2 The use of tactics and strategies by Chinese students in the Listening component of IELTS

Authors

Richard Badger
University of Leeds

Xiaobiao Yan
Guangdong University of Foreign Studies (GDUFS)

Grant awarded Round 12, 2006

This study is a comparative analysis of the strategies used in an IELTS Listening Test by first language users of English and Chinese learners of English.

ABSTRACT

This study investigates whether there are differences between the strategies used by native speakers/ expert users of English and those used by learners of English who are native speakers of Chinese when they take an IELTS Listening Test.

24 native speakers of Chinese (twelve pre-undergraduate and twelve pre-postgraduate), at an IELTS level for the Listening paper of between 5.5 and 6.5 and 8 native/expert speakers of English (three undergraduates, three masters level and two doctoral), took a sample listening test (from McCarter and Ash 2003).

Data were collected using a think-aloud protocol and then analyzed using a framework based on Goh (2002) adapted to include particular features of the data sets based on a grounded approach (Glaser & Strauss 1967; Glaser 1992; Senior 2006). This produced a three level system of coding, with an initial distinction between cognitive and meta-cognitive strategies, each of which was divided into sub-strategies and then again into the tactics used to carry out the strategies.

The result of an independent samples 2-tailed t-test revealed there were no significant differences between the two groups in terms of strategy use. At the level of sub-strategy there were differences on two out of thirteen metacognitive strategies. At the level of tactics there were significant differences for seven tactics (two cognitive and five meta-cognitive) out of fifty eight at $p \leq 0.005$. This suggests that the strategies and tactics adopted by native and non-native speakers of English in the IELTS Listening Module are not significantly different.

We also examined the differences between the twelve pre-undergraduate and twelve pre-postgraduate Chinese native participants but found no significant differences at strategy, sub-strategy or tactical levels.

The paper then discusses possible reasons for the results.

AUTHOR BIODATA

RICHARD BADGER

Richard Badger is a senior lecturer in the School of Education at the University of Leeds, Leeds, UK. He co-ordinates the MA TESOL programme and teaches modules in Teaching and Learning in TESOL, Investigating Language for TESOL and Learning and Teaching Vocabulary. His research interests include the teaching of academic writing, argument in academic contexts and academic listening. He has published in ELT Journal, the Journal of Second Language Writing, the Journal of Pragmatics, System and ESP Journal. He is currently working on a project investigating how undergraduates learn from biology lectures, focussing on the roles that PowerPoint play in this learning and on how teachers of ESP deal with topics where they lack disciplinary expertise.

XIAOBIAO YAN

Xiaobiao Yan is a lecturer in the College of Continuing Education at Guangdong University of Foreign Studies in Guangzhou, Guangdong, China. He has been engaged in IELTS teaching and research for several years and at present he is the coordinator for the IELTS Preparation. His research interests are language testing, particularly for listening and writing, and SLA. He has published in the Journal of Guangdong University of Foreign Studies, the Journal of Leshan Normal University, and a monograph entitled 'A Probe into Continuing Education Programme'. He is the author of 'Modern Business Writing', published by Zhongshan University Press. He is currently working on a university-funded project on the exploration and analysis of IELTS washback to IELTS teaching in speaking course.

IELTS
English for International Opportunity

The test that sets the standard

IELTS RESEARCH REPORTS
VOLUME 9, 2009

CONTENTS

1   INTRODUCTION

The IELTS Test is a high stakes test and relative success or failure can have a life changing impact on candidates. The language use which the test attempts to measure is associated very closely with cultural patterns. Many commentators argue that the Confucian background of native speakers of Chinese (Gieve and Clark 2005; Scollon 1999; Yao 2000) is significantly different from the cultural background most common in Australia, Canada and the UK. It is important therefore that we have confidence that the IELTS Test is proving an appropriate measure of the language ability of Chinese speaking students. A related question concerns the level of education of candidates for IELTS and whether the intellectual development typically associated with the completion of a degree may have an impact on the way in which those preparing for undergraduate and graduate study take the IELTS examination. This study is an attempt to address these issues.

The focus of this research is on listening, a key skill in language use, but much harder to test and research than speaking and writing because, like reading, most of the processes involved in listening happen within the minds of language users. Testing these skills requires the creation of a construct to understand what happens when language users read or listen and the adoption of an indirect means of assessment for these skills. Even compared with reading, listening presents additional difficulties to the test writer and researcher because it is "transient and occurs within limited capacity working memory" (Goh 2002, p 182).

IELTS is a test of communicative language use and, within the tradition of communicative language testing, the aim has generally been to evaluate whether candidates have the ability to communicate in the target-language use (TLU) domains (Bachman & Palmer 1996, p 18), that is "the real world situation in which the language will be used" (Buck 2001, p 83). Many commentators use the term 'task' to describe the activities that are carried out by language users outside the test situation. Bachman and Palmer define a target language use domain as "a set of specific language use tasks that the test taker is likely to encounter outside of the test itself" (Bachman and Palmer 1996, p 44). This notion means that one of the aims of test writers is to produce test tasks that are as similar as possible to TLU domain tasks. However, as Buck (2001, p 90) observes, "test tasks can never be entirely authentic replications of target language use tasks". For further discussion of the concept of 'authenticity', see Widdowson (2003).

Ellis (2003) addresses the impossibility of designing completely authentic test tasks by distinguishing between situational authenticity and interactional authenticity which may be taken as very similar to text and task authenticity (Guariento & Morley 2001; Skehan 1996). Situational authenticity is the extent to which the test task matches a real life situation. It would provide a rationale, for example, for including a listening text related to the task of filling in a form where filling in forms was part of the TLU domain. Interactional authenticity reflects the extent to which the test task elicits language behaviour which "corresponds to the kind of communicative behaviour that arises from performing real-world tasks" (Ellis 2003, p 6). For the form filling task, this would be the way in which users would use the listening text in completing the form.

1.1 Situational authenticity

An examination of Listening Test tasks in the IELTS shows that there is a plausible claim that they have some situational authenticity. For example, the test sample in *IELTS Testbuilder* (McCarter and Ash 2003), the commercial IELTS test practice book that we used in this research and which mirrors IELTS papers closely, included the following listening texts:

- A two person conversation on the phone between a credit card holder and a call centre employee

- A radio show in which a speaker discusses his success in giving up smoking with the radio presenter

- A conversation between a tutor and two undergraduate students about what one of their course mates is doing and the marks of the two undergraduate students.

- An extract from an academic lecture on bullying in the workplace.

All of these could be seen as coming from the TLU domains that candidates who are going to study in Higher Education Institutions in English speaking countries might encounter. There are some issues such as the intonation in the tutorial and the possibly inauthentic North American accent in the final text but it would be possible for test writers to use such texts as the basis for tasks with situational authenticity.

1.2 Interactional authenticity

Interactional authenticity is more problematic. The students have to complete a range of written multiple choice questions and gap filling exercises, neither of which are activities which would be carried out in relation to these kinds of listening texts outside an examination or language classroom and so do not have obvious interactional authenticity. However, it is possible to identify sufficiently strong links between non-examination and examination interactions to ground the validity of the examination. For example in the first section, candidates have to note down the post code (question 2) having heard the following extract.

O: And what's your post code?

C: <u>SE1 8PB</u>

O: <u>SE1 8PB</u>

C: That's it. [our underlining]

Similarly, in question 34 of section 4 the candidates have to complete with not more than three words the gap in the following phrase

Setting 34. _____ tasks.

The cue for this is:

The first item on the list: <u>giving people</u> tasks that managers themselves cannot do and which are therefore impossible to achieve. [our underlining]

This would seem to be fairly closely related to the task of taking notes in a pre-PowerPoint lecture and so to have interactional authenticity.

There are however several questions where the interactional authenticity is harder to justify. For example, in task two, which replicates an interview on the radio, candidates have to answer the following multiple choice question:

11 Mr Gold had problems because he

a   hated smoking

b   smoked

c   couldn't touch his toes

d   was very lazy.

The relevant extract from the tape script is:

Well I enrolled on a number of evening courses where I found I wasn't able to do the warm up sessions. Bending down to touch my toes made me breathless. Even though I hated to admit it my <u>problem was not so much my sitting around all the time but my fifteen to twenty a day smoking habit</u>. If I'd been able to limit myself to three or four cigarettes a day there would have been no problem but I was seriously addicted. And I'm talking about waking up at three a.m. and dying for a cigarette or in the days before twenty four hours shopping <u>driving across London to buy a packet of cigarettes when I ran out</u>. But above all my addiction meant making sure I never ran out at the expense of everything else including necessities. [our underlining]

It is quite difficult to see, first, what the interactionally authentic task would be for a radio interview, and, secondly, how the multiple choice format would relate to such a task. Similar issues arise with the tutorial situation, where again it is not immediately obvious what the interactional task should be.

The weakness of arguments based on interactional or task authenticity mean that claims about the ability of the IELTS Test to whether candidates can handle TLU tasks need support from elsewhere. In this paper, we explore the possibility that this may be found in the similarity of the behaviour of candidates taking IELTS to that of a group of people whose ability to handle the TLU can be assumed, that is native and expert users of English, and in particular we attempt to answer the following research questions:

■ What are the similarities and differences in the mental processes of native speakers of English and native speakers of Chinese when taking the IELTS Listening Test?

■ To what extent do the mental processes of Chinese speaking candidates preparing for undergraduate and postgraduate studies differ?

2   BACKGROUND TO THE RESEARCH

In the background literature section, we look at models of listening, the concept of strategies and talk-aloud protocols.

2.1 Models of listening

Researchers such as Anderson and Lynch (19880, Buck (2001), Rost (2002) and White (1998) have offered a range of models of listening. Here we discuss firstly top-down, bottom-up and interactive models and then Anderson's (2000) perception, parsing and interpretation model. Flowerdew and Miller (2005, p 85 ff) make a strong argument for saying that a model of listening should include a social element. However for the purposes of this piece of research and, in particular, the focus on listening within the socially constrained context of an examination, we have chosen to focus on psychological aspects of the listening process.

2.1.1 Top-down, bottom-up and interactive

A distinction is commonly made between top-down and bottom-up processes in listening. This is based on the view that there is a continuum of information that is needed for effective listening from phonetic and phonemic information at the bottom to schematic and world knowledge at the top.

Listening comprehension is the result of an interaction between a number of information sources, which include the acoustic input, different types of linguistic knowledge, details of the context, and general world knowledge and so forth (Buck 2001, p 3).

We regard this as an understatement of the degree of interaction required. Both top-down and bottom-up information require the interaction of listening text and the listener. To decode a series of sounds as being instances of particular phonemes, listeners need to have the raw data, that is, the listening text, but also need to bring to that data their knowledge of what counts as a phoneme in the language to which they are listening. The information that a particular sound represents, for example, /s/ in English, is not necessarily in the acoustic signal but in the acoustic signal as interpreted by listeners with the knowledge of what phones make up the /s/ phoneme in English.

Similarly, the relevant schemata that help listeners make sense of particular listening texts serve no purpose if they are simply stored in listeners' minds. The schemata need to be activated by the listening text. This is not to say that bottom and top information do not exist but that interaction is both between top and bottom information and between listener and listening text.

2.1.2 Perception, parsing, utilization

Anderson (2000) argues for a three stage view of comprehension: perception, parsing and utilization. When applied to listening, this means that listeners first store the input as a sound string (Anderson 2000, p 388). They then parse the sounds into the combined meaning of the words (Nagle and Sanders 1986). The third stage is when the listeners use the mental representation of the message. This may be simply a question of storing the meaning in memory or listeners may combine it with other elements in memory or context to make inferences.

While listening, listeners are not just involved in one of these stages.

> These three stages - perception, parsing and utilization - are by necessity partly ordered in time; however, they also partly overlap. Listeners can be making inferences from the first part of a sentence while they are already perceiving a later part. (Anderson 2000, p 388)

This also means that ambiguities at the perception stage may be resolved or rendered unimportant by information at the parsing or utilization stages.

If listeners are able to carry out the three processes of perception, parsing and interpretation without any difficulty, listening should be a straightforward process. However, listening is often not straightforward and most language users experience problems with comprehension. To gain an insight into the difficulties that listeners, and in particular L2 listeners, face, we need a model of how people learn to carry out skills such as listening.

2.1.3 Learning to listen

Information processing models of learning see the development of skills as having at least three stages. The first is the cognitive stage during which learners acquire knowledge about listening, sometimes called declarative knowledge. This would include, for example, information about the grammatical structure of the target language.

Secondly, at the associative or controlled stage, declarative knowledge is gradually proceduralized (Anderson 2000, p 282). For example, knowledge about grammatical structure becomes an ability to parse a listening text. At this stage, listening is a demanding activity.

> Learning of a skill initially demands learners' attention and thus involves controlled processing . . .
> Controlled processing requires considerable mental "space" or attentional effort (Saville-Troike 2006, p 73)

In the final stage, which Anderson terms autonomous (2000, p 282), listeners carry out the listening in a more and more automatic fashion.

Learners go from controlled to automatic processing with practice. Automatic processing requires less mental "space" and attentional effort (Saville-Troike 2006, p 73).

In this model, learning essentially involves development along a continuum from controlled to automatic use of the skills and sub-skills involved in listening, freeing learners' controlled capacity for new information and higher-order skills.

We draw the implication from this that controlled processes are more likely to be conscious, and thus we interpret the term 'automatic' as meaning that the processes at this stage are not under conscious control.

If this model is correct, people who are learning to listen in a second language are at least partially at the controlled stage and so have limited capacity for perceiving, parsing or interpreting the listening texts to which they are exposed. In a test situation, such people need to come up with some way of dealing with the problems they face. These solutions are often labelled "strategies" (Bialystok 1990; O'Malley & Chamot 1990; Oxford 1990).

2.2 Strategies and tactics

Strategies are frequently defined within a learning context. Oxford (1990, p 8) defines strategies as "specific actions taken by learners to make learning easier, faster, more enjoyable, more self-directed, more effective, and more transferable to new situations". Goh (2002, p 186), takes a broader view, saying strategies are "mental steps or operations carried out to accomplish cognitive tasks such as map-reading, memorization, processing information and problem solving."

While there is extensive discussion of strategies in the literature on learning (eg, O'Malley & Chamot 1990; Oxford 1990), here we are concerned with the processes that listeners go through in order to understand a listening text, and whether or not these lead to learning. Our concern is primarily with communication strategies but our understanding is informed by what people have written of learning strategies.

Although some writers suggest that strategies can be conscious or unconscious, for most authorities strategies are conscious steps taken by language users and this is coherent with the view of strategies being adopted to compensate for the fact that some part of the listening process has not become completely automatic. This is consistent with the research instrument we are using, think-aloud protocols, which assume that listeners can talk about the strategies they are using.

Goh (1998; 2002) makes a distinction between general and specific strategies. She describes tactics as 'individualized techniques through which a general strategy is operationalized' (Goh 2002, p 187). For example, a meta-cognitive sub-strategy such as directed attention can be operationalized through tactics, such as concentrating hard and identifying a failure in concentration.

2.3 A taxonomy for strategies and tactics

There is considerable disagreement about the best taxonomy for describing strategies and tactics in listening. For this study, we drew on Goh's (2002) taxonomy (see Appendix 3). This follows Purpura (1999) in identifying two broad strategies, cognitive and meta-cognitive, with cognitive strategies broadly covering the perception, parsing and interpreting process of listening, and metacognitive strategies covering problem solving activities. These two broad strategies were divided into sub-strategies which were partly drawn from the literature and partly derived from Goh's data in line with a grounded theory approach to data analysis (eg, Brown & Rodgers 2002; Glaser & Strauss 1967; Glaser 1992; Senior, 2006). One of the most significant differences between our research and that of Goh is that ours related to an examination paper, and this raised the question of the extent to which the strategies, sub-strategies and tactics used in an examination would be found to differ from a non-examination context.

Goh identified eight cognitive and six metacognitive strategies. Each sub-strategy was realized in a set of tactics. For example, within the cognitive strategy, she identified a sub-strategy labelled fixation which could be realized by the following four tactics:

- stop to think about the spelling of unfamiliar words,
- stop to think about the meaning of words,
- memorize/repeat the sounds of unfamiliar words,
- memorize words or phrases for later processing.

Again metacognitively, she labelled one sub-strategy, directed attention, which was realized through two tactics:

- concentrate hard,
- continue to listen in spite of difficulty.

A complete list can be found in Appendix 3.

2.4 Think-aloud protocols

It is common to investigate strategies using questionnaires. Oxford's (1990) development of an inventory of learning strategies has produced a range of questionnaire-based studies (eg, Phakiti 2003; Vanijdee 2003). However, we felt that this would not be appropriate with the kinds of learners we were investigating, particularly given the fact that we were not sure how accurately a questionnaire would capture strategy and tactic use. Instead, we drew on the research instrument of the think-aloud protocol (Brown & Rodgers 2002).

A verbal protocol is the data which is produced when a person 'is asked to either "talk aloud" or to "think aloud"' (Green 1998, p 1). It is made up of utterances made by an individual, either while or after the individual carries out a single task or a series of tasks; verbal protocols, thus, can be either concurrent or retrospective (Brown & Rodgers 2002). For listening the technical problems that arise in recording what listeners are saying at the same time as they listen to a text and the difficulty that listeners have in talking aloud while trying to comprehend a text meant that we had to adopt a retrospective approach. However, the nearer the protocol is to the event that the listeners are talking about the greater the validity and so we divided the IELTS Listening Test into sections at natural breaking points, and asked the listeners to think aloud about what they had just done.

Goh (2002, p 189) comments:

> Verbal data on listening processes are predominantly retrospective. Because of the rapid flow of information, the working memory has to be freed for processing continuous input. What listeners will typically do is to process the heeded input first before reporting through retrospective verbalization.

Bearing in mind Anderson's (2000) model of learning above, we hypothesised that native speakers/expert speakers of English would report fewer cognitive strategies than learners of English because they would have been automatized and so no longer accessible to the think-aloud protocol.

2.5 Research questions

Having reviewed the literature we were in a position to pose more specific research questions

1 What differences are there between native speakers of English and non-native speakers of English in terms of the strategies, sub-strategies and tactics they use when taking an IELTS Listening Test?

2 What differences are there between Chinese speaking candidates preparing for undergraduate and graduate studies in terms of the strategies, sub-strategies and tactics they use when taking a Listening Test?

3 To what extent are the strategies, sub-strategies and tactics used by native and non-native speakers of English in an IELTS Listening Test different from those reported in Goh's studies of listening?

3 THE STUDY

The study was carried out in Guang Dong University of Foreign Studies (GDUFS), Guang Zhou, China and the University of Leeds (UOL), Leeds, UK.

3.1 The participants

We collected data from twenty four volunteers on an IELTS preparation programme at GDUFS who had or were expected to obtain a score of between 5.5 and 6.5 on the Listening element. These bands were chosen because they are significant in deciding whether candidates are admitted to English medium tertiary education.

Twelve of the students were preparing for undergraduate studies through the medium of English (4 males and 8 females) and twelve were preparing for postgraduate studies (4 males and 8 females). We collected information about the participants' disciplinary background. Eight different majors and four different majors were expected to study for pre-postgraduate and pre-undergraduate groups respectively. Subjects' previous IELTS scores were collected at the same time. Information on the subjects is presented in Tables 1 to 4.

Accounting	5
Human Resources	1
Fashion Design	1
Tourism Management	1
Hotel Management	1
Management for Information System	1
Culture and Translation	1
Finance	1

Table 1: Subjects of pre-postgraduate study participants at GDUFS

IELTS Band scores	Number of students
5.5	3 (one score predicted by the teacher)
6.0	4
6.5	5

Table 2: IELTS scores of pre-postgraduate study participants at GDUFS

International Relationship and English	1
International Trade and English	1
International Business	6
Accounting	4

Table 3: Subject of pre-undergraduate study participants at GDUFS

IELTS Band scores	Number of students
5.5	3
6.0	4
6.5	5

Table 4: IELTS scores of pre-undergraduate study participants at GDUFS

We had hoped to investigate the impact of disciplinary background and gender but the numbers of students from particular disciplines and the relatively small overall sample meant that this was not practicable. The fact that the levels of the students as measured by IELTS were comparable between the pre-postgraduate and pre-undergraduate course meant that we were able to explore the impact of educational level on strategies, sub-strategies and tactics.

In addition, we collected data from eight self selecting participants with native levels of competence in Leeds (three undergraduates, three master's level and two doctoral). One of the doctoral students was not a native speaker of English but had a native-like command of the language. She had lived in the UK for over two years and prior to arrival had obtained a score of 8 on the IELTS Listening Test.

3.2 Ethical issues
The participants were all volunteers and saw and signed the consent forms, the English version of which appears in Appendix 5. The institutions in which the research was carried out are identified in this paper. This meant that if we linked information about gender, level of study or discipline to a particular think-aloud protocol, it would be possible to identify particular participants and so we decided not to include this information, where it was linked to what participants said or did, to ensure anonymity as far as we could.

3.3 Data collection
The data were collected from participants individually. We first gave the participants training tasks to accustom them to producing a protocol. These involved two mental arithmetic calculations and two anagram puzzles. The participants then took the attached test and completed a blank version of the answer sheet. We had asked the assistant director in Cambridge ESOL's Research and Validation Unit for permission to use an IELTS past paper in listening for this project but unfortunately this was not possible. Drawing on criteria proposed by Terry (2003, pp 66-76) and Saville and Hawkey (2004, pp 73-96), the sample test (McCarter & Ash 2003) was judged to be fairly close to an actual IELTS Test. It was also appropriate because of the test paper's unfamiliarity for the research participants.

At naturally occurring stages in the test (e.g. between sections, between reading the questions and listening to the recording) we asked the participants to say what mental processes they had gone through in arriving at or failing to arrive at answers. The researchers limited their contribution once the participants had started doing the tests to the following utterances:

- Keep talking
- Comment on what you have just heard or read / question XX, section XX

If participants said they had nothing to say about a particular section we asked them once to comment and, if they did not say anything at that stage, we continued to the next section. In the transcription for data analysis we removed all utterances from the researchers for ease of coding.

GDUFS participants were able to respond in English or Chinese. The think-aloud protocols were recorded on a mini-disk recorder or else directly on to a laptop computer by Xiaobiao Yan in GDUFS and Richard Badger in Leeds.

The recordings were transcribed and, if the think-aloud had been carried out in Chinese, translated into English. A sample non-native speaker protocol is provided in Appendix 1, and a sample native speaker protocol appears in Appendix 2.

3.4 Data analysis
The data were first chunked into what appeared to be plausible units that corresponded to Goh's tactics. The following extract from one GDUFS participant's protocol, was divided into two chunks.

> A and C is much...., um, A is certainly not the answer, so I just choose between B and C (C-I). He said he is free in, in, um....I am not quite sure about this question, because in the last section, the woman said, she will call. I don't remember what she said. She will call the man very soon (M-CM).

In the first chunk (ending C-I), the participant was trying to process utterances directly in order to infer the answer, which we treat as a cognitive strategy. In the second chunk (ending M-CM), comprehension monitoring tactics were used to check, and confirm understanding during listening. We classified this as metacognitive.

Initially we separately chunked data from two participants, discussed differences and then coded a further data set from another participant. Our chunking on the third data set agreed in over 95% of cases. We did not compare chunking on later data sets but did check each other's view on problematic instances.

3.4.1 Revising Goh's taxonomy

The data were analysed using Goh's categories (see Appendix 3). However, we had to make some changes at the level of sub-strategy and tactic. Our final taxonomy is given in Appendix 4.

We reorganised Goh's strategies so that the cognitive sub-strategies corresponded to Anderson's stages of perception (fixation), parsing (reconstruction) and utilisation (inferencing). We also treated the tactics that Goh classified as realising the cognitive strategy of prediction as a realisation of the meta-cognitive strategy of pre-listening preparation. Further changes were made to render the taxonomy more consistent with our understanding model of listening. For the sub-strategy of fixation, Goh identified four tactics:

- stop to think about the spelling of unfamiliar words,
- stop to think about the meaning of words,
- memorize/repeat the sounds of unfamiliar words,
- memorize words or phrases for later processing.

In our taxonomy, we added a further tactic to cover the situation where a listener focused on the sound of a phoneme (CFP in our taxonomy in appendix four), on the assumption that listeners would focus on the sounds in unknown words. This came up several times in our data for both UOL and GDUFS participants, not to do with individual phonemes, but related to the sounds of letters in a post code

> The postcode, I suppose that's 8PB (UOL participant).

The nearest our participants came to commenting on the processing of phonemes was in the following data.

> I just heard the pronunciation, but...Wahace. I don't know what word it is, may[be] it's a new word for me? Um 'Wahace' [Wales] (GDUFS participant).

This was treated as a fixation on a word rather than a phoneme (CFW- see appendix four). Generally, both groups of listeners had automatized their perception of individual sounds to the extent that they were no longer able to report on them.

These changes related to our views of the listening process. Most of the other changes related to the fact that we were working in an examination context.

We eliminated the sub-strategy of elaboration because it did not appear in our first three data sets and we did not require it in the remaining data sets, presumably because elaboration is not a common tactic in examinations. The sub-strategy of visualisation also did not appear in these three data sets although we had thought that learners might use visualisation in the examination.

We also eliminated the sub-strategy of prediction because it overlapped with the tactics under the sub-strategy of Inferring answer. For instance, the tactic 'anticipating details while listening' under the sub-strategy of Prediction seemed very similar to 'using co-text' from the sub-strategy of inferring.

At the level of tactic, we made several changes which related to the fact that our participants were taking an examination. So for example, under the sub-strategy of 'reconstruction', we added the tactic of 'reconstructing meaning from an examination question' and under the sub-strategy of 'inferring' added 'inferring the answer by using information from the text with the examination question paper'. These are discussed in more detail below where we address our third research question which relates to differences between the ways people in Goh's study listened as compared to those in an IELTS Test.

Our taxonomy uses letter codes such as CRQ and CIQ to describe strategies, sub-strategies and tactics. The C in CRQ stands for 'cognitive', the R for 'reconstruction' and the Q for 'examination question'. Similarly in CIQ, the C stands for 'cognitive', the I for 'inferring' and the Q for 'examination question'.

The changes in the metacognitive group were rather greater. First, we introduced the new sub-strategy of real time assessment of output (MAO, where M stands for 'meta-cognitive' and AO for 'assessment of output') because participants referred quite extensively to tactics such as making sure their answers had the right numbers of words.

We also made eleven changes at the tactical level, particularly realisations of comprehension monitoring (while listening) and comprehension evaluation (post listening).

3.4.2 Applying the new taxonomy

We jointly coded two data sets and discussed differences until we had reached agreement. We then coded a third data set independently and our coding agreed over 90% of the time.

3.5 Findings

In this section we address each of our research questions in turn.

3.5.1 Research Question 1

What differences are there between native speakers of English and non-native speakers of English in terms of the strategies, sub-strategies and tactics they use when taking an IELTS Listening Test?

At the level of strategy, the UOL participants reported an average of just over one hundred instances of strategy use compared to just below eighty for the GDUFS participants. This was almost all accounted for by differences in terms of cognitive strategies where the figures were just under thirty for the UOL participants and just over twenty for the GDUFS participants. We were surprised that the UOL participants were able to report this number of cognitive strategies However, at the level of strategy, the differences were not significant at $p \leq 0.005$ (see Table 5).

At the level of sub-strategy, the differences between the groups were again largely not significant. However, there were significant differences at $p \leq 0.005$ for two metacognitive strategies, directed attention (i.e. monitoring attention and avoiding distraction) and comprehension monitoring (i.e. checking interpretation for accuracy while listening), as shown in Tables 6 and 7.

	NSS	N	Mean	Std. Deviation	Std. Error Mean
C	NESE	8	49.13	33.753	11.934
	NC	24	20.21	10.384	2.120
M	NESE	8	56.00	21.824	7.716
	NC	24	59.17	20.459	4.176
T	NESE	8	105.13	32.520	11.498
	NC	24	79.45	26.493	5.408

C=Cognitive   M=metacognitive   T=total
NSS=Native speaker status   NESE=Native/Expert speaker of English   NC= Native speaker of Chinese

Table 5: Descriptive Statistics at the level of strategy

In both case the GDUFS participants used these strategies more frequently than the UOL participants. The UOL participants were probably less likely to need to calm themselves down or perhaps they did not engage in as much comprehension monitoring after listening to the listening text given their reduced commitment to scoring well on the test. It was surprising that the number of reports of the assessment of output metacognitive sub-strategy was not significantly different between the two groups, perhaps indicating that the Leeds participants were less familiar with the IELTS question types and were likely to spend more time on the process of listening in order to answer the answers than expected.

Sub-strategy	NSS	N	Mean	Std. Deviation	Std. Error Mean
Metacognitive: directed attention	NESE	8	0.75	1.39	0.49
	NC	24	4.13	2.42	0.49
Metacognitive: comprehension monitoring	NESE	8	6.88	4.58	1.62
	NC	24	20.63	6.16	1.26

NSS=Native speaker status   NSS= Native speaker status
NESE=Native/Expert speaker of English   NC= Native speaker of Chinese

Table 6: Descriptive statistics for significantly different sub-strategies

Sub-strategy	t	df	Sig	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						Lower	Upper
Metacognitive: directed attention	-3.720	30	.001	-3.375	.907	-5.228	-1.522
Metacognitive: comprehension monitoring	-5.780	30	.000	-13.750	2.379	-18.608	-8.892

Table 7: Independent Samples 2-tailed t-test for significantly different sub-strategies

At the level of tactics, there are significant differences at p≤0.005 for two cognitive tactics ('fixation on spelling', 'inferring information using world knowledge') and five metacognitive tactics ('identifying a failure in concentration', 'identifying a problem with the amount of input', 'identifying a problem with the process of answering a question', 'confirming that comprehension has taken place', 'identifying partial understanding'), as shown in Tables 8 and 9. We discuss each of these briefly.

'Fixation on spelling' (CFSP) was not reported at all by the UOL participants but this tactic does seem to be reported by several of the GDUFS participants (1.58) as a way of fixing, or not, what they have heard

I knew it was 'Wales', but I did not know how to spell it (GDUFS participant).

Tactic	Native speaker status	N	Mean	SD	SEM
Cognitive: fixation – spelling	NESE	8	0.00	0.00	0.00
	NC	24	1.58	1.79	0.37
Cognitive: Inferring answer – using world knowledge	NESE	8	5.63	3.96	1.40
	NC	24	0.71	1.12	0.23
Metacognitive: directed attention – failure of attention	NESE	8	0.38	0.74	0.26
	NC	24	3.38	2.06	0.42
Metacognitive: real-time assessment of input- problem with the amount	NESE	8	6.88	2.90	1.02
	NC	24	2.46	2.36	0.48
Metacognitive: real time assessment of output – process	NESE	8	6.25	4.27	1.51
	NC	24	2.50	2.41	0.49
Metacognitive: comprehension monitoring – confirm comprehension has taken place	NESE	8	0.88	2.48	0.88
	NC	24	7.25	3.63	0.74
Metacognitive: comprehension monitoring – partial understanding	NESE	8	0.63	0.74	0.26
	NC	24	3.13	1.77	0.37

NESE=Native/Expert speaker of English   NC= Native speaker of Chinese.
SD=Standard Deviation   SEM= Std. Error Mean

Table 8: Descriptive Statistics for significantly different tactics

Sub-strategy	t	df	Sig	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						Lower	Upper
Cognitive: fixation – spelling*	-4.329	23	.000	-1.583	0.37	-2.340	-.827
Cognitive: Inferring answer – using world knowledge	5.598	30	.000	4.917	0.88	3.123	6.710
Metacognitive: directed attention – failure of attention	-3.995	30	.000	-3.000	0.75	-4.534	-1.466
Metacognitive: real-time assessment of input – problem with the amount	4.335	30	.000	4.417	1.02	2.336	6.497
Metacognitive: real time assessment of output – process	3.111	30	.004	3.750	1.21	1.288	6.212
Metacognitive: comprehension monitoring – confirm comprehension has taken place	-4.602	30	.000	-6.375	1.39	-9.204	-3.546
Metacognitive: comprehension monitoring – partial understanding	-3.861	29	.001	-2.505	0.65	-3.833	-1.178

*=Equal variance not assumed (Levene's test for equality of variance).
CI=Confidence Interval. See Appendix 4 for an explanation of the tactic acronyms.

Table 9: Independent Samples 2-tailed t-test for significantly different tactics

Inferring information using world knowledge (CIW) was, rather surprisingly, used more by the UOL participants (5.63) than by GDUFS participants (0.71)

> *You actually have to use your own knowledge to think of the best answer, so it's different and strange in one set of questions, but I suppose that might be the object of it* (UOL participant).

Amongst metacognitive strategies, 'identification of a failure in concentration' (MDAF) was reported more by the GDUFS participants (3.38) than the UOL participants (0.38). Again this is probably related to the fact that the UOL participants were less concerned about their performance on the test.

> *I was absentminded at that time* (GDUFS participant).

'Identifying a problem with the amount of input' (MAIA) was rather surprisingly reported more by the UOL participants (6.88) than the GDUFS participants (2.46), perhaps because of the unfamiliarity with the exam format.

> *So I miss, I miss a lot of the blanks. Yeah. Yes, because I have to read and listen at the same time* (GDUFS participant).

> *I mean because those two are quite close together at least that's what I thought, I thought those two [questions] were answered quite quickly* (UOL participant).

This last comment reflected a common assumption among both UOL and GDUFS participants that the information needed for questions would be distributed relatively equally throughout the listening text.

'Assessment of output related to the process of answering a question' (MAOP) was reported an average of 6.25 times by UOL as opposed to 2.50 for GDUFS participants. The following comment from a UOL participant related to where the numbers appeared on the answer paper.

> *I mean I suppose in order to be able to fill it out in an official way you need some indication of where you have to write especially there, if someone wasn't confident about their own writing abilities in English it could make it difficult, could be confusing. It seems a bit needless because all the others are at the end of the sentence apart from that one* [question 15].

Differences in the amount of experience of IELTS style examination paper resulted in the UOL group commenting more on the layout of paper or question than their more practiced GDUFS colleagues.

> *I wanted to write the first of July, but that's four words* (UOL participant).

> *Oh, I think in this [section], um, gap-filling, I think it is very difficult* (GDUFS participant).

'Confirming that comprehension has taken place' (MCMC) was reported 0.88 times by UOL and 7.25 by GDUFS participants. Again, this is likely to reflect both the higher confidence of the UOL participants about their ability to answer questions and the lack of a felt need to check what they had done.

> *The interest question was fairly straightforward* (UOL participant).

> *And the name, and the first name, he said that slowly, so I can hear very . . . very clear* (GDUFS participant).

For the tactic of 'identifying partial understanding' (MCMP), the UOL figure was 0.63 as against 3.13 for the GDUFS participants. This is in line with the view that GDUFS participants were less likely to feel they had completely understood what they had heard.

> *I didn't quite remember clearly, only that the man grunted that when he was handing in fees in the bank, he had given some extra money* (GDUFS participant).

The data from the native/expert users was related to more than one question as in the example below related to the final part of the test.

> *Again quite a lot of, quite difficult I thought. I didn't get it all* (UOL participant).

While the differences between the groups in tactics usage, where these are significant, do raise some interesting issues, such as why inferring information using general world knowledge was not more widely used by the GDUFS participants, most of the differences are easier to account for in terms of attitudes to the examination rather than an issue with the validity of the IELTS examination.

Generally, there do not seem to be any significant differences between native speakers of English and non-native speakers of English in terms of the strategies, sub-strategies and tactics they use when taking an IELTS Listening Test.

3.5.2 Research Question 2

> What differences are there between Chinese speaking candidates preparing for undergraduate and graduate studies in terms of the strategies, sub-strategies and tactics they use when taking an IELTS Listening Test?

The pre-undergraduate students reported over one hundred and sixty strategies compared to just under one hundred and forty for pre-postgraduates with most of this difference accounted for by meta-cognitive strategies where the figures were about one hundred and twenty as against about one hundred respectively. However, the analysis of the protocols in terms of strategies, sub-strategies and tactics indicates that the difference between the means for undergraduate and postgraduate students were not significant.

3.5.3 Research Question 3

> To what extent are the strategies, sub-strategies and tactics used by native and non-native speakers of English in an IELTS Listening Test different from those reported in Goh's studies of Listening?

We address this question using the data from the differences between to Goh's taxonomy (Appendix 3) and the taxonomy we used on our data sets (Appendix 4). The process by which we altered Goh's taxonomy is described above in section 3.4.1. As noted there, some of the changes relate to differences in our conception of listening rather than the IELTS context and so are not relevant here.

A second group of changes concerns tactics which are typical of examinations rather than listening beyond the exam hall but which would be extremely difficult to eliminate. The relevant tactics are listed below:

- Comprehension monitoring: confirm that an exam question has been answered (MCMQA)
- Comprehension monitoring: identify examination questions not answered (MCMQN)
- Comprehension monitoring: Identify examinations skills not applied (MCMS)
- Comprehension evaluation against examination questions (MCEQ)
- Comprehension evaluation against experience of examinations (MCEP)

A third group of changes related to the ways students used skills other than listening in the examination. Four relate to reading:

- Reconstruct meaning from examination question (CRQ)
- Inferring information from the listening text and exam question paper (CIQ)
- Prepare using exam paper questions (MPQ)
- Pay selective attention to exam questions (MSAQ)
- Assess input in terms of links between elements in listening text and examination questions (MAIQ)

Three relate to writing:

- Real time assessment of output in terms of quantity required (e.g. one or two words) (MAOQ)
- Real time assessment of output in terms of process required (e.g. multiple choice vs., gap fill) (MAOP)
- Real time assessment of output in terms of intermediate processes (e.g. note taking) (MAOI)

It would be hard to design a listening examination which did not involve the use of other skills but it might be worth considering whether some of the reading could be replaced by further listening.

Finally at the level of sub-strategy we eliminated Goh's strategy of elaboration and, while we kept in the sub-strategy of visualisation, we found no instances of this in our data sets. The lack of elaboration reflects the fact that, unlike many other kinds of listening, exam listening rarely requires the listener to use the information obtained from a listening text in some other communicative activity. It is hard to see how this might be done if the focus is to remain on listening though a more holistic view of language use might permit this.

The absence of visualisation again seems to relate to the largely verbal nature of the examination paper. This may well be appropriate in a text which replicates a phone conversation, as in the first section on the examination paper we used, but seems less appropriate with the academic lecture in the final section. Academic lectures are increasingly multimodal (O'Halloran 2004) and the test writers might consider whether this could be built into future tests.

While many of these changes raise issues related to the examination, they can also be interpreted in a way which relates to the role of native or expert users in research into the effectiveness of the IELTS examination. This is illustrated in differing frequencies of the use of what we term examination tactics by UOL and GDUFS participant (see Table 10).

The difference between the means for the tactics for UOL and GDUFS participants were not significantly different. However, we were surprised that native/expert users often made more use of the examination specific tactics than did the potential candidates. This may reflect the fact that the relative unfamiliarity of native/expert users with this examination leads them to rely on general examination taking strategies and tactics.

Whatever the reason, it does raise some quite difficult issues about how data from native/expert users can be used to inform test design. The native/expert users are treating the IELTS as a specific kind of task in its own right, independent of the TLU tasks that test writers relate it to. In terms of the strategies and tactics, the test does not have task authenticity even for native speakers/expert users of English, though this may be seen less a critique of the IELTS Tests than of the use of task authenticity as a criterion for test evaluation. An exam is almost always perceived as an exam rather than as a replication of some other language task.

The aim of the IELTS Test is in some sense to evaluate the relationship between the competence of those taking the examination and expert users of English in the TLU. However how this relationship can be informed by the way expert users of English behave in an exam needs further exploration.

Tactic	Native speaker status	N	Mean	SD	SEM
Metacognitive: Comprehension Monitoring: confirm that an exam question has been Answered	NESE	8	0.63	0.92	0.32
	NC	24	1.75	2.21	0.45
Metacognitive: Comprehension Monitoring: identify examination Questions Not answered	NESE	8	3.38	3.78	1.34
	NC	24	4.75	2.36	0.48
Metacognitive: Comprehension Monitoring: Identify examinations Skills not applied	NESE	8	0.13	0.35	0.13
	NC	24	0.42	0.83	0.17
Metacognitive: Comprehension Evaluation against examination Questions	NESE	8	5.63	5.40	1.91
	NC	24	2.54	2.25	0.46
Metacognitive: Comprehension Evaluation against experience of Examinations	NESE	8	2.00	2.88	1.02
	NC	24	2.50	3.19	0.65
Cognitive: Reconstruct meaning from examination question	NESE	8	1.86	2.61	0.99
	NC	24	0.08	0.28	0.06
Cognitive: Inferring information from the listening text and exam question paper	NESE	8	3.88	6.14	2.17
	NC	24	2.50	2.41	0.49
Metacognitive: Prepare using exam paper Questions	NESE	8	11.88	11.28	3.99
	NC	24	1.58	2.13	0.43
Metacognitive: Pay Selective Attention to exam Questions	NESE	8	0.25	0.46	0.16
	NC	24	0.46	0.83	0.17
Metacognitive: Real time Assessment of Output in terms of Quantity required (e.g. one or two words)	NESE	8	2.38	4.10	1.45
	NC	24	0.42	0.78	0.16
Metacognitive: Real time Assessment of Output in terms of Process required e.g. multiple choice vs., gap fill	NESE	8	6.25	4.27	1.51
	NC	24	2.50	2.41	0.49
Metacognitive: Real time Assessment of Output in terms of Intermediate processes e.g. note taking	NESE	8	0.25	0.71	0.25
	NC	24	2.42	2.48	0.51

None of the differences are significant at p<0.05   SD=Standard Deviation. SEM=Std. Error Mean
NESE=Native/Expert speaker of English   NC= Native speaker of Chinese

Table 10: Descriptive statistics for "examination" tactics

4   DISCUSSION AND CONCLUSION

This study has shown that there are relatively few significant differences between the performances of native/expert users of English and potential IELTS candidates with Chinese as a mother tongue and the way in which they carry out the examination. This provides evidence for arguing that the IELTS Test does provide an accurate representation of candidates' language abilities.

The study found that there were no significant differences between pre-undergraduate and pre-postgraduate students taking the IELTS Test. Again, this supports the validity of the IELTS test. However, this lack of significant differences may have been partially related to the small numbers of participants in this study and it would be worth investigating this issue with a larger number of participants.

The study found that the range of tactics reported by participants taking the IELTS Examination differed somewhat from the strategies reported in a non-examination context. This raises at least two issues about the IELTS Examination which may need further investigation: firstly what texts are used in the IELTS test and, secondly, the use of native/expert users of English as one way of assessing the validity of the IELTS task

4.1 Choice of texts

At the moment, students are exposed to a listening text and a written text comprising the test rubric and questions. The form of the examination requires that candidates make use of the written text to answer question and it may be worth exploring whether some of the written texts may be replaced by additional listening texts.

For test taking purposes, one advantage of written questions in a listening examination is that it reduces reliance on memory. So, for example, in some sections of IELTS Listening, there may be ten or more items and candidates would need to have a very good memory to answer all ten items based on a single hearing of the text, particularly for questions where candidates have to fill in the blanks.

If a decision were made to have more spoken questions, this might be addressed by reducing the length of the sections, though this would reduce the text authenticity of what students hear. An alternative is to move to a system where the candidates hear the recording of both text and questions twice. This could be justified in terms of authenticity on the grounds that the second hearing compensates for the lack of contextual information that would be available to listeners outside a test or language learning situation.

However, our preferred solution is to maintain the convention that candidates only listen once to the spoken text but without seeing or hearing the questions. Although such a change would make the ability to take notes a more significant part of the listening construct, it would ensure that the listening was focussed more directly on understanding spoken input rather than on combining more or less authentic spoken and more or less inauthentic written text.

Candidates might then listen to, or read, the questions in their current format. Providing a way of staging access to the written text might well be difficult and would also make it harder to separate out listening and reading abilities so we would favour the questions being spoken rather than written. This would involve a major change to the IELTS test and so would need to be trialled on native/expert speakers of English.

In addition, many of the TLU tasks that the IELTS Test is based on, e.g. lectures, are now multi-modal events and it may be that test constructors need to consider the inclusion of other modalities such as still or moving visual images. There are clearly logistic issues for a paper-based testing system where replicating exposure to, for example, PowerPoint slides for relatively short periods of time are difficult but this would be less problematic for a computer-based testing system.

4.2 The use of native/expert users of English in test validation

The second issue relates to the use of native/expert users of English as one way of assessing the validity of the IELTS task. The data collected in this study suggests that native/expert users of English treat the IELTS not as derived from the TLU tasks to which it relates but as a specific kind of task in its own right. This makes it difficult to evaluate the relevance of the native/expert user data collected in this study to the validation of the IELTS Test.

One line of argument would be as follows: the IELTS examination is designed to judge to what extent candidates can perform as well as expert users of English in non-test contexts. At the moment, we have limited information about how expert users behave in these non-test contexts but we know how expert users behave in test contexts and this might be thought to relate in a systematic way to how they behave in non-test contexts. So, if candidates behave in a similar way to expert users in a test context, this is evidence for saying they will behave in a similar way to expert users in non-test contexts. At several points, this argument relies on plausibility rather than evidence.

In particular, the claim that expert users' behaviour in test contexts relates to their behaviour in non-test contexts, though not unreasonable, is largely unsupported. Indeed, it is hard to see how we might collect evidence for such a link in most non-test contexts. How would one collect data about how someone processes and interacts with a credit card call centre, without the possibly distorting effects of setting up an experimental context?

However, in at least one context that is used in IELTS, the lecture, it would be possible to carry out research into how native speaker /expert users of English reach understandings of what is going on. When students attend lectures in higher education institutions, they are not provided with explicit questions to which they must find answers (Badger, Sutherland, White and Haggis 2001). Instead, they annotate handouts or write notes which then contribute to answers to examination question or assignment tasks. It would not be very difficult to research using interviews (Sutherland, Badger and White 2002) or stimulated recall (Hodgson 1997). This research could also link fairly directly to the design of tests of academic listening in IELTS.

REFERENCES

Anderson, A, and Lynch, A, 1988, *Listening*, Oxford University Press, Oxford

Anderson, J R, 2000, *Cognitive psychology and its implications*, 5th ed, Macmillan, Basingstoke

Bachman, L F, & Palmer, A, 1996, *Language testing in practice*, Oxford University Press, Oxford

Badger, R, Sutherland, P, White, G, & Haggis, T, 2001, 'Note perfect: an investigation of how students view taking notes in lectures', *System*, vol 29, no 3, pp 405-417,

Bialystok, E, 1990, *Communication strategies a psychological analysis of second-language use*, Oxford: Basil Blackwell,

Brown, J, and Rodgers, T, 2002, *Doing second language research*, Oxford University Press, Oxford

Buck, G, 2001, *Assessing listening*, Cambridge University Press, Cambridge

Ellis, R, 2003, *Task-based language learning and teaching*, Oxford University Press, Oxford

Flowerdew, J, & Miller, L, 2005, *Second language listening theory and practice*, Cambridge University Press, Cambridge

Gieve, S, and Clark, R, 2005, 'The Chinese approach to learning: Cultural trait or situated response? the case of a self-directed learning programme', *System*, vol 33, no 2, pp 261-276

Glaser, B, and Strauss, A, 1967, *The discovery of grounded theory: strategies for qualitative research*, Aldine de Gruyter, New York

Glaser, B, 1992, *Basics of grounded theory analysis emergence vs forcing*, Sociology Press, Mill Valley CA

Goh, C, 1998, 'How ESL learners with different listening abilities use comprehension strategies and tactics', *Language Teaching Research*, vol 22, pp 124-147

Goh, C, 2002, 'Exploring listening comprehension tactics and their interaction patterns', *System*, vol 30, no 2, 185-206

Green, A, 1998, *Verbal protocol analysis in language testing research: a handbook*, Cambridge University Press/UCLES, Cambridge

Guariento, W, & Morley, J, 2001, 'Text and task authenticity in the EFL classroom', *ELT Journal*, vol 55, no 4, pp 347-353

Hodgson, V, 1997, 'Lectures and the experience of relevance', In *The experience of learning: Implications for teaching and studying in Higher Education*, 2nd ed, eds F Marton, D Hounsell and N Entwistle, Scottish Academic Press, Edinburgh, pp 159-171

McCarter, S, & Ash, J, 2003, *IELTS testbuilder*, Oxford: Macmillan,

Nagle, S, and Sanders, S, 1986, 'Comprehension theory and second language pedagogy', *TESOL Quarterly*, vol 20, no 1, pp 9-26

O'Halloran, K, 2004, *Multimodal discourse analysis: Systemic-functional perspectives*, Continuum, London

O'Malley, J, and Chamot, A, 1990, *Learning strategies in second language acquisition*, Cambridge University Press, Cambridge

Oxford, R, 1990, *Language learning strategies: What every teacher should know*, Heinle, Boston, Mass

Phakiti, A, 2003, 'A closer look at gender and strategy use in L2 Reading', *Language Learning*, vol 53, no 4, pp 649-702,

Purpura, J, 1999, *Learner strategy use and performance on language tests: a structural equation modeling approach*, Cambridge University Press, Cambridge

Rost, M, 2002, *Teaching and researching listening*, Longman, Harlow

Saville-Troike, M, 2006, *Introducing second language acquisition*, Cambridge University Press, Cambridge

Saville, N, and Hawkey, R, 2004, 'The IELTS impact study: investigating washback on teaching materials', in *Washback in language testing*, eds L Cheng, Y Watanabe and A Curtis, Lawrence Erlbaum Associates, New Jersey, pp 73-96

Scollon, S, 1999, 'Not to waste words or students; Confucian and Socratic discourse in the tertiary classroom', in *Culture in second language teaching and learning*, ed E Hinkel, Cambridge University Press, Cambridge, pp 13-27

Senior, R, 2006, *The experience of language teaching*, Cambridge University Press, Cambridge

Skehan, P, 1996, *A framework for the implementation of task-based instruction*, Applied Linguistics, vol 17, no 1, pp 38-62

Sutherland, P, Badger, R, and White, G, 2002, 'How new students take notes at lectures', *Journal for Further and Higher Education*, vol 26, no 4, pp 377-388

Terry, M, 2003, 'IELTS preparation materials', *ELT Journal*, vol 51, no 1, pp 66-76

Vanijdee, A, 2003, 'Thai distance English learners and learner autonomy', *Open Learning*, vol 18, no 1, pp 75-84

White, G, 1998, *Listening*, Oxford University Press, Oxford

Yao, X, 2000, *An Introduction to Confucianism*, Cambridge University Press, Cambridge

APPENDIX 1: NON-NATIVE SPEAKER PROTOCOL

The method I used to deal with the section 1 is to circle the key words while I was reading the introduction. And then I looked through the former questions roughly and judged the speaking speed and difficulty while he was reading the examples. Afterward I just did it according to the question order.

Shall I talk them one by one? You can make a brief analysis if you think there are some difficult questions.

I'm not very sure about the question 4 'day of birth' because I was a little bit absent-minded then.

And when it came to the tenth one, since I habitually do it according to the question order, he talked about the tenth one first during the ninth and tenth part which is different from my habit, therefore I just guessed it.

I've got some but they just flashed by.

Yes, I probably recalled a little and then I guessed them.

Section 2 is different from section 1 because of the faster speed and the gaping filling part. The fast speed is acceptable in the multiple choice part because I just circled the key words, but the gap-filling is relatively hard for it has word limit. Besides I think section 2 usually includes two parts and the second part is faster than the first part.

And question 17, he said 'he give up smoking' and another word which I can not catch, so I just guessed its meaning and filled in a familiar word.

And when it came to the eighteenth and the twentieth, I didn't catch them and do it orderly due to the faster speed. I just filled in a word probably because of my poor memory. So that's probably how I did them.

Maybe I have more confidence in the nineteenth and the twentieth but little in the seventeenth and the eighteenth.

Oh, one by one? In the first question [question 11] he said that he did some exercises while he went to take part in a class. Since he felt painstaking if he smoked and he needed to 'touch his toes' when he did that exercise. Then, he couldn't do that. I was a little bit hesitant to choose B or C that is 'smoked' or 'couldn't touch his toes'. I had thought about it after the listening of this question and I finally chose C because his problem was always about the 'smoke'.

And question 12 I got it clearly in which he said 'he travel across London in order to get cigarette'. It mentioned about his fancy for cigarettes.

I also had hesitation in question 13 on whether he gets to sleep or get up early in the morning because the former section has mentioned that he gets up at 3 am. to smoke, so I chose D according to my former memory.

Um, I didn't hear question 14 clearly, 'stopped smoking', I just heard a date—July 1st and I filled it in.

The whole sentence has mentioned about it in question 15, but I was not familiar with that word, so I can just spell it based on my experience.

And question 16, what he said seemed not like the original words and he seemed not to mention about the whole sentence 'work side effect'. As for this question, I filled in the answer for question 16 after I heard the key word—'giving up smoking' of question 17.

Because it mentioned 'habit' during the process of 'cut down smoking and 'give up smoking' of Q15. Then according to 'work side effect', I thought it should be 'habit'.

Just some words which I cannot spell out. Yes, spelling, especially in gap-filling.

Your attention will be distracted usually by two hesitating answers.

I think section 3 is a little bit simpler in content and slower in speed compared with section 2, may be we were more familiar with its content, so we catch it easily.

I'm quite sure about the question 21 and 22, in which 21 is 'past three years' and 22 is 'got a job'. I can do questions like gap-filling more fluently because I know something about abbreviation.

When it came to question 23, the first time I heard...the Lorraine said that she would 'turn to exam week', so I wrote 'exam week' firstly.

But later I heard 'turn to Wales', in which I saw the preposition 'to' and so I filled in Wales. I guess it's a place name.

And then for question 24 since he didn't said clearly that he had 'any mistake in his project', I just guessed that it's probably because he was too 'easy to make mistake', so I chose A.

I think the answer for question 25 is B because he has mentioned that......because I chose 'he has some mistake' as the answer for question 1 (Q24) and later I heard that there is something wrong with his end, so I chose B

Besides I often made some written mistakes, like B and D. Since there is time for us to fill in the answer sheet in IELTS and I would look through that question again while I was writing, so it's not a big problem in this part.

Question 26, it seems that there is an 'end' to modify, so I chose C.

Question 27 is just a copy of the original sentences.

I was hesitant to choose the answer for question 28 because both B and D were mentioned and I had to choose a better answer—B.

About this? Since he has mentioned that Frances' project is better than before and 'you can get a PhD' in the last question, therefore I thought if he has mentioned that he did better than other classmates as he kept studying. I think it should be a summary because I didn't get it from the original text

As for question 27, I just roughly heard a time and I filled it in the blank for there was little time to look through the former part.

In question 30, Steve has mentioned that not to 'go on his research' firstly because he had to 'turn to his work'. I was a little bit uncertain about answer D until he later mentioned that 'he didn't earn some money to do the things he would like to do'. So I finally chose C : 'He goes to earn some money'

There is still the problem about spelling in question 21 and 23, because I forgot the spelling of some words and I just spelt them out according to the pronunciations. And next, there are also ambiguous choices in section 3, so I cannot fill it in the black successfully by getting it from the original text and I had to think it over like I did question 28.

Compared with question 29, I did question 25 and 26 sooner for I got the answer in the text and their answers are shorter. Actually I often spend some time on those equivocal answers.

Section 4 is a little bit difficult for its speed is faster than the former three sections. According to my experience of taking IELTS, section 4 usually refers to something about geography, biology and technology, so there will be some unfamiliar words in this section.

I missed question 35 while I was thinking question 34 because I felt puzzled about his former words

Usually I would give it up usually in that case and turn to make sure the other questions. And the method is circling the key words as the former three questions. I heard 'technics and training' clearly in question 32 but I was not very sure about question 32, so I just wrote 'special economic'. What I heard was like 'ishal economic', and then I cannot catch it because it's a little bit fast.

Later, I circled 'experience, security and lack' in No. 33. According to the three key words I knew that its original text is based on this order, so I listened to it seriously when he mentioned 'insecurity' at the thought of the coming answer. Then I filled in 'a lack of awareness of the part of managers' for question 33

and 'set the management tasks' for question 34. Since he didn't read the whole sentence in No. 34, what I wrote is not the same as the original text, then...

probably. I missed No. 35 may be because the former two 'goalposts' are beyond my knowledge so that I could not catch the following part. Maybe it's related to my psychology to some extent. Speed is only one of the reasons but the unfamiliar words matters most. My mood will be influenced if I was not familiar with the word before the blank.

When it came to No. 36, I was slightly puzzled at first until I heard he 'replying to email' I recalled that the former part has mentioned that he 'mentioning the calling', so I wrote 'contact her' according to the text though I would change it a little while I wrote the answer on the answer sheet.

No. 37 I heard the answer directly from the text: 'you cannot expect your staff respect you'. Yes, I just very sure about this question.

I circled the key word for No. 38 which was 'technology'

Soon in No. 37 I heard 'your staff respect you' and then 'technology' was mentioned later. And I wrote the answer for this question according to its text that is 'company's strategy or practice'.

In No. 39......based on my key word ?In groups......? and it is ?tasks to...make...in groups or something else?? As for this question, I heard its text which is not always about this and he said ?other bullying strategies?, so I wrote a

very familiar word ?brainstorm?. At the first, I thought he would read the whole sentence directly but actually he divided it into two or three sentences. Afterwards I heard 'and way', since 'and ways in which' was directly mentioned in the latter part 'in which they can be counted', I just filled in 'counted'.

The major problem is the speed and the new words in section 4. Since the majority of the section 4 is gap-filling, the new words became a big obstacle. The only way is to spell them out according to their pronunciation even though they were still unknown.

Besides there are some traps in section 4 because of its requirement. Sometimes it requires 'no more than three words'. What I have met is 'no more than two words', so we have to pay attention to it or else you will fail.

APPENDIX 2: NATIVE SPEAKER PROTOCOL

Right so you want me to describe how I'm working it out ok,

well obviously it's an anagram,

so that one seems quite easy because instantly I'm seeing well I think it's table

and it's not that really separated it's just, they've just put the t in the middle rather than the a and the b.

I suppose because I've instantly seen those letters together and that's the most likely pattern, I've seen them before and instantly thought that

I'm having some trouble with the second one it seems a lot less immediate,

maybe they've scrambled the letters a lot more effectively,

possibly because it's, are there more vowels, yes I think there are, there's three vowels

Oh it's a mouse that's what it is. I just thought,

I started looking at it backwards then it becomes a lot easier

because I suppose the first bit which is O S U

if you look at that reading it backwards it becomes a lot easier,

I can't really describe it better than that.

It is important to blank temptation,

I would say shun because I suppose that just sounds like quite a formal sentence and shun possibly but I think shun would fit best in keeping with the character for that sentence

There is nothing so . . . practical

I suppose because possible doesn't make sense,

plausible makes sense very loosely but not really the correct use,

I think potential is a bit big, and doesn't make any sense

so I'm going to say practical.

Read the following text and choose the best answer, say why you reject the wrong answer,

right, what was the weather like?

I think brillig because, this is just a guess but I'm guessing that, I'm mean because it's not strictly a word obviously but um, I think he is sort of using that as a disambiguation to make that a word forming brilliant describing the weather, that's the best I can do,

I'll go down to six.

What came through the talgy woods,

um, the jabberwock, yes because it says the jabberwock with eyes aflame came wiffling through the talgy wood, yes so the jabberwock, so that's b for that question

Seven, the jabberwock is dangerous; because it says beware the jabberwock, the jaw that bites the claws that catch, which insinuates danger.

[subject listens to instructions for the test]

I thought the instructions were very clear,

I don't know I suppose it's one of those things where you are listening to it but sort of thinking more about what you are actually going to be doing

I mean it was very well phrased and everything and it seemed very clear

the only bit I don't really understand is why you have to transfer your answers at the end, that might be confusing.

So is that then the second I'm going to hear this conversation.

[subject listens to section one of the tapescript]

Seemed quite complicated for a, I'm just imagining if I wasn't speaking English.

Yes I suppose it's fairly clear, so they want you to basically write what you hear down, yes that's fine.

It's quite fast;

it's good that obviously because I did not know how to spell Moore

yes it is quite fast.

There is potentially the bit where it says the 13th of 7th

but I mean probably most people would get that.

It's strange it's very different to any other speaking and listening tests I've ever done, because it's very specific about interest payments and banking, which is very strange,

but I mean the questions all seem fine they are all fairly self explanatory to me.

Once again I think it's quite fast, but I suppose if you where, oh I don't know. I mean I think there fine, they're fairly, well they say what you are suppose to do and they say it in simple language so yes.

It is another payment to another restaurant but that's not really the right answer because the thing that he is really worried about,

but I suppose that's just a bit of a tricky question, there's nothing wrong with it, it's just a bit mean, fair enough.

Once again, I suppose there's no ambiguity him saying the interest has gone up and she's saying it has gone down,

so that means C,

but I wouldn't be entirely sure if it's clear whether we are saying what the caller thinks what's happened to the interest or what has actually happened.

The only thing about that was, this might sound a little bit silly, if I had been learning say, if this was in German, you know this thing where they say 020 I don't know if people would understand that because surely they would have learnt zero, I mean maybe not I suppose it would be how colloquial their teaching had been, just a thought anyway

[time for checking answer]

Seems kind of pointless, because you wouldn't, I mean that half minute thing because you wouldn't really, I mean you can't hear it again and you are going to have been concentrating hard on each question so you're not really going to have anything more than just worrying about did I get that one right, you know if you could check maybe having a summary of the conversation that might be more effective but I don't really see the point of that bit.

[subject listens to section 2]

Yes that looks fine simple.

I didn't get thirteen,

I think, I can't be sure that he mentioned all of those things actually

I mean because he did have trouble giving up smoking,

but obviously doing those things

so that didn't seem very clear.

I mean that one's more complicated than the previous one because it's sort of going through each section, bit by bit, it's giving you a whole conversation,

I mean because those two are quite close together at least that's what I thought, I thought those two

[QUESTIONS] were answered quite quickly,

but that one I was sort of listening for, because in comprehensions like this you often find like there's the little bit saying, 'did you have trouble giving up smoking'? 'Yes I did have trouble giving up smoking' (time 8.38) and so you look out for that bit,

but I didn't, but that just might be a habit from doing this kind of thing,

but I didn't, well the answer certainly didn't leap out at me, I didn't get 13

I suppose if I was doing the test I would just go right, well he's given up smoking so I'll put that because, it's probably most likely to be correct,

but I didn't actually get it.

Sorry is this, [subject looking at question 14-20] is there going to be more of him talking, or do we fill that out from what I've just heard.

[Subject listens to section 2b]

Oh better do all this quickly.

Well I didn't get the first one

it seemed very fast,

he just said a date

and it was like almost the first word he said and I was going, what was that?

But then he kept talking, so that was tricky apart from that as I was saying before he sort of, he says, it's basically paraphrasing the sentence, so that's quite easy,

but apart from that it seemed fairly …

apart from just because it's something people can panic about in exams because it's got a little line, I mean I suppose in order to be able to fill it out in an official way you need some indication of where you have to write especially there, if someone wasn't confident about their own writing abilities in English it could make it difficult, could be confusing. It seems a bit needless because all the others are at the end of the sentence apart from that one [question 15] . So it's not like testing really, I mean their ability in the middle of the sentence with only one, so it seems strangely done.

[30 second to check answers]

Once again I don't really see the point because you can't really check it unless you've got a really good memory or you hear it again, because, all you can do it brood on how well you think you did. I don't think that it's a good idea

that the question number is; I think that might be confusing because the question number is put immediately after the rest of the sentence, [in questions 21 to 23] so it seems very confusing. Just then I was looking at it and thinking have they tried to incorporate the question number into the question, it seems a bit strange. But then I realised it's just where they want you to write the answer, so I think it would be quite easy to just put 21 on the left of the question like it is and all the others rather than on the right, it seems a bit strange and probably a bit confusing I would think.

Yes apart from the number thing it's very simple.

The only think I would say about that one is, it's all fine apart from, there's quite a lot at the beginning where they don't really answer any of the questions

and then it say's 28, what the tutor thinks about Francis's chances of getting funding,

and it very quickly say's

like well three years ago,

and at that point you have to be going very quickly, and so you would just lose is through no fault of your own,

just because you have to concentrate on that, and by the time that you got back to the tape it had gone onto the next one,

or so it says, I mean for question 30 it says return to his job

which left me thinking if that was actually the case because he says he has got a job lined up,

I mean if you where just looking at it purely from a I suppose a tactical point of view he not really returning to a job, he's got a job, so that could be confusing.

Might make some think that it was C. but that's all I've got for that section.

Just because this one on each question you are getting more information than you are previously, I suppose the half minute could actually do a bit better because you are not having to rely on so much on your own memory you can think, well da, dum, da, dum, da, dum oh right, sorry that wasn't particularly well phrased.

Just because you get more information in the actual questions and in the answer you would stand a better chance of checking in that half minute

[subject reads questions 31-33]

Once again it's a bit confusing having the number, but not so much there because, having the number in the question it's not as confusing as it was previously there because it's., it seem more incongruous than it did before.

[subject lists to tape script for questions 31-33]

I don't need to say once again that could be confusing because I was starting to think, oh are we going straight on, because he kept talking. But no I suppose maybe if it was more clearly phrased it just the 31 to 33

Actually looking back it did but I suppose you could panic, I don't know that's ambiguous

Once again it's one where the answers aren't necessarily in order

they're a conversation well not a conversation a speech,

quite a realistic speech, not the most exiting one.

I think it says not blank or replying to e-mails,

and I thought that might be memoing

but I wasn't sure if that was a word, and obviously if I didn't know English that would be more of a problem.

Also with 35 it says stopping individuals blank to criticism

and for 34 and 38 and 37, I was waiting for the right word because I thought he would say it

but that one he didn't

so I put capacity because [that's the closest word I could think of]

from the gist of what he was saying that's the closest word I could think of

but it seems different to instead of picking out the word from a sentence which is what you do for most of the other questions, well 34, 37, 38, 39 and 40,

you actually have to use your own knowledge to think of the best answer, so it's different and strange in one set of questions, but I suppose that might be the object of it

that's all I've got for that section.

APPENDIX 3: GOH'S 2002 TAXONOMY

S	Sub-strategy	Tactic
Cognitive	Inferencing	use contextual clues, use familiar content words, draw on knowledge of the world, apply knowledge about the target language
	Elaboration	draw on world knowledge, draw on knowledge about the target language
	Prediction	anticipate general contexts, anticipate details while listening
	Contextualisation	place input in a social or linguistic context, find related information on hearing a key word, relate one part of text to another
	Translation	find L1 equivalents for selected key words, translate a sequence or utterance
	Fixation	stop to think about the spelling of unfamiliar words, stop to think about the meaning of words, memorize/repeat the sounds of unfamiliar words, memorize words or phrases for later processing
	Visualisation	imagine scenes events, objects etc being described, mentally display the shape of key words
	Reconstruction	reconstruct meaning from words heard, reconstruct meaning from notes taken
Meta-cognitive	Pre-listening	preview contents, rehearse sounds, encourage oneself to relax
	Selective attention	listen to words in groups, listen for gist, listen for content words, notice how information is structured pay attention to repetitions, notice intonation, listen to specific parts of the input, pay attention to visuals
	Directed attention	concentrate hard, continue to list in spite of difficulty
	Comprehension monitoring	confirm that comprehension has taken place, identify words or ideas not understood, check current interpretation with content of the message, check current interpretation with prior knowledge
	Real-time assessment of input	assess the importance of problematic parts that are heard, determine the potential value of subsequent parts of input
	Comprehension evaluation	check interpretation against some external sources, check interpretation using prior knowledge, match interpretation with the context of the message

S=Strategy

APPENDIX 4: ADAPTED TAXONOMY OF STRATEGIES, SUB-STRATEGIES AND TACTICS

Cognitive strategies

Fixation (F) (Focussing attention on decoding a small part of a text)

- Spelling (Sp)
- Phonemes or sounds (P)
- Words or parts of input (W)
- Repeat /Memorise sounds (RS)
- Repeat/ Memorise words phrases (RW)

Reconstruction answer from input (R) - using key words/phrase to recreate meaning including answering question (staying with literal meaning)

- Reconstruct meaning from word heard (W)
- Reconstruct meaning from phrase heard (P)
- Reconstruct meaning from notes taken (N)
- Reconstruct meaning from examination question (Q)

Inferring answer (I) by using information from the text plus one of the following(going beyond literal information)

- Co-text (C) - something from the listening text
- Familiar word (F)
- World Knowledge (W)
- Knowledge of target language (L)
- Visual clues (V)
- Exam question paper (Q)

Contextualization (C) (Relating new information to a wider familiar context)

- Social or linguistic context (SL)
- Find related information on hearing a key word (K)
- Relate one part of text to another (T)

Translation (T)

- Find L1 equivalents for key words (K)
- Translate a sequence of utterances (U)

Visualisation

- Imagine scenes being described (I)
- Mentally display the shape/sound of key words (K)

Metacognitive tactics

Pre-listening preparation (P)

- Preview contents (C)
- Rehearse sounds of potential key words (K)
- Encourage oneself to relax (R)
- Using exam paper questions to prepare (Q)

Selective attention (SA) (Decision to focus on specific aspects of input or identifying from something other than the listening text what to listen for).

- Listen to words in groups or phrases (P)
- Listen for gist or general meaning (G)
- Listen for key words/phrases (K)
- Listen for discourse markers (D)
- Listen for repetition (R)
- Intonation features (I)
- Pay attention to visuals and body language (V)
- Pay attention to exam questions (Q)

Directed attention (DA) (monitoring attention and avoiding distraction) (C)

- Concentrate hard (C)
- Concentrate hard in spite of difficulty (D)
- Identify a failure in concentration (F)

Real-time assessment of input (AI) (determining the value of specific parts of the input)

- Assess the importance of problematic parts that are heard (P)
- Determine the potential value of later parts of the input (L)
- Determine links between elements in listening text and examination questions (Q)
- A problem with the speed/clarity of the input (S) e.g. the speaker speaks too fast or mumbles
- A problem with the amount of input (A) e.g. too much to read or listen to

Real time assessment of output (AO)

- Quantity required (Q) (e.g. one or two words)
- Process required (P) (e.g. multiple choice vs., gap fill)
- Requirement for intermediate processes (e.g. note taking) (I)

Comprehension monitoring (CM) (Checking interpretation for accuracy while listening)

- Confirm that comprehension has taken place (C)
- Identify words or ideas not understood (N)
- Confirm that an exam question has been answered (QA)
- Identify examination questions not answered (QN)
- Check current interpretation with context of the message (M)
- Check current interpretation with world knowledge (W)
- Identify examinations skills not applied (S)
- Identify partial understanding (P)

Comprehension evaluation (CE) (Checking interpretation for accuracy after listening)

- Against examination questions (Q)
- Against world knowledge (W)
- Against context of message/rest of message (M)
- Against experience of examinations (P)
- Check interpretation using linguistic knowledge (L)

APPENDIX 5: CONSENT FORM FOR THE RESEARCH

Research background and aims

Thank you for agreeing to participate in this research, which focuses on investigation of how people do an IELTS listening test. Research findings will be used in the research project for the British Council.

As a researcher, I have an obligation to those taking part in the project to make sure that nothing negative arises from their involvement. The ethical principles governing my research are set out below.

Code of ethics

- What participants tell me will be treated in the strictest confidence.
- No individual will be identified by name.
- Any data which I might use when reporting the findings of this research will be anonymised.
- Participation in the project is entirely voluntary.
- Participants are under no pressure to answer any question they may feel uneasy about.

Consent

I would very much value your participation in this project and am happy to answer any further questions you may have about it. RMB 150 (£10) will be paid to each participant in this study.

If you would like to take part in the project, please sign below:

Signed _____   Date_____