

A Review of the IELTS Listening Comprehension Module

Edward Chaffin

Introduction

It is the initial purpose of this review to provide a brief look at the format and procedures involved in taking the Listening Module of the International English Language Testing System (IELTS). Following an outline of test (Module) content, a more evaluative discussion of differing aspects of the Module is included. Though the scope of this paper is limited, I would nonetheless like to look at a variety of issues in IELTS' Listening Module regarding task authenticity and test validity.

The Academic version of the test includes Listening, Reading, Writing and Speaking Modules. I have chosen to focus on the Listening Module in an attempt to see just how accurate Yi'an is when he states "in the area of listening comprehension we are still rather ignorant" (27, Yi'an, 1998). While listening comprehension sections seem standard inclusions within a host of high stakes language tests, the amount of corresponding research into the field of listening seems quite limited (Read, 2002). This may be due to a certain inability to accurately account for what Buck and Tatsuoka refer to as the "complex, multidimensional cognitive process" responsible for listening comprehension yet it would still seem the issue deserves more scrutiny (119, Buck & Tatsuoka, 1997). Particularly so when listening sections are included in such very high stakes proficiency tests as the IELTS.

General Format and Procedure

While there are separate Academic and General Training Versions of the IELTS, the Listening Module is the same in both versions. The Listening Module is the first of three IELTS Modules that must be completed in one sitting. As the IELTS Handbook explains, the Listening Module of the test contains 40 questions in four sections and takes approximately thirty minutes to complete. Candidates are allowed to listen once to a CD recording composed of both monologues and conversations. Before the recording begins, candidates are allowed one minute to pre-read information found on the "Question Paper". Candidates record answers to questions found on the "Question Paper" while the recording is being played.

After the recording is completed, candidates are allowed ten minutes to transfer their answers on the "Question Paper" to the "Answer Sheet". Since examiners neither score nor see answers recorded on the "Question Paper", it is imperative for candidates to transfer answers to the "Answer Sheet". Failure to do so, as the Handbook further

explains, will result in a candidate receiving a band score of 0 for the module. Answers recorded on the “Answer Sheet” that use “poor spelling and grammar” are “penalized” (11, IELTS, 2005). “Poor” on p. 11 is clarified as “incorrect” on p. 14 while it goes largely assumed “penalized” translates to the answer being marked completely incorrect. Both UK and US spelling conventions are accepted.

Listening Sections

Content of the initial two listening sections is referred to in the Handbook as “concerned with social needs” (11, IELTS, 2005). A conversation between two people is given first followed by a monologue. Examples of conversation topics include travel arrangements and deciding on a night’s activities while examples of a possible monologue include speeches given within a university or business context.

Sections one and two are differentiated from sections three and four in that the later “are concerned with situations related more closely to educational or training contexts” (11, IELTS, 2005). Examples of conversations to be expected in section 3 include an exchange between a teacher and student and a conversation between three students. Section four consists of a monologue with the example of a general university lecture given.

Scoring

Purported by the Handbook to provide “a profile of a candidate’s ability to use English”, candidates receive a score for the Listening Module between 1 and 9 on IELTS’ banded scoring system. A score of 0 is reported for candidates who skip the Module altogether. Whole and half band scores are possible. A score of 1 indicates a candidate has no real ability to use the language other than a few isolated words while a score of 9 indicates an “Expert User” with full operational command of English and complete understanding (4, IELTS, 2005).

Band scores derive from a “confidential” system that translates the number of correct answers into the band system. Examples of raw score translation to band scores may be found on IELTS website at:

<http://www.ielts.org/teachersandresearchers/scoreprocessingreportingandinterpretation/default.aspx>

Generally speaking, an overall band score of between 6 and 7.5 is a benchmark for students wanting to study within English medium schools in the UK, Australia, New Zealand and various other locations. A score of 4.5, indicating a candidate is a “limited user” of English, gains students’ access into the university where I teach.

Test Authenticity

While authenticity in tests as a concept is well understood (McNamara, 2000), IELTS' domain and examples of criterion need explanation. Since the ultimate goal of IELTS regarding the Academic Version of its test is to "assess the language ability of candidates who need to study or work where English is the language of communication" (2, IELTS, 2005), one would largely expect listening activities to mimic those found in academia and/or the business community. Listening to aural input while answering questions about the content of that input can be seen as a task found within the domain IELTS relates to. Though IELTS attempts to include a high degree of authenticity within the Listening Module, it can be seen that a number of compromises are necessary to, on the one hand, engender authenticity and, on the other, recognize "the practicalities imposed by the test situation" (28, McNamara, 2000).

Tasks

The IELTS Listening Module's input channel is both aural and visual while its response channel is entirely visual. Candidates are allowed to view and write on the "Question Paper" while listening to the four recordings of the Module. While candidates are tasked with both the relatively difficult job of answering questions and listening simultaneously throughout the Module, Read (2002) indicates that this practice effectively simulates an actual "target language use situation" for university students. Experiential knowledge indicates simultaneously listening and answering questions simulates tasks expected within the business community as well. Since IELTS' primary concern is to determine if a candidate is "ready to study or train in the medium of English at an undergraduate or postgraduate level", it seems apt to include test-tasks which mimic authentic academic or work related tasks candidates will be expected to perform (2, IELTS 2005).

The non-reciprocal nature of the interaction between recording and candidate is a limitation not necessarily seen within the criterion, however. It may be argued that in actuality, a range of devices based on the ability to interact with a speaker or speakers would be at the disposal of participants within the domain outlined by IELTS. Students in a classroom often ask for clarifications of points covered within a lecture, greater explanations, repetition of key points, and even the spelling of new or difficult vocabulary. Ferris' needs analysis study indicates that from her pool of university level respondents, "79% reported that students always or often asked questions before, during or after class" (298, Ferris, 1998). Exchanges in a less formal context where the

listener is a de facto participant allow a similar battery of tools to be used to clarify or confirm comprehension.

While IELTS' input channel does in many ways mimic the criterion, clearly a compromise between authenticity and the practicalities imposed by the testing context is seen here. The restrictions practicality places on the Listening Module are somewhat assuaged in the Speaking Module where a high level of interaction between interviewer and candidate is allowed but, if viewed independently, the inability of candidates to utilize common listening comprehension aids tends to cast a pall over IELTS' attempt at authenticity.

Listening Scripts

While a general, positive consensus may be reached as to the advisability of including recorded lectures and conversations in listening comprehension tests, the actual form the scripts used might take is a contentious issue. The debate over the use of "oral" or "written" language in creating listening recordings is ongoing. Research intending to determine whether recordings that use spontaneous and impromptu conversations or monologues are easier or more difficult for test subjects to understand than scripts that are planned beforehand and effectively "read" has thus far produced conflicting results. Though Read's study tends to indicate "oral" language may be easier for candidates to understand, he reports that Shohamy's study indicated "written" language was easier for her test subjects to understand (Read, 2002). Since both types of language can be found in IELTS' domain, more research is necessary to determine both the frequency of use of the two and the exact effect the use of "oral" or "written" language has on test-takers ability to complete listening tasks. However, Ferris argues convincingly based on the outcome of her study that "students need to hear authentic lectures" to prepare them for actual lecturers who "mumble, talk quickly, do not provide visuals, and use inaccessible vocabulary or slang" (310, Ferris, 1998).

Perhaps in an attempt to ameliorate calls for the inclusion of either all "oral" or all "written" language, empirical knowledge suggests that IELTS employs a mixture of both scripted and improvised material for the Listening Module. While a concern for organizational characteristics and what IELTS refers to in its website as "constraints of time" restrict their recordings to containing "certain features" of university level materials and not, therefore, necessarily "simulations" as such, a certain amount of realism within listening samples can be discerned. Pauses, corrections, fillers, and natural repetitions occur most notably in conversations but can also be heard in more "lecture" oriented scripts. This would appear to be yet another example of IELTS

compromising between aural input authenticity (“oral” language) and the need to design, as Bachman (1990) describes, accurately organized, lengthy recordings (“written” language).

Accentuated Language

The Handbook emphasizes that since IELTS is expected to be used internationally, “a range of English accents and dialects are used in the recordings” (11, IELTS, 2005). Empirical evidence shows that IELTS does make use of recordings from both native and non-native speakers of English in the Listening Module and the practice seems to cause a fair amount of confusion and additional stress for speakers unaccustomed to the particular accent or dialect being used. While Major, Fitzmaurice, Bunta and Balasubramanian (2002) would view this as lending authenticity to the test, the use of recordings that employ non-native speakers is contentious. Where test takers share the same native language as the speaker in a particular listening section, claims of preferential bias have been made. Whether these claims may be born out remains a difficult question to fully answer and seems to largely depend on individual language groups. While Major et al. report native Spanish speakers may perform better on a listening test that uses a recording of a native Spanish-speaker, their research indicates a similar scenario including Chinese-listeners and speakers resulted in lower scores for the test-takers. “Sometimes” would appear to be the most current, definitive answer as to whether accentuated English has a significant impact on test-taker’s performance in Listening Comprehension tests (185, Major et al., 2002). More research would seem to be in order regarding this issue.

Test Validity

As mentioned before, while a full-scale analysis of the validity of IELTS’ Listening Module is well beyond the scope of this review, I would nonetheless like to briefly introduce statements and research that both support and question how well IELTS can predict the suitability of a candidate for study at English medium universities. This is done within IELTS’ own assertion that correlations between academic success and IELTS scores are “often relatively weak” due to the complex relationship between all factors involved in academic success

(<http://www.ielts.org/teachersandresearchers/faqs/default.aspx#PredictAcademicSuccess>).

While the Assessment Coordinator at the University where I teach responded that IELTS does “reasonably well” in determining candidates’ suitability for study, I focus specifically on the Listening Module below.

Interaction

Although the issue of lack of interaction between candidate and recording has been discussed previously, it merits review here in light of validation. Buck (1997), at least, sees visually responding to recorded input as probably the “most useful” (and certainly expedient) way of assessing a candidate’s broad listening comprehension abilities (71, Buck, 1997). In agreement, Bachman posits that having to immediately interpret and reproduce discourse, exemplified by rapid question answering, is one way to “probe’ the test takers highest level of ability” (135, Bachman, 1990). In addition, regardless of the amount of oral interaction involved, Ferris’ study indicated that both professors and students viewed the ability to take accurate notes as “essential” for success in their courses (Ferris, 1998). While it may be argued that candidates answer more questions than take notes on the aural input given in the Listening Module, an IELTS Coordinator at my own institution suggested “note completion”, as she phrased it, did share some common characteristics with actual note-taking. Thus, it would seem the ability to record non-interactional responses to essentially transactional language is a useful, valid skill that has quantifiable applications beyond the test environment.

However, questions arise over the exact method IELTS uses to ascertain exactly what candidates have comprehended from the given input. A brief evaluation of response formats and test items used to actually score the Listening Module illustrates, yet again, IELTS’ catering to the requirements of a variety of informed stakeholders and test critics that hold disparate views regarding the validity incumbent within particular response formats and test items.

Response Formats and Test Items

Using McNamara’s (2000) definition of both “response format” and “test item”, it may be seen that IELTS uses a significant combination of formats and items related to those formats in the Listening Module. While IELTS’ distinction between “integrated” responses within the Listening and Reading Modules as opposed to the “communicative”, constructed response necessitated within the Speaking and Writing Modules may suggest that the balance of scored responses in the Listening Module could be viewed as “fixed”, it is worth noting that integrated tasks by definition may include moderately high levels of actual response construction (McNamara, 2000). However, MC questions are definitive of a fixed response type item while short answer questions serve as a good example of a constructed item type. Matching is largely a fixed item while sentence completion satisfies most requirements for a constructed item.

As seen below, IELTS' uses a combination of each of these in addition to others to gauge candidates' comprehension.

The validity debate over formats and items

If anything, the research community seems unified by their disagreement as to what constitutes "the" valid test format or item within listening comprehension tests (Brindley & Slater, 2002; Buck & Tatsuoka, 1998; Dunkel, Henning, & Chaudron, 1993; McNamara, 2000; Yi'an, 1998). Perhaps as a method of fending off attack from any number of possible critics, IELTS uses a variety of response formats and items to evaluate the Listening Module. A list, taken from the 2005 Handbook, is provided below:

- multiple choice
- short-answer questions
- sentence completion
- notes/summary/diagram/flow-chart/table completion
- labeling a diagram which has numbered parts
- classification
- matching

(6, IELTS, 2005)

Research indicates that problems can be found with each of these formats and the items used within them. In brief summary; Yi'an suggests MC questions " favored the more advanced listener, but put the less able at a disadvantage; and...allowed much uninformed guessing and resulted in the subjects' giving the correct answers for the wrong reasons" (40, Yian 1998). Bachman seems to share the concern that a "speededness" in having to respond immediately to input "may lead to guessing" (128, Bachman, 1990). Bachman sites Samson, 1983, and Shohamy, 1984, in support of the idea that constructed responses are generally more difficult for test candidates to answer. Brindley and Slayter seem to concur citing a study done by Bern (1993) suggests test candidates "performed better on multiple-choice questions than on either an open-ended or cloze task" but, they include the validity of "short answer questions or information transfer tasks remain largely unexplored" (377, Brindley & Slayter, 2002). In addition, their own study indicated that the use of a "'table' item format" resulted in scores 30% higher than a similar test which employed a sentence completion format (380, Brindley & Slayter, 2002). Read (2002) cites Freedle and Kostin's conclusion that "while characteristics of the test items themselves have some effect, it is text and

text/item overlap variables which are the most significant predictors of item difficulty". Finally, McNamara seems to question the overall validity of expecting candidates to uniformly agree on the salient parts of an aural lecture thereby raising doubt over the general validity of employing constructed responses as tools which serve to answer fixed response items (27, McNamara, 2000). This brief summary illustrates the veritable tip of the iceberg to the debate raging over response and item type.

So, while the research continues to accumulate yet remain, to date, inconclusive, IELTS has opted to include as many response formats and items as possible. The increasingly wide spread use of IELTS seems to indicate this is a profitable course to take. Personally, it is difficult for me to make an informed opinion in the face of such well-argued research. If validity is defined by frequency of use, it may be argued that IELTS' Listening Module is a valid test vehicle. However, a critical approach might argue that few incontrovertible facts could be produced to support IELTS and others' claim that the test is acceptably valid.

Conclusion

I had hoped through the course of this review to provide significant information to either support or refute Yi'an's claim that the language testing community remains "rather ignorant" on their understanding of what constitutes a valid assessment of a test candidate's comprehension of aural input. While IELTS argues somewhat convincingly that they have succeeded in designing a valid testing vehicle, the amount of debate over the validity of response types and items casts a shadow over IELTS claims. Issues of authenticity further muddy waters that should be, close to ten years after the release of IELTS as a benchmark test, by now relatively clear. It may be argued that necessary test practicalities will forever distance the test environment from actual criterion and as such, language comprehension tests will always remain very loose indicators of success within the domain targeted.

Perhaps with an unrealistic rejection of the real constraints test developers face, my concluding assessment of IELTS' Listening Module is that while it may be of use as a very general indicator of academic preparedness, other factors should be equally well taken into consideration before making a definitive decision on a candidate's suitability for a given academic context. An assessment of a candidate's language learning history might well have a part to play in this decision. In addition, technology makes personal interaction between institute and candidate a much more realistic option than

it may have been a decade ago. Experience living in a foreign culture may be a good indicator of a candidate's ability to succeed, as well. While these options may be more time consuming and expensive for institutes to employ, the research indicating limitations within IELTS' test would suggest some type of equally weighted alternate assessment scheme be used to finally evaluate a candidate's possibility of success.

Works Cited

- Bachman, L. (1990). *Fundamental Considerations in Language Testing*. Chap. 5. (pp. 111 - 159). Oxford : Oxford University Press.
- Brindley, G. & Slatyer, H. (2002). Exploring task difficulty in ESL listening assessment. *Language Testing* [Electronic Version], 19(4), 369-394.
- Buck, G. (1997). The testing of listening in a second language. In C. Clapham (Ed.) Vol 7. *Language testing and assessment. Encyclopedia of language and education*. Dordrecht, (pp. 64-74) Netherlands: Kluwer.
- Buck, G. & Tatsuoka, K. (1998). Application of the rule-space procedure to language testing: examining attributes of a free response listening test. *Language Testing* [Electronic Version], 15(2), 119-157.
- Ferris, Dana. (1998). Students' Views of Academic Aural/Oral Skills: A Comparative Needs Analysis. *TESOL Quarterly* [Electronic Version], 32(2), 289-318.
- Dunkel, P., Henning, G. & Chaudron, C. (1993). The Assessment of an L2 Listening Comprehension Construct: A Tentative Model for Test Specification and Development. *Modern Language Journal* [Electronic Version], 77(2), 180-91.
- International English Language Testing System. (2005). IELTS Handbook 2005 [Electronic Version], London: The British Council.
- Major, R. C., Fitzmaurice, S. F., Bunta, F., & Balasubramanian, C. (2002). The Effects of nonnative Accents on Listening Comprehension: Implications for ESL Assessment. *TESOL Quarterly* [Electronic Version], 36(2), 173-190.
- McNamara, T. (2000). *Language Testing*. Oxford, England: Oxford U.P.
- Read, J. (2002). The use of interactive input in EAP listening assessment. *Journal of English for Academic Purposes* [Electronic Version], 1, 105-119. (no page numbers are given in the electronic version so page numbers are not seen when citing material from this source)
- Yi'an, W. (1998). What do tests of listening comprehension test? -A retrospection study of EFL test-takers performing a multiple-choice task. *Language Testing* [Electronic Version], 15(1), 21-44.