

TESTING LISTENING

It may seem rather odd to test listening separately from speaking, since the two skills are typically exercised together in oral interaction. However, there are occasions, such as listening to the radio, listening to lectures, or listening to railway station announcements, when no speaking is called for. Listening may also be tested for diagnostic purposes.

The special problems in constructing listening tests arise out of the transient nature of the spoken language. Listeners cannot usually move backwards and forwards over what is being said in the way that they can a written text. The one apparent exception to this, when a tape-recording is put at the listener's disposal, does not represent a typical listening task for most people.

CONTENT

Operations

Some operations may be classified as *global*, inasmuch as they depend on an overall grasp of what is listened to. They include the ability to:

- obtain the gist;
- follow an argument;
- recognise the attitude of the speaker.

Other operations may be classified whether what is to be understood is explicitly stated or only implied.

Informational:

- obtain factual information;
- follow instructions (including directions);
- understand requests for information;
- understand expressions of need;
- understand requests for help;
- understand requests for permission;
- understand apologies;
- follow sequence of events (narration);
- recognise and understand opinions;
- follow justification of opinions;
- understand comparisons;
- recognise and understand suggestions;
- recognise and understand comments;

- recognise and understand excuses;
- recognise and understand expressions of preferences;
- recognise and understand complaints;
- recognise and understand speculation.

Interactional:

- understand greetings and introductions;
- understand expressions of agreement;
- understand expressions of disagreement;
- recognise speaker's purpose;
- recognise indications of uncertainty;
- understand requests for clarification;
- recognise requests for clarification;
- recognise requests for opinion;
- recognise indications of understanding;
- recognise indications of failure to understand;
- recognise and understand corrections by speaker (of self and others);
- recognise and understand modifications of statements and comments;
- recognise speaker's desire that listener indicate understanding;
- recognise when speaker justifies or supports statements, etc. of other speaker(s);
- recognise when speaker questions assertions made by other speakers;
- recognise attempts to persuade others.

Texts

Text type might be first specified as monologue, dialogue, or multi-participant, and further specified: conversation, announcement, talk or lecture, instructions, directions, etc.

Text forms include: description, exposition, argumentation, instruction, narration.

Length may be expressed in seconds or minutes. The extent of short utterances or exchanges may be specified in terms of the number of turns taken.

Speed of speech may be expressed as words per minute (wpm) or syllables per second (sps). Reported average speeds for samples of British English are:

	wpm	sps
Radio monologues	160	4.17
Conversations	210	4.33
Interviews	190	4.17

(Tauroza and Allison, 1990)

Dialects may include standard or non-standard varieties.

Accents may be regional or non-regional.

If authenticity is called for, the speech should contain such natural features as assimilation and elision (which tend to increase with speed of delivery) and hesitation phenomena (pauses, fillers, etc.).

Intended audience, style, topics, range of grammar and vocabulary may be indicated.

SETTING THE TASKS

Selecting samples of speech (texts)

Passages must be chosen with the test specifications in mind. If we are interested in how candidates can cope with language intended for native speakers, then ideally we should use samples of authentic speech. These can usually be readily found. Possible sources are the radio, television, spoken-word cassettes, teaching materials, the Internet and our own recordings of native speakers. If, on the other hand, we want to know whether candidates can understand language that may be addressed to them as non-native speakers, these too can be obtained from teaching materials and recordings of native speakers that we can make ourselves.

If recordings are made especially for the test, then care must be taken to make them as natural as possible but it is better to base the passage on a genuine recording, or a transcript of one. If an authentic text is altered, it is wise to check with native speakers that it still sounds natural. If a recording is made, care should be taken to ensure that it fits with the specifications in terms of speed of delivery, style, etc.

Writing items

For extended listening, such as a lecture, a useful first step is to listen to the passage and note down what it is that candidates should be able to get from the passage. We can then attempt to write items that check whether or not they have got what they should be able to get. This note-making procedure will not normally be necessary for shorter passages, which will have been chosen (or constructed) to test particular abilities.

In testing extended listening, it is essential to keep items sufficiently far apart in the passage. If two items are close to each other, candidates may miss the second of them through no fault of their own, and the effect of this on subsequent items can be disastrous, with candidates listening for 'answers' that have already passed. Since a single faulty item can have such an effect, it is particularly important to trial extended listening tests, even if only on colleagues aware of the potential problems.

Candidates should be warned by key words that appear both in the item and in the passage that the information called for is about to be heard. For example, an item may ask about 'the second point that the speaker makes' and candidates will hear 'My second point is . . . '.

The wording does not have to be identical, but candidates should be given fair warning in the passage. It would be wrong, for instance, to ask about 'what the speaker regards as her most important point' when the speaker makes the point and only afterwards refers to it as the most important. Less obvious examples should be revealed through mailing.

Other than in exceptional circumstances (such as when the candidates are required to take notes on a lecture without knowing what the items will be, see below), candidates should be given sufficient time at the outset to familiarise themselves with the items.

Possible techniques

Multiple choice

There is the problem of the candidates having to hold in their heads four or more alternatives while listening to the passage and, after responding to one item, of taking in and retaining the alternatives for the next item. If multiple choice is to be used, then the alternatives must be kept short and simple.

Short answer

This technique can work well, provided that the question is short and straightforward, and the correct, preferably unique, response is obvious.

Gap filling

This technique can work well where a short answer question with a unique answer is not possible.

Information transfer

This technique is useful in testing listening since it makes minimal demands on productive skills. It can involve such activities as the labelling of diagrams or pictures, completing forms, making diary entries, or showing routes on a map. *Note taking*

Candidates take notes during the talk, and only after the talk is finished do they see the items to which they have to respond. When constructing such a test, it is essential to use a passage from which notes can be taken successfully.

Partial dictation

While dictation may not be a particularly authentic listening activity (although in lectures at university, for instance, there is often a certain amount of dictation), it can be useful as a testing technique. As well as providing a 'rough and ready' measure of listening ability, it can also be used diagnostically to test students' ability to cope with particular difficulties (such as weak forms in English).

Since it is listening that is meant to be tested, correct spelling should probably not be required for a response to be scored as correct. However, it is not enough for candidates simply to attempt a representation of the sounds that they hear, without making sense of those sounds. To be scored as correct, a response has to provide strong evidence of the candidate's having heard and recognised the missing word, even if they cannot spell it. It has to be admitted that this can cause scoring problems.

Presenting the texts (live or recorded?)

The great advantage of using recordings when administering a listening test is that there is uniformity in what is presented to the candidates. This is fine if the recording is to be listened to in a well-maintained language laboratory or in a room with good acoustic qualities and with suitable equipment (the recording should be equally clear in all parts of the room). If these conditions do not obtain, then a live presentation is to be preferred. If presentations are to be live, then greatest uniformity (and so reliability) will be achieved if there is just a single speaker for each (part of a) test. If the test is being administered at the same time in a number of rooms, more than one speaker will be called for. In either case, a recording should be made of the presentation, with which speakers can be trained, so that the intended emphases, timing, etc. will be observed with consistency. Needless to say, speakers should have a good command of the language of the test and be generally highly reliable, responsible and trustworthy individuals.

Scoring the listening test

It is probably worth mentioning again that in scoring a test of a receptive skill there is no reason to deduct points for errors of grammar or spelling provided that it is clear that the correct response was intended.

References

Arnold, J. 2000. Seeing through listening comprehension anxiety. *TESOL Quarterly* 34: 777-786

Buck, G. 1991 The testing of listening comprehension: an introspective study. *Language Testing* 8:67-91

Buck, G. and K. Tatsuoka. 1998. Application of the rule-space procedure to language testing: examining attributes of a free response listening test. *Language Testing* 15:119-157

Hughes, A. 2003. *Testing for Language Teachers*, Cambridge University Press, 160-171