ЧУЖДОЕЗИКОВО ОБУЧЕНИЕ

FOREIGN LANGUAGE TEACHING

TESTING ORAL SKILLS: ELICITATION TECHNIQUES, CERTIFICATE EXAM FORMATS, ISSUES AND CHALLENGES

Silvia Apostolova[1]

Abstract: *Out of the four language skills, speaking is thought to be the most difficult to teach and test. The reason for this is the nature of the speech act and the way foreign languages are learned away from the native speaking country. There are also other factors to add, such as personality, memory, and talent for the language. Testing the skills properly is time-consuming and difficult, and there are a number of elicitation techniques that can be used both to teach and test speaking skills. These techniques are suitable for the classroom and for official certification. World-renowned certification examiners have come up with specific techniques to measure foreign language oral abilities at given proficiency levels. They are now used as part of the enrolment process in many universities.*

Keywords: *speech act; elicitation techniques; speaking skills; certificate formats.*

Introduction

With the development of new approaches for teaching a foreign language, more and more comprehensive tests that measure the level of proficiency of all four language skills have been developed. Still, the skills regarding oral abilities are, to this day, the hardest to test and assess accurately. There are many factors that affect speech besides good command of the language: the type of speech act (small talk or a public speech), personal eloquence and presentation skills, topical knowledge, background, tiredness, daily routine, personality, the conversation partner, etc. And, in a test environment, there is also the factor *stress of being tested*. Moreover, every assessor (if more than one is involved) can award different scores to one and the same person, depending on their personal likes and dislikes and on the general impression made by the person being tested. This ambiguity poses an additional challenge to the assessment accurateness.

---

[1] **Silvia Apostolova** – PhD student at St. Cyril and St. Methodius University of Veliko Tarnovo (Faculty of Modern Languages, Department of Language and Literature Teaching Methodology and Literary Theory) and English teacher at the Vocational School of Mechanical and Electrical Engineering in Lovech; email: d1277@sd.uni-vt.bg

Main Issues
Testing Speaking

When it comes to oral skills testing, one of the main difficulties is that there is hardly any elicitation technique that evaluates the level of proficiency only of the speaking skill, i.e. integrated skills are tested most of the time, such as listening and speaking or reading and speaking, as there has to be some kind of a stimulus for the test taker to produce speech. For example, if the test participants are involved in a discussion (one of the least assessor-controlled elicitation techniques), each person's listening skills are correlated with his/her good performance on the test; there is a constant change between being a listener and being a speaker, and speaking appropriately depends on comprehending the message being spoken. So, good listening skills necessarily become a factor in the testing of speaking. As Brown and Abewichrama (2019) wrote:

Although speaking is a productive skill that can be directly and empirically observed, those observations are invariably colored by the accuracy and effectiveness of a test-taker's listening skill, which necessarily compromises the reliability and validity of an oral production test. (p. 156)

According to them, it is possible to isolate and test only listening, whereas this is impossible for the speaking part. Moreover, there are very limited tasks where the oral proficiency can be assessed without the aural participation of an interlocutor.

Elicitation Techniques

The design and selection of elicitation techniques is yet another challenge simply because, when speaking, every person uses the grammatical structures and vocabulary that he/she is most comfortable with, which prevents the assessor from having an objective view of the exact language knowledge of the test taker. To induce the test taker to use certain structures, specially designed tasks must be chosen, and most of the time this is not so simple.

Assessment and Assessment Scale

The scale of assessing oral skills is not that easy to compile as well, and the score given often depends completely on the assessor's personal opinion. For instance, there are no True/False or multiple-choice tasks where all assessors would award exactly the same score to one and the same task. The judgement here is more subjective and sometimes requires mediation, just like in the assessment of writing skills. One of the most effective ways to get around this problem is to have more than one assessor. They may be present in the test or mark from recorded tapes, or there may be a mixture of these – one assessing live and one from a record. A very important aspect here is that all assessors must be specifically trained for their role in the examination.

As for the marking system, assessors can use a holistic or analytic scale, i.e. they can give a single mark based on the overall performance (holistic) or an average mark based on the separate marks for the various categories of language proficiency: grammar, pronunciation, fluency, content, etc. (analytic). Barry O'Sullivan (2008) conducted a research test where students sat a Cambridge First Certificate in English (FCE) examination. They were first given a single mark according to the overall impression of their speaking skills (a holistic mark). Then, they were tested again and given marks on each category separately – pronunciation, grammar, vocabulary, fluency, and comprehension; an average mark based on the scores in each category (an analytical mark) was calculated. When both marks were compared, it turned out that the holistic and analytic were the same. Therefore, according to O'Sullivan, the students' achievements were exactly the same no matter what scale was used.

Nunan (1999) pointed out that good speaking skills require someone to be linguistically competent in terms of articulating the sound well, having sufficient vocabulary, and mastering structural or grammatical components. He went further and added other skills involved in the speaking proficiency, such as answering questions completely and logically, being able to repair the conversation after a breakdown, and demonstrating sociolinguistic/cultural competence (i.e. using the language appropriately to the context). The latter is hard to be included in a general format scale. So, most world-renowned certificate

exams like Cambridge or TOEFL iBT® have a scale where separate scores are awarded for grammar, pronunciation, fluency, content, organization, and vocabulary.

Speaking Tasks

There are different categories of speaking tasks that can be included in tests. Brown and Abewick-rama (2019: 157–158) distinguish five types of speaking tasks:

Imitative – test takers are required to copy a word, phrase, or sentence. Pronunciation is the main aspect of the assessment although grammar is also considered in the scoring criteria. The communicative competence of the language is not essential here. The listening and short-term storage abilities are key factors for success.

Intensive – pronunciation or phonology are not emphasized. The tasks consist of producing short pieces of oral language designed to demonstrate proficiency in narrow grammatical, phrasal, lexical, or phonological terms (e.g. intonation, stress, rhythm, juncture, etc.). The interaction with the interlocutor or test administrator is kept at minimum. Examples of such tasks are reading aloud, sentence and dialogue completion, picture-cued storytelling, and translation up to a simple sentence level.

Responsive – the test taker is stimulated to speak promptly in very short conversations, standard greetings and small talk, simple requests, comments, etc., with only one or two follow-up questions or comments.

Interactive – compared to the responsive tasks, the test taker here is expected to produce long and complex sentences and dialogues/conversations with multiple exchanges and sometimes multiple participants. This category can be further broken down into two types:

a) *transactional language*, the purpose of which is to exchange specific information;

b) *interpersonal exchanges*, the purpose of which is to maintain social relationships – small talk, using humour, slang, or colloquial language.

Extensive – this type of speaking tasks involves a wide range of speech production: speech, presentation, storytelling, etc. During the oral production of speech, interruption from listeners is either highly limited or ruled out altogether. Such speaking tasks usually require advance planning and preparation. This type of speaking can be referred to as the ultimate speaking skill and relies on strong language competencies.

Oral Production Elicitation Techniques

There is such an array of oral production tasks that a complete list is hard to be compiled. Some techniques resemble others and there is only a small difference between them; so, in the descriptions below, they are put down as a variation. Reviewing academic literature leads the author of this paper to believe that there is no natural classification of speaking test techniques. In some cases, it is apparent that the techniques call on other, more general skills, as well as on oral proficiency like general knowledge of the language in combination with general knowledge on a topic, culture, the world, especially when it comes to discussions and conversations which are classified as authentic communicative tasks. Special attention should be paid to the scoring system, so that the oral proficiency is promptly rewarded and not the general presentation skills of the test taker.

The elicitation techniques below follow the order of decreasing predictability of response and direction from most assessor-controlled to least assessor-controlled test tasks (Underhill 1987):

A. Sentence Repetition

The test taker hears a series of sentences or utterances and repeats them as accurately as possible. Sentence length can vary from one word to fifteen or twenty. With short and grammatically simple sentences, short-term memory and recall by imitation are often sufficient, but up to a certain length.

Advantages of the technique:
- quick and effective;
- easy to score – right and wrong answers.

Disadvantage:

- the tasks are neither authentic in terms of language, nor communicative.

A variation can be **repeating sentences of increasing length or within a specific language area** (specific items of vocabulary, points of expected first-language interference, phonetically similar words, different stress and intonation patterns) (Underhill 1987: 86–87).

B. Sentence Transformation

The test taker is asked to manipulate sentences to demonstrate knowledge of specific language structures. He/she is given a stimulus sentence, usually in writing, and is asked to transform it into a different grammatical pattern. Usually, there is an example of the pattern to follow.

Transformations include tense or concord changes, active-passive voice transformations, or sentence-joining tasks. Further difficulty can be added if the test taker is asked to transform the sentence by beginning with certain words or a phrase.

The advantages and disadvantages of this elicitation technique are similar to the previous one. The tasks are neither authentic, nor communicative, but the technique does allow rapid testing of particular structural areas and an estimation of the test taker's ability to correct himself/herself (Underhill 1987: 84–85).

C. Sentence Correction

The test taker is given a sentence, orally or in writing, which contains an error. The task is to identify the error and provide a correct version. The errors may be of different kinds graded by level of difficulty: grammatical errors such as concord, tense, choice of prepositions or word order are relatively easy, other errors such as use of idioms or collocations are more difficult (Underhill 1987: 84).

To make the task more difficult, a sentence with no errors at all can be included. Also, the test taker can be asked to provide more than one answer.

D. Sentence Completion from an Aural or Written Stimulus

A series of sentences is prepared with one, two, or a few words missing from each, for example in dialogue form. The test taker has the task to read through the sentences, one at a time, and suggest a possible way of completing each one. Some preparation time can be given if needed, but for checking oral ability at a more advanced level it may be better not to allow any preparation (Underhill 1987: 81–83).

One of the variations is to use **conventional cloze or multiple-choice tests**. The tests can be aimed at testing particular structures or vocabulary items. The interlocutor can choose different questions for each test taker in order to prevent the test from being compromised. The advantage of doing such a test orally is that it offers the opportunity for self-correction and allows the interlocutor to find out why the test taker has chosen a particular wrong answer. This is a suitable technique for diagnostic testing.

Another variation of the technique is to ask the test taker to **complete a text where the gaps are sentences** instead of words. This is a more open-ended task. The stimulus sentences are carefully prepared to require comprehension and production of discourse reference markers.

Variation 3: **using spoken cues.** The interlocutor offers a sentence to be completed, e.g. "I'm sorry, I can't join you for lunch, but…", and asks the test taker to finish it. The task is more difficult than with a written cue. After the first answer is given and marked as right or wrong, the interlocutor can move on to the next sentence.

The task can be made more difficult if the test taker is asked for **appropriate responses to questions** (Underhill 1987: 59–60). The learner is given a few short, unrelated situations that occur in everyday life. He/she may be given the situations written out on a sheet of paper, they can be read to him/her, or both in combination. Then, the test taker is asked to imagine himself/herself in each situation and to provide what he/she thinks would be an appropriate spoken response. The situations should be chosen such that are easily described and understood and require an answer of one or two sentences at the most. The aim is to check the test taker's command of functional language like asking for information, apologizing, accepting or refusing invitations, etc.

For more advanced speakers, the technique can be further developed into a **Question-and-Answer** session (a variation) (Underhill 1987: 61–63). This consists of a series of unrelated questions which are graded in order of increasing difficulty, starting with short and simple questions, such as "What is your name?" and "Where do you come from?", and ending with long and complex questions like "If you had not been here this morning, what would you have been doing instead?". Since the questions are not arranged in a logical order, this makes the technique suitable for tests carried out in language laboratories. It should be made clear, though, how much time the test taker has to answer each question, and that he/she is to speak for the full amount of time. There is no limitation on the types of questions that can be included – yes/no questions, question tags, descriptive questions, hypothetical questions, justification-of-opinion questions, or questions only for specific grammatical structures.

E. Translating and Interpreting

Both the interlocutor and the test taker have a native-language text in front of them which the latter may be or may be not familiar with. The interlocutor chooses a short passage or series of passages from the text and asks the test taker to translate it into the foreign language. The direction of the translation can also be from foreign into native language, or the directions of translation can change during the test to increase task difficulty. In the scoring system, there has to be a mark for accuracy and correctness of the translation, as well as for style and feeling for the original (Underhill 1987: 79–80).

Translation is often regarded as old-fashioned and unsuitable for use in tests as it relates to the direct method of language teaching. Translating well is a different skill from speaking well in the foreign language and, on its own, it does not make a satisfactory test of oral ability. However, it is a quick and easy test to administer and can conveniently be planned as part of a sequence of test techniques.

F. Retelling a Story from a Written Stimulus

The test taker reads a passage or series of short passages and is asked to retell what he/she has learned in his/her own words immediately afterwards. The test taker is not allowed to refer to the written text once the retelling has begun. It is possible to delay the recall by some question-and-answer activities between the reading and retelling stages to test the test taker's memory and mental organization (Underhill 1987: 73–75).

This technique can be used at all levels. At lower levels, it may be necessary to offer help with vocabulary and to check the understanding, so that the success in retelling is not just a reflection of the degree of comprehension. As an authentic activity, it necessarily involves comprehension and processing skills as much as speaking skills.

At elementary and intermediate levels, instructions can be given to keep the retelling concise and to point out only the important points in order to discourage test takers to repeat words and phrases in a parrot-like fashion.

At higher levels, emphasis can be put on the use of conjunctives and other sentence-connecting devices and on rewarding the use of words, not present in the original text. Accurate and natural use of this discourse vocabulary is a skill that distinguishes fluent speakers from advanced learners.

The skills being tested here are again not purely speaking.

A variation of the technique is to **use note cards from which test takers have to produce a short exposition or description** – a mini-presentation, anecdotes, or narration. The marking system has to allow for fluent but less imaginative test takers to be awarded more generously.

G. Giving Instructions/Description/Explanations

The test taker has to describe an object, a system, or an everyday procedure at some length (six to ten sentences). For this he/she has little time to prepare (Underhill 1987: 69–70). The topics should be carefully selected, i.e. the objects or procedures should be well-known, so that the test taker is encouraged to produce connected discourse. At the same time, the topic has to allow considerable freedom of choice of expression without requiring extensive preparation. Here are some examples of suitable topics:

- How do you prepare Spaghetti Bolognese?

- Describe your room.
- Give instructions for using a top-up phone voucher.

A variation of the task is to **ask** test takers **about their personal opinions or attitudes** on popular topics or an **explanation of job-related or academic terms**.

H. Using a Picture or Picture Story

Before the tests start, the test taker is given a picture or a sequence of pictures to look at. Then, the interlocutor asks him/her to describe the picture or to make up the story at his/her own speed and time and with his/her own words of choice. The interlocutor may ask questions aimed at focusing on a particular piece of information or further elaborating a point that may have been missed or not made clear. This could be further developed into a discussion as the interlocutor asks general questions intended to elicit the test taker's attitudes or opinions on a topic related to the subject of the picture, but not directly about it (Underhill 1987: 66–69).

A picture story usually consists of four to twelve drawings which tell a story. The story should be simple and allow the test taker to add his/her own interpretation of the people or events involved. The pictures are specially selected to exclude objects or actions that are difficult to describe or are very culture specific.

Advantages

Visual stimuli are taken in quicker than written ones, so even sophisticated topics are easy to be introduced. Since the stimuli are the same, the speech produced by the test takers is directly compared, and it is easy to distinguish the best communicators. Still, the language used by different test participants is not completely predetermined, and there is a lot of room for personal interpretation.

Disadvantages

There is the danger that the test taker will miss the point of a picture or story for personal or cultural reasons. As no vocabulary items are given, those with richer vocabulary will clearly have a huge advantage. And this can be quite demoralizing to test takers with a poorer vocabulary. Also, it is quite often difficult to find suitable materials for the test, particularly in situations where the test security can be compromised, so several alternative pictures/picture stories are needed to be used in rotation.

I. Interview

The interview is the most popular of all oral tests. It is a direct, face-to-face oral exchange between the test taker and the interviewer (interlocutor) (Underhill 1987: 54–56). It follows a predetermined pattern, but still allows both people a degree of freedom to say what they genuinely think. Unlike discussions, the interview is structured and the interviewer maintains firm control of the question-and-answer process – whatever the test taker says is more or less a direct response to the questions asked. However, the test taker still has the freedom to answer as he/she likes, or to develop his/her comments and opinions. When the topic is exhausted, it is up to the interviewer to move forward with the questions. There may be several topics raised in an interview, but each is explored in enough detail with follow-up questions that allow the test takers to develop it and to show their proficiency of the language.

The oral interview technique is well suited for testing learners at the intermediate level and below, where, with detailed rating scales, it is easy to distinguish the test takers' skills and level of proficiency. It often fails to discriminate effectively at higher levels of language proficiency.

A tight-controlled interview will not easily elicit the test taker's best language performance, and it is therefore more important at higher level that the interviewer stand back a little and allow the participant to demonstrate his/her conversational fluency skills like filling conversational pauses naturally and correcting one's own errors. These are two of the features that distinguish higher-level from intermediate-level foreign language speakers.

J. Role-Play

An imaginary situation is presented to the test taker where he/she has to take on a particular role and converse with the interlocutor in an appropriate way (Underhill 1987: 51–54). The situation could cover all topics from our daily lives or formal and informal events, particular language functions, hobbies, etc. The technique can also be used to check for specific vocabulary items or to check the test taker's ability to get around in unfamiliar situations, such as the following ones:

- Two people meet at a party and introduce each other;
- A policeman and a member of the public reporting an accident to him;
- A shop manager and a complaining customer;
- A hotel receptionist and a tourist asking for landmarks and places of interest;
- A doctor and a patient with a problem.

The test taker is usually given a set of instructions, just before the test, that explain in simple language exactly what he/she is expected to do. For students at a higher level of proficiency, the instructions may be expressed in terms of the general situation.

For example, *Imagine you have found a job in another city away from your parents. You have a friend at work who wants to visit your hometown. Talk to your workmate and recommend to him/her a place to stay at, how to get there and what to see.*

The instructions may be made more specific like particular cities or sites to visit, or problems the person is experiencing, or more general like a given country or a favourite cuisine.

Through this technique, the ability to ask questions and to narrate is tested. It can also be applied to small talk or everyday language, to more formal or academic situations.

If test takers are unfamiliar with role-playing or their level of language proficiency is elementary or intermediate, the procedure and purpose should be well explained beforehand, and the instructions for each role-play should be given in writing (in the native language if necessary). The interlocutor has to make sure the test taker understands the task and what is required that they do before starting the test.

Personal reluctance to participate has always been a possible problem. By definition, role-playing implies pretending to be someone else and many people can be uncomfortable with this. In certain cultures, role-playing is neither well accepted, nor a usual way to behave, so students will be personally unhappy to pretend to be someone they are not. Moreover, if the role-play is between two test takers with a different level of oral proficiency, the person with poorer command may feel threatened and reluctant to participate.

K. Oral Report

The test taker is given a topic and prepares an oral presentation lasting from five to ten minutes. He/she is expected to refer to notes, but reading aloud is not desirable. The use of technical aids, such as multimedia devices, an interactive board, a white board, or a flipchart, is welcome if appropriate. At the end of the presentation, the speaker has to face questions from the interlocutor. Making presentations is an authentic and communicative activity for both professional and academic purposes (Underhill 1987: 47–49).

In formal test procedure circumstances, the test taker makes the presentation directly to the interlocutor. In less formal situations, mini-presentations may be a routine part of the daily teaching in the classroom and be used for testing purposes at the same time. The presentation can be taped either for marking or for follow-up classroom analysis.

Choosing the topic is very important. It should neither be so specialized that only the speaker him-/herself is interested, nor should it be so general that it has no other service but as a language exercise. Ideally, the student should choose the topic after consulting his/her teacher who will help match his/her ability with the difficulty of the topic. In formal test situations, however, this is not possible. The assessor has to be careful and take into consideration whether the topic is a difficult one in itself, irrespective of the fluency of the speaker.

When marking presentations, specific mark categories need to be introduced – e.g. explaining factual data, expressing opinions or arguments, dealing with questions, summarizing, etc. If there is an

audience, then the assessor can mark the impression of the presentation on the listeners in terms of comprehension, reaction, and answering questions from the audience.

L. Discussions and Conversations

The test taker is given a topic on which his/her opinion is needed. So, the conversation starts with short statements from the participants sharing views and thoughts (Underhill 1987: 45–46).

This task is authentic and communicative as it is the most natural thing in the world – two people having a conversation on a topic of common interest. It is also the hardest to be carried out within the framework of a formal language test; it is necessary that both parties are relaxed and confident, and then the activity (a conversation) is allowed to become dominant, while its actual purpose (the language test) is temporarily subordinated.

A variation is to start a **joint discussion** or **decision-making** on various issues among a few test takers, i.e. a group discussion. The organization and the rules are slightly different in comparison with a discussion with just one test taker.

During the discussion, the interlocutor maintains overall control but is willing to delegate the leading position to one or more of the test participants from time to time to steer the conversation or bring up a new topic. There is a lot more to watch out for in this technique than just the logical strings of words – e.g. tone of voice, pitch and intonation, face expressions and body language, which together work and contribute to the successful negotiation. These are features of natural conversation that make this elicitation technique authentic and communicative when it succeeds.

In practice, the ability of the interlocutor to create the right atmosphere and the personal characteristics of the participants are of great importance to the outcome of the test. Of course, it is a challenge for the interlocutor to create the right atmosphere in a very short time, especially if he/she is not familiar with the participants, so special training is required.

Most of the time, an interview can naturally develop into a discussion, but this takes a lot of interviewing skills on the interlocutor's part to be able to predispose the test taker(s) and lead the interview in this direction. As a rule, only students with quite a high level of proficiency will feel confident enough to take this path.

Taking the initiative, asking questions, and expressing disagreement are all functional language features that require good oral command of the language and a high level of proficiency. Functional language can be learned just like everything else but, in a test environment, people are sometimes reluctant to take risks. The natural instinct of many is to keep quiet, to speak only when asked, and not to try to do anything *clever* if they do not feel confident enough or know they will be penalized for silly mistakes. Therefore, there is the danger that a discussion/conversation technique will reward extrovert and talkative personalities rather that those who are less forthcoming.

English Proficiency Tests

These days, universities and colleges worldwide offer numerous courses taught exclusively in English, making a certain level of English language proficiency a requirement for international students to enrol in their study programmes. For this reason, they have to provide an English language certificate which states they have enough knowledge to participate in classes, understand the textbooks, and complete the assignments. Universities usually have a score for each acceptable certificate above which the candidate is admitted.

Below is a short review of some of the most famous certificates and what their part that tests speaking skills consists of.

1. TOEFL®: Test of English as a Foreign Language

TOEFL® is a very popular English certificate around the world. In most cases, it is offered as an internet-based test (TOEFL® iBT) where test takers interact with a computer all the time, but there is also a paper-based test designed only for locations without an internet connection, which does not include the Speaking part. Normally, both the internet and paper-based versions can only be taken in specialized test centres.

TOEFL® iBT includes 4 parts:

- Reading;
- Listening;
- Speaking;
- Writing.

For the Speaking part, which lasts about 17 minutes, the test taker speaks into a microphone, his/her answers are recorded, sent to an online scoring network, and evaluated by three to six examiners. The speaking practice consists of four tasks that resemble a real-life situation in and out of the university classroom.

For the first one, called *Independent Speaking*, the test taker hears a question and has 15 seconds to prepare a response – an opinion on a familiar topic. Then, he/she has to speak for 45 seconds. The student is expected to draw ideas, share experience, and express viewpoints with supporting statements.

The next three tasks test integrated skills – Reading and Speaking, and Listening and Speaking, or just what the student would have to do in a classroom. For the reading and speaking tasks, the test taker is usually given a piece of text (e.g. an article, a passage from a book or an extract from an academic paper), and after reading it (for about 45 seconds), he/she has 60 seconds to express opinions, give reasons for certain events, or elaborate on a particular piece of information. For the listening and speaking task, the test taker listens to a lecture or a speech and answers questions on the passage he/she has heard. The preparation time is 20 seconds, and the speaking time is 60 seconds.

The Speaking section includes native speaker English accents from the United States, the United Kingdom, Australia or New Zealand to better reflect the variety of accents a student might encounter while studying abroad.

As the test is carried out in a laboratory and there is no present interlocutor, test takers should be aware that they have no opportunity to correct themselves. They have to state opinions clearly and give reasons to support their statements. There is no *correct* answer, which eliminates tension to a certain extent.

As with all speaking tasks, the test takers' responses should be intelligible, they should demonstrate effective use of grammar and vocabulary, and should be well developed and coherent.

2. IELTS: International English Language Testing System

IELTS offers an academic and a professional version of their test, designed for academic and professional environments. Like TOEFL®, IELTS is one of the most popular English language tests in the world. IELTS test results use a 0–9 grading scale. For candidates to be admitted in English-language programmes, the commonly required scores range between 5.5 and 6.5. IELTS is also accepted as language proof for certain work permits, as well as for immigration (IELTS for general training). IELTS Academic lasts for 2 hours and 45 minutes and has four parts:

- Listening;
- Academic Reading;
- Academic Writing;
- Speaking.

IELTS Academic is available both online and on paper. In either case, the test is carried out at an official test centre and the Speaking part is a face-to-face interaction with a certified examiner. It takes approximately 11 to 14 minutes; it is recorded to avoid ambiguity and increase test reliability; it consists of three tasks.

The first task is an interview with the examiner for about four or five minutes. The questions cover all kinds of general topics, such as home, family, work, studies, and interests.

The second task is a mini-presentation asking for opinion on a particular topic. The test taker receives a card with a question and instructions on what points he/she has to focus on; the preparation time is one minute, while the speaking time is up to two minutes. The examiner then asks one or two questions.

The third task is a continuation of the second one – the test taker is asked further questions on the topic in task two. This provides an opportunity for more abstract ideas and issues to be discussed. The session lasts between four and five minutes.

3. PTE Academic: Pearson English Language Test

The PTE Academic test has recently gained a lot of popularity; an increasing number of universities accept it as relevant proof of English knowledge from aspiring international students. Unlike TOEFL® iBT, IELTS Academic, and the Cambridge exams, which have four parts, the PTE test is divided into three sections:

- Speaking and Writing;
- Reading;
- Listening.

The PTE Academic test takes three hours to complete and is only available in a computer-based format, i.e. the Speaking part is recorded and sent to assessors. The first part is Speaking and Writing, and it takes between 77 and 93 minutes to complete all questions split in seven question types. Test takers can record answers only once. It tests both speaking and writing skills, and the language is what the students might hear in an academic environment.

The speaking tasks are:

A. Reading aloud: text (up to 60 words) appears on the screen and the test taker has to read it into the microphone. The preparation time is 30–40 seconds;

B. Repeating sentences: the test taker listens to a recording of a sentence and then has to repeat it into the microphone. Preparation time: 3–9 seconds. The time to answer is 15 seconds;

C. Describing an image: an image appears on the screen and the test taker has to describe it in detail. There is no fixed time for preparation, but the time to answer is only 40 seconds;

D. Retelling a lecture: after listening to or watching a lecture, the test taker has to retell the lecture in his/her own words. He/she has 90 seconds to prepare and 40 seconds to speak. The student is allowed to take notes on an erasable board while the recording is playing.

E. Answering short questions: after listening to a question, the test taker has to answer it with a few words. For each question, he/she has 3–9 seconds to prepare and up to 10 seconds to answer.

4. Cambridge English Qualifications

Cambridge English Language Assessment (formerly known as the University of Cambridge ESOL Examinations) is a non-profit organization offering English language certificates for different language levels. They offer many types of exams, but the most popular ones and the ones that universities recognize or require are:

- B2 First (formerly known as FCE: First Certificate in English);
- C1 Advanced (formerly known as CAE: Certificate in Advanced English);
- C2 Proficiency (formerly known as CPE: Certificate of Proficiency in English).

All these English exams are divided into four parts:

- Reading and Use of English;
- Writing;
- Listening;
- Speaking.

Cambridge certificates are mainly used in the United Kingdom, but they are also accepted by numerous universities and colleges worldwide.

Unlike other certificates, Cambridge qualifications do not have an expiry date. However, universities and other organizations can choose whether to accept results older than two-three years.

B2 First Speaking Test

The Speaking test takes 14 minutes to complete, and test takers are split into pairs. Candidates are expected to be able to respond to questions and to interact in conversational English. There are two

examiners present at the test – one asks questions and the other awards scores. The test consists of four tasks:

- Task 1: A conversation between the examiner and each candidate (spoken questions);
- Task 2: An individual *long turn* for each candidate. The test taker is given a pair of photographs to talk about for a minute without interruption. Then, the second candidate follows with a brief response or comment;
- Task 3: A two-way conversation between the candidates where they have to decide on something; they are usually given a written or a picture stimulus;
- Task 4: A discussion on topics related to Task 3 (spoken questions).

C1 Advanced Speaking Test

The settings here are similar to the B2 First test – the speaking tasks are done in pairs and there are two examiners present: one to ask questions and lead discussions and the other to award scores. The test takes 15 minutes and includes four tasks:

- Task 1: Short questions and answers between each of the test takers and the examiner;
- Task 2: *Long turn*. The examiner gives each test taker between two and five photographs and asks them to talk about them. The first candidate has to speak for one minute without interruption and then the other candidate comments on what the first one has said for about 30 seconds. Then the roles switch, i.e. the second candidate talks about a different set of photographs, and the first one has to listen and comment;
- Task 3: A collaborative task. The two candidates have a conversation, and they have to take a decision on an issue; usually photographs are used as a stimulus;
- Task 4: Discussion. Task 3 is further developed into a discussion on the topics or issues raised in the task. The discussion is again between the two candidates.

C2 Proficiency Speaking Test

This test has three parts and again test takers take it in pairs. There are two examiners. One of the examiners (the interlocutor) conducts the test and the other examiner (the assessor) listens and awards scores.

The test takes 16 minutes per pair of candidates and includes the following tasks:

- Task 1: An interview for about two minutes. The interlocutor asks each of the candidates a series of questions, addressing a question to each of them in turn, so that they are given an opportunity to talk about themselves, express opinions, or speculate on various topics;
- Task 2: A collaborative task (4 minutes). The interlocutor gives the test takers some spoken instructions and one or more pictures to look at. First, each candidate has to answer a question which focuses on his/her reaction to aspects of one or more pictures (1 minute). The second part involves decision-making where both candidates discuss a matter and agree on a particular action to be taken;
- Task 3: Long turn and discussion (10 minutes). The interlocutor gives each candidate a card with a question and some ideas on it and the latter has to speak for about two minutes. After he/she finishes, the partner has to comment, and the interlocutor then asks both candidates a question on the same topic. The interlocutor follows the same procedure with the other candidate and then leads a discussion with both of them.

Conclusion

In practice, when learning a foreign language (whether in a formal school or in environments where the language is spoken as native), speaking skills are usually the last to develop. This is natural as the learner needs to acquire a certain amount of vocabulary items and perfect the listening comprehension skills first, then master grammatical structures, and, most of all, gather the courage to utter his/her first sentence, no matter how wrong it may be in terms of grammar and vocabulary.

From the author's experience as a teacher, when students are preparing for certificate exams, everyone is frightened by the Speaking part. They can write long essays and deal with complex grammat-

ical structures, but speaking seems so difficult that, given the choice, they would rather not go through it. With sufficient practice, this reluctance can easily be overcome. The purpose of learning a language is not to take a certain exam, but to be able to use the language properly when needed.

Sometimes, though, poor performance in oral tests does not correspond to a student's inability to communicate in the foreign language. On the contrary, in informal situations, they can actually be very good communicators, skilled in small talk, giving directions, explaining viewpoints, and narrating events (even though their grammar and vocabulary may not always be flawless). There are many reasons for this: not everyone can perform well under pressure, knowing that he/she is being assessed; not enough test experience; low confidence in the spoken language in formal settings; etc. This is where the role of the interlocutor is very important – to predispose the test taker and to lead him/her into a friendly chat or small discussion, so that he/she feels relaxed and is able to ignore the test environment and demonstrate his/her abilities at best.

So, when designing speaking skill test tasks, all of the above needs to be taken into consideration. Tasks have to be able to get the most out of the test participant: to facilitate and encourage him/her to speak, to resemble real communication situations and feel natural, and to give something in return (i.e. to teach the test taker something). There is an abundance of elicitation techniques, all aimed at predisposing and encouraging test takers to *rise and shine* and to do their best.

As for the certificate exams, most of them have a format that resembles what the test candidates will be using the language for: in class, working on assignments, meeting and chatting with friends. The elicitation techniques are chosen such that they are least assessor-controlled and allow candidates to show their skills as speakers and communicators, apart from PTE Academic which has a slightly different structure. Of course, most candidates go through thorough training before sitting a test, and they have a lot of practice on all sections of the chosen certificate exam, which means that these tests have a positive backwash on what is taught in the classroom and how it is taught.

REFERENCES

Brown, H. D., & Abeywickrama, P. (2019). Language assessment: Principles and classroom practice (3ʳᵈ ed.). NY: Pearson Education.

Nunan, D. (1999). Second language teaching & learning. Boston, MA: Heinle & Heinle.

O'Sullivan, B. (2008). Notes on assessing speaking. Ithaca, NY: Cornell University Language Resource Centre.

Rahmawati, Y., & Hidayatullah, E. S. (2014). Developing assessment for speaking. *Indonesian Journal of English Education*, *1*(2), 199–210. https://doi.org/10.15408/ijee.v1i2.1345

Underhill, N. (1987). *Testing spoken language: A handbook of oral testing techniques*. Cambridge: Cambridge University Press.

English Certificate Exams – Online Resources

IELTS certification examination: https://www.cambridgeenglish.org/exams-and-tests/ielts/

Cambridge B2 First certification examination: https://www.cambridgeenglish.org/exams-and-tests/first/

Cambridge C1 Advanced certification examination: https://www.cambridgeenglish.org/exams-and-tests/advanced/

Cambridge C2 Proficiency certification examination: https://www.cambridgeenglish.org/exams-and-tests/proficiency/

PTE Academic certification examination: https://www.pearsonpte.com/pte-academic/test-format