

7

Oral production tests

7.1 Some difficulties in testing the speaking skills

Testing the ability to speak is a most important aspect of language testing. However, at all stages beyond the elementary levels of mimicry and repetition it is an extremely difficult skill to test, as it is far too complex a skill to permit any reliable analysis to be made for the purpose of objective testing. Questions relating to the criteria for measuring the speaking skills and to the weighting given to such components as correct pronunciation remain largely unanswered. It is possible for people to produce practically all the correct sounds but still be unable to communicate their ideas appropriately and effectively. On the other hand, people can make numerous errors in both phonology and syntax and yet succeed in expressing themselves fairly clearly. Furthermore, success in communication often depends as much on the listener as on the speaker: a particular listener may have a better ability to decode the foreign speaker's message or may share a common nexus of ideas with him or her, thereby making communication simpler. Two native speakers will not always, therefore, experience the same degree of difficulty in understanding the foreign speaker.

In many tests of oral production it is neither possible nor desirable to separate the speaking skills from the listening skills. Clearly, in normal speech situations the two skills are interdependent. It is impossible to hold any meaningful conversation without understanding what is being said and without making oneself understood at the same time. However, this very interdependence of the speaking and listening skills increases the difficulty of any serious attempt to analyse precisely what is being tested at any one time. Moreover, since the spoken language is transient, it is impossible without a tape recorder to apply such procedures as in the marking of compositions, where examiners are able to check back and make an assessment at leisure. The examiner of an oral production test is working under great pressure all the time, making subjective judgements as quickly as possible. Even though samples of speech can be recorded during a test, the tape-recording, by itself, is inadequate to provide an accurate means of reassessing or checking a score, since it cannot recapture the full context of the actual situation, all of which is so essential to any assessment of the communication that takes place.

Yet another (though not insuperable) difficulty in oral testing is that of administration. It is frequently impossible to test large numbers of students because of the limited time involved. It is not difficult to appreciate the

huge problems relating to a test situation in which thousands of students have to be examined by a handful of examiners, each student being tested for a period of, say, ten or fifteen minutes. Although the use of language laboratories for such tests has made it possible in some cases to administer more reliable oral production tests to large numbers of students, the actual scoring of the tests has not been so easily solved.

Excluding tests of reading aloud and one or two other similar tests, oral tests can have an excellent backwash effect on the teaching that takes place prior to the tests. For example, in one country the oral test was retained as part of a school-leaving examination simply to ensure that at least some English would be spoken in the last two years of the secondary school – even though the test itself was considered an unreliable measuring instrument as a result of the large number of unqualified examiners who had to administer it. For this reason, and indeed, because oral communication is generally rated so highly in language learning, the testing of oral production usually forms an important part of many language testing programmes.

The following sections in this chapter will give an idea of the range of possible types of oral tests. Some of the exercises (e.g. picture descriptions) have proved very useful in many tests while others (e.g. pencil-and-paper tests) have met with varying degrees of success. In spite of its high subjectivity, an extremely good test is the oral interview. In many cases, one or two sub-tests (or oral activities) are used together with the oral interview to form a comprehensive test of oral production skills.

7.2 Reading aloud

Many present-day oral tests include a test of reading aloud in which the student is given a short time to glance through an extract before being required to read it aloud. The ability to read aloud differs greatly from the ability to converse with another person in a flexible, informal way. Although reading aloud may have a certain usefulness, only a few newsreaders and teachers may ever require training and testing in this particular skill. The majority of students will never be called on to read aloud when they have left school. It is a pity, therefore, that students are required to sacrifice their enjoyment of silent reading in order to practise reading aloud. We read primarily for information or enjoyment, and the silent reading skills so necessary for this purpose differ greatly from those of reading aloud. The backwash effects of this kind of test may be very harmful, especially in areas where the reading skills are misguidedly practised through reading aloud. Finally, how many native speakers can read aloud without making any errors?

Tests involving reading aloud are generally used when it is desired to assess pronunciation as distinct from the total speaking skills. In order to construct suitable tests of reading aloud, it is helpful to imagine actual situations in real life in which the testees may be required to read aloud. Perhaps one of the most common tasks is that of reading aloud directions or instructions to a friend, colleague or fellow-worker: e.g. how to wire a plug, how to trace faults in a car engine, how to cook certain dishes. For example, the following instructions relate to a situation in which a teacher or class monitor may be asked to read aloud:

First put the headset on. Make sure it is in its most comfortable position with the headband over the centre of the head. The microphone should be about 1½ inches from the mouth.

To record, put the white switch to the position marked *Work*. Put the red switch to *Speak* and press the red recording button, which will now light up.

(etc.)

Another situation which might occur in real life is that in which the student is asked to read aloud (part of) a letter he has received. For all the different extracts, however, it is advisable to draw up certain features which must be included in each passage: e.g. one Yes/No question, one Wh- question, two sentences each containing a subordinate clause, one question tag, the phoneme contrasts i:-ɪ, p-b, ɒ-ɔ:, etc. In this way, some degree of consistency can be achieved.

A test more useful in many ways than reading aloud is the retelling of a short story or incident. In this type of examination, the students are required to retell a story they have just read. If carefully constructed, such a test can assess most of the phonological elements which are otherwise tested by reading aloud. Unfortunately, it often measures other skills such as reading comprehension, memory and organisation, too.

7.3 Conversational exchanges

These drills are especially suitable for the language laboratory and can serve to focus attention on certain aspects of the spoken language, especially in those countries where English is taught as a foreign language and the emphasis is primarily on the reading skills. However, several of the test items themselves are far from communicative in any sense at all and do not allow for authentic interaction of any kind. The essential element of constructive interplay with unpredictable stimuli and responses is absent from all these items as a result of the attempt to control the interaction taking place. The item types range from items presenting the testees with situations in which they initiate conversations to incomplete conversations with the part of one speaker omitted (i.e. a one-sided dialogue). Tests containing such item types are on the whole reliable, but they cannot be described as being valid tests of speaking. If an opportunity is provided in other parts of the test for real oral interaction (i.e. genuine conversation and discussion), however, these controlled test items can be of some use in directing the attention of the students to specific language areas and skills.

Type 1 The testees are given a series of situations and are required to construct sentences on the lines of a certain pattern or group of patterns. Again, it is essential that two or three models be given to the testees so that they know exactly what is required. (The testees read or hear the situation and then make the appropriate responses, shown in the brackets.)

Examples:

Mrs Green lives in a flat. She doesn't like living in a flat and would like to live in a small house with a garden. (*She wishes she lived in a small house with a garden.*)

It's raining heavily. Tom and Anna are waiting impatiently at home to set off on their picnic. (*They wish it would stop raining.*)

1. Mr Black has a small car but his neighbours all have large cars. He would like a large car, too.
2. Anna hasn't learnt how to swim yet but most of her friends can swim.
3. Tom is waiting for Bill outside the cinema. The show is just about to start but Bill has not arrived yet.

4. Mrs Robinson doesn't like living in towns; she wants to live in the country.
(etc.)

Type 2 This type of test item is similar to the previous type but not as strictly controlled.¹ No model-responses are given by the examiner and the students are free to use whatever patterns they wish.

A friend of yours has forgotten where he has put his glasses. He cannot see too well without them. What will you say to him? (*Let me help you to look for them, etc.*)

You are on your way to school when it starts to rain heavily. Unfortunately, you and your friend have no raincoats. There is nowhere to shelter but your school is only a hundred yards away. What do you say to your friend? (*Shall we make a dash for it?/Let's run the rest of the way.*)

1. You are trying to get to the public library but you are lost. Ask a police officer the way.
2. Your friend has just returned from a holiday abroad. What do you say to him?
3. A waitress has just brought you the bill but has totalled it up incorrectly. What do you say to her?
4. A friend of yours wants to see a film about a murder. You have already arranged to see it another evening, but you know she would be hurt if she knew. Make up an excuse.

Type 3 The students hear a stimulus to which they must respond in any appropriate way.¹ (This test often relies on conventional greetings, apologies, acceptable ways of expressing polite disagreement, etc.)

Do you mind if I use your pencil for a moment?
(*Not at all/Certainly/Please do/Go ahead, etc.*)

What about a game of tennis?
(*Yes, I'd love a game/All right. I don't mind/Don't you think it's a bit too hot?, etc.*)

1. Please don't go to a lot of trouble on my behalf.
2. Oh dear, it's raining again. I hope it stops soon.
3. We shan't be late, shall we?
4. Karen asked me to say she's sorry she can't come tonight.

Type 4 This is similar to the previous type of item, but the stimuli and responses form part of a longer dialogue and the situation is thus developed. Because of its total predictability, however, this type of item is sometimes referred to as a dialogue of the deaf! The man in the dialogue below continues regardless of what the testee says.

You're on your way to the supermarket. A man comes up and speaks to you.

MAN: Excuse me. I wonder if you can help me at all. I'm looking for a chemist's.

PAUSE FOR TESTEE'S REPLY

MAN: Thank you. Do you know what time it opens?

PAUSE FOR TESTEE'S REPLY

MAN: Thanks a lot. Oh, er, by the way, is there a phone box near here?

PAUSE FOR TESTEE'S REPLY

MAN: Oh dear, I'll need some coins. Do you have any change for a £5 note?

PAUSE FOR TESTEE'S REPLY

MAN: Well, thanks a lot. You've been most helpful.

This dialogue clearly becomes absurd if, when asked where there is a chemist's, the testee replies, 'I'm sorry, I don't know,' and the man promptly thanks him and asks what time it opens. Nevertheless, the use of pre-recorded material of this kind makes it possible to use the language laboratory to test large numbers of students in a very short time.

Type 5 This item² takes the form of an incomplete dialogue with prompts (shown in brackets in the following example) whispered in the student's ear.

You are at the reception desk of a large hotel. The receptionist turns to address you:

RECEPTIONIST: Can I help you?
(You want to know if there is a single-room available.)

YOU:

RECEPTIONIST: Yes, we have a single room with an attached bathroom.
(Ask the price.)

YOU:

RECEPTIONIST: Thirty-four pounds fifty a night.
(You want to know if this includes breakfast.)

YOU:

RECEPTIONIST: Yes, that's with continental breakfast.
(You have no idea what 'continental breakfast' is.)

YOU:

RECEPTIONIST: It's fruit juice, coffee or tea and bread rolls.
(The receptionist is speaking too quickly. What do you say?)

YOU:

RECEPTIONIST: Fruit juice, coffee or tea, and bread rolls.
(Book the room for two nights.)

YOU:

RECEPTIONIST: Certainly. Room 216. The porter will take your bag and show you where it is.
(Thank the receptionist.)

YOU:

7.4 Using pictures for assessing oral production

Pictures, maps and diagrams can be used in oral production tests in similar ways to those described in the previous chapter on testing the listening skills. Pictures of single objects can be used for testing the production of significant phoneme contrasts, while a picture of a scene or an incident can be used for examining the total oral skills. This section will concentrate on the use of pictures for description and narration.

The students are given a picture to study for a few minutes; they are then required to describe the picture in a given time (e.g. two or three minutes). Occasionally, the number of words each student speaks is counted by one examiner in the room, while the other examiner counts the number of errors made. The score is thus obtained on the basis of the number of words spoken and the errors made (but this procedure is very unreliable.) Separate scores for general fluency, grammar, vocabulary,