Introduction

"If we are to develop and use language tests appropriately, for the purposes for which they are intended, we must base them on clear definitions of both the abilities we wish to measure and the means by which we observe and measure these abilities." (Bachman 1990: 81)

Bachman's statement, on the face of it seemingly self-evident, in fact touches upon a very complex issue as clear definitions of what constitutes language ability and how we are to measure it are not easy to arrive at.

In recent years, the influence of the communicative language teaching (CLT) movement, which made such a rapid and indelible mark on the fields of language teaching and learning, has begun to make itself felt in the field of language testing.

It is no longer assumed that the testing of language ability need not extend much beyond the assessment of grammatical (in the broadest sense) competence; any language test which claims to be communicative in purpose will also be aiming to observe and measure a whole range of inter-related language skills which, according to the particular model of communicative competence on which the test is based, will be defined and labelled in different ways, but generally include such notions as sociolinguistic, textual, pragmatic and strategic competencies.

Perhaps the area of testing in which the CLT movement has been most influential is that of the oral proficiency test. It would doubtless be difficult nowadays to find an oral proficiency test which does not lay claim to being essentially 'communicative', whether the test is set by a large, external examination board, or is a small-scale, internal test in any one of the countless educational, academic or commercial institutions around the world which need to assess oral proficiency in some way.

Much progress has been made in the development of testing methods and formats which allow samples of oral performance to be elicited in specific,

communicative contexts, bearing in mind the inherent constraints of a test situation.

The fear that the quest for internal validity (content and face validity) would lead to the inevitable sacrifice of reliability due to the necessarily subjective, impressionistic means of assessment that a communicatively oriented oral test would seem to impose is slowly being overcome with the development and widespread use of criterion-referenced rating scales which aim to enable assessors to match performance against corresponding band descriptors.

However, despite the increasing sophistication of both theoretical models of communicative competence and testing methods, it still seems to be the case that many oral tests are ***not*** based on clear descriptions of what they are supposed to be measuring.

This study originates from a concern that the definition of the concept of *grammatical* competence in the assessment of oral proficiency is often extremely vague, and moreover the relationship of grammatical competence with other communicative language skills is still uncertain.

A question of central importance here is whether grammatical competence can, or indeed should, be assessed as an independent component or as an integral part of language ability. Related to this of course are the questions of how grammatical competence is to be defined, to what extent it can be observed and measured, and indeed *how* it can be observed and measured.

The study reported here sets out to attempt to shed some light on these issues in general, but the particular focus is the effect of task type on the assessment of grammatical competence in oral proficiency tests.

Chapters 1 and 2 deal with some of the background issues involved in the study. Chapter 3 explains the purpose, design and procedure of the study. Chapter 4 presents the results, and finally Chapter 5 summarises the results, discusses the conclusions and implications drawn from the findings, and suggests some areas for further research.

1. Some issues involved in the assessment of oral proficiency

1.1 The nature of language proficiency

Until recently, the implicit conception of language proficiency in both language teaching and testing was that it entailed little more than a mastering of grammar, lexis, and, in the case of spoken language, phonology. Before the 1970's, the method of testing discrete linguistic points was widely accepted and implemented, particularly as the data yielded in this way was easily quantifiable and could be objectively scored, thus allowing a high rate of reliability (Weir 1990: 2).

However, in the 1970's there was an increasing awareness that proficiency in a language involved much more than a knowledge of linguistic forms and that any theory of language ability needed to take into account a much broader view of how the various elements in language combine in order to perform its basic function; that of communication.

This shift in emphasis, which was to become known as the Communicative Approach, originated from many sources, including the work in the 1960's on speech acts (Austin 1962, Searle 1969), the Hymesian model of communicative competence (Hymes 1972) and the early development of systemic functional linguistic theory (Halliday 1976, 1978), as well as the growing field of second language acquisition, particularly the work known as the 'morpheme order studies' carried out during the decade (Brown 1973, Dulay and Birt 1974) which led to a questioning of the traditional methods of teaching grammar.

By the end of the decade, discrete-point testing was coming under increasing attack as the call came for a more integrative method of testing language proficiency (Oller 1979, Morrow 1979) and theories of communicative competence, most notably that of Canale and Swain (1980) which included

sociolinguistic, strategic and, later, (Canale 1983) discourse competencies as well as grammatical competence, began to influence the field of applied linguistics. In short, the concept of language proficiency was greatly expanded, and much debate ensued as to the best methods for measuring it.

Harley *et al.* (1990) note two principal stages in the way that language proficiency has been viewed and investigated since this time. First came Oller's (1979) somewhat controversial work, based on factor analysis. He claimed that the results of his research supported the 'unitary competence hypothesis', according to which language proficiency consists of a single, measurable ability, and furthermore is largely indistinguishable from the general trait of intelligence (Oller 1981). The second stage involved the use of confirmatory factor analysis, "a much more sophisticated form of factor analysis" (Harley *et al.* 1990: 9), which drew on Canale and Swain's model of communicative competence (Bachman and Palmer 1982). The results of Bachman and Palmers' study suggested that more than one factor can be distinguished in language proficiency. They also indicated that the components of what they termed as *grammatical* competence (morphology and syntax) and *pragmatic* competence (lexis, cohesion, organisation) are closely related, whilst the components which make up sociolinguistic competence (sensitivity to register, naturalness, cultural references) are distinct.

More recently, the highly influential model of language competence developed by Bachman (1990) is divided into organisational competence (which includes grammatical and textual competence) and pragmatic comptence (which includes illocutionary and sociolinguistic competence).

The study designed and reported by Harley *et al.*(1990), which aimed to determine whether the traits of *grammatical, discourse* and *sociolinguistic* competence could be empirically distinguished, used confirmatory factor analysis to test hypotheses about the relationship between these traits, but in addition took into account the ways in which comparisons between native and non-native

speaker performances and the relationship between first and second language proficiency could also shed light on the conceptualisation of the nature of proficiency (*ibid*.:10). Their results suggest that

> "different aspects of proficiency seem to be differentially related to attributes of individuals. Cognitive variables appear to be more strongly related to discourse aspects of proficiency and to written aspects of proficiency than they are to oral grammatical skills" (*ibid*.:25).

This would seem to indicate, then, that the practice in the testing of oral proficiency of assessing grammatical skills as an independent component is likely to enhance the construct validity of testing methods based on theories of communicative competence, although Weir's comment (1990:13), that despite a general recognition that linguistic competence is an essential part of communicative competence, more empirical research is needed to establish the way in which they relate to each other, still appears to be the case.

1.2 Authenticity and the format of the oral proficiency test

Although it may be true that recent theories of communicative competence have not led to the abandoning of the assessment of grammatical ability in oral testing, it is certainly the case that grammar is now seen in a different light and there is a need to view it in relation to other skills, such as linguistic appropriacy, discourse and interaction skills, task management and negotiation of meaning. This has meant that much attention is now given to the *context* in which language is produced and the notion of *authenticity* in the testing situation. As Weir (1993:31) points out:

> "the more direct we can make a test and the more we can incorporate contextual features and interactional features of real-life activity into our tests, the more confidently we can extrapolate and make statements about what candidates should be able to do in that real-life context."

The first move in this direction came with the use of the oral interview format, whose claim to validity (in terms of content) was on the grounds that it reflected

conversation and therefore incorporated features of natural, spoken discourse, such as turn-taking, topic management, and the unpredictability of content.

When the British Council launched the English Language Testing Service (ELTS) in 1980, it was claimed that the interview format used, in combination with assessment scales based on a taxonomy of skills achieved a satisfactory compromise between "the constructive interplay with unpredictable stimuli" and "scientific measurement" (Carroll 1980: 53). This led the way for what Shohamy and Reves (1985) noted as a 'wave of enthusiasm' for the development and implementation of authentic language tests. However, they cautioned against an unquestioning belief that so-called authentic tests truly are authentic and further suggested that reliability can easily be overlooked by concentrating too much on content validity, a reservation echoed by Cameron (1987), for example, in his review of Carroll and Hall's popular handbook *Make Your Own Language Tests* (1985). Whilst admitting that "an authentic test can measure something adequately, provided that its criteria are relevant to a certain universe of language use" (Cameron 1987:150), he expresses his concern that two questions remain unanswered: how much consistency between raters is necessary to ensure adequate reliability, and, moreover, how much consistency can reasonably be expected (*ibid*.).

The debate over the interview format was taken up by van Lier (1989) who asked if the oral proficiency interview (OPI) really *is* an example of conversational language use (*ibid*.: 489). He concludes that the OPI is *not* an example of conversational language use, mainly because of the asymmetry of the exchanges; the interviewer, in conducting and controlling the interview is violating one of the basic principles of conversation, that there should be "potentially equal distribution of rights and duties of talk" (*ibid*.:495). This has also been shown empirically (Lazaraton 1992), and test developers have recently begun to show an interest in an alternative format - the paired or group format - which allows the assessor / interlocutor to take a more unobtrusive role in the proceedings.

This format is currently used, for example, in the Speaking part of the Certificate in Advanced English (CAE), set by the University of Cambridge Local Examinations Syndicate (UCLES) and launched in 1991, and as part of the Cambridge Assessment of Spoken English (CASE).

Many writers in the testing literature have outlined the potential advantages and disadvantages of this format (Carroll and Hall 1985: 52-53; Underhill 1987: 49-50; Heaton 1988: 102-103; Hughes 1989: 105; Weir 1990: 78-79). The advantages usually referred to highlight the *interactive, purposeful* and *contextualised* potential of the format, whilst the greatest disadvantage is deemed to be the effect of the variable of *domination* by one or more of the participants.

On the other hand, relatively little empirical research seems to have been carried out to test hypotheses about the benefits and drawbacks of this method compared to the OPI, although Folland and Robertson (1976) and Lombardo (1984) have described their experiences of group and paired formats respectively at university level and found the results to be successful in terms of face and content validity and, to a certain extent, reliability. Shohamy *et al*. (1986) provide optimistic results to support the inclusion of a 'group discussion' as part of a test battery for assessing oral proficiency, as does Fulcher (1996).

We might tentatively conclude, then, as there is as yet no empirical research to prove otherwise, the paired or small group method may be preferable to a one-to-one interview in terms of the testing of oral communicative abilities and skills, in that it is more likely to elicit natural features of spoken discourse.

As regards the notion of authenticity, Wood (1991:236) points out that "any language test is, by its very nature, inauthentic" if we are to compare it to non-institutionalised discourse. However, we can surely aim to achieve *authentic test language* by striving to validate our methods by as many means as possible.

1.3 Task types

If it is decided that a paired or group format is to be used for the assessment of oral proficiency, the question remains as to what sort of task, or tasks, will be selected.

Within the communicative paradigm, it is generally accepted that the choice of task should reflect real-world tasks as far as possible (Brumfit 1984, Johnson 1982, Littlewood 1992). For the testing of *oral* proficiency, it should further be ensured that the tasks reflect theories of spoken discourse (Weir 1990:12).

Much of this theory comes from the fields of discourse analysis, conversation analysis, pragmatics and other related fields, although there are also accounts of the nature of spoken discourse from the field of applied linguistics which have been influential in test development, notably that of Bygate (1987).

Nonetheless, as no task is ever likely to elicit *all* the characteristics of spoken discourse, an informed decision must be taken on which type of task(s) will best serve the testing purpose.

There are many ways of categorising task types, according to their nature (and sometimes level of difficulty), including transactional/interpersonal (Nunan 1991); static/dynamic/abstract (Brown *et al.* 1984); problem-solving (Littlewood 1991); information gap (with the added dimension of optional vs. required exchange of information) (Doughty and Pica 1986); convergent/divergent (Duff 1986).

What is of central importance, however, as regards task selection for use in oral assessment procedures, is the question of what the effect of task type might be on the performance elicited. Courtney (1996:321) puts it this way:

> "If we can categorise oral communication tasks in terms of type, how does this help us with the question of justifying their use? We can select them on the basis of their input configurations, but can we expect to be able to select them for their expected performance as well?"

The research that has been done on this matter to date has been mostly concerned with the effects of task type on language *learning* (see, for example, Brown and Yule 1983; Long and Porter 1985; Doughty and Pica 1986; Duff 1986; Courtney 1996). Although test developers can draw a good deal of useful information from this work, it would appear that further empirical research is needed which is situated more directly in a testing situation, taking into account considerations of test methods and constraints, washback effects, and, of course, theoretical constructs concerning the measurement of language proficiency.

2. Assessing grammatical competence in oral proficiency tests.

To return to the issues raised in section 1.1, regarding the conceptualisation of language proficiency and its measurement; if we are to accept that grammatical competence can and should be assessed as an independent component of oral proficiency, two questions we then need to ask are:

- What does grammatical competence involve?
- How can we assess it?

2.1 Components of grammatical competence

Within Canale and Swain's framework of communicative competence, grammatical competence "will be understood to include knowledge of lexical items and rules of morphology, syntax, sentence-grammar semantics, and phonology." (Canale and Swain 1980: 29). Bachman's (1990) framework of communicative language ability differs somewhat in its overall categorisation of competencies, but nevertheless his model of grammatical competence is similar: knowledge of vocabulary, morphology, syntax and phonology/graphology (*ibid*.:87), as is the model used operationally in the development of CASE: syntax, morphology, vocabulary, pronunciation (Milanovic *et al*. 1996:16).

The decision as to whether or not to separate the various components varies. Some examination boards opt to assess structural, lexical and phonological accuracy together, (eg. IELTS), others put vocabulary and structure together, with pronunciation separately (eg. CAE), others still consider all three separately (eg. CASE).

2.2 Assessing grammatical competence

In order to discuss how grammatical competence might be assessed, it is useful to refer directly to some of the most widely known and used tests of oral proficiency to see how grammatical competence is treated. It should be noted that as the study which will be presented as part of this dissertation is primarily concerned with the *structural* features of grammar, rather than the lexis, from here on the term *grammar* will be used in the restrictive sense of structure (syntax and morphology), whilst acknowledging that lexis and indeed phonology are, strictly speaking, parts of grammatical competence.

In recent years, there has been a tendency to use scales, or bands, to describe levels of performance (Alderson 1991:71) following the pioneering example of the Foreign Service Institute (FSI) of the USA. The FSI scales comprise both holistic and analytical scales, the latter being six-point scales used to rate accent, grammar, vocabulary, fluency and comprehension. The band descriptors for grammar (reproduced in Hughes 1989:111) are based on error frequency and the subsequent effect on communication. Thus band 2 describes "constant" errors "frequently preventing communication"; band 4 indicates that errors are merely "occasional" but do not cause misunderstanding, whilst band 6 candidates must produce "no more than 2 errors during the interview".

One of the problems with basing an assessment of grammatical competence solely on accuracy, or the lack of it, is that this may encourage candidates to limit themselves to 'simple' syntactic structures which they feel safe with in order to control the amount of errors made. In other words, this system, by allowing only for the penalising of errors, can have a very restrictive effect on the language output of a student who may be concentrating on producing 'accurate' language rather than communicating in a more natural fashion.

Most band descriptors these days allow for a balance between both penalising errors and rewarding performance, and indeed often emphasise the positive skills

of the candidate rather than the negative side of the performance, particularly at the higher levels. The ESU Framework's Yardstick for Speaking (Carroll and West 1989:28-29) describes the oral proficiency of a level 7 student as follows:

> "Handles a wide range of speech operations with good confidence and competence. Message is clearly conveyed and with interest. Presentation and interaction relevant and appropriate to listener's knowledge of topic and language. Spoken text is clearly organised with suitable sequencing and cohesion. Occasionally lacks fluency and flexibility, with some lapses of appropriacy and linguistic uncertainty. Uses coping strategies effectively. Uses a wide language repertoire with occasional lapses of accuracy. Speech features influenced by L1 but these in no way affect communication."

The IELTS band descriptor for level 7 has similarly 'positive' criteria, again with little mention of errors:

> "Communicates effectively on a wide range of general, academic, vocational or leisure topics. Errors in vocabulary and structure may occur without inhibiting communication. Communicates readily and fairly precisely using complex sentence forms and a wide range of modifiers, connectives and cohesive features. Displays some flexibility in the use of speculative, argumentative, descriptive and narrative language."

In holistic scales such as the above, grammatical competence is viewed as being bound up with other components of proficiency such as effective communication and flexibility of style and genre and so does not stand out as an independent construct, although grammar is explicitly referred to in all but the top and bottom levels.

Other examination boards prefer to separate out their parameters across the various levels in a more analytical form. The Speaking part of the CAE (UCLES) examination, for example, stipulates criteria for fluency, accuracy, pronunciation, task management and interactive communication. A mark of 5/6 for accuracy is given to a performance which meets the following criteria[1]:

[1]The assessment criteria for the CAE (UCLES) Speaking test are reproduced in Hashemi (1994: 21)

> "Evidence of a good range of structures and vocabulary. Errors few in number and minor in gravity. These errors do not impede communication"

The bands for the UCLES/RSA Certificate in Communicative Skills in English: Oral Interaction (reproduced in Weir 1990:177-178) specify what the candidate must be able to do to achieve a pass at a given level in terms of five components. Grammar is dealt with in two components; accuracy and range. At level 3, "grammatical / lexical accuracy is high though occasional errors which do not impede communication are acceptable", whilst "a wide range of language must be available to the candidate".

It would appear, then, that there is a general trend towards both (i) recognising grammar as a distinct component which, although it can be assessed independently to a certain extent must also be viewed as being interwoven with other components of communicative competence such as appropriacy, flexibility, communicative effectiveness, discourse competence and so on, and (ii) trying to assert a balance between grammatical accuracy (in terms of error gravity and frequency) and range (in terms of the variety and complexity of structures).

However, the question of how to assess grammatical accuracy and range is still largely unanswered. How exactly do we discriminate between the different band levels? A closer look at the ESU Framework's yardstick for Speaking reveals that whilst a level 8 student has a "good" language repertoire with "few slips of the tongue", a level 7 student uses a "wide language repertoire" with "occasional lapses of accuracy" and at level 6, the student has a "fair" language repertoire and "noticeable" lapses in accuracy. This is indeed confusing as the final decision must therefore rest on an interpretation of the differences between *good/wide/fair* repertoires and *few/ occasional/noticeable* lapses. In fact most band descriptors are found to be similarly vague, with the CAE assessment of grammatical range, for instance, resting on the distinctions between *wide*, *good*, *restricted* and *limited*.

Alderson (1991:82), describing the problems encountered by the ELTS Revision Project in arriving at descriptors for the revised IELTS test, admits that for grammatical accuracy, the difference in levels came down to a choice in quantifiers, but concludes that

> "Ultimately, since we are not creating an equal interval scale, what will matter is whether assessors can use the scales and agree on their understanding of the descriptions that define the levels."

This is obviously an important point; it should of course be assumed that training sessions will be given to raters who are to use a particular set of descriptors, and it can reasonably be expected that a fair degree of agreement can be sought and found on what constitutes a 'serious' error or how frequent errors must be in order to be considered 'frequent'.

Nonetheless, the problem of grammatical range may not be so easy to resolve, for even if we can agree on what might constitute a 'complex' structure as opposed to a 'simple' structure and even if we can demonstrate that student X has produced a 'wider range' of structures than student Y, two problems still remain unsolved.

Firstly, can we be sure that the task set to assess performance will actually elicit a wide and / or complex range of structures? Will *all* communicative tasks be likely to elicit a similar range?

Secondly, assuming that student X does in fact produce the range and complexity of structures to meet the criteria set at a specific level, was it actually *necessary* to produce these structures? To put it in another way, student Y might perform the same task in a perfectly natural and appropriate way, with a great deal of communicative effectiveness and interaction, delivering competent messages with approximately the same meaning as student X and managing to complete the task successfully, but *without* reaching the same level of linguistic complexity or producing the same variety of structures as student X. To hypothesise even further, what if student Y was not a student of English at all,

but a native speaker? This is not unfeasible, if we are discussing higher levels of proficiency, and, as these levels are often assessed with native speaker-like performances in mind (Pollitt and Murray 1996), it would not seem fair to expect second language students to perform at a 'higher' level than native speakers themselves.

Fulcher (1987) compares the transcript of an informal conversation amongst four highly educated native speakers with the (former) ELTS band descriptors for repetition, hesitation, stumbling, propositional development and grammatical accuracy and concludes that the participants "would not meet the appropriate criteria embedded in the ELTS assessment scales if they are measured on the recorded material sampled from real life" (Fulcher 1987:290). He suggests that the scales, along with "the majority of communicative tests in current use" (*ibid*.) were basically derived from Munby's (1978) taxonomy of skills and are therefore a reflection of what communicative theorists *believe* to happen in communicative situations rather than on empirical evidence; in other words, perceived content validity is being sought at the expense of construct validity. He underlines the need for data-based criteria, a proposal which he himself later takes up in a study aimed at constructing rating scales for fluency and accuracy (Fulcher 1993). Interestingly though, his database consists of actual *student* speech, and his approach to developing a rating scale for accuracy is based on identifying and categorising student errors, therefore taking neither native speaker performance nor the crediting of complexity and range of structures into consideration.

Chambers and Richards, in a series of studies, looked at the assessment of oral proficiency in French at GCSE level. By carrying out a detailed analysis of recorded GCSE candidates' examination interviews and comparing these with teachers' marks (Richards and Chambers 1992), they found that there was an 'alarmingly low level of agreement' between teachers' marks and their own measures of structural complexity. In a further study (Chambers and Richards 1993), they interviewed teachers (who also act as GCSE assessors) to ascertain

their judgements of what actually constitutes linguistic complexity and found their opinions to be extremely subjective, and, furthermore:

> "A study of our transcripts of GCSE candidates reveals that hardly any instances of the syntactic structures suggested by the teachers can be found in the conversations that they had been asked to mark" (*ibid*:26)

This then led the authors to make a comparison of native and non-native speaker performance on the same task (the 'free conversation', which is basically a one-to-one interview) on the basis that "a mismatch between examiners' expectations of candidates and what native speakers would achieve is ... a threat to validity" (Chambers and Richards 1995:6). Basing their analysis of linguistic complexity on examination board specifications and teachers' opinions of what they would give credit for, they found that in some areas the results suggest that "traditional priorities need to be examined" (*ibid*.:8) and conclude that:

> "native performances on the various assessment tasks used in examinations have much to offer. This is not to say that the levels of achievement for the foreign-language learners should be modelled entirely on native performances. Rather the latter offer a yardstick against which the validity of tasks and the features which discriminate between levels of foreign-language proficiency can be cross-checked" (*ibid*.:9)

Chambers and Richards' study is interesting, not least because of their method of comparing native speaker to non-native speaker performance *on the same tasks.* This would seem to be appropriate than simply measuring non-native speech against informal native speaker conversation, or indeed measuring informal native speaker discourse against descriptors which have been developed to assess oral proficiency in a testing situation (cf. Fulcher 1987). However much a task set reflects a real world situation, the crux of the matter is that it is situated in an artificial context and the resulting discourse is always produced within the test constraints, such as unfamiliarity of surroundings or persons present, time

pressure, perhaps the presence of a tape recorder or video camera; and most of all the inevitable sense that one's performance is going to be judged and evaluated in some way, which could moreover have implications for one's future. Thus it is unlikely that any performance will truly match the sort of discourse produced in an informal, non-institutional setting.

3. The purpose of the study and methodology

3.1 The purpose of the study

Before stating the purpose of the study reported in this dissertation, the main points made in the previous sections will be summarised:

1. There is a case for assessing grammatical competence as an independent component of oral proficiency, although grammar should be viewed within a broader framework of communicative competence;

2. The use of a 'paired' format in oral proficiency testing is likely to enhance the content validity of the test;

3. Different task types may be selected for the use in paired formats, and empirical research is needed to determine to what extent task type will affect linguistic output;

4. The assessment of grammatical competence is often based not only on accuracy (frequency / gravity of errors) but also on the range and complexity of structures used. However, the extent of the range and complexity of structures may well be affected by the task variable;

5. At higher levels of oral proficiency, native speaker (NS) performances on test tasks could be used as a yardstick against which to measure non-native speaker (NNS) performance, and may be able to inform us in the development of criteria for the assessment of structural range and complexity.

With these points in mind, the main purpose of this study is to investigate the variable of task type on linguistic output in a test situation, with particular reference to grammatical structures. To this effect, samples of NS and NNS

performances on two different task types were analysed and compared, as described in the following section.

3.2 Methodology

3.2.1 Task design

Two tasks were prepared for use with paired students. Both tasks are essentially communicative in nature in that they are interactive, contextualised and reflect real-world communicative acts. The tasks were designed so as to ensure that the students would be able to perform without the participation of an interlocutor, once the instructions had been given, thus minimising the variables caused by interlocutor input.

Task 1

For this task, the two students each have a task sheet which is headed with written instructions explaining that they are friends who work together and have decided to rent accommodation together. Student A's task sheet contains two advertisements for accommodation, whilst Student B's task sheet has two *different* advertisements for accommodation. The instructions explain that they have each visited the accommodation shown on their own task sheet, and they must work together to discuss the four possibilities and try to reach an agreement on which property/properties would be the most suitable and which would not be suitable. (See appendix I - Task 1).

Task 2

Both students have the *same* task sheet, which consists of the title of the topic for discussion ('Homelessness') and two quotations which present differing view points on the subject. The students are instructed that the quotations may be of help to them in their discussion but they do not necessarily have to refer to them. (See appendix II - Task 2).

The tasks differ in various ways. Task 1 is an *information gap* activity, with a *required* exchange of information (Doughty and Pica 1986), but simply exchanging information is not enough - the participants are also required to

give/seek opinions, reach an agreement and make decisions. Therefore the task has a definite goal and can be described as *convergent* (Duff 1986) but with many possible outcomes.

Task 2 is basically a *discussion* activity, which has no particular goal and no *required* exchange of information, although obviously if information is not exchanged (along with opinions) the resulting dialogue will be very poor indeed.

Task 1 also requires the students to enact a situation, whereas in Task 2 they are not required to project themselves into an imaginary situation, but just 'be themselves'. In this way, the task differs from Duff's (1986) *divergent* debate task, where participants are assigned a viewpoint on a topic which they have to defend with as many arguments as possible. This option was not selected as imposing an opinion on someone is somewhat inauthentic and could bias the performance of a participant whose own views do not correspond with those assigned to him/her.

The level of difficulty of both tasks is assumed to be fairly high and correspond roughly to the type of tasks which might be set in phase C of the CAE (UCLES) examination.

An **Instruction sheet** was also prepared which states exactly how the interlocutor should proceed with the test. (see appendix III - Interlocutor's instructions).

The two tasks were then used to elicit performances from both native speakers of English and non-native speakers of English.

3.2.2 Participants

The **native speaker** participants were selected from volunteers from a class of PGCE (Education and Training) students from the University of Wales, College of Newport. The volunteers were asked to present themselves in pairs, rather than

being assigned partners at random. This was intended to ensure that the participants felt relaxed with each other and therefore the interaction produced was assumed to be as natural as possible under the conditions imposed by the 'test' situation. Three pairs were selected for participation; one female-female pair, one male-male pair and one female-male pair, thus eliminating any gender bias.

The **non-native speakers** were selected from volunteers in a class of first year students at the Universidade do Algarve, Portugal. These students are studying for a first degree in Tourism, for which foreign languages, notably English, play an important part. The students selected all had eleventh or twelfth grade English from Portuguese Secondary school and one year of English as a Foreign Language at University. They were assumed to correspond roughly to level 6/7 of the English Speaking Union's Nine Level Scale.

The students had not been 'prepared' for the tasks given, i.e. they had not been given any particular teaching input before participating in the study.

Again, the students were asked to self-select their own pairings. It was hoped that the gender distribution would be the same as the native speaker participants, but unfortunately the recording of the only male-male pair was inaudible in parts, so a 'reserve' pair was used, thus making the distribution of gender as follows: one pair female-female, two pairs female-male.

3.2.3 Data collection

The groups of students were told that they were taking part in a research project for an MA dissertation in the area of oral proficiency testing, but they were not given any further details before undertaking the tasks - therefore they did not realise that their performances were going to be observed and studied with *grammatical* competence in mind.

They were then briefed on the procedure: they would be given task sheets for two tasks and allowed a short silent preparation time, after which they would be recorded performing the tasks.

The preparation time allowed for the **non-native speakers** was ten minutes (approximately five minutes per task). This time was spent in a classroom under the supervision of their English class teacher. They were not allowed to discuss the tasks with each other, but were permitted to check their understanding of the vocabulary used in the task sheets with their teacher if necessary.

After ten minutes, they were shown to another classroom where they were seated facing each other on one side of a desk. The interlocutor, who was previously unknown to them, was seated on the other side of the desk, and a tape recorder was placed on the desk, thus simulating a typical test situation.

Having spent a minute or so putting the students at ease, the interlocutor gave the task instructions and checked that the students understood what was required of them. The order of the tasks was alternated from pair to pair, to prevent bias from the effects of warming up and / or fatigue.

The students were given a maximum of four minutes to complete each task.

The same procedure was used for the **native speakers**, except that it was felt to be unnecessary to give ten minutes' preparation time. Therefore the participants were merely given the time necessary to absorb the information given on the task sheets before beginning the task; this was generally about two minutes per task.

It was hoped that by following this procedure, test conditions would be replicated sufficiently to produce a similar type of interaction to that produced in a real test situation on the part of the non-native speakers, and that the native speakers would also produce language that was 'authentic' under the circumstances, thus creating a yardstick of native speaker performance against which to compare non-native speaker performance which had been produced in the same conditions.

The tapes were subsequently transcribed, using transcription conventions adapted from Eggins and Slade (1997). As extracts from the transcripts are used illustratively throughout, the transcription key is provided (Appendix IV).

The transcriptions are hereafter referred to by the following tags:

NS P1 T1 = Native Speakers Pair 1 Task 1

NS P2 T1 = Native Speakers Pair 2 Task 1

NS P3 T2 = Native Speakers Pair 3 Task 2

etc..

NNS P1 T1 = Non-native Speakers Pair 1 Task 1

etc..

The speakers are identified by an initial as follows:

NS P1	F = female speaker
	J = female speaker
NS P2	A = female speaker
	P = male speaker
NS P3	S = male speaker
	T = male speaker
NNS	M = female speaker
P1	R = male speaker
NNS	N = male speaker
P2	P = female speaker
NNS	S = female speaker
P3	T = female speaker

3.2.4 Linguistic analysis

In order to investigate whether or not each task is likely to elicit different grammatical features, a quantitative analysis was made to determine the frequencies of certain grammatical structures in the discourse produced by the participants in the study in each task. However, it is not particularly helpful to study grammatical forms without considering the meaning that they are

expressing. Therefore a qualitative analysis was also made to attempt to relate grammatical forms to their functions. This analysis is largely based on the theory of systemic functional grammar developed by Halliday, which seeks to demonstrate how choices made as to the meanings we want to express are related to grammatical choices.

The question as to which grammatical structures to choose proved rather difficult to resolve. One possible source would be test specifications, but in fact most examination boards who have tests designed to assess higher levels of oral proficiency do not specify particular linguistic structures (Alderson et al. 1995:30). Another source could be the grammatical syllabuses of materials designed to teach for specific examination purposes, but these syllabuses generally cover a very wide grammatical range and no indication is made of priority areas. The question was partly resolved by asking ten experienced teachers of English (both native and non-native speakers) to state which grammatical features they would look for in order to give credit for grammatical competence in an oral test of proficiency at an upper-intermediate / advanced level. The eventual selection of linguistic features (outlined below) was helped by the results of this small survey.

Mood

An important feature of interaction is the selection of mood, which basically determines the role of the speaker as giver or seeker of information. At the clause level, the selection of mood determines the "presence and configuration of certain 'negotiable' elements of the clause structure" (Eggins and Slade 1997:74). The mastering of both the configuration and the presence/absence of these elements is essential for any competent performance in spoken English. The 'structure' of mood types is often one of the criteria for assessment of accuracy in spoken English, particularly in the case of question-forming which can be

problematic for many students of English, due to the Subject-Verb 'inversion' and the use of auxiliaries. Errors in interrogative structures may well be penalised; conversely, proficient use of question forms is considered by many as a feature of linguistic competence worth rewarding.

Therefore it is useful to determine the mood patterns which might be expected from each task, with particular reference to the potential for eliciting questions, along with another feature of the mood system also mentioned by some of the teachers asked: the 'tag' question.

Verbal groups

Verbs, (which we shall refer to as verbal groups (VGs) due to the fact that many 'verbs' have more than one element), are central to any message that we may wish to communicate, as they allow us to express both events in the external world around us and in the internal world of our thoughts, feelings and perceptions (Downing and Locke 1992:110). It is of little wonder, therefore, that so much attention is given to them by both learners and teachers of a language, and indeed by assessors of linguistic competence.

Verbal groups can be either finite or non-finite, or a combination of both (verbal group complex). Finite VGs carry within their structure the temporal location of the event being expressed in relation to the present time of the speech situation, whereas non-finite VGs do not. In other words, finite VGs are selected for tense, and mastering the structures of the various tenses and their appropriate uses can be problematic for learners of English, as indeed for any language. Besides this, finite VGs can express modality. Quirk *et al.* (1985:20) note that "the use of modal verbs is one of the most problematic areas of English grammar", not least because of the way that each of the central modal auxiliaries can express different meanings. Finite VGs also show polarity - whether a proposition is positive or negative - and again, the use of finite operators (auxiliaries) to

indicate negation is generally considered to be problematic for students of English.

As non-finite clauses are necessarily dependent on finite clauses, their production in speech involves more planning and processing than simply stringing together finite clauses, and so are often considered to be a measure of careful, linguistically complex speech amongst both native and non-native speakers. For students of English, the correct use of non-finite verb forms in complex VGs is likewise considered to be a sign of linguistic proficiency.

It was therefore decided to look at the use of finite VGs in terms of tense, (mentioned by all the teachers consulted), polarity and modality, (mentioned by some) and also the various functions of non-finite VGs

Subordination (Hypotaxis)

Messages can be extended and elaborated by means of paratactic (coordinating) and hypotactic (subordinating) relationships between clauses. It is often argued that subordination does not occur as frequently in speech as in writing, as spoken discourse is more likely to be linked by paratactically linked clauses (eg. Chafe 1982; Brown and Yule 1983: 4; Bygate 1987:15). The reason usually given is that the processing constraints involved in producing spontaneous (as opposed to planned) speech means that ideas cannot be as economically ordered as in writing. It may well be that the nature and context of the speech needs to be taken into account. Akinnaso (1982: 109) argues that a study such as Poole and Field's (1976), which found that more subordinate clauses were used in speech than in writing, may have been due to the nature of the oral data, as individual speech samples were obtained from freshman students "who probably took the experiment very seriously", the resultant data being "highly structured and, in certain areas, more complex than the written data".

It was predicted, then, that features of subordination would be likely to occur in the data used in this study as the participants are under the artificial constraints of

a testing situation and are likely to take more 'care' about the language they produce.

Although only one of the teachers consulted referred directly to features of subordination (relative clauses), another mentioned 'linking ideas', which is assumed to mean in a hypotactic sense as credit is unlikely to be given for the use of paratactic coordinators such as *and, but, or*. Yet another referred to 'the correct formation of complex structures', which (possibly) can be construed as including subordination. Although, as Biber (1988: 229) notes, subordination can *generally* be taken as an index of structural complexity, this is open to debate, as we shall see later in the discussion of the results of the analysis.

3.2.5 Clause division

For the purposes of this study, the basic unit of analysis will be the clause, which, according to the principles of functional grammar, is the main resource through which meanings are expressed (Halliday 1994:19, Thompson 1996: 35). As a structural unit, the clause is relatively easily identifiable because it has its own internal structure, unlike the 'sentence' which is unsuitable as a unit of analysis for spoken data as it is often difficult to determine where a sentence begins or ends (Downing and Locke 1992: 277). By taking the clause as the basic unit, we can also look at what is happening *within* the clause at the rank of group, and *beyond* the clause to consider the logical function of grammar which accounts for the way in which clauses can be combined.

Having transcribed the data, then, each turn was divided into clauses, according to the following principles[2]:

- a clause may be *full*, *elliptical*, *abandoned* or *minor*
- a **full** clause generally consists of a Subject, a Finite element and a Predicator, possibly with one or more Complement(s) and Adjunct(s).

  Repetitions and **false starts** are considered as part of a full clause as long as the elements remain essentially the same, thus the following examples are counted as full clauses:

  (1) `I met a- I met a homeless guy the other day (NSP3T2)`

  (2) `I've seen- watched programmes on the TV (NSP1T2)`

[2]The principles for clause division and the grammatical labels used throughout this dissertation are consistent with those generally associated with systemic functional grammar (Eg. Downing and Locke 1992, Eggins 1994, Halliday 1994, Thompson 1996)

- an **elliptical** clause is a potential full clause with one or more elements which are left out, but can be recovered either from the text or the context, as in example (3) which has an ellipted Subject:

(3) `didn't like that (NSP2T1)`

However, ellipsis may occur *within* an element of the clause; in this case the clause is considered to be full. In example (4), the ellipsis of the non-finite verb (?*being* or ?*living*) within the Complement element of the clause does not affect the status of the clause as **full**:

(4) `Yeah but I don't like the idea of seven floors up (NSP2T1)`

- an **abandoned** clause is a clause which has not been completed by the speaker, perhaps because of another speaker's interruption, as happens to speaker N in the following example:

(5)
```
N: The price is=
P:  =is the SAME, as mine  (NNSP2)
```

or because of the speaker changing her mind about the content or organisation of the message, (example 6):

(6)
```
(i)they are only- (ii)they only have the
streets                (NNSP1T2)
```

- a **minor** clause does not have a mood structure, that is, it does not have a potential Subject and Finite combination. Eggins and Slade (1997:94-5) identify three types of minor clause:

Lexicalised minor clauses: words such as *Okay, Right, Fine, Exactly*, when used in non-elliptical contexts. They are typically used as responding contributions. This category also includes *yes (yeah, yep etc.*) and

no (nope etc.) when functioning as a response to a previous declarative clause. However, it should be noted that yes / no and their derivatives also frequently have a role in conversation as continuity markers and in this case are classified as textual Adjuncts. The difference can generally be distinguished by the intonation used; as a continuity marker yes / no words will be unstressed and do not occur in their own tone group, but are immediately linked to further clause elements. Example (7) shows the use of *yeah* as a minor clause, and (8) shows how *yeah* can be a textual Adjunct:

(7) T: ...we never go quite as far as they go
S: yeah (NSP3T2)

(8) T: Yeah I'd definitely say so (NSP3T2)

Formulaic expressions: Typically of greeting and thanks, such as *Hi, Thanks, Bye*

Non-lexical items: Typical feedback and backchannel indicators such as *Mm, Mm hmm, Uh huh.*

Other clause types: non finite, projected, embedded, dependent

Non finite clauses are recognised in systemic functional grammar, but as Thompson (1996:191) points out, the boundaries between complex verbal groups (finite verb + non finite verb) and clause complexes (finite clause + non-finite clause) are "inherently blurred". Halliday (1994: 290) suggests that two clauses should be recognised when the second verb involves a change of Subject, so he prefers *Mary wanted to go* as one clause with a complex verbal group, and *Mary wanted John to go* as a clause complex.

For the sake of convenience, we shall count both types as one clause, considering the second part to be a Complement filled by a clause. Therefore non-finite clauses will only be considered separately if (a) they are obviously not filling an element within another clause, eg:

(9) (i) And sitting down with him, (ii)I was amazed at his situation (NSP2T2)

or (b) in the case of a projected clause (see below, example 11).

Projected clauses

Halliday (1994:250) sums up the notion of a projected clause as being "the logical-semantic relationship whereby a clause comes to function not as a direct relationship of (non-linguistic) experience but as a projection of a (linguistic) representation". In the simplest terms, a projected clause corresponds to reported speech or reported mental ideas. For example:

(10) (i) and I said (ii) oh can I help you? (NSP2T2)

(11) (i)Well...they're saying (ii) having a home is a basic human right (NSP1T2)

(12) (i) I thought (ii) I'd just take some time to speak to him (NSP2T2)

However, it should also be noted that in conversation expressions such as *I think, I reckon, I guess* are frequently used as a modal Adjunct expressing epistemic possibility, and therefore do not introduce a projected clause. For example:

(13) (i)I think you've gotta take each case as individual (NSP3T2)

Embedded clauses

A clause can function as an element or part of an element within another clause. In this case it is considered to be embedded and therefore is not counted as a clause within its own right. Example (14) shows an embedded clause at Subject and (15) shows an embedded clause at Complement:

(14) The first one I saw was quite nice (NSP3T1)

(15) It's a little bit more than we budgeted for (NSP2T1)

Embedded clauses identified in the data are mostly either a) defining relative clauses (with or without the relative pronoun- see example 14) or b) wh-clauses (sometimes with wh-word omitted as in example 15).

Dependent Clauses

A clause is considered to be dependent when it is hypotactically related in some way to the main clause:

(16) (i)Unless he's got a permanent address, (ii)he can't claim benefit (NSP2T2)

(17) ... (i)okay one was uhm a COTTAGE (ii)and it's got a large garden and orchard (iii)which I'm not too happy about (iv) uhm even though the property was very nice (NSP1T1)

3.2.6 The consideration of errors

Although errors are not the focus of this study, it is useful to note the frequency of errors in the NNS discourse where appropriate, so as to give a more complete picture. The consideration of errors will not extend much beyond this, and no attempt will be made to analyse error gravity. It should be noted that 'errors' are only counted as such when there is a clear case for arguing that a grammatical

form is 'incorrect' (according, at least, to the rules of Standard English) or 'inappropriate' in the sense that it can impede communication.

4. Results

4.1 General profile of the data

Length of the samples

The recorded performance of each pair on each task was timed and it was noted that the average length of the samples for Task 1 was around 4 minutes, whilst the Task 2 samples were rather shorter, averaging around 3 minutes. Although it would be possible to determine a 'cut-off point' for use in the analyses, whereby all the samples are made to be exactly the same length, this is not what would happen in a real test situation. Therefore frequencies have been determined by calculating percentages of occurrences of a particular feature in individual samples.

Presentation of results

The results for the **native** speakers are mostly presented as frequencies *per pair*, as we are interested in the *potential* frequency of a particular feature rather than individual performance and this format has the advantage of making the results more accessible in terms of economical presentation of figures. The **non-native** speaker results however are generally presented as frequencies *per speaker*, as individual performances are sometimes referred to. In all cases, the **mean frequency** is given, and occasionally the mean frequency is the only figure presented.

Basic clause analysis

Tables 4.1.1 - 4.1.4 on the following pages show the relative mean frequencies for **completed**, **abandoned** and **minor** clauses (see section 3.2.5) for NS and NNS for each task.

It can be seen that there is little difference between the NS figures for both tasks. The NNS figures reveal that there was a higher frequency of abandoned clauses in Task 2, reflecting the difficulties experienced in expressing the more complex ideas required by this task. In Task 1, the frequency of abandoned clauses is similar to that of the native speakers.

The NS also showed a greater tendency to use minor clauses than the NNS.

Clause type	**%**
completed	74.5
abandoned	8
minor	17.5
(Total clauses)	(100% = 459 clauses)

Table 4.1.1 ***Relative frequency of clause types (NS - Task 1)***

Clause type	**%**
completed	73
abandoned	10
minor	17
(Total clauses)	(100% = 292 clauses)

Table 4.1.2 ***Relative frequency of clause types (NS - Task 2)***

Clause type	%
completed	79
abandoned	10
minor	11
(Total clauses)	(100% = 373 clauses)

Table 4.1.3 ***Relative frequency of clause types (NNS - Task 1)***

Clause type	%
completed	71.5
abandoned	17.5
minor	11
(Total clauses)	(100% = 260 clauses)

Table 4.1.4 ***Relative frequency of clause types (NNS - Task 2)***

4.2 Analysis of mood

For this analysis, only clauses which select independently for mood were counted. This means that **dependent** clauses, which do not select independently for mood, were not included, nor were **minor** or **abandoned** clauses as minor clauses do not show mood in their structure, and abandoned clauses generally do not show a complete structure.

The basic mood types which can be identified at the clause level are **declarative**, **interrogative**, **imperative** and **exclamative**. In order to discover if there appears to be a general difference in mood selection between the two tasks, frequencies for each mood type were calculated by determining percentages using the total number of clauses which carry mood selection produced by each pair of native speakers for each task (tables 4.2.1, 4.2.2) and by each of the non-native speakers (tables 4.2.3, 4.2.4).

4.2.1 Mood selection by the native speakers

Mood	**NS P1**	**NS P2**	**NS P3**	**Mean**
Declarative	88.2	74.3	77.1	**79.9**
Interrogative	10.7	20.9	21.2	**17.6**
Imperative	1.1	4.8	1.7	**2.5**
Exclamative	0	0	0	**0**
(Total clauses)	(93)	(105)	(118)	**(316)**

Table 4.2.1 ***Mood selection (%) by NS pairs - Task 1***

Mood	**NS P1**	**NS P2**	**NS P3**	**Mean**
Declarative	96.4	95.8	94.8	**95.7**
Interrogative	3.6	2.8	5.2	**3.8**
Imperative	0	1.4	0	**0.5**
Exclamative	0	0	0	**0**
(Total clauses)	(56)	(72)	(58)	**(186)**

Table 4.2.2 ***Mood selection (%) by NS pairs - Task 2***

As is to be expected, the declarative mood is dominant in both tasks. The exclamative does not appear at all and the imperative features very little, although it is used by all three pairs in Task 1 (a total of 8 occurrences). Its functions include suggesting (e.g. `Let's try it (NSP2T1)`), persuading (eg. `imagine being part of a small village community! (NSP3T1)`), and requesting further information (e.g. `Now tell me again about the flat (NSP1T1)`).

The most noticeable difference between the tasks is that a substantially higher proportion of interrogatives was produced in Task 1. The questions asked in this task (a total of 55), function to ask for both information (58%) and (31%). The remainder can be classified as 'confirmation-seeking', a natural feature of spoken interaction which is often realised by an elliptical interrogative form, for example:

```
J: Which would you rather live in, a flat or a
house?
F: Uhm I think I'd rather have a house?=
J: =Would you?=
F: =mm
J: Because the other property I saw is in fact
a- a modern two bedroom FLAT
     (NSP1 T1)
```

However, 68% of the questions in Task 1 are realised by full clauses, thus displaying their full structures.

Both polar interrogatives and wh-interrogatives are used, and, although it is interesting that the wh-words used are limited to *what..?*, *how...?*, and *which...?* (and a single instance of *who...?*), a wealth of questions covering a wide range of structures are elicited from the native speaker participants in this task.

In Task 2, on the other hand, the questions asked (a total of 8), which function largely to ask for opinions, are not a prominent feature of the discourse. In fact three of them occur together in this opening exchange in which the first speaker

is initiating the discussion and the second speaker, by repeating the question, is merely hedging for time to arrange his thoughts:

```
S:  Why  do  you  think  there's  this  homeless
problem then? What d'you think causes it?
T:  What causes it, well... it's er...  (NSP3T2)
```

The difference between the two tasks as regards the frequency of interrogatives is doubtless due to the nature of the tasks. In order to successfully complete Task 1, a great deal of specific information must be sought and given, and opinions must be ascertained, whereas Task 2 does not require a solution to a problem; opinions, arguments and experiences are readily offered and the speakers react to each others' contributions with further contributions of their own, rarely challenging or probing for more information.

4.2.2 Mood selection by non-native speakers

The mood analysis for the non-native speakers, on the other hand, seems to contradict these findings. The results (see tables 4.2.3, 4.2.4 below) show that there is virtually no difference in the mean frequency of interrogatives between the tasks. Only two speakers in Task 1 are close to NS mean, whilst in Task 2 five of the speakers produce a *higher* frequency of questions than the NS mean.

Mood	**NNS P1**		**NNS P2**		**NNS P3**		**Mean**
	M	R	N	P	S	T	
Declarative	92.1	98	82.5	98.5	82.9	96	**91.7**
Interrogative	7.9	2	17.5	1.5	17.1	4	**8.3**
Imperative	0	0	0	0	0	0	**0**
Exclamative	0	0	0	0	0	0	**0**
(Total clauses)	(38)	(50)	(63)	(65)	(41)	(25)	**(282)**

Table 4.2.3 ***Mood selection (%) by non-native speakers - Task 1***

Mood	**NNS P1**		**NNS P2**		**NNS P3**		**Mean**
	M	R	N	P	S	T	
Declarative	95	89.5	91.2	81.5	80.6	100	**89.6**
Interrogative	5	10.5	5.9	14.8	19.4	0	**9.2**
Imperative	0	0	2.9	3.7	0	0	**1.1**
Exclamative	0	0	0	0	0	0	**0**
(Total clauses)	(20)	(38)	(34)	(27)	(31)	(17)	**(167)**

Table 4.2.4 ***Mood selection (%) by non-native speakers - Task 2***

Although a closer examination of the actual questions asked shows that there is a difference in the function of the interrogative, with 44% of the 25 questions elicited in Task 1 seeking specific information compared to only 12% (of 17 questions) in Task 2, it is very much evident that none of the speakers display the range of questions found amongst the native speakers.

Only 28% of the questions asked in Task 1 are full clauses (compared to 68% NS). Whilst elliptical questions such as `why three bedrooms? (NNSP1T1), like what? (NNSP2T1)` show that a high level of interaction is being achieved, they can hardly be said to be demonstrating a high level of grammatical proficiency; even more so when the questions are ellipted to the extent that we can only predict their mood from the context and intonation (eg. `no smokers? (NNSP2T2) nothing? (NNSP3T1)`).

The majority of questions in Task 2 are opinion-seeking and indeed often begin with `What do you think...?` This seems to be due to the fact that the less linguistically proficient students find it hard to maintain the flow of a conversation that requires complex ideas to be expressed and so opinion-seeking questions are used to manipulate a shift in topic and to manoeuvre the other speaker into taking a turn. Moreover, some of the interrogatives in Task 2 can be

said to be demonstrating *strategic* competence, when the speaker is struggling to express complex ideas, e.g.:

> I think that's a mistake for them to feel pity of themselves, I think that they have to- ...er to- ...to aaargh! how can I say this? to work, and have a... have a fight (NNSP2T2)

or comprehension checks such as the question in the following exchange:

> T: ...but I think the people is not very er TOGETHer, it is very...
> S: They- they- they just- they just want to know of their OWN lives=
> T: =Yes... ==yes... yes
> S: ==Is this what you are saying? Okay.
> (NNSP3T2)

Errors in the formation of interrogatives are very few; only one speaker in fact makes errors. This may be due to the fact that the range and complexity of the questions asked is very limited and few questions are produced with auxiliary forms.

4.2.3 Question tags

The linguistic feature of putting a question tag at the end of a statement allows a speaker to produce a clause which is midway between the declarative and interrogative mood, thus serving to seek confirmation of facts, opinions, hypotheses and so on, or simply to establish a kind of 'shared' view of things, not necessarily demanding a reply. Carter and McCarthy (1997:18) note that "tags are an essential feature of grammar in use in informal ... contexts of interaction and are particularly appropriate to contexts in which meanings are not simply stated but are negotiated and re-negotiated". In this case, tagged declaratives might be expected to appear in both tasks, but particularly in Task 1.

The frequency of the clause type **declarative:tagged** was calculated in relation to the total number of declarative clauses. The results (table 4.2.5) show that tags were used in all of the native speaker dialogues, with a slightly higher frequency in Task 1.

	NS P1	**NS P2**	**NS P3**	**MEAN**
TASK 1	6.1	6.8	2.2	**5**
TASK 2	3.7	2.9	1.8	**2.8**

Table 4.2.5 ***Overall frequency (%) of declarative:tagged clauses - native speakers***

In Task 1 tags are particularly used to negotiate opinions referring to specific details of the accommodation being discussed, for example:

```
F: It sounds a bit noisy doesn't it?
J: Mm it does     (NSP1T1)

P: Well there you go the seventh floor flat
would be useless wouldn't it?
A: Yeah I know    (NSP2T1)
```

In Task 2, tags are used more rhetorically, with no particular response expected, perhaps to indicate a shared view and sympathy for the other speaker's contribution, for example:

```
F: ...they've been sexually abused ==in their
hou-homes and things=
J:                                  ==Mm  hmm...  mm
hmm...
F: =and been kicked OUT=
J: =Mm hmm that hasn't felt like a HOME to them
at all has it they've got to get away
```

The non-native speaker results are not presented in table form as it was found that there were no occurrences of tags in either task.

4.3 Analysis of Verbal Groups

As already stated (section 3.2.4), Verbal Groups (VGs) were analysed from the aspects of tense, finiteness, modality and polarity. The unit of analysis here is the rank of *group* rather than clause, and all VGs produced have been counted in the analysis, including those found in abandoned (where the complete VG is produced) and embedded clauses, as well as those found in main and dependent clauses. Obviously, ellipted VGs are not counted.

4.3.1 Finite Verbal Groups - native speakers

Finite VGs were counted and coded for tense selection, according to the following categories:

Present Simple

Present Continuous

Simple Past

Past Continuous

Present Perfect

Past Perfect

Future: (*be* ***going to*** *do..*)

Future: (with auxiliary ***will***)[1]

All finite VGs that have a modal auxiliary were categorised as **Modal** VGs.

The frequencies presented in tables 4.3.1 and 4.3.2 on the following page were obtained by calculating percentages of the total count for finite VGs in each case.

[1] This category includes all verbs conjugated with the auxiliary *will* in clauses which make explicit reference to future time, i.e. an Adjunct expressing future time is contained in the clause. Thus many uses of *will* are excluded, for *will* is also used as modal operator to express for example epistemic certainty (eg. He'll be at home now), intention (eg. What will you do now?), or willingness (eg. I'll come with you if you like).

TASK 1	NS P1 T1	NS P2 T1	NS P3 T1	Mean Frequency
Present Simple	78.5	60	52	**63.5**
Present Continuous	0	2	0	**0.6**
Simple Past	8	19.5	35	**20.8**
Past Continuous	0	0	0	**0**
Present Perfect	0	3	2	**1.6**
Past Perfect	0	0	0	**0**
Future: going to	1	0	0	**0**
Future: will	0	0	0	**0**
Modal VGs	12.5	15.5	11	**13**
TOTAL	100 (=103 verbs)	100 (=97 verbs)	100 (=89 verbs)	

Table 4.3.1 ***Relative frequency (%) of verb tenses and modal VGs (NS - Task 1)***

TASK 2	NS P1 T2	NS P2 T2	NS P3 T2	Mean frequency
Present Simple	54	44.5	59.5	**52.5**
Present Continuous	8	1.5	4	**4.5**
Simple Past	5.5	32.5	7.5	**15.2**
Past Continuous	0	4	0	**1.3**
Present Perfect	15.5	0	5	**6.8**
Past Perfect	0	1.5	0	**0.5**
Future: *going to*	0	0	0	**0**
Future: *will*	0	0	1	**0.3**
Modal verbs	17	16	23	**18.7**
TOTAL	100 (=72 VGs)	100 (=74 VGs)	100 (=79 VGs)	

Table 4.3.2 ***Relative frequency (%) of verb tenses and modal VGs (NS - Task 2)***

Task 1 does not appear to elicit a wide range of verb tenses; for each pair more than 80% of the VGs are expressed in either the Present Simple or Simple Past tenses, and if we discount the modal VGs (which do not necessarily show tense aspect) this figure rises to around 95%.

A close look at the lexical choice of verbs reveals that the verbs ***be*** and ***have (got)*** account for the majority, indeed the mean frequency of these two verbs taken together is just over 60% (See table 4.3.3 below). This is due to the focus on describing the features of the accommodation using the structures *it's..., there's...., it's got.... / it was...., there was...., it had....* and so on. Some speakers elected to describe predominantly in the present tense, others in the past tense, others in fact switched between the two. This explains the fact that the frequencies for the Present Simple and Past Simple tenses are so unevenly distributed in this task.

TASK 1	**NS P1 T1**	**NS P2 T1**	**NS P3 T1**	**MEAN**
Be	47.5	47.5	49.5	**48.5**
Have (got)	16.5	7	14.5	**12.6**
(OTHER)	(36)	(45.5)	(36)	**(39.2)**

Table 4.3.3 Lexical choice of verbs by native speakers - Task 1 (% of occurrences)

It is also interesting that all the speakers made considerable use of the cohesive device of ellipsis when describing the accommodation, producing clauses with ellipted verbs which can presumably be recovered as forms of *be* and *have*, although often not directly from the text. This type of situational ellipsis, in which the hearer must recover the verbs from the context, is doubtless due to the fact of having written notes which were frequently referred to directly, resulting in exchanges such as the following:

T: Kitchen?
S: Yeah
T: Yeah, ... bathroom?
S: Yep.
T: Small backyard... has yours got a- == a garden?
S: == oh no
(NS P3 T1)

The other tenses in Task 1 occur almost incidentally; it would certainly be unwise to claim on the basis of this sample that any tenses except the Present Simple and Simple Past are likely to occur, and it should also be borne in mind that only a limited lexical range of verbs seems to be used. This could mean that, from the point of view of *structure*, two 'problem' areas might not be sufficiently tested: irregular verbs (in the Simple Past) and the third person singular in the Present Simple (except for *it is*, *it's got* etc.).

Task 2, on the other hand, seems to elicit a greater range of verbal tenses, and the distribution across speakers is much more varied as the subject matter gives rise to a much wider range of topics and the content of the discussion is obviously much less predictable. Although the Present Simple tense is dominant, it expresses a wider range of functions than in Task 1, serving to comment on current states of affairs, personal feelings, abstract matters and so on.
The Simple Past and Present Perfect tenses also figure quite strongly, functioning to describe personal experiences and anecdotes, for example:

I mean I know I've seen- watched programmes on the TV where they've interviewed people like this (NSP1T2)

When I used to go to Italy I mean there were REALLY I mean you know... er the homeless there had nothing! (NSP3T2)

The recounting of experiences and anecdotes is a prominent feature of the discourse; all the speakers use this as a means of contextualising their opinions and arguments. Other tenses are also used in this way; one speaker, who gives a long description of an experience she had when she encountered a homeless person in the street, makes use of the Past Continuous tense to describe the background to the situation:

> uhm a guy was lying on the side of the road, he was being sick... (NSP2T2)

Another speaker comments on the current situation of her daughter, using the Present Continuous:

> J: My daughter is working in Romania=
> F: =Yeah=
> J: =and uhm she is dealing with children.... (NSP1T2)

Indeed there are occurrences of the Present Continuous in all three pairs.
The Past Perfect, however, occurs only once and there is only one explicit reference to future time.

Another feature of the verb forms in Task 2 is that all three pairs use the **passive** form at least twice, whereas in Task 1 only one example of the passive form can be found, (and that is taken almost directly from the notes about one of the properties: `it's been restored`).

There is not a great difference in the frequencies of modal VGs in Task 1 and Task 2, although the mean frequency in Task 2 (18.7%) is slightly higher than that of Task 1 (13%). (Modality is further discussed in section 4.3.4 below)

4.3.2 Errors in Verbal Groups - non-native speakers

For analysis of the VGs produced by the non-native speakers, only VGs which are used correctly and appropriately were counted. Thus VGs containing errors of Subject-Verb agreement, inappropriate selection of tense, errors in the use of non-finite forms, or errors in the choice or omission of auxiliaries were not included in the scores.

Table 4.3.4 shows the frequency of errors (calculated as a percentage of the total VGs produced by each speaker). The mean frequency of errors in Task 2 is higher than in Task 1, and all speakers except one make a higher proportion of mistakes in Task 2.

Speaker	**TASK 1**	**TASK 2**
M	0	7
R	10	13
N	4	7
P	5.5	8
S	26	15.5
T	8.5	34.5
MEAN	9	14.2

Table 4.3.4 ***Frequency (%) of errors in VGs (non-native speakers)***

The frequencies of VGs for non-native speakers presented in the sections below are therefore calculated as percentages of the total of **correctly formed and appropriately used VGs.**

As a point of interest, the NS performances were not totally 'error-free'; some errors in Subject-Verb agreement were noted, for example. In NS speech, this would be considered a 'slip of the tongue', whereas in NNS speech it would be seen as a lack of grammatical accuracy.

4.3.3 Finite Verbal Groups - non-native speakers

Tables 4.3.5 and 4.3.6 show non-native speaker frequencies for tense and modal VGs.

TASK 1	**NNS P1**		**NNS P2**		**NNS P3**		**Mean**
	M	**R**	**N**	**P**	**S**	**T**	
Present Simple	41	77	80	81	60	90.5	**71.5**
Present Continuous	0	0	0	0	0	0	**0**
Simple Past	41	9	4.5	2	20	9.5	**14.4**
Past Continuous	0	0	0	0	0	0	**0**
Present Perfect	0	0	4.5	2	0	0	**1.1**
Past Perfect	0	0	0	0	0	0	**0**
Future:going to	0	0	0	4	0	0	**0.6**
Future:will	0	0	0	0	0	0	**0**
Modals	18	14	11	11	20	0	**12.4**
TOTAL	100 (=39 VGs)	100 (=43 VGs)	100 (=44 VGs)	100 (=48 VGs)	100 (=30 VGs)	100 (=21 VGs)	

Table 4.3.5 ***Relative frequency (%) of verb tenses and modal VGs (NNS - Task 1)***

TASK 2	NNS P1		NNS P2		NNS P3		Mean
	M	R	N	P	S	T	
Present Simple	89.5	72.5	76.5	83.5	75	64.5	**77**
Present Continuous	0	7.5	3	0	8.5	0	**3.2**
Simple Past	0	2.5	0	0	0	21.5	**4**
Past Continuous	0	0	0	0	0	0	**0**
Present Perfect	0	5	0	0	0	0	**0.8**
Past Perfect	0	0	0	0	0	0	**0**
Future:going to	0	0	0	0	0	0	**0**
Future:will	0	0	0	0	0	0	**0**
Modals	10.5	12.5	20.5	16.5	16.5	14	**15**
TOTAL	100 (=19 VGs)	100 (=40 VGs)	100 (=34 VGs)	100 (=30 VGs)	100 (=36 VGs)	100 (=14 VGs)	

Table 4.3.6 Relative frequency (%) of verb tenses and modal VGs (NNS - Task 2)

The results from **Task 1** are very similar to those of the native speakers, with the Present Simple and Simple Past dominating the tense selection, and the reason is the same - the main body of the dialogue is focused on a description of the accommodation under consideration, with four of the speakers using predominantly the Present Simple tense for description and two speakers alternating between the Present Simple and the Simple Past tenses. In fact, the pattern of frequencies of the verbs *to be* and *have (got)* (table 4.3.7) is remarkably similar to that of the native speakers, the combined mean frequency being almost identical at 61.5%.

TASK 1	**M**	**R**	**N**	**P**	**S**	**T**	**MEAN**
BE	54	50	52.5	69	27	57	**51.5**
HAVE (GOT)	10	14	5	4	26.5	0	**10**
(OTHER)	(36)	(36)	(42.5)	(27)	(46.5)	(43)	**(38.5)**

Table 4.3.7 ***Lexical choice of verbs by non-native speakers - Task 1 (% of occurrences)***

No other tenses occurred in the dialogues of either P1 or P3, although the P2 speakers both used the Present Perfect, firstly to introduce the conversation:

N: So, P---, have you checked... any apartments or houses for us?=
P: =Yes, I've seen two (NNSP2T1)

and later, speaker N negotiates a change in topic (to the accommodation that *he* has seen):

N: Well I've seen two myself (NNSP2T1)

Speaker P also uses the future:*going to* form twice, eg.

I think that's gonna be... very expensive! (NNSP2T1)

It is interesting that the native speakers made no use of this 'real' future time sense, preferring instead to discuss the situation as a largely hypothetical one, thus reflecting the 'artificial' context of the task. (See section 4.3.4 below on modality).

Task 2, however, did not elicit a much wider range of tenses amongst the non-native speakers than did Task 1, with the Present Simple accounting for a mean frequency of 77%. Two speakers used *only* the Present Simple (one of them being speaker P who actually used the widest range of tenses of all the speakers

in Task 1). The Present Continuous was used by three of the speakers, and the Present Perfect by only one. The Simple Past was relatively unused, except by one speaker - the only one in fact who recounted a personal experience, a feature which, as may be recalled, was exploited by all the native speakers in their discussions of the topic.

The relative frequency of modal VGs is similar to that of the native speakers, with the mean frequency for Task 2 (15%) being slightly higher than that of Task 1 (12.4%).

As for the use of the passive form, there was a single occurrence in Task 1, and only two occurrences in Task 2.

4.3.4 Modality

The forms identified as modal VGs (modals) are VGs which express modality in their internal structure by the inclusion of one of the following:

- Modal auxiliary (eg, can, could, might, etc.)
- Lexico-modal auxiliary (eg. have got to)
- Modal idiom (eg. would rather)

As we have seen, modals are used in both tasks, although with a slightly higher frequency in Task 2 (both NS and NNS).

Because both tasks seem likely to elicit modals, the discussion of modality will be rather limited. However it is useful to highlight the main differences and similarities between the use of modals in each task and between the native and non-native speakers.

The **native speaker** use of modals is dominated by the expression of epistemic (extrinsic) possibility[2] in both tasks. In Task 1, the speakers switch primarily between giving categorical statements of facts (describing the accommodation according to the information given) and hypothesising on the likelihood of what

[2] Expressions of epistemic possibility are assessing the likelihood of a proposition being true and include expressions of certainty, belief, possibility, impossibility, reasonable inference, expectation.

renting a particular property would entail. In Task 2, the focus is very much on delivering opinions, and the use of epistemic possibility therefore enables speakers to give opinions without committing themselves to the truth of their assertions.

Nevertheless, there is a noticeable difference in the forms used to express epistemic possibility. In Task 1, the hypothetical nature of the context is reflected in the use of *would*, which besides expressing epistemic possibility (`eg. Yeah I s'pose it'd be a pain moving everything in as well wouldn't it really? (NSP2T1)`) is also used to speculate about necessity (*would have to*) and ability (*would be able to*) and to express preferences (*would like, would rather*). All of these forms together account for 65% of modals in this task. In Task 2, *can('t)* and *could(n't)* are the most frequent forms used, expressing not only epistemic but also intrinsic[3] possibility and ability.

Apart from these dominant meanings and forms, another use of modals which stands out in Task 1 is expressing suggestions (eg. `Shall we go for that one first? (NSP2T1)`), and in Task 2 the expression of necessity (eg. `folk've got to be willing to help themselves (NSP2T2)`).

Modal auxiliaries that are noticeable for their absence are *may, must* and *ought to* - there are no occurrences of these auxiliaries in either task.

[3] Expressions of intrinsic possibility include expressions of permission, prohibition and social unacceptability / undesirability

The **non-native** speaker use of modal *meaning* is similar to that of the NS, but the forms used are rather different. In both tasks, the dominant modal is *can('t)*, accounting for 69% of all modals used in Task 1 and 48% in Task 2. *Can* is used to perform a variety of functions, and, particularly in Task 1, is often used in place of the more 'natural' sounding (according to the NS performances and the hypothetical nature of the task) auxiliary *could* (or *would*, *might*), for example:

...er so, we can- we can rent the first one because it's cheaper (NNSP3T1)

The over (and sometimes possibly inappropriate) use of this form means that these speakers do not demonstrate such a broad range of modals as the native speakers.

It should be noted that modality can also be expressed by **Modal Adjuncts**. The two forms of Modal Adjuncts identified in the data were adverbs (eg. *possibly*, *perhaps* etc.), and mental activity clauses[4] such as *I (don't) think, I reckon, I suppose*.

In the NS discourse, 7.3% of all completed clauses in Task 1 and 11.3% in Task 2 contained a Modal Adjunct. In both cases, the majority of Modal Adjuncts were realised by mental activity clauses.

In the NNS discourse, no adverbs expressing epistemic possibility were used, but *I think* was used by all the speakers, particularly in Task 2 where it appears in 15.7% of all completed clauses.

[4]Mental activity clauses are recognised by Eggins and Slade (1997:101) as being Modal Adjuncts which realise "incongruent modalisation with explicit subjective source", although traditionally grammar references do not recognise this form of modalisation. (Holmes (1988:32-33) discusses the lack of attention given to these forms in EFL materials.) Eggins and Slade (1997:115: note 2) suggest that the best means of determining whether or not these clauses are to be considered as Adjunct is by listening to the intonation (as Adjunct the clause will fall into the intonation contour of the larger clause) and the tag-test, which will show whether or not the *I* is the Subject of the clause.

4.3.5 Non-finite Verbal Groups - native speakers

Non-finite VGs, which do not show tense, were analysed separately and were identified as being either *to*-infinitives or participles. Many participles have a rather uncertain status, as they can be classified as *adjectives* deriving from verbs, so only those participles which are clearly expressing a process were included in the frequency count.

By referring to tables 4.3.8 and 4.3.9, which show the comparative frequency of non-finite VGs compared to finite VGs in both tasks, it can be seen that the frequency of non-finite VGs elicited in Task 2 is almost three times that of Task 1.

TASK 1	**NS P1 T1**	**NS P2 T1**	**NS P3 T1**	**MEAN**
non-finite	4.6	7.5	5.2	**5.8**
finite	95.4	92.5	94.8	**94.2**
TOTAL	100%	100%	100%	
	(= 109 VGs)	(= 106 VGs)	(= 96 VGs)	

Table 4.3.8 ***Relative frequency (%) of non-finite and finite VGs (NS- Task 1)***

TASK 2	**NS P1 T2**	**NS P2 T2**	**NS P3 T2**	**MEAN**
non-finite	15.1	16.5	16	**15.9**
finite	84.9	83.5	84	**84.1**
TOTAL	100%	100%	100%	
	(= 86 VGs)	(=91 VGs)	(= 94 VGs)	

Table 4.3.9 ***Relative frequency (%) of non-finite and finite VGs (NS - Task 2)***

A more detailed analysis of non-finite VGs showed that in both tasks there were more occurrences of the infinitive (V-*to*-infinitive) than the participle form (table 4.3.10) but the frequencies of both types increased from Task 1 to Task 2. The only participle form identified as being a VG was the *present* participle (-*ing* form).

	Task 1	**Task 2**
V-to-infinitive	69	55
Participle	31	45
(TOTAL)	100% = 16 VGs	100% = 41 VGs

Table 4.3.10 ***Relative frequency of V-to-infinitive and participle form (all NS)***

In both tasks, the functions of the **V-*to*-infinitive** fall almost entirely into two identifiable categories, each one accounting for around 50% of occurrences:

i) as part of a Verbal Group complex, where the V-*to*-infinitive expresses a process which is modified in some way by another, preceding verb (which carries the finiteness)

e.g.: `the argument does tend to fall on the side of the carers (NSP2T2)`

ii) as a qualifier in a Nominal or (Adjectival) Group

e.g.: `I don't think we've got time to do that (NSP1T1)`

`It's difficult to generalise (NSP3T2)`

Therefore although there was a higher frequency of V-*to*-infinitive in Task 2, there does not appear to be a difference its function between the tasks.

In Task 1, there are only 5 occurrences of the **participle** form in total. In Task 2, however, there is a large increase in the use of the participle. The functions of the participle in Task 2 fall mostly into 3 categories:

i) following a preposition (part of a prepositional Complement)

e.g.: they're scared of being with people, other people (NSP1T2)

ii) as part of a Verbal Group complex,

e.g: I remember working in Switzerland (NSP2T2)

ii) as the main VG of a dependent or embedded clause,

e.g: and sitting down with him, I was amazed at his situation (NSP2T2)

4.3.4 Non-finite Verbal Groups - non-native speakers

Tables 4.3.11 and 4.3.12 show the comparative frequency of finite and non-finite verbal groups produced by the non-native speakers in both tasks.

TASK 1	**NNS P1 T1**		**NNS P2 T1**		**NNS P3 T1**		**MEA**
	M	**R**	**N**	**P**	**S**	**T**	
non-finite VGs	4.9	0	10.2	4	3.2	0	**3.7**
finite VGs	95.1	100	89.8	96	96.8	100	**96.3**
TOTAL	100 (=41 VGs)	100 (=44 VGs)	100 (=49 VGs)	100 (=50 VGs)	100 (=31 VGs)	100 (=21 VGs)	

Table 4.3.11 ***Relative frequency (%) of finite and non-finite VGs (NNS - Task 1)***

TASK 2	NNS P1 T1		NNS P2 T1		NNS P3 T1		MEA
	M	**R**	**N**	**P**	**S**	**T**	
non-finite VGs	32.1	11.1	15	8.8	5.3	6.7	**13.2**
finite VGs	67.9	88.9	85	91.2	94.7	93.3	**86.8**
TOTAL	100	100	100	100	100	100	
	(=28 VGs)	(=45 VGs)	(=40 VGs)	(=34 VGs)	(=38 VGs)	(=15 VGs)	

Table 4.3.12 ***Relative frequency (%) of finite and non-finite VGs (NNS - Task 2)***

As with the native speakers, the mean frequency (and indeed each individual speaker frequency) of non-finite VGs increases in Task 2. The mean frequencies for both tasks are only slightly lower than the NS mean frequencies.

In **Task 1**, the actual occurrences of non-finite VGs are few. Two speakers do not produce any at all and the participle form is particularly lacking - there are only three occurrences in total, all in fact occurring together in the following exchange:

P: ... it's only FIVE minutes from the centre of the city
N: Five minutes WALKing? or by- by bus or
P: Walking
N: Walking (NNSP2T1)

The V-*to*-infinitives used in Task 1 mostly function as part of a Verbal Group complex, e.g:

we need to do something about it (NNSP1T1)

In **Task 2**, there is a significant use of the V-*to*-infinitive (table 4.3.13) but again the participle form is not very frequent.

	Task 1	**Task 2**
V-to-infinitive	70	84.5
Participle	30	15.5
(TOTAL)	100% = 10 VGs	100% = 26 VGs

Table 4.3.13 ***Relative frequency of V-to-infinitive and participle form (all NNS)***

The V-*to*-infinitive functions mostly as a qualifier of a Nominal or Adjectival group (64% of occurrences), but also features as part of a Verbal Group complex (18%), and as the VG in a non-finite clause filling the Complement slot of a main clause (13.5%)

4.3.5 Negation in VGs

Finally, we will look at the frequencies of negative VGs to see if one task seems more likely to elicit negative forms than the other.

For the native speakers, there was a slight increase in the mean frequency of negative VGs in Task 2. This increase was also reflected across each pair (see table 4.3.14).

	NS P1	**NS P2**	**NS P3**	**MEAN**
TASK 1	13.6	9.3	6.7	**9.9**
TASK 2	15.3	12.2	8.9	**12.1**

Table 4.3.14 ***Overall frequencies (%) of negative VGs - native speakers***

The non-native speakers repeated this pattern, with the mean frequency increasing in Task 2, but in fact the increase was much greater (see table 4.3.15),

with the frequency at least doubling for each speaker, although it should be remembered that the total number of VGs produced in Task 2 was smaller.

	NNS P1		**NNS P2**		**NNS P3**		**MEAN**
	M	**R**	**N**	**P**	**S**	**T**	
TASK 1	18	14	7	6.5	10	5	**10.1**
TASK 2	37	27.5	20.5	26.5	22	14.54	**24.6**

Table 4.3.15 ***Overall frequencies (%) of negative VGs - non-native speakers***

No apparent reason could be found for the increase in frequency of negatives in Task 2, and there is no significant difference in the forms of the negative used.

4.4 Analysis of Subordination

In this study, subordination was investigated by examining both **embedded** and **dependent** clauses. The types of subordinate clauses identified in the data are: **adverbial clauses**, **relative clauses, *that*-clauses** and ***wh*-clauses**, all of which will be further defined below, and **non-finite clauses**.

All embedded and dependent clauses were counted and frequencies determined by calculating the percentage of occurrences in the total number of **completed** clauses (i.e. discounting minor and abandoned clauses) produced. The actual forms of subordination are then discussed in more detail.

4.4.1 Overall frequency of subordinate clauses - native speakers

Tables 4.4.1 and 4.4.2 show the overall frequency of subordinate clauses for both tasks.

TASK 1	**NS P1**	**NS P2**	**NS P3**	**MEAN**
Dependent clauses	11.4	6.4	5.5	**7.8**
Embedded clauses	6.7	6.4	4	**5.7**
TOTAL	**18.1**	**12.8**	**9.5**	**13.5**

Table 4.4.1 ***Overall frequency (%) of subordinate clauses - NS Task 1***

TASK 2	**NS P1**	**NS P2**	**NS P3**	**MEAN**
Dependent clauses	15	8.9	13.4	**12.4**
Embedded clauses	15	2.5	10.4	**9.3**
TOTAL	**30**	**11.4**	**23.8**	**21.7**

Table 4.4.2 ***Overall frequency (%) of subordinate clauses - NS Task 2***

It can be seen that there is a higher frequency of subordinate clauses, of both the embedded and dependent type, in Task 2.

4.4.2 Overall frequency of subordinate clauses - non-native speakers

Tables 4.4.3 and 4.4.4 show that the NNS also used subordinate clauses with a higher frequency in Task 2 (the mean frequency is almost the same as the NS mean frequency for Task 2) but the frequency for Task 1 is lower than that of the native speakers.

TASK 1	**NNS P1**		**NNS P2**		**NNS P3**		**MEAN**
	M	**R**	**N**	**P**	**S**	**T**	
Dependent clauses	0	2	4.5	1.5	12.8	3.8	**4.1**
Embedded clauses	2.6	5.9	1..5	3	0	3.8	**2.8**
TOTAL	**2.6**	**7.9**	**6**	**4.5**	**12.8**	**7.6**	**6.9**

Table 4.4.3 ***Overall frequency (%) of subordinate clauses - NNS Task 1***

TASK 2	**NNS P1**		**NNS P2**		**NNS P3**		**MEAN**
	M	**R**	**N**	**P**	**S**	**T**	
Dependent clauses	4.8	17.4	8.1	6.9	3	19	**9.7**
Embedded clauses	4.8	10.9	16.2	13.8	33.3	0	**13.2**
TOTAL	**9.6**	**28.3**	**24.3**	**20.7**	**36.3**	**19**	**22.9**

Table 4.4.4 ***Overall frequency (%) of subordinate clauses - NNS Task 2***

4.4.3 Types of subordinate clauses identified in the data

Adverbial clauses

Schleppegrell (1992) argues that although subordination is often used as a measure of linguistic complexity, in spoken discourse many clauses introduced by so-called 'subordinating conjunctions' (adverbials) are not actually subordinate at all, and she cautions against simply counting tokens of 'subordinators' such as *because* in order to draw conclusions about syntactic complexity. She points out that many uses of *because* are not examples of hypotaxis, but rather paratactic markers of explanation and elaboration (Schleppegrell 1992:125) thus functioning in a broader, cohesive role in a piece of discourse.
A close look at the transcripts of both the native and non-native speakers reveal many instances of *because* which support Schleppegrell's claim, for example:

J: Which would you rather live in, a flat or a house?
F: Uhm I think I'd rather have a house?=
J: =Would you?=
F: =mm
J: **Because** the other property I saw is in fact a - a modern two bedroom FLAT (NSP1T1)

Yes! but institutions like... uhm CHARITY institutions they do what they can but it's not enough **because** people... uhm... it's selfish people in Portugal (NNSP1T2)

Another adverbial which *can* be used as a subordinator, particularly in written discourse (to show a 'result' clause), but often functions as a coordinator in speech, is *so*. Indeed, the majority of occurrences of *so* in the data are clearly not examples of subordination, e.g:

P: but it's more CENTRAL, you know, and I think for er == what we want
N: ==**so** you- **so** you LIKE this one (NNSP2T1)

For this reason, care was taken to count only those examples of *because* and *so* which clearly introduce subordinate ideas of reason and result respectively, eg:

> Yeah, but we wouldn't need to use the car **because** there's- there IS a railway station nearby (NSP2T1)
>
> It's got central heating **so** it's nice and warm (NSP1T1)

Other adverbial conjunctions identified can be classified as **concessive** (*although, even though*), **conditional** (*if, unless, whether or not, as long as*), **temporal** (*before, when*) and **spatial** (*where*).

Relative clauses

Relative clauses were counted as a single category, whether defining, non-defining, or commenting on the whole content of the main clause, and whether or not the relative pronoun was actually used.

***That*-clauses**

That-clauses typically introduce reported speech or ideas, and may or may not include the conjunction *that*.

Another type of structure found in the native speaker discourse is the **prefacing** structure used to introduce an idea; *the (...) thing (..) is (that).....x.* Although the *that*-clause is sometimes arguably the main idea of the clause, all structures of this type have been included in the count for subordination as *that*-clauses.

This type of prefacing often creates a highly complex structure, as can be seen in this example (which also contains a further *that*-clause (with *that* omitted) and a relative clause, both embedded at Subject):

I've always thought the one thing that's set this country aside from America was **that** we- we never GO quite as far as they go (NSP3T2)

***Wh*-clauses**

The *wh*-clauses identified in the transcripts are generally introduced by *what* or *how* and are used at Complement, eg:

I don't know **what** I would do (NSP1T2)

The *wh*-pronoun is sometimes omitted:

it's a little bit more than we budgeted for (NSP2T1)

Non-finite clauses

Non-finite clauses are counted here if they are dependent clauses with a non-finite Main verb. Therefore not all the non-finite VGs identified in section 4.3 are included.

4.4.4 Relative frequencies of different types of subordinate clause - native speakers

Table 4.4.5 shows the mean frequencies of the different types of subordination identified, in relation to the number of completed clauses produced, in both tasks:

	TASK 1	**TASK 2**
adverbial clauses	5	8
relative clauses	3.6	8
that-clauses	1.8	2.3
wh-clauses	2.2	1.9
non-finite clauses	0.9	1.5

Table 4.4.5 ***Mean frequencies (%) of different types of subordinate clauses NS Tasks 1 and 2***

The relative clauses in **Task 1** tend to be variations of the same idea: `the two properties that I saw; the first one I saw,` and so on, although there are also examples of 'comment' clauses introduced by *which*.

The most frequent type of adverbial clause is **conditional**, once again reflecting the hypothetical nature of the task, but examples of reason, result and concessive clauses can also be found.

Task 2 seems to elicit a higher frequency of adverbial clauses, but again the most frequent type is the conditional, followed by reason clauses introduced by *because*.

A higher frequency of relative clauses is also apparent in Task 2, and these clauses fulfil a much wider range of functions than in Task 1.

There is no significant difference between the two tasks as regards the use of *that*-clauses, *wh*-clauses or non-finite clauses.

4.4.5 Relative frequencies of different types of subordinate clause - non-native speakers

Table 4.4.6 shows that a pattern similar to that of the NS can be found as regards the frequencies of both adverbial and relative clauses:

	TASK	**TASK 2**
adverbial clauses	4	8
relative clauses	1.4	6.8
***that*-clauses**	0	0.5
***wh*-clauses**	1.4	6.8
non-finite clauses	0	1

Table 4.4.6 ***Mean frequencies (%) of different types of subordinate clauses - NNS Tasks 1 and 2***

The use of relative clauses in Task 1 is once again fairly limited to expressions of the same kind of idea (`what about the other one that you saw? /the first one that I saw`), whereas the use is more varied in Task 2.
It is noticeable that the adverbial category of subordinate clause does not include as many examples of the conditional as the NS in either task.
The frequency of *wh*-clauses appears to increase quite significantly in Task 2, but this rise is mostly due to one speaker.
That-clauses and non-finite clauses do not appear in Task 1, but their occurrence in Task 2 is also very infrequent.

5. Summary and conclusions

5.1 Summary of results

The results presented in the previous chapter are briefly summarised in quantitative terms below, by indicating where there appears to be a difference between the two tasks (< >) and where there does not (=) based on the mean frequency figures for native speaker and non-native speaker performance.

Interrogatives:	T1 > T2 (NS)	T1 = T2 (NNS)
Tagged declarative:	T1 > T2 (NS)	T1 = T2 (NNS)
Range of tenses:	T1 < T2 (NS)	T1 = T2 (NNS)
Modals:	T1 < T2 (NS)	T1 < T2 (NNS)
Non-finite VGs:	T1 < T2 (NS)	T1 < T2 (NNS)
Negative VGs:	T1 < T2 (NS)	T1 < T2 (NNS)
Subordinate clauses	T1 < T2 (NS)	T1 < T2 (NNS)

When the results are represented in this way, they seem to suggest that, judging by NS performance, there is a case for arguing that *all* the features selected are affected by the task variable.

This summary is extremely simplistic, however, and does not really present a true picture of the results. It is important to remember, for example, that 'less than' (<) does not mean that the feature is not likely to appear at all. Nonetheless, it does seem to be the case that certain structures were affected by strongly by the task variable. Amongst the native speakers, this is particularly true for the use of the interrogative form and verbal tenses, and for the frequency of subordinate clauses.

Non-native speaker performance showed similar results for modals, non-finite VGs, negative VGs and subordinate clauses, but not for interrogatives where the

frequencies of occurrence were very similar for both tasks, or tagged declaratives, which did not occur at all, or range of tenses, where little difference was found between the two tasks.

However it is the NS results that are most significant in terms of obtaining a task 'profile' which could then be useful to inform decisions as to which task(s) to select for observing and measuring certain features of grammatical competence and to indicate to assessors which sort of structures they might look for. Largely on the basis of the NS results, then, a more qualitative profile of the two tasks could be drawn up along the following lines:

	Task 1	**Task 2**
Interrogative forms	Speakers should be able to demonstrate a wide range of interrogative forms, both full and elliptical, in order to seek and confirm specific information and opinions. Tag questions may also occur.	Questions may occur, mostly to ascertain opinions or initiate a new topic. Interrogative forms will perhaps also be used as a strategy to negotiate meaning during exchanges of complex ideas. Tag questions may also occur.
Verbal forms	Tense selection will be largely restricted to the Present Simple and Simple Past, with a fairly limited lexical range (*it is / it's got* etc. will doubtless be frequently used). A great deal of verbal ellipsis is likely to occur when describing the accommodation. Negative forms should occur. Non-finite forms will not necessarily occur, particularly the participle form.	A wide range of tenses can be expected if the speaker uses the resource of recounting personal experiences and anecdotes to contextualise and illustrate ideas and arguments. Negative forms should occur. Non-finite forms may well occur, both infinitives and participles

Modal forms	Modals should occur, particularly to express hypothetical meanings of possibility, likelihood, preferences, and also suggestions. Modal Adjuncts such as *I think, I suppose* may be frequent	Modals should occur, typically to express likelihood, necessity, advisability. Modal Adjuncts such as *I think, I suppose* may be frequent.
Subordinate clauses	Some subordinate clauses may occur, for example the conditional, but generally ideas are linked paratactically.	Complexity of ideas and thoughts may well entail the use of subordinate clauses, particularly adverbial and relative clauses.

5.2 Conclusions

The results of this study seem to suggest that the linguistic output in terms of grammatical structures is likely to be different between tasks. This finding is certainly not surprising; intuitively we might guess this to be so, and indeed other empirical studies have also suggested this to be the case in *other* areas of linguistic output. Duff (1986) found that convergent (problem solving) tasks generate more interaction than divergent (debate) tasks. Courtney (1996) finds that the preliminary results of his study designed to test the effect of task type on output in terms of the strategies associated with negotiation of meaning seem to support the hypothesis that "there is a statistically significant relationship between task type and learner performance" (*ibid.*:323).

Any conclusions drawn from the results presented here are, of course, extremely tentative, due to the limitations of the study. The data sample is small, and because of this no statistical tests such as t-tests have been used. Furthermore, the population is restricted, particularly the NNS population who all have similar learning backgrounds as well as cultural, linguistic, age and social backgrounds. Obviously, then, further research involving a larger and more varied population would be needed to support the hypothesis that T1 differs from T2 in terms of grammatical output.

However, it can be argued that only relatively small databases are needed to reveal grammatical (as opposed to lexical) patterns in spoken English as grammar consists of a relatively small number of items and patterns are frequently repeated (Carter and McCarthy 1995: 143). The results of this study do seem to suggest that grammatical patterns are fairly predictable according to task type. Different tasks can be considered to elicit different *genres* of spoken discourse, and I would therefore support Carter and McCarthy's (*ibid.*: 144) argument that a genre specific description of the spoken language is an extremely

useful resource for teachers and learners of English, and, in this case, those involved in developing tests of oral proficiency.

If we are to accept the limited evidence of the findings of this study, the implications then seem to be two-fold; both for the testing and the teaching of task-based spoken English.

The principal implications for testing would appear to be firstly, that if we wish to measure grammatical competence as an independent component of communicative competence, we need to be sensitive to the role of grammar *in context* and to the different functions that grammatical structures are called upon to realise in different communicative tasks. Secondly, care should be taken with band descriptors which make general statements about range and complexity of grammatical structures. What exactly is meant by these terms must be carefully considered and clearly defined. It may well be that the definitions need to be adjusted to suit a particular task type.

Let us consider again the task types used in this study, according to their genre. **Task 1** has a focused, practical goal and participants will therefore be unlikely to be concentrating on producing complex, varied language. The nature of the task requires participants to concentrate on the economical, efficient exchange of information and opinions under pressure of time and in order to bridge an information gap. Thus it is not surprising that the task does not seem to elicit a wide or complex range of structures, as traditionally defined in second language pedagogy. The complexity of this task does not lie in the exchange of complex and varied ideas (which we might reasonably expect to be reflected in complex and varied language), but in the high degree of interaction and negotiation of meaning that is necessarily involved.

Task 2, on the other hand, requires participants (whether NS or NNS) to demonstrate a greater degree of eloquence in the expression of information; they need to be articulate and possess a wider and more complex range of grammatical structures in order to get their ideas and opinions across.

All of this implies, then, that band descriptors for grammatical competence which are not developed with task type in mind could be confusing and even misleading to those involved in assessment procedures. For example, to assess performance on Task 1 with complexity and range of verbal tenses in mind would inevitably lead to problems. Furthermore, it also suggests that an attempt to tap grammatical competence by means of one particular type of task may not be sufficient to gain an overall picture. The use of a battery of task types in a test of oral proficiency could well be a more appropriate alternative. This case has been well argued by Shohamy *et al* (1986). In this case, a task such as Task 1 could be used to assess the appropriate use of grammatical structures to convey and negotiate the necessary information, with a greater emphasis on overall task management and interactive skills, whereas a task such as Task 2 may need more weighting given to complexity and range of grammatical structures which reflect the more complex and varied *content* of the discussion and the discourse competence necessary for the textual organisation within longer turns.

Although the main objective of this study was to determine the effect of task type in a test situation, inevitably the results have implications for the teaching of spoken English, whether as a general task-based classroom activity or as specific preparation for a test (test washback). The points of interest here lie especially in the results which show distinct *differences* in the NS and NNS performances, which perhaps indicate that the NNS are approaching the tasks with a different conceptualisation of what each task involves. The discrepancies in the use of the interrogative form are particularly interesting. The fact that using the interrogative form in Task 1 as both a grammatical and interactive tool would be likely to enhance performance was only exploited by two of the NNS (and even these speakers did not tend to use full forms). The use of NS performances on specific tasks as a teaching resource can serve to highlight features such as this, as well as other identifiable features of 'natural' spoken discourse, for example

the ways in which tags, modals and hypotaxis are (or are not) used, or the way that the recounting of personal experiences (Task 2) can be exploited grammatically. An interesting research experiment would be to determine if the use of NS performances as teaching materials for specific task-based activities actually has an effect.

The implications for both testing and teaching extend beyond this, however. Recently, there has been interest in the concept of a spoken grammar of English which is distinct from a written grammar (Brazil 1995; Carter and McCarthy 1995, 1997; McCarthy and Carter 1995). It is beginning to be realised that there are many frequently occurring features of informal, spoken discourse which are not given much attention (or sometimes not even recognised) in EFL / ESOL pedagogy. Examples of this are situational ellipsis (briefly discussed in section 4.3.1), and the 'fronting' or prefacing of topical information, for example: `this other one, this three bedroomed house, that sounds quite reasonable (NSP2T1)` (see also section 4.4.3, page 65). As further work is carried out on the nature of spoken grammar, no doubt the whole concept of oral grammatical competence will need to be redefined.

In this study, grammatical features which are traditionally considered to be complex, or problematic, were deliberately chosen as in this way the study better reflects the reality of the current situation of oral proficiency testing. Nonetheless, the selection of certain grammatical structures by no means reflects the entire picture of range and complexity in spoken discourse.

Halliday (1994: xxiv) notes that

> "I have often pointed out that speech is no less complex than writing, but that the two gain their complexity in different ways. The complexity of writing lies in its density, the packing together of lexical content, but in rather simple grammatical frames (....) The complexity of spoken language is more like that of a dance; it is not static and dense but mobile and intricate. (....) Much more of the meaning [of spoken language] is expressed by the grammar than by the vocabulary. As a

> consequence, the ... structure is highly complex, reaching degrees of complexity that are rarely attained in writing."

It might well be asked if the concept of grammatical competence in spoken language is not relying too heavily on the grammar of *written* language. Out of interest, three experienced EFL teachers were asked to listen to the tapes of the non-native speakers and comment on their performances. The comments made in relation to grammatical competence seem to support the idea that assessments are often made according to a sort of internal syllabus of grammar based on pedagogical practices which originate from the more traditional 'sentence' grammars, and give priority to typically 'problematic' areas of grammar. Features picked out as being examples of 'complex' language included the conditional form and the present perfect form. References were made to the fact that only 'simple' structures were used to link ideas, and to the fact that some of the speakers did not 'try' to use verbs other than 'is' and 'has' (Task 1). One teacher did not like the fact that 'double subjects' were being used (eg. 'the price it's...'), but in fact the native speakers also used this structure, eg: `the house well it's a TERRACE house (NSP1T1)` (although we should note that intonation is important here; NS use of this structure would have the repeated subject as unstressed). Another useful area for further research would in fact be to investigate the actual criteria used for judging spoken grammar by individual assessors, using introspective techniques.

In conclusion, then, it is hoped that the study reported here has shed at least a little light on the relationship between task type, grammatical competence and oral proficiency testing by identifying some of the problems in this complex relationship and attempting to suggest ways in which these problems might be approached in the present and in the future.

REFERENCES

Akinnaso, F.N. (1982) On the differences between spoken and written language. *Language and Speech* 25/2: 97-125

Alderson, J.C. (1991) Bands and scores. In Alderson, J.C. and B. North (eds.) *Language Testing in the 1990's*. London: MacMillan

Alderson, J.C., C. Clapham and D. Wall (1995) *Language Test Construction and Evaluation*. Cambridge: Cambridge University Press

Austin, J.L. (1962) *How to Do Things with Words*. Oxford: Clarendon Press

Bachman, L.F. (1990) *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press

Bachman, L.F. and A.S. Palmer (1982) The construct validation of some components of communicative proficiency. *TESOL Quarterly* 16/4: 449-465

Biber, D. (1988) *Variation Across Speech and Writing.* Cambridge: Cambridge University Press

Brazil, D. (1995) *A Grammar of Speech.* Oxford: Oxford University Press

Brown, G., A. Anderson, R. Shillcock and G.Yule (1984) *Teaching Talk: Strategies for Production and Assessment.* Cambridge: Cambridge University Press

Brown, G. and G. Yule (1983) *Teaching the Spoken Language*. Cambridge: Cambridge University Press

Brown, R. (1973) *A First Language*. London: Allen and Unwin

Brumfit, C.J. (1984) *Communicative Methodology in Language Teaching.* Cambridge: Cambridge University Press

Bygate, M. (1987) *Speaking*. Oxford: Oxford University Press

Cameron, D. (1987) Review of Carroll, B.J. and P. Hall (1985) Make Your Own Language Tests. *ELT Journal* 41: 149-151

Canale, M. (1983) From communicative competence to communicative language pedagogy. In J.C. Richards and R.W.Schmidt (eds) *Language and Communication.* London: Longman

Canale, M. and M. Swain (1980) Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics* 1:1-47

Carroll, B. J. (1980) *Testing Communicative Performance*. Oxford: Pergamon

Carroll, B. J. and P. Hall (1985) *Make Your Own Language Tests*. Oxford: Pergamon

Carroll, B.J. and R. West (1989) *ESU Framework: Performance Scales for English Language Examinations*. Harlow: Longman

Carter, R. and M. McCarthy (1995) Grammar and the spoken language. *Applied Linguistics* 16/2: 141-158

Carter, R. and M. McCarthy (1997) *Exploring Spoken English*. Cambridge: Cambridge University Press

Chafe, W.L. (1982) Integration and involvement in speaking, writing and oral literature. In D. Tannen (ed.) *Spoken and Written Language*. Norwood, NJ: Ablex

Chambers, F. and B. Richards (1993) Oral assessment: the views of language teachers. *Language Learning Journal* 7: 22-26

Chambers, F. and B. Richards (1995) The 'free conversation' and the assessment of oral proficiency. *Language Learning Journal* 11: 6-9

Courtney, M. (1996) Talking to learn: selecting and using peer group oral tests. *ELT Journal* 50/4: 318-326

Doughty, C. and T. Pica (1986) Information gap tasks: do they facilitate second language acquisition? *TESOL Quarterly* 20/2: 305-325

Downing, A. and P. Locke (1992) *A University Course in English Grammar*. Hemel Hempstead: Prentice Hall

Duff, P. (1986) Another look at interlanguage talk: Taking tasks to task. In R. Day (ed.) *Talking to Learn.* Rowley, Mass.: Newbury House

Dulay, H. and M. Burt (1974) Errors and strategies in child second language acquisition. *TESOL Quarterly* 8:129-136

Eggins, S. (1994) *An Introduction to Systemic Functional Linguistics.* London: Pinter

Eggins, S. and D.Slade (1997) *Analysing Casual Conversation*. London: Cassell

Follard, D. and D. Robertson (1976) Towards objectivity in group oral testing. *ELT Journal* 30: 156-167

Fulcher, G. (1987) Tests of oral performance: the need for data-based criteria. *ELT Journal* 41/4: 287-291

Fulcher, G. (1993) The construction and validation of rating scales for oral tests in EFL. Unpublished PhD thesis: University of Lancaster

Fulcher, G. (1996) Testing tasks: issues in task design and the group oral. *Language Testing* 13/1: 22-51

Halliday, M.A.K. (1976) *System and Function in Language* (ed. G.R.Kress) London: Oxford University Press

Halliday, M.A.K. (1978) Language as Social Semiotic: The Social Interpretation of Language and Meaning. London: Edward Arnold

Halliday, M.A.K. (1994) *An Introduction to Functional Grammar*. London: Edward Arnold

Harley, B.; J. Cummins, M. Swain and P. Allen (1990) The nature of language proficiency. In B. Harley, P.Allen, J. Cummins and M. Swain (eds.) *The Development of Second Language Proficiency.* Cambridge: Cambridge University Press

Hashemi, L. (1994) *CAE Practice Tests 2: Teacher's Book.* Cambridge: Cambridge University Press

Heaton, J.B. (1988) *Writing Language Tests* (2nd edn.) London: Longman

Hughes, A. (1989) *Testing for Language Teachers*. Cambridge: Cambridge University Press

Hymes, D. (1972) On Communicative Competence. In J.B. Pride and J. Holmes (eds) *Sociolinguistics*. Harmondsworth: Penguin

Johnson, K. (1982) *Communicative Syllabus Design and Methodology*. Oxford: Pergamon Press

Lazaraton, A. (1992) The structural organisation of a language interview: a conversation analytic perspective. *System* 20/3: 373-386

Littlewood, W. (1992) *Teaching Oral Communication.* Oxford: Blackwell

Lombardo, L. (1984) Oral testing: getting a sample of real language. *English Teaching Forum* 22/1: 2-6

Long, M. and P. Porter (1985) Group work, interlanguage talk and second language acquisition. *TESOL Quarterly* 19: 207-228

McCarthy, M. and R. Carter (1995) What is spoken grammar and how should we teach it? *ELT Journal* 49/3: 207-218

Milanovic, M.; N Saville, A. Pollitt and A. Cook (1996) Developing rating scales for CASE: theoretical concerns and analyses. In A. Cumming and R, Berwick (eds.) *Validation in Language Testing.* Clevedon: Multilingual Matters

Morrow, K. (1979) Communicative language testing: revolution or evolution? In Brumfit,C.J. and K. Johnson (eds) (1979) *The Communicative Approach to Language Teaching.* London: Oxford University Press

Munby, J. (1978) *Communicative Syllabus Design*. Cambridge: Cambridge University Press

Nunan, D. (1991) *Language Teaching Methodology.* Hemel Hempstead: Prentice Hall

Oller, J.W. Jnr (1979) *Language Tests at School*. London: Longman

Oller, J.W.Jnr (1981) Language as intelligence? *Language Learning* 31: 465-492

Pollitt, A. and N.L. Murray (1996) What raters *really* pay attention to. In M.Milanovic and N. Saville (eds.) *Studies in Language Testing 3: Performance Testing, Cognition and Assessment.* Selected papers from the 15th Language Testing Research Colloquium (LTRC). Cambridge: Cambridge University Press

Poole, M. and T. Field (1976) A comparison of oral and written code elaboration. *Language and Speech* 19: 305-312

Richards, B. and F. Chambers (1992)Assessing open-ended oral tasks in foreign languages: an empirical study of reliability in the GCSE examination. University of Reading: Working Paper of the Department of Arts and Humanities in Education

Schleppegrell, M.J. (1992) Subordination and linguistic complexity. *Discourse Processes* 15: 117-131

Searle, J.R. (1969) *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press

Shohamy, E. and T. Reves (1985) Authentic language tests: where from and where to? *Language Testing* 2: 48-59

Shohamy, E., T. Reves and Y. Bejarano (1986) Introducing a new test of oral proficiency. *ELT Journal* 40/3: 212-220

Tarone, E. and G. Yule (1989) *Focus on the Language Learner*. Oxford: Oxford University Press

Thompson, G. (1996) *Introducing Functional Grammar*. London: Arnold

Underhill, N. (1987) *Testing Spoken Language.* Cambridge: Cambridge University Press

Van Lier, L. (1989) Reeling, writhing, drawling, stretching and fainting in coils: oral proficiency interviews as conversation *TESOL Quarterly* 23/3: 489-508

Weir, C. (1990) *Communicative Language Testing*. Hemel Hempstead: Prentice Hall

Weir, C. (1993) *Understanding and Developing Language Tests.* Hemel Hempstead: Prentice Hall

Wood, R. (1991) *Assessment and Testing: a Survey of Research*. Cambridge: Cambridge University Press

APPENDIX I

TASK 1

STUDENT A

You and your partner work at the same place in the centre of a small city. You have decided to rent accommodation together.

Below, you will find two suggestions (which you have already visited - be prepared to fill in some details from your imagination!)
Your partner also has two suggestions (which **you** have **not** visited).

Discuss the possibilities together and try to agree which properties you will consider and which you will not.

C17th cottage, large garden and orchard; living room, small kitchen, 2 small bedrooms,open fire place in every room, bathroom at back, roof recently restored but interior needs some attention. Partially furnished. Situated in small village 10 miles from city.
Railway station nearby. £450 pcm*.

Modern 2-bedroomed flat (1st floor); spacious living room; gas central heating; washing machine;
5 mins. from city centre.
Off-street parking.
No smokers.
£425 pcm*. References required.

* pcm = per calendar month

STUDENT B

You and your partner work at the same place in the centre of a small city. You have decided to rent accommodation together.

Below, you will find two suggestions (which you have already visited - be prepared to fill in some details from your imagination!)
Your partner also has two suggestions (which **you** have **not** visited).

Discuss the possibilities together and try to agree which properties you will consider and which you will not.

3 bedroomed terrace house; living room, dining room, kitchen, bathroom, small back yard. Fully furnished. Needs redecorating.

Situated in lively area of city with many shops and amenities nearby. Bus route to city centre (10 mins.)
£425 pcm*

*pcm = per calendar month

Luxury 2-bedroomed apartment (7th floor - superb views) in new riverside development. Fully equipped kitchen (including washing macine, dishwasher, microwave). One large double and one small single bedroom, living room with balcony. Partially furnished. Private garage. Satellite TV. 10 min. walk from city centre.
£700 pcm* No pets.
Suit young professionals.

APPENDIX II

TASK 2

TOPIC FOR DISCUSSION: **Homelessness**.

(The quotations below may help you in your discussion)

" Having a home is a basic human right. The government of a country has an obligation to provide housing for all its citizens"

"Homeless people have only themselves to blame. They should not expect society to provide solutions to their problems"

APPENDIX III

INSTRUCTIONS TO BE GIVEN BY INTERLOCUTOR

TASK 1

Imagine that you both work together in the centre of a small city.

You have decided to rent accommodation together.

Each of you has notes about two possibilities for renting.

Yesterday, each of you viewed the properties that you have notes about, but you have NOT visited your partner's possibilities.

I want you to discuss the four possibilities together, using the notes you've been given and filling in missing details and answering your partners questions from your imagination.

Try to agree which of the properties you will consider and which you won't.

[Check for comprehension of instructions]

OK, you have about four minutes for discussion. Remember, I will not be participating in the discussion.

When you're ready, you can begin.

[.....]

[Bring discussion to a close after 4 minutes if necessary]

TASK 2

I'd like you to discuss the topic of homelessness.
You can refer to the quotations given on your task sheet if you want to, but you don't have to - you can discuss the topic in any way you like.

[Check for comprehension of instructions]

Remember, I won't be joining in your discussion.

When you're ready, you can begin. [Allow a maximum of 4 minutes for discussion]

APPENDIX IV

TRANSCRIPTION KEY

(Adapted from Eggins, S. and D. Slade (1997) *Analysing Casual Conversation.* London: Cassell)

Punctuation

, Comma indicates parcelling of talk, natural pause/breathing time (silent beats in Halliday's (1985/94) system)

. Full stop marks a definite termination. (Its absence at the end of a turn marks lack of definite termination, e.g. trailing off, interruption, no final intonation)

? Interrogative with rising tone, uncertainty, rising inflection

! Surprised intonation, animated tone, surprise, shock, amazement etc.

" " Quotation marks show a change in voice quality in reported speech

Emphasis

CAPS Emphatic stress is shown by using capital letters. Also, a noticeable increase in volume compared to surrounding talk

Intervals

... Short hesitation within a turn (less than 1 second)

[2 secs] Timed pause (of more than 1 second)

- false start / restart

Overlaps / Contiguous utterances

= = Double equals sign shows: i) simultaneous utterances when two entire turns begin together e.g:

A: = = Yes I think so
B: = = I'm not sure

ii) overlapping utterances- e.g:

A: I really don't know = = do you?
B: = = I'm not sure

= Single equals sign marks contiguous utterances: there is no interval between adjacent utterances e.g.:

A: I don't really know=
B: =I'm not sure

Other transcription conventions used

( ) Empty parentheses show inaudible segments of talk

(said) Words in parentheses show transcriptioner's guess / doubt

[laughs] Non-verbal information is shown in square brackets

Analytical symbols used in transcriptions

01 Turn numbers are shown in numbers down the left-hand column
NV1 Non-verbal moves are shown in this way
(i) Clause divisions are shown in roman numerals