesson is to use modals of present deduction in a discussion'. On paper this is all straightforward; in practice not so, particularly at the activate stage. Therefore I was curious to discover why learners 'deviated' from the objectives. How was it that the activate stage, unless very tightly controlled, differed from the language contextualising activity? The students were told to discuss, they had the linguistic prompts, they were talking but not 'on task', i.e. using the language they had been given.⁴⁹ The plan laid it out all so clearly and the teaching sequence was supposed to provide a seamless lesson from context to context.

Referring to discussions from anthropology, ethnography and linguistics I showed how context is more than an 'ornamental backdrop' to communicative events. With specific reference to Hymes' SPEAKING mnemonic and Halliday's contextual variables I discussed how context in part shapes and is shaped by the participants and communicative resources. Context is thus not a given, but dynamic. Thus whilst a discussion is likely to contain certain lexicogrammatical features such as the present tense and modality, the way participants interact during a discussion has the potential to continually reshape it; and accordingly the language used. A discussion may turn into an argument and a discussion may have instances of deduction, informing, requesting, warning, advising, threatening, suggesting, etc. realised through different speech functions and lexicogrammatical structures. So in terms of the language classroom we cannot simply hand learners 'a context' and a set of linguistic structures and consider the job done. Learners create their own contexts and the effect of this will often take them on different paths through an activity and recourse to different communicative resources – all of which is to be expected.

Having observed numerous classroom activities I became aware that learners frequently engaged with them in ways that differed from those set out on lesson plans. Not only did they go about activities in different ways, due to what Kumaravadivelu (2003) refers to as 'mismatches' between how the activities are conceived by the teacher and the learners, but they often used different language than planned and drew on a variety of different communicative resources to express meaning. Using social semiotic multimodality as a methodological framework, I was interested to observe what communicative modes learners used, in addition to the verbal, to engage with each other and the activity. Communicative events comprise modal ensembles with meaning made through

⁴⁹ I can be argued here that the learners were 'on task' in relation to the focus on meaning but not 'on task' as far as the focus on form(s) is concerned.

configurations of modes. I argue that a mode such as gesture, for example, has the potential to convey ideational, interpersonal and/or textual meaning. In a given communicative event configurations of modes integrate to make meaning. Meaning is not always evenly distributed, however, and one, or more, function may come to the fore. I gave the example of asking for and giving directions in which pointing largely carried ideational content whilst spoken language carried interpersonal in addition to ideational and textual content. It became evident from my data that learners used combinations of modes to express meaning and that, in the main, these were attended to by the other participants. As my research has focused on modality, I explored how modes such as posture and eye gaze are able to modalize spoken language.

I have indicated above (1.5 *A note on the written style of this thesis*) that my research is practice-derived, intended to provide valuable insights into multimodal communication and English language teaching, and to develop new ideas for language teaching pedagogy. It is with this in mind that I set out the teaching strategies that now follow.

7.1 Teaching strategies: research into practice

Here I feel I should remind the reader that my research was motivated by observations of teaching pre-sessional students at a university in London. The students were all planning to study at the university, or similar institution, at BA or MA level. *Language Leader* formed the pre-sessional syllabus. In the Teacher's Book we are informed the course book is for 'learners who may go on to, or are already in, further education' (Albery, 2008: 4). The teaching strategies I now discuss are with this in mind. I am not proposing an alternative syllabus or suggesting the strategies here entirely replace the current approach. My research has been confined to epistemic modality and whilst certain aspects of my findings might be generalizable to other areas of English language teaching and learning,⁵⁰ it is not my intention to make any claims in that regard. What I am arguing for is additions to teaching epistemic modality as a means of raising awareness of how it functions interpersonally in spoken discourse.

I have critiqued the description of modality and the choices of modal structures in a lesson of modals of present deduction in *Language Leader Intermediate*. Whilst acknowledging the importance of studying the rules of usage, i.e. linguistic form using an the ESA model (see *Appendix 2* for the language focus of the lesson) the ideas I present below are intended to provide an opportunity for learners to study and use modal structures in spoken discourse. They are intended to develop a communicative competence that incorporates linguistic, discourse, functional and multimodal competence

⁵⁰ Data-driven learning, task-based instruction and a Three Is approach are, of course, adaptable to teaching a wide range of lexicogrammatical structures.

developing intercultural awareness and personal identity. They foster an inductive and exploratory approach in which the process of interaction is part of the learning outcomes. Language learning is thus treated as more than the delivery of language as ‘product’ (see Skehan, 1998; Ellis, 2003 for a critique on language as product).

My thesis has focused on epistemic modality and context and how the two interrelate. I have argued that modality is more than simply discussing degrees of certainty. As part of the interpersonal metafunction, modality is the means through which participants in interaction are able to express affinity with their own and their interlocutor’s propositions; to that end modality is bound up with issues of power: social status and roles. Thus for a fuller understanding of modality, we need to look beyond the rules of form and meaning pertaining to usage to how contextual variables affect speakers’ choices of modality. Moreover, I have argued that modality is signalled by a variety of structures and not limited to modal auxiliary verbs. My analysis of classroom data also revealed that modality can be signalled through modes of communication other than the verbal. And so the teaching strategies I now propose offer ways of exploring modality in context.

7.1.1 Data-driven learning (see *Appendix 9.1* for sample lesson)

Must is used mostly to talk about the feelings and wishes of the speaker and hearer – for example, to give or ask for orders. *Have (got) to* is used mostly to talk about obligations that come from ‘outside’ – for example from laws, regulations, agreements and other people’s orders.

(Swan, 1980: 345 emphasis in original)

One of the most frequent questions I am asked by students is to clear up the following grammar point: ‘What is the difference between *must* and *have (got) to*?’⁵¹ Of course, being a fully qualified TEFLer I am able to reel off the rule above and hope that suffices. Yet this question is formulated by learners who are living, studying and often working in London and come into regular contact with English texts. The question is often asked on the back of a discrepancy in the rule they have themselves observed. The rule comes from a widely used grammar reference ‘intended for intermediate and advanced students, and for teachers of English’ (ibid: xi). It is widely used and referred to on teacher training courses and a commonly found resource on the book shelves in English language schools; indeed as Hughes points out ‘for the teacher the mainstay of the reference grammar shelf is still Michael Swan’s *Practical English Usage*’ (Hughes, 2010: 402). There is no information in the introduction of Swan’s

⁵¹ For this example I am discussing modals with personal (intrinsic) meaning, i.e. permission, obligation or volition. Also referred to as deontic modality (Coffin et al., 2009).

book as to what has informed the rules of usage, but he does tell us ‘the examples are as real as I can make them’ (ibid: xii).⁵² Some examples are as follows:

- I **must** stop smoking. (I want to.)
- I’**ve got to** stop smoking. Doctor’s orders.
- This is a terrible party. We really **must** go home.
- This is a lovely party, but we’**ve got to** go home because of the baby-sitter’

(Swan, 1980: 345 emphasis in original)

The description of the same modal forms is similar in *Language Leader Intermediate*:

[w]e use both *have to* and *must* to talk about something that is necessary and important, but there are some differences in meaning.

- Use *have to* to say something is essential or that it is a general rule.

When you develop a new design you **have to** try it out a number of times.

It **has to** be strong enough to carry eight people.

- Use *must* to say something is necessary or important in your personal opinion.

I feel that we **must** make the design more modern.

It **must** be on the desk by the end of the day

(Cotton et al., 2008: 146).

Swan and Cotton et al. seem to apply similar rules of usage, i.e. *have to* for rules, laws, etc. coming from ‘outside’; *must* for something necessary, important and coming from the speaker.

The rules are, of course, pedagogic rules i.e. rules for teaching and learning and thus simplified to meet L2 learners’ needs (Kumaravadivelu, 2003). Thornbury says pedagogic rules ‘should be easily applicable, have a wide coverage and few exceptions, as well as being short and memorable’ (1997: 145). At a first glance the excerpts from Swan and Cotton et al. arguably meet the requirements of pedagogic rules: they are short and memorable, clear and concise. Yet when it comes to using the forms in gap-fill exercises, I find, as a native speaker, I am less than satisfied with the rules, see below.

‘A soldier _____ obey orders’ (Swan and Walter, 1997: 112).

‘I _____ start working harder or I’ll fail the exams’ (Cotton et al., 2008: 146)

The ‘answers’ are ‘has to’ and ‘must’ respectively but I do not feel comfortable with this. And a similar uncertainty occurs when I substitute the modals for each other in the examples I have cited from Swan

⁵² Hughes notes ‘although the most recent editions pay some attention to corpus data and users’ comments, [*Practical English Usage*] remains largely based on isolated examples and the intuition of the author as to correctness; Swan 2005’ (2010: 402).

and Cotton et al. above; the reader might like to try this. Perhaps it is my own shaky grammar as a result of being one of the teachers Thornbury refers to when he says ‘they may never have formally studied the subject they are teaching’ (1997: xiv) or perhaps it is my intuition as a native speaker that is awry, regardless I find the examples and rules do not tally.⁵³

The purpose of the above is to bring me to the approach I want to discuss here, that is data-driven learning (DDL). DDL ‘consists in using tools and techniques of corpus linguistics for pedagogical purposes’ (Gilquin and Granger, 2010: 359). Questions such as the difference between *must* and *have to* become part of an exploratory approach to learning in which learners (and teachers) directly consult corpora to answer questions. To go about this Kennedy and Miceli (2001) propose a four-stage search strategy that can be incorporated or form the basis of a language lesson.

1. Formulate the question.
2. Devise a search strategy.
3. Observe the examples and select the relevant ones.
4. Draw conclusions.

(cited in Chambers, 2010: 355).

So a question might be ‘What is the difference between *must* and *have to*?’ A search strategy would be to access the British National Corpus (BNC). A selection of concordance lines would then be collected and observed and from any patterns that emerged, conclusions of use could be drawn based on empirical evidence.

Data-driven learning is not meant as a challenge to pedagogic grammar, but as a supplement to language teaching through the analysis of actual language use. Simplified pedagogic rules and invented examples can only tell us so much. If we want to explore questions such as the difference between *must* and *have to*, it is helpful to observe how they are distributed across registers, spoken and written language, their respective degrees of formality, their interpersonal function, their use with passive voice and personal pronouns or how institutional voices are recontextualised in texts using these modal forms. DDL can address questions of a qualitative kind as well as quantitative.

⁵³ I refer the reader to the earlier discussion on the distinction between ‘usage’ and ‘use’ in 6.3 *Interpersonal meaning: uncovering meaning through discourse*. Pedagogic rules are rules of usage. The discrepancy noticed by the learners comes from examples of use. Use includes applications that are not covered by usage but are accessible in context.

Online language corpora such as the BNC or the Brigham Young University British National Corpus (BYU-BNC) are an easily accessible and valuable resource for use in the classroom or for homework tasks. DDL necessitates a shift from teacher-led learning to one in which the learner is more active in the learning process. The teacher takes one step back and an inductive approach to rule forming is facilitated. Using corpus data and concordance lines encourages discovery and requires the use of critical skills such as analysing, interpreting and theorising in addition to language processing skills such as noticing and consciousness raising.

7.1.2 Task-based instruction (see Appendix 9.2 for sample lesson)

In *Methodology and Terminology* I referred to task-based instruction. Using a definition by Skehan (1998) I showed how it differed from a language activity, in which the focus is primarily a practice-oriented means of displaying language form. To recap a task-based approach is designed to ‘encourage naturalistic acquisitional mechanisms, cause the underlying interlanguage system to be stretched, and drive development forward’ (Skehan, 1998: 95). As a definition of tasks in a task-based approach, Skehan proposes the following:

- Meaning is primary;
- There is some communication problem to solve;
- There is some sort of relationship to comparable real-world activities;
- Task completion has some priority;
- The assessment of the task is in terms of outcomes.⁵⁴

Whilst Skehan puts an emphasis on meaning, it should not be assumed that task-based instruction does away with a focus on form. Indeed Willis (1996), Littlewood (2004) and Willis and Willis (2010), for example, discuss the integration of a focus on form, i.e. specific lexicogrammatical features in task-based instruction. Attention is also drawn to the importance of this in the task-based model proposed below.

I have argued that the ESA model adopted in *Language Leader Intermediate* is limited to the study and practice of linguistic form. And whilst the study of form and usage is necessary, the activation stage does not really cater for a discussion in which participants are expected to exchange points of view. Indeed, as I have pointed out, giving learners a set of pictures and asking them to ‘Work with a partner and discuss what you think the designs below are’ (Cotton et al., 2008: 75) invariably leads to rather stilted discourse in which single utterance guesses are made. When we consider what the rubric is asking of the learners, this is hardly surprising. For example if presented with the following picture (*Figure 11*) a reasonable response is arguably ‘It’s a fishing net’ (picture reproduced from *Figure 2* in 2.3.1 *Problematizing a lesson on modal verbs*).

⁵⁴ A distinction is made between lesson ‘objectives’ and lesson ‘outcomes’ in my conclusion. See also D’Andrea (1999).

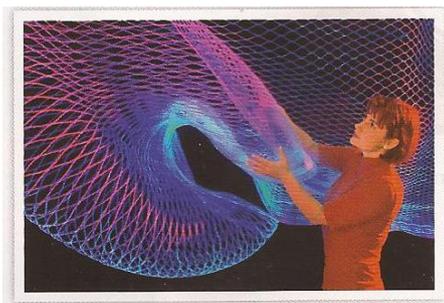


Figure 11 (Cotton et al., 2008: 75)

A slightly more bizarre response might be the invented example below:

It's a fishing net or it's a large spring, no, it might be a sculpture, or it could be a dress or maybe it's a fish, no, I think that it might be a new kind of bed, or perhaps it's a balloon, but I don't think it's a sofa. What do you think Roy?

My point is if you ask someone what they think something is, they are likely to tell you what they think it is, and that might reasonably be curtailed to one or two 'guesses', not repeated guesses and not a discussion, as my data has shown.⁵⁵ So what is required is a means by which the learners can enter into a discussion and a language task by which this can come about; hence task-based instruction.

For this purpose I draw on Willis' task-based model. Willis offers five principles to provide input, use, and reflection on the language input and use.

1. There should be exposure to worthwhile and authentic language.
2. There should be use of language.
3. Tasks should motivate learners to engage in language use.
4. There should be a focus on language at some points in the task cycle.
5. The focus on language should be more or less prominent at different times.

(Willis, 1996 in Skehan, 1998: 126).

⁵⁵ It is also worth considering how frequent a speech function such as deduction using a modal verb occurs in a discussion. For example in the following news clip, on the possible eruption of the Bardarbunga volcano in Iceland, the seismologist uses only two modal verbs for deduction based on present evidence <http://www.bbc.co.uk/news/science-environment-28854809>. Compare this with the following comedy sketch; a useful teaching resource: <http://www.youtube.com/watch?v=eKQOk5UIQSc>

In terms of lesson staging Willis (1996) and Willis and Willis (2010) propose the following sequence of pre-task, task cycle and language focus:

Pre-task

Introduction to topic and task

Task cycle

Task

Planning

Report

Language focus

Analysis

Practice

During the pre-task stage the learners are introduced to the topic and task. A possible way of doing this would be to present the class with a text of a similar kind. So for example if the task is a discussion, then a short recording of a discussion would be appropriate. At this stage learners activate schematic knowledge, are presented with authentic language, and attention is drawn to form.

The task stage consists of three parts: the task, planning and report. Firstly the learners are given the task to complete at which stage the focus is on meaning, i.e. completion of the task. The next stage is for the learners to prepare to report back to the teacher or the class and so the planning stage allows the learners time to put their report together focusing on the language, structure, sequencing, etc. they need to do so. The report stage provides an opportunity for learners to practise the language.

The final stage is the language focus. The language comes after the task 'with the intention that any language which is focused upon is relevant to learners and required for a communicative purpose, rather than introduced because a syllabus dictates that it should be covered at a particular point' (Skehan, 1998: 128). Drawing attention to form, explicit focus on a particular structure or practice-oriented work.

7.1.3 Three Is (see *Appendix 9.3* for sample lesson)

The task-based approach above sits well at the crossroads between form-dominated and communication dominated approaches. Another such approach is the Three Is methodology

proposed by McCarthy and Carter (1995). The Three Is stand for illustration, interaction and induction. They explain as follows:

[i]llustration means wherever possible examining real data which is presented in terms of choices of forms relative to context and use. Interaction means that learners are introduced to discourse-sensitive activities which focus on interpersonal uses of language and the negotiation of meanings, and which are designed to raise conscious awareness of these interactive properties through observation and class discussion. Induction takes the conscious-raising a stage further by encouraging learners to draw conclusions about the interpersonal functions of different lexical-grammatical options, and to develop a capacity for noticing such features as they move through the different stages and cycles of language learning.

(McCarthy and Carter, 1995: 217)

McCarthy and Carter put forward the methodology as an alternative to PPP for an inductive study of the interpersonal implications of spoken grammar, and to promote greater language awareness (*ibid*). As we have seen, PPP and ESA are based on a deductive method of grammar teaching in which the teacher presents learners with explicit grammatical rules before giving learners the opportunity to practise them in oral or written exercises and activities. In such a method the learner is put into a receptive role with little emphasis on exploring or discovering language 'rules' and patterns for themselves. Thus the argument runs that it leads to a 'superficial knowledge' of the language. Alternatively an inductive approach 'is based on the premise that the essence of grammar teaching lies in helping learners discover what the grammatical rules are' (Kumaravadevelu, 2003: 185), and thus encourages learners to observe and notice patterns. Typically learners study contextualized language with the teacher avoiding the presentation of explicit description and explanation of grammatical rules prior to learners engaging with texts. Rules are developed explicitly or implicitly by the learners through the analysis of patterns in texts. Such a method is thought to provide lifelong analytical and heuristic language learning skills.

Whereas a task-based approach puts an emphasis on using language in meaningful communication, a Three Is approach is more analytical and an emphasis is put on classroom discussion. Kramsch notes the importance of metatalk in the language classroom, i.e. talk about talk. She argues that it is 'what the language classroom does best' (1993: 246) and that 'most of the sociocultural appropriateness that communicative approaches to language teaching have shown are important in communication cannot be taught only by doing' (*ibid*). So as I have indicated above with communicative resources such as gesture or sounds such as 'ugh' the purpose of study is not to present, practice and produce,

but to become aware of how they function in the construction of meaning. As Kramsch points out '[c]ontexts are too varied, too changing; behaviours can only be observed, analysed, and talked about' (ibid). A Three Is approach is therefore well suited for exploring multimodal texts.

The language classroom provides a purposeful setting in which discussion can take place, encouraging an exploratory, inductive and heuristic approach to learning and language study. Such an approach might use a range of texts for a comparative study of how linguistic structures and communicative modes are used across different registers, genres, modes as well as raising awareness of how they are used across cultures.

7.2 Handling real spoken data for teaching purposes

Throughout this section I make reference to the use of spoken language corpora in language teaching and learning. However, examples taken from corpora are not unproblematic. Transcribed spoken data can be extremely difficult to read with or without the addition of conventions to show prosodic features, for example. Language which has not been 'culturally disinfected' (Carter, 1995: 50) for use in the language classroom can abound with cultural references and values (ibid). Concordance lines present isolated decontextualized examples which might arguably be less suitable for pedagogic purposes than the invented examples in course books and pedagogic grammars precisely because they have been decontextualized. Concerns such as these raise questions over the pedagogic validity of using spoken data retrieved from corpora. However, recourse to corpora yields examples that are not as off-putting as these concerns suggest. Consider the concordance line below.

Get hungry. (SP:PS0M4) I won't get hungry. (SP:PS0M5) You **must be hungry**, you hardly had any dinner last night. # (SP:PS0M4) There we are.

Whilst this is by no means the most opaque example, it presents a clear example of *must be* + collocate. It comes from a search of collocations with *must be* in the BYU-BNC. By clicking on 'the speakers' more information pertaining to the 'context' is given. SP:PS0M5 is Nicola, a housewife from the home counties aged between 25-34. SP:PS0M4 is Oliver, Nicola's son aged between 0-14.⁵⁶ By clicking on the concordance line the following extended text is given (continued overleaf):

baseball bat? (SP:PS0M5) Well I should imagine they're playing baseball. (SP:PS0M4) Ohh.
(SP:PS0M4) Mummy, who's that? (SP:PS0M5) What darling? (SP:PS0M4) Who's that? And the Joker. (SP:PS0M5) Oh! Batman and the Joker. (SP:PS0M4) And the Joker. I don't like the Joker. (SP:PS0M5) Don't you? Alright which cereal would you like for breakfast? (pause) (SP:PS0M4) (unclear) (SP:PS0M5) Come and choose one then. (SP:PS0M4) I don't like it. (SP:PS0M5) You've got to have some cereal. Get **hungry**. (SP:PS0M4) I won't get **hungry**. (SP:PS0M5) You **must be hungry**, you hardly had any dinner last night.
(SP:PS0M4) There we are. (SP:PS0M5) Mummy look. They got saucer in this. (SP:PS0M4) The saucer

⁵⁶ See Appendix 10 for full information.

in that one. You've got that story already haven't you? (SP:PS0M5) Mm. (unclear) two ones aren't I. (SP:PS0M4) No. There we are. Sit down in the corner (unclear) nice breakfast. You **must be** a **hungry** boy. (SP:PS0M5) (unclear) (SP:PS0M4) Well it doesn't matter you can look at it afterwards. (unclear). Be careful.

Arguably it does not make the easiest reading. It is not sequenced clearly and speaker references are given in a cryptic system. And whilst there is more of the co-text it perhaps becomes more opaque as is not easy to grasp what the baseball bat, the saucer and Batman and the Joker have to do with eating cereal. Further these nouns contain cultural references that might add to its lack of clarity. Yet all is not lost. Carter proposes re-modelling such data for pedagogic purposes. He writes: '[t]he attempt here by the materials developer is to achieve clarity, tidiness, and organization for purposes of learning, but at the same time to ensure that the dialogue is structured more authentically and naturalistically by modelling on real corpus-based English' (1995: 52-53). As an example I have remodelled the above (continued overleaf):

Mother: Alright which cereal would you like for breakfast?

Tim: Oh, mum.

Mother: Come and choose one then.

Tim: I don't like it.

Mother: You've got to have some cereal. You'll get hungry.

Tim: I won't get hungry.

Mother: You **must be hungry**, you hardly had any dinner last night. (Gives Tim the cereal) There we are.

(Dialogue recycled from 6.3 above *Interpersonal meaning: uncovering meaning through discourse*.)

Carter points out the need for caution when remodelling to avoid 'distortion' (ibid). I argue the above text presents the learners with something more pedagogically real than the extract taken straight from the corpus and does so without much distortion.

7.3 Rounding up

My research into modality, context and classroom interaction suggested to me that teaching language usage does not develop learners communicative competence, i.e. 'not only the linguistic forms of the language but also its social rules, the knowledge of when, how, and to whom it is appropriate to use these forms' (Paulston, 1992: 98). Moreover, that communicative competence needs to reach beyond

linguistic and functionally oriented goals to include discourse, pragmatic and multimodal competence. To meet these needs I proposed three teaching strategies.

Data-driven learning encourages an inductive approach to exploring modality in actual language use using corpora. My argument being that learners can discover which modal structures are used and how they function in discourse. Task-based instruction puts an emphasis on meaning in the language classroom and promotes language use. Learners carry out 'real world' tasks before focusing on form, based on the argument that the forms are relevant to the task. In a Three Is approach discourse is the starting point for study. Multimodal texts are explored and discussed in relation to how context and communicative resources configure to make meaning.

Of course, such classroom strategies of this kind reshape the classroom dynamic. It is no longer teacher-centred as in PPP or ESA and so there are different demands on the learners and the teacher. In data-driven learning, for example, the teacher becomes a facilitator in the learning process (Chambers, 2010). Inductive and exploratory learning also requires a rethinking in terms of stipulating lesson objectives. As D'Andrea notes:

[d]efining learning objectives requires teachers to make conscious choices about a wide range of teaching and learning considerations. The process of identifying teaching/learning objectives essentially defines what it is the teacher wants the student to learn. (1999: 43)

Rather the inclusion of learning outcomes takes a more student-centred and less prescriptive approach. It is said to 'facilitate the students orientation to the subject being studied' (ibid) in which the 'outcomes equals outputs' (ibid).

Whilst I argue advantages with the approaches I have proposed, there are limitations too. The use of corpora can be time consuming and laborious. Moreover, drawing conclusions from concordances is not an easy task for those not trained to do so. Gilquin and Granger (2010) found research in to DDL giving 'mixed results' from learners (p.365) with some finding it 'fascinating' and others finding it 'frustrating'. They also suggest that more research is needed 'to test the efficiency of DDL'. From my own experience of using corpora in the classroom I have found mixed responses from learners. Some learners enjoy the inductive analytical approach whereas others prefer deductive instruction. I would suggest that to some extent this depends on the learners' previous educational culture.

Using native speaker corpora is also up for debate as a means of providing target language for learners. It could be argued that text frequencies do not indicate anything about the variation of frequency of items between individuals or of the same individual on different occasions. Learners who over- or underuse certain items might not be seen to deviate from the norm as some native speakers also do.

There are also questions as to whether learners should mirror native speakers and if the need is to be understood is it important whether *might* or *I think* is used.

Skehan raises questions in regard to Willis' model of task-based instruction. He claims there is no link to 'wider-ranging theory' (p.129) and 'no explicit connection with research' (ibid). DDL, task-based instruction and a Three Is approach do not form a cohesive methodology, with no mention of larger pedagogic goals such as language acquisition or the development of interlanguage. As such they make syllabus design and language courses less than straightforward to plan and write. We have seen how course books are designed around a structural approach to grammar and for many years this has become a standard way of structuring courses, levels and syllabuses in language teaching. PPP and ESA are tidy packages that promote a specific kind of teaching and learning. And, as Skehan points out, they put the teacher in a secure role, they are easily trainable and easily assessed, and lend themselves well to accountability (1998: 94); and whilst criticised, they are still a mainstay in TEFL.

Turning to multimodality, whilst the literature clearly informs us that verbal language is synchronised with other communicative modes, documents such as the CEFR citing the importance of raising intercultural awareness of 'nonverbal' communication, and with multimodal corpora available for study purposes becoming a reality, it is doubtful whether the inclusion of other modes of communication into 'language' courses will happen in the near future. An immediately foreseeable problem is the question of what learners need to know. In a similar vein the question has been raised in relation to teaching the grammar of conversation (Thornbury and Slade, 2010). Thornbury and Slade refer to features of spoken grammar for which there is a 'natural tendency' for learners to use (p.209); the same would apply to other communicative resources. More research is needed in this area.

Not wishing to overgeneralize, the arguments I have put forward and the teaching strategies I have outlined in this thesis are with specific groups of learners in mind. They are enrolled on a language course with the intention of entering higher education in the UK. I argue from the perspective of a language teacher preparing learners for university study and as a former learner on an MA with an international cohort, that an interpersonal awareness and use of modality is necessary for academic study at university. I argue too that recourse to corpora can better inform teaching materials of actual language use. A perspective that I have formed as a teacher and researcher. And I argue for a more inductive and exploratory approach to the study of how language and context interrelate. In my experience as an EFL and EAP teacher, learners often come to class with their own questions about language, the approaches here give the learners the tools to answer them for themselves – an invaluable skill for higher education. And finally I have expressed the need for raising awareness of

how communicative modes integrate to make meaning – towards what I have referred to as a multicompetence. Of course, more research is needed to establish the effectiveness of the teaching strategies I propose.

References

- Aijmer, K. (2009) (ed.) *Corpora and Language Teaching*. John Benjamins Publishing Company.
- Albery, D. (2008) *Language Leader Intermediate Teacher's Book*. Pearson Longman.
- Argyle, M. (1972) 'Non-verbal communication in human social interaction' in Hinde, R.A. (ed.) *Non-Verbal Communication*. Cambridge University Press.
- Austin, J.L. (1962) *How to Do Things with Words*. London: Clarendon Press.
- Baker, W.D., Green, J.L. and Skukauskaite, A. (2008) 'Video-enabled research: A microethnographic perspective' in Walford, G. (ed.) *How to do educational ethnography*. The Tufnell Press, London.
- Bezemer, J. and Jewitt, C. (2010) 'Multimodal Analysis: Key Issues' in Litosseliti, L. (ed.) *Research Methods in Linguistics*. Continuum.
- Bezemer, J. and Mavers, D. (2011) 'Multimodal transcription as academic practice: a social semiotic perspective'. *International Journal of Social Research Methodology* Vol. 14, No. 3: 191-206. Routledge, Taylor Francis Group.
- Biber, D., Conrad, S. and Reppen, R. (1998) *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. (1999) *Longman Grammar of Spoken and Written English*. Longman.
- Biber, D., Conrad, S. and Leech, G. (2002) *Longman Student Grammar of Spoken and Written English*. Longman.
- Bilbrough, N. (2007) *Dialogue Activities*. Cambridge University Press.
- Block, D. (2002) 'McCommunication' A problem in the frame for SLA' in Block, D. and Cameron, D. (eds.) *globalization and Language Teaching*. Routledge, Taylor Francis Group, New York and London.
- Block, D. (2003) *The Social Turn in Language Acquisition*. Edinburgh University Press.
- Blommaert, J. (2005) *Discourse*. Cambridge University Press.
- Bloomaert, J. and Dong, J. (2010) *Ethnographic Fieldwork: A Beginner's Guide*. Multilingual Matters.
- Bohle, U. (2014) 'Body posture and movement in interaction: Participation management' in Muller, C., Cienki, A., Fricke, E., Ladewig, S.H., McNeill, D. and Bressemer, J. (eds.) *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction. Volume 2*. De Gruyter Mouton, Berlin and Boston.
- Breen, M.P. (2001) 'The social context for language learning: a neglected situation?' in Candlin, C.N. and Mercer, N. (eds.) *English Language Teaching in its Social Context*. Routledge, Taylor Francis Group, New York and London.
- Brigham Young University British National Corpus (BYU-BNC) <http://corpus.byu.edu/bnc/>
- Brown, G. and Yule, G. (1983) *Teaching the Spoken Language*. Cambridge University Press.

- Brumfit, C.J. (1979) 'Communicative' Language Teaching: and Educational Perspective' in Brumfit, C.J. and Johnson (eds.) *The Communicative Approach to Language Teaching*. Oxford University Press.
- Brumfit, C.J. and Johnson, K. (1979) editorial comments on Halliday, M.A.K. 'Towards a Sociological Semantics' (extracts) in Brumfit, C.J. and Johnson, K. (eds.) *The Communicative Approach to Language Teaching*. Oxford University Press.
- Cambridge.org/corpus (accessed 20/12/2014)
- Carter, R. (1998) 'Orders of reality: CANCODE, communication, and culture'. *ELT Journal* 52/1: 43-56. Oxford University Press.
- Chambers, A. (2010) 'What is data-driven learning?' in O'Keeffe, A. and McCarthy, M. (eds.) *The Routledge Handbook of Corpus Linguistics*. Routledge, Taylor Francis Group, New York and London.
- Cheng, W. (2010) 'What can a corpus tell us about language teaching?' in O'Keeffe, A. and McCarthy, M. (eds.) *The Routledge Handbook of Corpus Linguistics*. Routledge, Taylor Francis Group, New York and London.
- Chick, J.K. (2001) 'Safe-talk: collusion in apartheid education' in Candlin, C.N. and Mercer, N. (eds.) *English Language Teaching in its Social Context*. Routledge, Taylor Francis Group, New York and London.
- Coffin, C., Donohue, J. and North, S. (2009) *Exploring English Grammar from Formal to Functional*. Routledge.
- Conrad, S. (2005) 'Corpus linguistics in L2 teaching' in Hinkel, E. (ed.) *Handbook of Research in Second Language Teaching and Learning*. L. Erlbaum Associates, Mahwah, N.J.
- Conrad, S. (2010) 'What can a corpus tell us about grammar?' in O'Keefe, A. and McCarthy, M. (eds.) *The Routledge Handbook of Corpus Linguistics*. Routledge, Taylor Francis Group, New York and London.
- Cook, G. (1989) *Discourse*. Oxford University Press.
- Cook, V. (2001) *Second Language Learning and Language Teaching*. London: Arnold.
- Cotton, D., Falvey, D. and Kent, S. (2008) *Language Leader Intermediate*. Pearson Longman.
- Council of Europe (2001) *Common European Framework for Reference for Languages*. Council of Europe and Cambridge University Press.
- Coxhead, A. (2010) 'What can corpora tell us about English for Academic Purposes' in O'Keeffe, A. and McCarthy, M. (eds.) *The Routledge Handbook of Corpus Linguistics*. Routledge, Taylor Francis Group, New York and London.
- Criado, R. (2013) 'A Critical Review of the Presentation-Practice-Production model (PPP) in Foreign Language Teaching' in Monroy, R. (ed.) *Homenaje a Francisco Gutierrez Diez*. Murcia: Edit.um.
- Cullen, R. and Kuo, I. (2007) 'Spoken Grammar and ELT Teaching Materials: A Missing Link?' *TESOL Quarterly* 41/4: 361-386.

- D'Andrea, V.M. (1999) 'Organizing teaching and learning: outcomes based-planning' in Fry, H., Ketteridge, S. and Marshall, S. (eds.) *A Handbook for Teaching and Learning in Higher Education*. Routledge, Taylor and Francis.
- Delamont, S. (2008) 'For lust of knowing – observation in educational ethnography' in Walford, G. (ed.) *How to do educational ethnography*. the Tufnell Press, London.
- Dicks, B., Flewitt, R., Lancaster, L. and Pahl, K. (2011) 'Multimodal and ethnography: working at the intersection'. *Qualitative Research* 11 (3): 227-237. London: Sage.
- Dornyei, Z. and Murphy, T. (2003) *Group Dynamics in the Language Classroom*. Cambridge University Press.
- Duranti, A. and Goodwin, C. (1992) *Rethinking context: language as an interactive phenomenon*. Cambridge University Press.
- Eggs, S. (1994) *Introduction to Systemic Functional Linguistics*. Continuum.
- Eggs, S. and Slade, D. (1997) *Analysing Casual Conversation*. London: Equinox.
- Ekman, P. and Friesen, W.V. (1969) 'The repertoire of non-verbal behaviour: categories, origins, usage, and coding'. *Semiotica* 1: 49-98.
- Ellis, R. (1994) *The Study of Second Language Acquisition*. Oxford University Press.
- Ellis, R. (2003) *Task-based Language Learning and Teaching*. Oxford University Press.
- English, F. (2012) *Student Writing and Genre: reconfiguring academic knowledge*. London and New York: Bloomsbury Academic.
- Erlam, R. (2008) 'What do teachers know about language teaching? Bridging the gap between research and language pedagogy'. *Innovations in Language Learning and Teaching* 2, No. 3: 163-179.
- Fairclough, N. (1995) *Critical discourse analysis: the critical study of language*. Longman.
- Finch, G. (2000) *Key Concepts in Language and Linguistics*. Palgrave MacMillan.
- Flewitt, R. Hampel, R., Hauck, M. and Lancaster, L. (2010) 'What are multimodal data and transcription?' in Jewitt, C. (ed.) *The Routledge Handbook of Multimodal Analysis*. Routledge, Taylor Francis Group, New York and London.
- Flowerdew, L. (2010) 'Using corpora for writing instruction' in O'Keeffe, A. and McCarthy, M. (eds.) *The Routledge Handbook of Corpus Linguistics*. Routledge, Taylor Francis Group, New York and London.
- Forgas, J.P. (1985) *Interpersonal Behaviour. The Psychology of Social Interaction*. Pergamon Press.
- Fowler, R. and Kress, G. (1979) 'Critical Linguistics' in Fowler, R., Hodge, B., Kress, G. and Trew, T. *Language and Control*. London: Routledge & Kegan Paul.
- Gairns, R. and Redman, S. (2002) *Natural English Intermediate*. Oxford University Press.
- Gammidge, M. (2004) *Speaking Extra*. Cambridge University Press.

- Gilquin, G. and Granger, S. (2010) 'How can data-driven learning be used in language teaching?' in O'Keeffe, A. and McCarthy, M. (eds.) *The Routledge Handbook of Corpus Linguistics*. Routledge, Taylor Francis Group, New York and London.
- Gordon, T., Holland, J. and Lahelma, E. (2001) 'Ethnographic Research in Educational Settings' in Atkinson, P., Coffey, A., Delamont, S., Lofland, J. and Lofland, L. (eds.) *Handbook of Ethnography*. Sage.
- Granger, S. (2009) 'The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation' in Aijmer, K. (ed.) *Corpora and Language Teaching*. John Benjamins Publishing Company.
- Gray, J. (2002) 'The global Coursebook in English Language Teaching' in Block, D. and Cameron, D. (eds.) *Globalization and Language Teaching*. Routledge, Taylor Francis Group, New York and London.
- Graziano, M. (2014) 'Gestures in Southern Europe: Children's pragmatic gestures in Italy' in Muller, C., Cienki, A., Fricke, E., Ladewig, S.H., McNeill, D. and Bressemer, J. (eds.) *Body – Language – 2*. De Gruyter Mouton, Berlin and Boston.
- Halliday, M.A.K. (1975) *Learning How to Mean – Explorations in the Development of Language*. Edward Arnold.
- Halliday, M.A.K. (1978) *Language as Social Semiotic*. London: Arnold.
- Halliday, M.A.K. (1985/1994) *An Introduction to Functional Grammar*. 2nd edition. Edward Arnold.
- Halliday, M.A.K., (1989) *Spoken and Written Language*. Oxford University Press.
- Harmer, J. (1991) *The practice of English Language Teaching*. Longman.
- Harmer, J. (1996) 'Is PPP dead?' *Modern English Teacher*. 5 (2): 7-14.
- Harmer, J. (1998) *How to teach English: an introduction to the practice of English language teaching*. Longman.
- Haight, J.R. and McCafferty, S.G. (2008) 'Embodied language performance: drama and the ZPD in the second language classroom' in Lantolf, J.P. and Poehner, M.E. (eds.) *Sociocultural Theory and the Teaching of Second Languages*. London: Equinox.
- Hewings, A. and Hewings, M. (2004) 'Impersonalizing Stance: A Study in Anticipatory 'it' in Student and Published Academic Writing' in Coffin, C., Hewings, A. and O'Halloran, K. (eds.) *Applying English Grammar*. The Open University.
- Hinde, R.A. (1972) *Non-Verbal Communication*. Cambridge University Press.
- Hodge, R. and Kress, G. (1988) *Social Semiotics*. Cornell University Press.
- Hornby, A.S. (1950a) 'The Situational Approach in Language Teaching (I)'. *ELT Journal* IV (4): 98-103. Oxford University Press.
- Hornby, A.S. (1950b) 'The Situational Approach in Language Teaching (III)'. *ELT Journal* IV (6): 150-156. Oxford University Press.
- Howard, M. (2008) *Spoken Dialogues in English Language Teaching Text Books: The Speech Genre of Asking for and Giving Directions*. MA Dissertation, London Metropolitan University.

Michael John Howard A Multimodal Perspective on Modality in the English Language Classroom

Howard, M. (2010) 'A discourse approach to teaching modal verbs of deduction'. *Folio* Vol. 14 No. 2: 21-23.

Howatt, A.P.R. (2004) *A History of English Language Teaching*. Oxford University Press.

<http://www.bbc.co.uk/news/science-environment-28854809> (accessed on 19 August 2014)

<http://www.cambridgeenglish.org/teaching-english/> (accessed 2012)

Hughes, R. (2010) 'What a corpus tells us about grammar teaching materials' in O' Keeffe, A. and McCarthy, M. (eds.) *The Routledge Handbook of Corpus Linguistics*. Routledge, Taylor Francis Group, New York and London.

Hymes, D. H. (1972) 'On Communicative Competence' in Pride, J.B. and Holmes, J. (eds.) *Sociolinguistics. Selected Readings*. Harmondsworth: Penguin.

Hymes, D. H. (1979) 'On Communicative Competence' (extracts) in Brumfit, C.J. and Johnson, K. (eds.) *The Communicative Approach to Language Teaching*. Oxford University Press.

Jacobs, G.M. and Ball, J. (1996) 'An investigation of the structure of group activities in ELT coursebooks'. *ELT Journal* 50/2: 99-107. Oxford University Press.

Jaworski, A. and Thurlow, C. (2009) 'Gesture and movement in tourist spaces' in Jewitt, C. (ed.) *The Routledge Handbook of Multimodal Analysis*. Routledge, Taylor Francis Group, New York and London.

Jewitt, C. (2009) 'An introduction to multimodality' in Jewitt, C. (ed.) *The Routledge Handbook of Multimodal Analysis*. Routledge, Taylor & Francis Group, New York and London.

Johansson, S. (2009) 'Some thoughts on corpora and second-language acquisition' in Aijmer, K. (ed.) *Corpora and Language Teaching*. John Benjamins Publishing Company.

Johnstone, B. and Marcellino, W. (2010) *Dell Hymes and the Ethnography of Communication*. Carnegie Mellon University, Research Showcase @ CMU
<http://repository.cmu.edu/cgi/viewcontent.cgi?article=1013&context=english> (accessed September 2013).

Kendon, A. (1990) *Conducting Interaction. Patterns of Behaviour in Focused Encounters*. Cambridge University Press.

Kendon, A. (1995) 'Gesture as illocutionary and discourse structure markers in southern Italian conversation'. *Journal of Pragmatics* 23: 247-279.

Kendon, A. (2004) *Gesture: Visible Action as Utterance*. Cambridge University Press.

Kennedy, G. (1998) *An Introduction to Corpus Linguistics*. Longman: London and New York.

Kerr, P. and Jones, C. (2007) *Straightforward Upper Intermediate*. MacMillan.

Kidwell, M. (2014) 'The role of gaze in conversational interaction' in Muller, C., Cienki, A., Fricke, E., Ladewig, S.H., McNeill, D. and Bressemer, J. (eds.) *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction*. Volume 2. De Gruyter Mouton, Berlin and Boston.

Kramersch, C. (1993) *Context and Culture in Language Teaching*. Oxford University Press.

- Kreidler, C.W. (1998) *Introducing English Semantics*. London, Routledge.
- Kress, G. and Fowler, R. (1979) 'Interviews' in Fowler, R., Hodge, B., Kress, G. and Trew, T. *Language and Control*. London: Routledge Kegan and Paul.
- Kress, G. and Hodge, R. (1979) *Language as Ideology*. London: Routledge Kegan & Paul.
- Kress, G. (1982) *Learning to Write*. Routledge, Taylor Francis Group, New York and London.
- Kress, G. and Van Leeuwen, T. (2006) *Reading Images. The Grammar of Visual Design*. 2nd Edition. Routledge.
- Kress, G., Jewitt, C., Ogborn, J. and Tsatsarelis, C. (2001) *Multimodal Teaching and Learning: the Rhetorics of the Science Classroom*. Continuum.
- Kress, G. (2009) 'What is mode?' in Jewitt, C. (ed.) *The Routledge Handbook of Multimodal Analysis*. Routledge, Taylor and Francis Group, New York and London.
- Kumaravadivelu, B. (2003) *Beyond Methods: Macrostrategies for Language Teaching*. Yale University Press.
- Lave, J. and Wenger, E. (1991) *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.
- Lee, D.Y.W. (2010) 'What corpora are available?' in O'Keeffe, A. and McCarthy, M. (eds.) *The Routledge Handbook of Corpus Linguistics*. Routledge, Taylor Francis Group, New York and London.
- Leech, G. and Svartvik, J. (1975) *A Communicative Grammar of Spoken English*. Longman.
- Lewis, M. (1993) *The Lexical Approach*. Hove: LTP.
- Liddicoat, A.J. (2011) *An Introduction to Conversation Analysis*. Continuum.
- Lillis, T. (2001) *Student Writing: Access, Regulation, Desire*. Routledge, London.
- Littlejohn, A.P. (1992) *Why are English Language Teaching materials the way they are?* Unpublished PhD thesis.
- Littlewood, W. (1981) *Communicative language teaching: an introduction*. Cambridge University Press.
- Littlewood, W. (2004) 'The task-based approach: some questions and suggestions'. *ELT Journal* 58 (4): 319-326. Oxford University Press.
- Malinowski, B. (1923) 'The Problem of Meaning in Primitive Languages' in Ogden, C.K. and Richards, I.A. *The Meaning of Meaning*. New York: Harcourt, Brace and World, Inc.
- Martin, J.R. and Rose, D. (2003/2011) *Working with discourse*. 2nd edition. Continuum.
- Masuhara, H., Hann, N., Yong, Y. and Tomlinson, B. (2008) 'Adult EFL courses'. *ELT Journal* 62 (3): 294-312. Oxford University Press.
- McCarten, J. (2010) 'Corpus-informed course book design' in O'Keeffe, A. and McCarthy, M. (eds.) *The Routledge Handbook of Corpus Linguistics*. Routledge, Taylor Francis Group, New York and London.

Michael John Howard A Multimodal Perspective on Modality in the English Language Classroom

McCarthy, M. (1991) *Discourse Analysis for Language Teachers*. Cambridge University Press.

McCarthy, M. (1998) *Spoken Language and Applied Linguistics*. Cambridge University Press.

McCarthy, M. (2008) 'Accessing and Interpreting Corpus Information in the Teacher Education Context'. *Language Teaching* 41 (4): 563-574.

McCarthy, M. and Carter, R. (1995) 'Spoken grammar: what is it and how can we teach it?' *ELT Journal* 49/3: 207-218. Oxford University Press.

McCarthy, M. and O'Dell, F. (2008) *Academic Vocabulary in Use*. Cambridge University Press.

McConachy, T. (2009) 'Raising sociocultural awareness through contextual analysis: some tools for teachers'. *ELT Journal* 63 (2): 116-125. Oxford University Press.

McEnery, T. and Wilson, A. (1996/2005) *Corpus Linguistics*. 2nd Edition. Edinburgh University Press.

McNeill, D. (2005) *Gesture & Thought*. University of Chicago Press.

Murphy, R. (1994) *English Language in Use*. Cambridge University Press.

Nitta, R. and Gardner, S. (2005) 'Consciousness-raising and Practice in ELT Coursebooks'. *ELT Journal* 59/1: 3-13. Oxford University Press.

Norris, S. (2004) *Analysing Multimodal Interaction*. Routledge, Taylor Francis Group, New York and London.

Norris, S. (2009) 'Modal density and modal configurations. Multimodal actions' in Jewitt, C. (ed.) *The Routledge Handbook of Multimodal Analysis*. Routledge, Taylor Francis Group, New York and London.

Norton, B. (2000) *Identity and Language Learning*. Longman.

Nunan, D. (March 2000) *Seven Hypotheses about Language Teaching and Learning*. Plenary presentation, TESOL convention, Vancouver.

Ong, W.J. (1982/2002) *Orality and Literacy*. Routledge, Taylor Francis Group, London and New York.

Parrott, M. (2000) *Grammar for English Language Teachers*. Cambridge University Press.

Paulston, C.B. (1992) *Linguistic and Communicative Competence*. Multilingual Matters.

Philpot, S. and Curnick, L. (2011) *Headway Academic Skills*. Oxford University Press.

Rampton, M.B.H. (1990) 'Displacing the "native speaker": expertise, affiliation, and inheritance'. *ELT Journal* 44 (2): 97-101. Oxford University Press.

Redston, C. and Cunningham, G. (2007) *Face 2 Face Upper Intermediate*. Cambridge University Press.

Ricci Bitti, Pio E. (2014) 'Facial expression and social interaction' in Muller, C., Cienki, A., Fricke, E., Ladewig, S.H., McNeill, D. and Bressemer, J. (eds.) *Body – Language – Communication. An Interactional Handbook on Multimodality in Human Interaction. Volume 2*. De Gruyter Mouton, Berlin and London.

Richards, J.C. and Rodgers, T.S. (2001) *Approaches and Methods in Language Teaching*. Cambridge University Press.

Rivers, W.M. (1968/1981) *Teaching Foreign Language Skills*. 2nd edition. University of Chicago Press.

Michael John Howard A Multimodal Perspective on Modality in the English Language Classroom

Roach, P. (1983) *English Phonetics and Phonology*. Cambridge University Press.

Ruhlemann, C. (2010) 'What can corpus tell us about pragmatics?' in O'Keeffe, A. and McCarthy, M. (eds.) *The Routledge Handbook of Corpus Linguistics*. Routledge, Taylor Francis Group, New York and London.

Saeed, J.I. (2009) *Semantics*. 3rd edition. Oxford: Wiley-Blackwell.

Samuda, V. (2001) 'Guiding relationships between form and meaning during task performance: The role of the teacher' in Bygate, M., Skehan, P. and Swain, M. (eds.) *Researching Pedagogic tasks. Second Language Learning, Teaching and Testing*. Longman. Pearson Education.

Scollon, R. and Scollon, S.W. (2009) 'Multimodality and language. A retrospective and prospective view' in Jewitt, C. (ed.) *The Routledge Handbook of Multimodal Analysis*. Routledge, Taylor and Francis Group, New York and London.

Scott, M. (1999) 'Agency and subjectivity in student writing', in Jones, C., Turner, J. and Street, B. (eds.) *Students Writing in Higher Education*. Benjamins, Amsterdam.

Scrivener, J. (1994) *Learning Teaching*. Heinemann.

Sieloff Magnan, S. (2008) 'The unfulfilled promise of teaching for communicative competence' in Lantolf, J.P. and Poehner, M.E. (eds.) *Sociocultural Theory and the Teaching of Second Languages*. London: Equinox.

Sinclair, J.M. (1991) *Corpus, Concordance, Collocation*. Oxford University Press.

Skehan, P. (1998) *A Cognitive Approach to Language Learning*. Oxford University Press.

Soars, L. and Soars, J. (2005) *New Headway Upper Intermediate*. 3rd Edition. Oxford University Press.

Soars, L. and Soars, J. (2009) *New Headway Intermediate*. 4th edition. Oxford University press.

Stenglin, M. (2009) 'Space and communication in exhibitions. Unravelling the nexus' in Jewitt, C. (ed.) *The Routledge Handbook of Multimodal Analysis*. Routledge, Taylor Francis Group, New York and London.

Swan, M. (1980) *Practical English Usage*. Oxford University Press.

Swan, M. and Walter, C. (1997) *How English Works*. Oxford University Press.

Tavakoli, P. and Howard, M.J. (2012) 'TESOL teachers' views on the relationship between research and practice'. *European Journal of Teacher Education* Vol. 35 No. 2: 229-242. Routledge, Taylor and Francis Group.

Tavakoli, P. 'Connecting Research and Practice in TESOL: A Community of practice Perspective'. *RELC Journal*. February, 20, 2015, DOI: 10. 1177/0033688215572005.

Tesco Food, Family Living Magazine. April 2014.

Ten Have, P. (1999) *Doing Conversation Analysis*. Sage.

Thornbury, S. (2006) *An A-Z of ELT*. MacMillan Education Australia.

Thornbury, S. (1997) *About Language*. Cambridge University Press.

Michael John Howard A Multimodal Perspective on Modality in the English Language Classroom

- Thornbury, S. and Slade, D. (2010) 'The Grammar of Conversation' in Cook, G. and North, S. (eds.) *Applied Linguistics in Action*. Routledge, Taylor Francis Group, New York and London.
- Tilbury, A. and Hendra, A. (2011) *English Unlimited Upper Intermediate*. Cambridge University Press.
- Tomlinson, B. (1998) *Materials Development in Language Teaching*. Cambridge University Press.
- Ur, P. (1996) *A Course in Language Teaching. Practice and Theory*. Cambridge University Press.
- Van Ek, J. (1979) 'The Threshold Level' (extracts) in Brumfit, C.J. and Johnson, K. (eds.) *The Communicative Approach to Language Teaching*. Oxford University Press.
- Van Leeuwen, T. (1999) *Speech, Music, Sound*. MacMillan.
- Van Leeuwen, T. (2005) *Introducing Social Semiotics*. Routledge.
- Volosinov, V.N. (1929/1986) *Marxism and the Philosophy of Language*. (trans. Matejka, I. and Titunik, I.R.) Cambridge, MA and London: Harvard University Press.
- Walford, G. (2008) (ed.) *How to do educational ethnography*. the Tufnell Press, London.
- Widdowson, H.G. (1978) *Teaching Language as Communication*. Oxford University Press.
- Widdowson, H.G. (1979) 'The Teaching of English as Communication' in Brumfit, C.J. and Johnson, K. (eds.) *The Communicative Approach to Language Teaching*. Oxford University Press.
- Widdowson, H.G. (2004) *Text, Context, Pretext*. Blackwell.
- Wilkins, D.A. (1976) *Notional Syllabus*. Oxford University Press.
- Wilkins, D.A. (1979) 'Grammatical, Situational and Notional Syllabuses' in Brumfit, C.J. and Johnson, K. (eds.) *The Communicative Approach to Language Teaching*. Oxford University Press.
- Willis, J.R. (1996) *A Framework for Task-based Learning*. Longman.
- Willis, D. and Willis, J. (2010) 'Six propositions in search of a methodology: applying linguistics to task-based language teaching' in Hunston, S. and Oakley, D. (eds.) *Introducing Applied Linguistics. Concepts and Skills*. London/New York: Routledge.
- Wilson, J.J. (2007) *Total English Advanced*. Pearson Longman.
- Wittenburg, P., Brugman, H., Russel, A., Klasmann, A. and Sloetjes, H. (2006) *ELAN: a Professional Framework for Multimodality Research*. Max Planck Institute.
- Woodward, T. (2001) *Planning Lessons and Courses*. Cambridge University Press.
- Young, R.F. (2009) *Discursive Practice in Language Learning and Teaching*. Wiley-Blackwell.

Appendices

Appendix 1 *Language Leader Intermediate* audio scripts

Anna, Barbara

A: Wow, look at that. What is it?

B: It's a drawing. I think it's a sort of flying machine.

A: Yes, it could be that. Who do you think it's by?

B: Mmm, it might be by da Vinci, I believe he did that sort of thing. Have a look at the sign. What does it say?

A: Erm ... yes, you're right, it is by da Vinci.

B: Goodness! It's in very good condition ... it says here he was born in 1452, so it must be over 500 years old.

A: Yes, and it's an amazing drawing.

B: Yeah, I read somewhere he was fascinated by birds and flying. Perhaps that's where he got his ideas for the drawing.

A: Yeah, you're probably right.

Elias, Freddie

E: Freddie, I want to see the racing cars. Where are they?

F: I've got a guide, but I'm sure they're in the large space at the back. Let's go there first.

E: Here they are. Look at that one over there. It must be a Ferrari, surely.

F: Yes, it can't be anything else. It's so red and stylish. But it's a pretty old one – look it's quite high off the ground. What year do you think it is?

E: It could date back to the 1930s.

F: That's impossible. It can't be that old. Look, here it says it was made in the 1940s.

E: You're right. It's amazing that they were able to go so fast at that time.

George, Sally

G: What's that?

S: I've no idea. What could it be?

G: It might be a spaceship. Well, a toy spaceship. It's the right shape.

S: No, I'm certain that it's not that. It wouldn't really be in a museum of design.

G: Mmm, I see what you mean. And maybe it's too heavy to be a toy. What else could it be?

S: Mmm, I don't know really. What does it say on the notice?

G: It says it's a lemon squeezer. Apparently it's the designer Philippe Starck's best-known design.

S: OK, I see it now. Would you like something like that?

G: No way! It just doesn't look practical.

Appendix 2 Language Leader Intermediate lesson

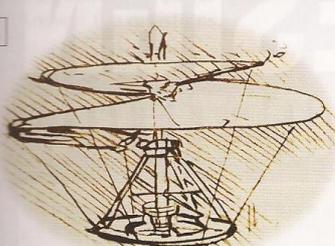
7.3

A LIFE OF DESIGN

LISTENING

3a 1.41 Listen to three pairs of students at a museum of design. In which order do they talk about the things below?

A



B



C



3b Listen again. Are these sentences true, false or not given?

- 1 The first design is by da Vinci.
- 2 This design is over 500 years old.
- 3 Elias and Freddie are looking at a Ferrari.
- 4 The car was made in the 1950s.
- 5 The third object is the designer's best-selling design.
- 6 George would like to buy one of these.

GRAMMAR: modals (present deduction)

4a Look at Track 1.41 on page 175 and underline the modal verbs *must*, *can't*, *might* and *could*. Look at the words around them.

4b Match each modal verb with one of these meanings.

1 It can't be true.	a) I think this is possible.
2 It might/could be true.	b) I'm certain that this is true.
3 It must be true.	c) I'm certain that this is not true.

4c Look at the modal verbs in Exercise 4b. What is the opposite of *must be* when we are talking about deduction?

GRAMMAR TIP

We can also use modal verbs with a continuous form:

She might be feeling ill. (= I think she's feeling ill.)

You must be joking! (= I'm certain you're joking.)

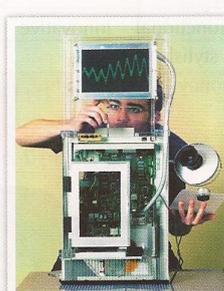
➔ Language reference and extra practice, pages 146–147

5 Rewrite the sentences below using *must*, *can't*, *could* or *might*.

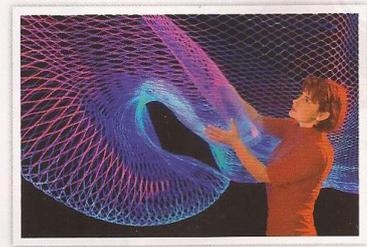
- 1 Raymond Loewy is definitely one of the most influential designers of all time.
- 2 I'm sure this design is by Armani.
- 3 I'm sure this painting is not by Leonardo.
- 4 Maybe this painting is by Picasso, but I'm not sure.
- 5 Not many people are attending the exhibition. I'm sure they aren't promoting it very well.
- 6 This product looks dangerous to me.
- 7 People are not sure if it is a Loewy design.
- 8 I'm sure that designer is working very hard. I've seen a lot of her designs.

6 Work with a partner to discuss what you think the designs below are. Try to use these words and phrases.

must might could can't
 I'm sure/certain ... maybe/perhaps
 It's possible that ...
 It's not possible that ...







SPEAKING

7 If you had the skills, what would you like to design/re-design?

Design UNIT 7 75

Appendix 3 Transcription conventions

(number)	timed pause
(.)	short pause
()	empty brackets indicate uncertain/inaudible speech
(())	double brackets add additional information
(h)	indicates breathiness
?	rising intonation
↑	indicates a higher pitch within an utterance
↓	indicates a lower pitch within an utterance
.	indicated a stopping fall in intonation
–	underscoring indicates some form of stress
=	at the end of one line and the beginning of another indicates no gap between utterances
[indicates a point of overlap
><	at the beginning and end of an utterance indicates slowing down
<>	at the beginning and end of an utterance indicates speeding up
:	indicates a longer sound – multiple colons indicate more prolonged sound
°°	utterances beginning and ending with degree signs are spoken more quietly
S1, S2 etc.	speaker 1, speaker 2, etc.

Appendix 4 You must be joking! BYU-BNC corpus

good sleep. (SP:PSOWS) Oh yeah! So you're gon na get rid of me early today. (SP:PSOWN) Yippee!
 (pause) No I just want (SP:PSOWS) Why mum? (SP:PSOWN) I just think you should have a good sleep.
 (SP:PSOWS) I'll go to bed the usual time. (SP:PSOWN) Eight o'clock. (SP:PSOWS) I don't wan na
 (pause) I'm not! (SP:PSOWN) You are! (SP:PSOWS) I said, I'm not! (SP:PSOWN) You can go up and get
 a bit of reading done in your bed at eight o'clock. (SP:PSOWS) Must be da-- **you must be joking!**
 (SP:PSOWN) I'm not joking, I'm serious! (SP:PSOWS) Tt! (banging) (SP:PSOWN) Wish I was going to my
 bed at eight o'clock. (banging) (SP:PSOWS) But mum (SP:PSOWN) Do that again (SP:PSOWS) yo-- will
 you leave us alone! (SP:PSOWN) You go to bed eight o'clock. (SP:PSOWS) (sigh) (SP:PSOWN) You need
 the rest. (SP:PSOWS) (screaming) I couldn't give a blooming damn! (SP:PSOWN) I don't care. You
 want yo-- (pause) the rest. (SP:PSOWS) (crying) But I'm not! (SP:PSOWN) You are. Now do

Speaker information for PSOWS	
Name:	David
Number of words:	4,669
Number of turns:	681
Standard header information:	
Sex:	Male
Age:	0-14
Social Class:	C1
Education:	n/a
First Language:	n/a
Dialect/Accent:	Scottish
Additional information:	
Age:	n/a
Occupation:	student (state primary)
Role:	Son

Speaker information for PSOWN	
Name:	Valerie
Number of words:	6,845
Number of turns:	1,120
Standard header information:	
Sex:	Female
Age:	35-44
Social Class:	C1
Education:	n/a
First Language:	n/a
Dialect/Accent:	Scottish
Additional information:	
Age:	n/a
Occupation:	staff nurse (pt)
Role:	Self

Appendix 5 Cirio advertisement

(Tesco Food April, 2014)

CIRIO, the true taste of Italy.



Autentico Italiano dal 1856

Francesco Cirio was the canned vegetable industry pioneer, starting back in 1856. Nowadays his Cirio brand is loved throughout Italy along with 70 countries all over the world. With top quality produce from a huge farmers Cooperative, Cirio products are controlled from "seeds to table".

The exquisite taste of our juiciest Italian tomatoes is created with care and expertise, by processing them in just 24 hours. We bring true Italian flair to your cooking and you will taste the difference!

 www.cirio1856.com

 FIND US ON FACEBOOK
[Facebook.com/CirioUK](https://www.facebook.com/CirioUK)

SILVANO GUIDONE & ASSOCIATI

Worldwide Limited. ©2013 Twentieth Century Fox Home Entertainment LLC. All Rights Reserved. Artwork subject to change.

[Appendix 6 Classroom data \(CD ROM\)](#)

Appendix 7 Student consent form

I am a PhD student at London Metropolitan University. As part of my PhD research I am collecting video recorded data from English language classes. I intend to use this data to analyse how students interact with each other and their teachers in a typical English language lesson. The data I collect will be used for the purposes of my research only and no one else will have independent access to it. All names of those who wish to participate will be changed to ensure anonymity.

If you would like more information about my research, you can email me at [REDACTED] where I would be more than happy to discuss any aspect of my research. My PhD supervisor is Dr Fiona English of London Metropolitan University and she can be contacted at [REDACTED] if you have any queries about me or my research.

I agree/do not agree to take part in this research.

Signed _____.

Name

Nationality

Languages spoken

Number of years studying English

Any additional comments

Appendix 8 Speech functions

Move + time	Speaker T	Speaker N	Speech function	Discourse purpose
1 24:38	what's this?		question: open: fact	Initiating conversation; demanding factual information
2 24:39		yeah what's this?	question: responding: support: engaging (p.204)	Reacting and showing willingness to interact
3 24:42		looks like a diagram	respond: support: reply: answer	Providing information to Move 1
4 24:44	uh is this correct direction or rotation	(laughs 24:48)	question: closed: opinion	Demanding agreement with opinion information
5 24:49	or something		Sustain: continue: prolong: extend	Offering contrasting information
6 24:50	you think?		Sustain: continue: monitor	Checking participant engagement
7 24:52		I have no idea what that is	Respond: support: supply: answer	Indicating inability to provide information demanded in Move 1
8 24:54		some equipment?	Rejoinder: support: track: probe	Volunteering information for confirmation
9 24:57	some equipment sss		Respond: support: register	Displaying attention to Move 8
10 25:00	maybe that way's good		Sustain: continue: append: extend	Offering addition information to Moves 4,5,& 6
11 25:03	Uh?		Sustain: continue: monitor	Checking participant engagement
12 25:04	yeah I have no idea as well		Respond: support: register	Displaying attention to Move 7

13	25:07		uh this looks like it rotates	Sustain: continue: prolong: elaborate	Clarifying Move 8
14	25:08		some kind of fan or	Sustain: continue: prolong: elaborate	Exemplifying Move 13
15	25:12		I don't know	Sustain: continue: prolong: enhance	Modifying Move 14
16	25:13	Uh no		Respond: confront: reply: disagree	Providing negative response to Move 14
17	25:15	lamp?		Open: initiate: information: opinion	Seeking agreement with opinion
18	25:18	no?	(inaudible)	Sustain: continue: monitor	Checking participant engagement
19	25:19	(inaudible)		-	-
20	25:21	I have no idea of this		Continue: prolong: extend	Offering contrasting information
21	25:29	yeah		Continue: prolong: elaborate	Clarifying Move 20
22	25:31	I don't know what is it		Continue: prolong: elaborate	Restating Move 20
23	25:32	what it is		Continue: prolong: elaborate	Clarifying grammatical error in Move 22
24	25:32		no	Respond: support: register	Showing support to participant

Appendix 9 Sample lessons

The sample lessons I put forward here I have used on different occasions. I have found them useful for teaching and raising awareness of modality in terms of how the interpersonal and discourse function of modality relates to context.⁵⁷ I point out that the language focus is *modality* and not modals of present deduction as in the lesson I have critiqued from *Language Leader Intermediate*. (The task-based lesson can give rise to any amount of language, I refer to modality as this has been the focus of my thesis.) As with all lessons they need to be adapted to the particular group of learners and the teaching and learning situation: the teaching notes provide guidance only. Furthermore a certain amount of ground work is required to individual lessons such as introducing learners to the concept of interpersonal meaning and communicative modes such as gesture and eye gaze; once this is done it can be extended to other tasks and lessons.

I have tried to make the sample lessons as varied as I can to show the scope of what can be done. The lesson on epistemic modality in academic writing is largely text-based; the business proposition lesson promotes discussion and introduces role play; multimodal spoken interaction uses multimedia for discursive purposes. Of course, whilst the lessons may prove workable, more research is needed to establish their effectiveness for developing interlanguage and how the study of multimodality might integrate into a methodological framework for teaching and learning.

9.1 Data-driven learning: Epistemic modality in academic writing

Rationale

I have argued above that the selection of modals in *Language Leader Intermediate* is not referenced to spoken or written English. The authors of the course book inform us that as a course, *Language Leader* is suitable for those already in or aiming to enter higher education (Cotton et al., 2008). Arguably then a knowledge of which modal structures are favoured in writing and speech is useful. A frequent question I am asked is ‘What is the difference between *might* and *may*?’ Additionally the learners are interested to know how they can be used in academic writing. The inductive approach taken by data-driven learning is an ideal way of addressing this question. Learners become familiar with language use as well as with the academic texts they explore to answer their question.

Outcomes

- Formulating a research question, carrying out research and drawing conclusions.
- Using inductive skills to raise language awareness.
- Interacting with language related technology to explore academic texts to enhance written English.

Procedure

1. Formulate the question:
For example, which modal form is more frequent in academic writing ‘must’ or ‘may’? And is there any difference in meaning?
2. Devise a research strategy:
Learners access the free online corpus Brigham Young University-British National Corpus <http://corpus.byu.edu/bnc/> or British National Corpus <http://www.natcorp.ox.ac.uk/> and

⁵⁷ I have also received positive feedback from Professor Keith Folse of University of Central Florida who uses my A Discourse Approach to Teaching Modal Verbs of Deduction to teach ‘the interpersonal value of modals’ (personal correspondence).

consult corpora-based references such as *The Longman Grammar of Spoken and Written English* (Biber et al., 2002).

3. Observe the examples and select the relevant ones:

Learners collect information relevant to their research question. For example, *may* occurs much more frequently than *might*. Both forms are used to express logical possibility (epistemic meaning) although some instances of *may* are used for permission.

4. Draw conclusions:

E.g. both modal forms can be used in academic writing although *may* is preferred.

Additions

- Learners can work in groups and research questions collectively. These can then be used in presentations in front of the class.
- The above example may seem relatively straight-forward to anyone familiar with using corpora; however, for the learners I have worked with it is not so. It is therefore worthwhile getting the most out of the task. Corpora are able to tell us about collocations and colligations and so a useful addition to the above is for learners to observe how the modal forms collocate and colligate.
- The following exercise comes from *Headway Academic Skills*:
'Rewrite the statements to make them less certain, using the words in brackets'
 1. Buying more fair trade food leads to the increase of use in places. (may)
 2. Globalization has an effect on local culture and traditions. (could)
 3. An increase in exports leads to an increase in the number of jobs. (might)'

(Philpot and Curnick, 2011: 39)

As I have argued in my thesis, rewriting sentences of this kind does little more than practice language usage. It is questionable as to whether learners have learnt anything about the interpersonal discourse function of modality from the above exercise. Using corpora, learners can access the texts in which utterances are made to explore how the writer uses modality to position him/herself vis-à-vis the reader. Learners can then explore how meaning and stance change with the use of different modal structures. This can be done through rewriting texts to express different viewpoints.

- Exploring corpora for other modal structures commonly used in academic writing.

9.2 Task-based instruction: A business proposition

This lesson has been adapted from Gammidge (2004).

Rationale

In this lesson there is a focus on meaning with learners engaged in a task of discussion. The purpose of the task is for learners to speculate, deduce, negotiate, propose, suggest, persuade, agree and disagree as to which business would work best in a given location. This could give rise to any amount of language but here the focus is on modality. I have argued that modality is signalled using a range of lexicogrammatical structures and so this needs to be taken into account when it comes to the language focus – this is not a lesson on modal auxiliary verbs. During the group discussion and feedback, notes or audio/video recordings taken by the teacher on examples of the language the learners use become the language focus of the lesson; this could be errors or areas to develop interlanguage. The emphasis is not on the immediate production of linguistic structures, but raising

awareness through study and discussion, and then providing a communicative space for discussion in which there is a focus on meaning before form.

Outcomes

- Completion of the task through group discussion.
- The use and study of language functions such as negotiation, persuasion, agreement and disagreement.
- The use and study of the interpersonal function of modality in spoken language.
- Raising awareness of modal structures beyond modal auxiliary verbs.

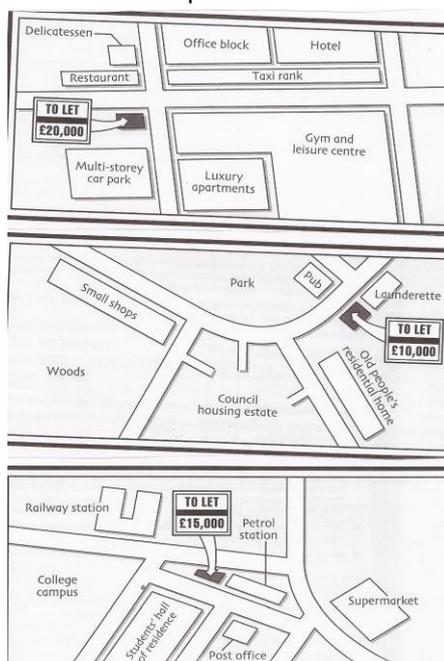
Procedure

1. Pre-task

- Learners discuss where they live and what local amenities there are in their area.
- I give an anecdote: I live in East London. There is a big Asian community there so it's easy to eat out in Asian restaurants and buy Asian food in local shops. But there is something really important missing: there is absolutely nowhere to buy a good cup of coffee...or a nice piece of cake! I think the area would really benefit from a coffee shop selling cakes and light snacks. I reckon it would generate good business and would probably be a great success. If it was big enough it could even have live music. I seriously think it would be a sound investment and I know it would pull in customers. I might even open a café there myself!
- The learners then discuss what amenities are missing from their part of town and how the area would benefit.

2. Task

- Hand out the pictures below one at a time and ask the learners what they think would be a good business to set up at each location.



- The learners form groups and are given the following rubric:

Imagine that you are business people who are going to open a shop or start a business together. There are three possible shops available that you could rent, shown in the maps below. Discuss the alternatives. Decide what type of shop or business would be successful and choose the best neighbourhood for it.

- The groups discuss and decide what kind of business would be successful and choose the best neighbourhood for it.
 - The groups feedback to the whole class explaining their ideas and giving reasons. (During the task stage the teacher collects, by making notes or recording, examples of modal structures that the learners use in their discussions, these then becomes the language focus of the lesson.)
3. Language focus
- The modal structures that have been collected during the discussion and feedback session become the focus of a focus on form. This can be carried out by using learners' errors or eliciting/suggesting alternatives as a means of developing learner interlanguage.
 - The teacher recaps each group's ideal shop or business and offers the other groups a chance to comment in an open forum. Here the learners have the opportunity to recycle the language focus.

Alternatives

- A more 'formal' alternative could have the learners prepare a formal presentation on their chosen shop or business. This works well if there is some kind of incentive. I have done this task in the style of 'Dragon's Den'.⁵⁸ I tell the class I have money to invest in the most viable business proposition. Groups can comment on each other's proposals to add to the competitive aspect.
- The following link can be used as a means of providing structure to a group discussion. Here there is opportunity for language study as well as looking at how participants use other communicative modes in social interaction:
<https://www.youtube.com/watch?v=eyWxjNECRBE>

9.3 Three Is: Multimodal spoken interaction

Rationale

Verbal language is one of many modes of communication in spoken interaction; studying how communicative modes integrate to make meaning aims to develop learners' multicompetence. Using audio and visual multimedia, learners explore how participants in a panel discussion integrate the modes of gesture, posture, eye gaze and spoken language to express interpersonal meaning.

Outcomes

- Raised awareness of how communicative modes integrate to make meaning.
- Study and discussion of the interpersonal function of communicative modes in spoken interaction.
- Use of metadiscursive, inductive and heuristic communication skills.

Procedure

1. Illustration

⁵⁸ See <http://www.bbc.co.uk/programmes/b006vq92>

- Play the excerpt following programme asking the class to make notes on the gist of each person's contribution to the discussion. Have a feedback session to establish what the learners have gleaned. (This task is not a listening activity in the traditional sense and does not lead to a comprehension exercise of form or meaning. The purpose of this stage is to familiarize the learners with the content of the discussion and the participants.)

http://www.channel5.com/shows/the-wright-stuff/episodes/friday-05-september?related=clips#related_contents (minutes: 17.00-23.44)

2. Interaction

- Play the excerpt again with the sound down and ask the class to attend to the modes of gesture, posture and eye gaze. (This can be done in a number of ways, i.e. assign a learner to a participant or a mode, or ask the class to note down anything that catches their attention.)
- Feedback as a class and discuss what emerged from the listening. Encourage the learners to think about what meanings the speakers assign to the modes; encourage them to think about how the modes integrate with spoken language.
- Play the excerpt again and ask the class to make notes on what spoken language was used alongside the modes they had previously focused on. (Again this is not a comprehension exercise: the learners do not need to write utterances down word for word.)
- Feedback as a class and discuss how the modes integrated and how they helped to convey interpersonal meaning. Contextual factors such as the participants and their social roles and relationships can be explored and what effect these have on language choice.

3. Induction

- For out of class work learners work in groups or individually to collect ethnographic data on how participants in social interaction use multimodal communication. Observations are then discussed or presented in class.

Alternatives

- The focus here has been on spoken communication, however, this approach can be adapted to written texts. Texts used for illustration and interaction can then be rewritten by learners at the induction stage to explore how stance and interpersonal meaning are affected by different choices of language and vice versa.
- Drama activities are a way exploring embodied actions (Haught and McCafferty, 2008). Learners assign themselves roles and engage in a panel discussions. The purpose of this activity is not to 'produce/activate' multimodal embodied actions in the sense of a PPP or an ESA type lesson, but to allow a creative exploration of multimodal communication.⁵⁹

⁵⁹ Dialogue Activities (Bilbrough, 2007) outlines some useful ways for learners to create their own personas and roles for drama, role play and discussion activities.

Appendix 10 Speaker information BYU-BNC corpus

Speaker information for PSOM5	
Name:	Oliver
Number of words:	7,767
Number of turns:	1,080
Standard header information:	
Sex:	Male
Age:	0-14
Social Class:	C2
Education:	n/a
First Language:	n/a
Dialect/Accent:	London
Additional information:	
Age:	n/a
Occupation:	student (private pre)
Role:	Son

Speaker information for PSOM4	
Name:	Nicola
Number of words:	14,770
Number of turns:	1,359
Standard header information:	
Sex:	Female
Age:	25-34
Social Class:	C2
Education:	n/a
First Language:	n/a
Dialect/Accent:	Home Counties
Additional information:	
Age:	n/a
Occupation:	Housewife
Role:	Self

Appendix 11 Transcripts

Transcript A

- T (Presents picture)
(3 secs)
- S1 What she's doing?
- S2 He=
- S1 =He ((laughter))
- S1/S2 (3 secs) uhm/uh
- S2 So here is the mouse (.) so he's searching something (=
- S1 =yeah research=
- S2 =research
(2secs)
- S2 ↓ maybe he's ↑ scientist and trying to find out something (2 secs) here's the diagram so it shows (.) maybe how fast runs ↓ the mouse and it affects that and that and there is the result of that=
- S1 =You have a very good imagination [S2 ((laughs))] () I don't know (.)
(3secs)
- S1 He's researching he's researching something
(2 secs)
- S2 And checking how it works=
- S1 =Yes ↑ recording ↓ maybe maybe he's recording something
(4 secs)

Transcript B

- T What about the second picture?
- S1 This B
- S2 Uhm
- S1 I think massage for head
- T Ah, ok
- S1 Yeah ((gestures to head)) I can see before (.) yeah? ((shrugs))
- T Yutsuke?
- S1 ((gestures with open hand)) I don't know

S3 I think that is hair massager
T OK (.) It's it's a lemon squeezer
S3 ((breathes out))=
S1 =↑wha::?
S2 Lemon?
T Lemon=
S2 =ah ((general laughter))=
T: =ah=
S2 I can't believe it ((laughter))

Transcript C

T What about the second picture?
S1 (.) This one yes?
S2 Signal transmission [S1 yes] something like that
T So uh an aerial you mean? (.) is [S2 yeah ((gestures))] like an a
S1 TV
T TV aerial
S1 TV TV aerial
T ok
S1 ↓something like that
T uh it's actually actually it's um (.) (.) a drawing by Leonardo Da Vinci [S1 ((breaths out)) oh yes] on his flying machine=
S1 =oh flying machine
S2 uh::: () ya
S1 ((pics up picture))

Transcript D

T ((Presents picture)) Picture 1
(4 secs)
S1 ((Looks at T, pointing at picture, beats with index finger)) I know
(5 secs)

- S1 This is (.) This is first (.) first (.) airplane
(7 secs)
- S1 It makes (.) ah (3 secs) it (.) it is made (2 secs) Leonardo Da Vinci (.)=
S2 =(Looks at T) This is looking satellite
S3 (.) I don't know [S2 ((laughs))] what it is [S2 strange] >< yes strange
(2 secs)
- S2 This person?
- S1 It is (.) it made Leonardo Da Vinci
(2 secs)
- S2 persons () ((gestures))
- S1 flying (.) [S3 oh] ((gestures)) (.) first first this is first (.) airplane (2 secs) first airplane
S3 maybe
- S1 () °°yeah
(4 secs)
- S3 I don't know >< what it is
S2 strange picture
S3 yes

Transcript E

- T There's picture 2
- S1 Oh=
S2 =Octopus= ((laughs with S1))
(2 secs)
- S3 This is a:::=
S2 =Yes=
S3 =Spider (.) f(h)ork
(2.8 secs)
- S1 It's like a (1.5) a:: (2.1) extra::
(1.5 secs)
- S2 Fork (.) spaceship ((laughs))
(2.3 secs)

- S2 I dunno
S3 It's like fork
(laughter)
(15 secs)
S2 ((inaudible))

Transcript F

- T Number 3
S1 Oh ↑uh
S2 Oh it frog
(1 sec)
S2 She wanna (.) like frog (.) so (.) [S1 ah] ya=
S1 =Lady Gaga ((laughs))
S2 =((laughs)) ha Lady Gaga ah ((inaudible))=
S1 =She could be::: >a mother of °Lady Ga°< ((laughs))
S2 ((laughs)) (1.5) but she looks (.) very old
S1 ((inaudible)) or maybe (.) she (.) she might be::: (2) she might be cleaning (1.5) her house (.)
↓or
S2 ↓ah

Transcript G

- T ((whispers)) third picture
(2 secs)
S1 [S2 car] l(h)ts (h)a c(h)ar=
S2 °car°
(5 secs)
S2 traditional car ((laughs)) (1.0) yeah
(11 secs)
S2 ((inaudible)) ((looks at teacher)) (2.0) °car yeah°
(10 secs)

Appendix 12 Multimodal moments



Plate A Posture – sitting upright.

S2 So here is the mouse (.) so he's searching something ()=

S1 =yeah research=

S2 =research

(2secs)

S2 ↓maybe he's ↑scientist and trying to find out something (2 secs) here's the diagram so it shows (.) maybe how fast runs↓the mouse and it affects that and that and there is the result of that=

S1 =You have a very good imagination [S2 ((laughs))] () I don't know (.)

In the first frame both speakers are silent. In the second frame (top right) S1 (frame right) takes S2's lead and as she straightens her body says 'yeah research'. Next there follows a two second pause before S2 begins on her long spoken turn. Whilst doing so she sits upright and S1 sits back to listen. Here we can see how posture co-ordinates with spoken turns.



Plate B Posture – F-formation. In the first frame S1 (left of frame) is discussing with S2. In terms of F-formation her body is somewhat turned away from S2. Also whilst the right shoulder of S2 is

inclined towards S1, her body position is somewhat rotated away from her. The second frame can perhaps throw some light in this when we learn that in this frame S1 and S2 are speaking with the teacher. The F-formation taken up by S1 and S2 includes the presence of the teacher – they are forming a triangle. This is an interesting perspective on their interaction as it shows they are conscious of the teacher, almost as part of the discussion even when they are speaking as a pair in the first picture, and most likely the presence of the camera.



Plate C Pointing – correction.

S1 What she's doing?

S2 He=

S1 =He ((laughter))

S2 (on the right) points and beats at the picture as she corrects S1 and says 'He'. While she does this she gazes towards S1.



Plate D Pointing - 'I think that is hair massager'. S2 (second from left) co-ordinates pointing and beats with his spoken utterance. S2's uses pointing as deixis and beats to mark prominence. The other students gaze attends to his gesture which extends into the gesture space.



Plate E Gaze – (gesture space). S1 (first left) extends her hand into the gesture space and uses an iconic gesture co-ordinated with ‘helicopter’. The two other students in the group attend to the gesture rather than the speaker’s face which is turned in the direction of the teacher.



Plate F Iconic gesture - ('massage for head'). In this frame S3 (third from left) is using an iconic gesture to show 'massage for head'. This gesture is not, however, co-ordinated with the spoken utterance 'massage for head'. Instead it is co-ordinated with her next utterance, the prolonging and clarifying move 'ah yeah?' as she looks to the teacher.



Plate G Iconic gesture – ‘extra’. In this frame the student partially out of shot is spreading her arms open at the same time as saying ‘extra’. Here the iconic gesture ‘extra’ is illustrated using the dimension of size indicated by the spreading of the arms.