**University of Latvia**

**Faculty of Modern Languages**

**English Department**

**Types of Tests Used in English Language.**

**Bachelor Paper**

**Anželika Ozerova**

**Riga**

**2004**

**Declaration of academic Integrity**

I hereby declare that this study is my own and does not contain any unacknowledged material from any source.

Signed:

12 May, 2004

**Abstract.**

The present paper attempts to investigate various types of tests and their application in the language classroom. The theoretical part deals with the basic data about testing, the comparison of such issues as assessment and valuation, reasons for testing, types of tests, such as diagnostic, progress, achievement, placement and proficiency tests; test formats and ways of testing.

It relates theory to practice by analyzing two proficiency tests: TOEFL and CFC tests. They are carefully discussed and compared to find any similarities or differences in their structure and design. The conclusions drawn are based on the theory and analyses of the tests. The data obtained indicate that the both tests though being sometimes different in their purpose, design and structure, are constructed according to the universally accepted pattern.

**Table of Contents**

**Introduction** …………………………………………………........................1

**Chapter 1**

What is test?……………………………………………………………………3

**Chapter 2**

2.1 Inaccurate tests……………...…………………………………………….7

2.2 Validity……………………..……………………………………………..8

2.3 Reliability………….. ……………………………………………………11

**Chapter 3**

3.1 Diagnostic tests………………………………. ………………………….13

3.2 Placement tests…………………………...……………………………….15

3.3 Progress tests……………………………………………...........................17

3.4 Achievement tests………………………..……………………………….18

3.5 Proficiency tests…………………………………………………………..20

**Chapter 4**

4.1 Direct and Indirect testing…..…………………………………………....22

4.2 Discrete point and integrative testing……………………………………..24

4.3 Criterion-refernced and Norm-referenced testing…………………………25

4.4 Objective and Subjective testing...………………………………………..26

4.5 Communicative language testing…………………………………………26

**Chapter 5**

5.1 Multiple choice tests………………………………………………………29

5.2 Short answer tests…………………………………………………………32

5.3 The Cloze tests and Gap-filling tests……………………………………..33

5.4 C-Test……………………………………………………………………..35

5.5 True/false items……………………………………………………………36

5.6 Dictation…………………………………………………………………...36

5.7 Listening Recall……………………………………………………………38

5.8 Testing Grammar through Error-recognition Items……………………….38

5.9 Controlled Writing…………………………………………………………39

5.10 Free Writing………………………………………………………………40

5.11 Test Formats Used in Testing Speaking Skills…………………………..41

**Chapter 6**

Analysis of the Test of English as a Foreign Language and Cambridge First

Certificate test according to test design criteria………………………………..43

 **Conclusions**…………………………………………………………………...55

 **Theses**. ………………………………………………………..........................57

**Bibliography**…………………………………………………….......................59

**Appendix**

**Introduction**

Among all words used in a classroom there is the only word that usually makes the students shudder: “test”. There is hardly a person who would claim that s/he favours tests and finds them very motivating. However, tests cannot be avoided completely, for they are inevitable elements of learning process. They are included into curriculum at schools and are to check the students’ level of knowledge and what they are able to do; they could be accomplished at the beginning of the study year and at the end of it; the students could be tested after working on new topics and acquiring new vocabulary. Moreover, the students are to face the tests in order to enter any foreign university or reveal the level of their English language skills for themselves. For that purpose they take specially designed tests that are Test of English as a Foreign Language, or TOEFL test (further in the text) and CFC (further in the text), or Cambridge First Certificate. Although, these tests can sometimes serve for different purposes and are unrelated, they are sometimes quite common in their design and structure. Therefore, the author of the paper is particularly interested in the present research, for she assumes it to be of a great significance not only for herself, but also for the individuals who are either involved in the field or just want to learn more about TOEFL and CFC tests, their structure, design and application. Therefore, the present research will display various aspects of the theory discussed, accompanied with the practical part vastly analyzed.

Thus, ***the goal*** of the present research is to investigate various types of test formats and ways of testing, focusing particularly on TOEFL and CFC tests, in order to see how the theory is used and could be applied in practice.

***The hypothesis*** is as follows: Serving for almost similar purpose, however being sometimes different in their design and structure, the TOEFL and CFC tests are usually constructed according to the accepted universal pattern.

***The enabling objectives are as follows:***

* To review literature on the nature of tests in order to make theoretically well-motivated discussions on the choice of testing types;
* To analyse the selected types of tests, such as TOEFL and CFC tests;
* To draw relevant conclusions.

***Methods of Research:***

Theoretical:

1. Analytical and selective study of the theory available;
2. Juxtaposition of the ideas selected from theory and tested against practical evidences;
3. Drawing conclusions.

Practical:

* Selecting and adapting appropriate tests types, such as TOEFL and CFC, to exemplify the theory.

The paper consists of six chapters each including sub-chapters. Chapter 1 discusses the general data about tests. Chapter 2 describes reliability and validity. Chapter 3 focuses on various types of tests. Chapter 4 deals with ways of testing. Chapter 5 speaks on four language skills. Chapter 6 offers the practical part of the paper.

***Chapter 1***

***What is test?***

Hicks (2000:155) considers that the role of tests is very useful and important, especially in language learning. It is a means to show both the students and the teacher how much the learners have learnt during a course. The author of the paper agrees with the statement, for she believes that in order to see whether the students have acquired the material and are making constant progress, the teacher will inevitably have to test his/her learners. It does not mean that a usual test format with a set of activities will be used all the time. To check the students’ knowledge the teacher can apply a great range of assessment techniques, including even the self-evaluation technique that is so beloved and favoured by the students. Moreover, according to Heaton (1990:6), tests could be used to display the strength and weaknesses of the teaching process and help the teacher improve it. They can demonstrate what should be paid more attention to, should be worked on and practised. Furthermore, the tests results will display the students their weak points, and if carefully guided by the teacher, the students will be even able to take any remedial actions.

Thompson (Forum, 2001) believes that students learn more when they have tests. Here we can both agree and disagree. Certainly, preparing for a test, the student has to study the material that is supposed to be tested, but often it does not mean that such type of learning will obligatory lead to acquisition and full understanding of it. On the opposite, it could often lead to the pure cramming. That, consequently, will result in a stressful situation the student will find her/himself before or during the test, and the final outcome will be a complete deletion of the studied material. We can base that previous statement on our own experience: when working at school, the author of the present research had encountered such examples for many times.

However, very often the tests can facilitate the students’ acquisition process, i.e.: the students are to be checked the knowledge of the irregular verbs forms. Being constantly tested by means of a small test, they can learn them successfully and transfer them to their long-term memory, as well. Although, according to Thompson tests decrease practice and instruction time. What he means is that the students are as if limited; they are exposed to practice of a new material, however, very often the time implied for it is strictly recommended and observed by a syllabus. That denotes that there will be certain requirements when to use a test. Thus, the students find themselves in definite frames that the teacher will employ. Nevertheless, there could be advantages that tests can offer: they increase learning, for the students are supposed to study harder during the preparation time before a test.

Thompson (ibid.) quotes Eggan, who emphasises the idea that the learners study hard for the classes they are tested thoroughly. Further, he cites Hilles, who considers that the students want and expect to be tested. Nonetheless, this statement has been rather generalized. Speaking about the students at school, we can declare that there is hardly a student who will truly enjoy tests and their procedure. Usually, what we will see just sore faces when a test is being mentioned. According to Thompson, the above-mentioned idea could be applied to the students who want to pass their final exams or to get a certificate in Test of English as a Foreign Language (TOEFL) or First Certificate (FCE). Mostly this concerns adults or the students who have their own special needs, such as going abroad to study or work. This again supports the idea that motivation factor plays a significant role in the learning process.

Moreover, too much of testing could be disastrous. It can entirely change the students’ attitude towards learning the language, especially if the results are usually dissatisfying and decrease their motivation towards learning and the subject in general.

Furthermore, as Alderson (1996:212) assumes, we should not forget that the tests when administered receive less support from the teacher as it is usually during the exercises in a usual language classroom. The students have to cope themselves; they cannot rely on the help of the teacher if they are in doubt. During a usual procedure when doing various activities the students know they can encounter the teacher’s help if they require it. They know the teacher is always near and ready to assist, therefore, no one is afraid to make a mistake and try to take a chance to do the exercises. However, when writing a test and being left alone to deal with the test activities, the students panic and forget everything they knew before. The author of the paper believes that first what the teacher should do is to teach the students to overcome their fear of tests and secondly, help them acquire the ability to work independently believing in their own knowledge. That ability according to Alderson is the main point, “the core meaning” of the test. The students should be given confidence. Here we can refer to Heaton (1990:7) who conceives, supported by Hicks, that students’ encouragement is a vital element in language learning. Another question that may emerge here is how to reach the goal described above, how to encourage the students. Thus, at this point we can speak about positive results. In fact, our success motivates us to study further, encourages us to proceed even if it is rather difficult and we are about to lose confidence in ourselves. Therefore, we can speak about the tests as a tool to increase motivation. However, having failed for considerable number of times, the student would definitely oppose the previous statement. Hence, we can speak about assessment and evaluation as means for increasing the students’ motivation.

Concerning Hicks (2000:162), we often perceive these two terms – evaluating and assessment – as two similar notions, though they are entirely different. She states that when we assess our students we commonly are interested in *“how and how much our students have learnt”*, but when we evaluate them we are concerned with *“how the learning process is developing”*. These both aspects are of great importance for the teacher and the students and should be correlated in order to make evaluation and assessment *“go hand in hand”*. However, very frequently, the teachers assess the students without taking the aspect of evaluation into account. According to Hicks, this assessment is typically applied when dealing with examinations that take place either at the end of the course or school year. Such assessment is known as *achievement test*. With the help of these tests the teacher receives a clear picture of what his/her students have learnt and which level they are comparing with the rest of the class. The author of the paper agrees that achievement tests are very essential for comparing how the students’ knowledge has changed during the course. This could be of a great interest not only for the teacher, but also for the authorities of the educational establishment the teacher is employed by. Thus, evaluation of the learning process is not of the major importance here. We can speak about evaluation when we deal with “small” tests the teachers use during the course or studying year. It is a well-known fact that these tests are employed in order to check how the learning process is going on, where the students are, what difficulties they encounter and what they are good at. These tests are also called “diagnostic” tests; they could be of a great help for the teacher: judging from the results of the test, analysing them the teacher will be able to improve or alter the course and even introduce various innovations. These tests will define whether the teacher can proceed with the new material or has to stop and return to what has not been learnt sufficiently in order to implement additional practice.

With respect to Hicks, we can display some of her useful and practical ideas she proposes for the teachers to use in the classroom. In order to incorporate evaluation together with assessment she suggests involving the students directly into the process of testing. Before testing vocabulary the teacher can ask the students to guess what kind of activities could be applied in the test. The author of the paper believes that it will give them an opportunity to visage how they are going to be tested, to be aware of and wait for, and the most important, it will reduce fear the students might face. Moreover, at the end of each test the students could be asked their reflections: if there was a multiple choice, what helped them guess correctly, what they used for that – their schemata or just pure guessing; if there was a cloze test - did they use guessing from the context or some other skills, etc. Furthermore, Hicks emphasises that such analysis will display the students the way they are tested and establish an appropriate test for each student. Likewise, evaluation will benefit the teacher as well. S/he not only will be able to discover the students’ preferences, but also find out why the students have failed a particular type of activity or even the whole test. The evaluation will determine what is really wrong with the structure or design of the test itself. Finally, the students should be taught to evaluate the results of the test. They should be asked to spot the places they have failed and together with the teacher attempt to find out what has particularly caused the difficulties. This will lead to consolidation of the material and may be even to comprehension of it. And again the teacher’s role is very essential, for the students alone are not able to cope with their mistakes. Thus, evaluation is inevitable element of assessment if the teacher’s aim is to design a test that will not make the students fail, but on the contrary, anticipate the test’s results.

 To conclude we can add alluding to Alderson (1996:212) that the usual classroom test should not be too complicated and should not discriminate between the levels of the students. The test should test what was taught. The author of the paper has the same opinion, for the students are very different and the level of their knowledge is different either. It is inappropriate to design a test of advanced level if among your learners there are those whose level hardly exceeds lower intermediate.

Above all, the tests should take the learners’ ability to work and think into account, for each student has his/her own pace, and some students may fail just because they have not managed to accomplish the required tasks in time.

Furthermore, Alderson assumes (ibid.) that the instructions of the test should be unambiguous. The students should clearly see what they are supposed and asked to do and not to be frustrated during the test. Otherwise, they will spend more time on asking the teacher to explain what they are supposed to do, but not on the completing of the tasks themselves. Finally, according to Heaton (1990:10) and Alderson (1996:214), the teacher should not give the tasks studied in the classroom for the test. They explain it by the fact, that when testing we need to learn about the students’ progress, but not to check what they remember. The author of the paper concurs the idea and assumes that the one of the aims of the test is to check whether the students are able to apply their knowledge in various contexts. If this happens, that means they have acquired the new material.

## Chapter 2

***Reliability and validity***

* 1. ***Inaccurate tests***

Hughes (1989:2) conceives that one of the reasons why the tests are not favoured is that they measure not exactly what they have to measure. The author of the paper supports the idea that it is impossible to evaluate someone’s true abilities by tests. An individual might be a bright student possessing a good knowledge of English, but, unfortunately, due to his/her nervousness may fail the test, or vice versa, the student might have crammed the tested material without a full comprehension of it. As a result, during the test s/he is just capable of producing what has been learnt by tremendous efforts, but not elaboration of the exact actual knowledge of the student (that, unfortunately, does not exist at all). Moreover, there could be even more disastrous case when the student has cheated and used his/her neighbour’s work. Apart from the above-mentioned there could be other factors that could influence an inadequate completion of the test (sleepless night, various personal and health problems, etc.)

However, very often the test itself can provoke the failure of the students to complete it. With the respect to the linguists, such as Hughes (1989) and Alderson (1996), we are able to state that there are two main causes of the test being inaccurate:

* Test content and techniques;
* Lack of reliability.

The first one means that the test’s design should response to what is being tested. First, the test must content the exact material that is to be tested. Second, the activities, or techniques, used in the test should be adequate and relevant to what is being tested. This denotes they should not frustrate the learners, but, on the contrary, facilitate and help the students write the test successfully.

The next one denotes that one and the same test given at a different time must score the same points. The results should not be different because of the shift in time. For example, the test cannot be called reliable if the score gathered during the first time the test was completed by the students differs from that administered for the second time, though knowledge of the learners has not changed at all. Furthermore, reliability can fail due to the improper design of a test (unclear instructions and questions, etc.) and due to the ways it is scored. The teacher may evaluate various students differently taking different aspects into consideration (level of the students, participation, effort, and even personal preferences.) If there are two markers, then definitely there will be two different evaluations, for each marker will possess his/her own criteria of marking and evaluating one and the same work. For example, let us mention testing speaking skills. Here one of the makers will probably treat grammar as the most significant point to be evaluated, whereas the other will emphasise the fluency more. Sometimes this could lead to the arguments between the makers; nevertheless, we should never forget that still the main figure we have to deal with is the student.

***2.2. Validity***

Now we can come to one of the important aspects of testing – validity. Concerning Hughes, every test should be reliable as well as valid. Both notions are very crucial elements of testing. However, according to Moss (1994) there can be validity without reliability, or sometimes the border between these two notions can just blur. Although, apart from those elements, a good test should be efficient as well.

According to Bynom (Forum, 2001), validity deals with what is tested and degree to which a test measures what is supposed to measure (Longman Dictionary, LTAL). For example, if we test the students writing skills giving them a composition test on Ways of Cooking, we cannot denote such test as valid, for it can be argued that it tests not our abilities to write, but the knowledge of cooking as a skill. Definitely, it is very difficult to design a proper test with a good validity, therefore, the author of the paper believes that it is very essential for the teacher to know and understand what validity really is.

Regarding Weir (1990:22), there are five types of validity:

* Construct validity;
* Content validity
* Face validity
* Wash back validity;
* Criterion-related validity.

Weir (ibid.) states that construct validity is a theoretical concept that involves other types of validity. Further, quoting Cronbach (1971), Weird writes that to construct or plan a test you should research into testee’s behaviour and mental organisation. It is the ground on which the test is based; it is the starting point for a constructing of test tasks. In addition, Weird displays the Kelly’s idea (1978) that test design requires some theory, even if it is indirect exposure to it. Moreover, being able to define the theoretical construct at the beginning of the test design, we will be able to use it when dealing with the results of the test. The author of the paper assumes that appropriately constructed at the beginning, the test will not provoke any difficulties in its administration and scoring later.

Another type of validity is content validity. Weir (ibid.) implies the idea that content validity and construct one are closely bound and sometimes even overlap with each other. Speaking about content validity, we should emphasise that it is inevitable element of a good test. What is meant is that usually duration of the classes or test time is rather limited, and if we teach a rather broad topic such as “computers”, we cannot design a test that would cover all the aspects of the following topic. Therefore, to check the students’ knowledge we have to choose what was taught: whether it was a specific vocabulary or various texts connected with the topic, for it is impossible to test the whole material. The teacher should not pick up tricky pieces that either were only mentioned once or were not discussed in the classroom at all, though belonging to the topic. S/he should not forget that the test is not a punishment or an opportunity for the teacher to show the students that they are less clever. Hence, we can state that content validity is closely connected with a definite item that was taught and is supposed to be tested.

Face validity, according to Weir (ibid.), is not theory or samples design. It is how the examinees and administration staff see the test: whether it is construct and content valid or not. This will definitely include debates and discussions about a test; it will involve the teachers’ cooperation and exchange of their ideas and experience.

Another type of validity to be discussed is wash back validity or backwash. According to Hughes (1989:1) backwash is the effect of testing on teaching and learning process. It could be both negative and positive. Hughes believes that if the test is considered to be a significant element, then preparation to it will occupy the most of the time and other teaching and learning activities will be ignored. As the author of the paper is concerned this is already a habitual situation in the schools of our country, for our teachers are faced with the centralised exams and everything they have to do is to prepare their students to them. Thus, the teacher starts concentrating purely on the material that could be encountered in the exam papers alluding to the examples taken from the past exams. Therefore, numerous interesting activities are left behind; the teachers are concerned just with the result and forget about different techniques that could be introduced and later used by their students to make the process of dealing with the exam tasks easier, such as guessing form the context, applying schemata, etc.

The problem arises here when the objectives of the course done during the study year differ from the objectives of the test. As a result we will have a negative backwash, e.g. the students were taught to write a review of a film, but during the test they are asked to write a letter of complaint. However, unfortunately, the teacher has not planned and taught that.

Often a negative backwash may be caused by inappropriate test design. Hughes further in his book speaks about multiple-choice activities that are designed to check writing skills of the students. The author of the paper is very confused by that, for it is unimaginable how writing an essay could be tested with the help of multiple choices. Testing essay the teacher first of all is interested in the students’ ability to apply their ideas in writing, how it has been done, what language has been used, whether the ideas are supported and discussed, etc. At this point multiple-choice technique is highly inappropriate.

Notwithstanding, according to Hughes apart form negative side of the backwash there is the positive backwash as well. It could be the creation of an entirely new course designed especially for the students to make them pass their final exams. The test given in a form of final exams imposes the teacher to re-organise the course, choose appropriate books and activities to achieve the set goal: pass the exam. Further, he emphasises the importance of partnership between teaching and testing. Teaching should meet the needs of testing. It could be understand in the following way that teaching should correspond the demands of the test. However, it is a rather complicated work, for according to the knowledge of the author of the paper the teachers in our schools are not supplied with specially designed materials that could assist them in their preparation the students to the exams. The teachers are just given vague instructions and are free to act on their own.

The last type that could be discussed is criterion-related validity. Weir (1990:22.) assumes that it is connected with test scores link between two different performances of the same test: either older established test or future criterion performance. The author of the paper considers that this type of validity is closely connected with criterion and evaluation the teacher uses to assess the test. It could mean that the teacher has to work out definite evaluation system and, moreover, should explain what she finds important and worth evaluating and why. Usually the teachers design their own system; often these are points that the students can obtain fulfilling a certain task. Later the points are gathered and counted for the mark to be put. Furthermore, the teacher can have a special table with points and relevant marks. According to our knowledge, the language teachers decide on the criteria together during a special meeting devoted to that topic, and later they keep to it for the whole study year. Moreover, the teachers are supposed to make his/her students acquainted with their evaluation system for the students to be aware what they are expected to do.

* 1. ***Reliability***

According to Bynom (Forum, 2001) reliability shows that the test’s results will be similar and will not change if one and the same test will be given on various days. The author of the paper is of the same mind with Bynom and presumes the reliability to be the one of the key elements of a good test in general. For, as it has been already discussed before, the essence of reliability is that when the students’ scores for one and the same test, though given at different periods of time and with a rather extended interval, will be approximately the same. It will not only display the idea that the test is well organized, but will denote that the students have acquired the new material well.

A reliable test, according to Bynom, will contain well-formulated tasks and not indefinite questions; the student will know what exactly should be done. The test will always present ready examples at the beginning of each task to clarify what should be done. The students will not be frustrated and will know exactly what they are asked to perform. However, judging form the personal experience, the author of the paper has to admit, that even such hints may confuse the students; they may fail to understand the requirements and, consequently, fail to complete the task correctly. This could be explained by the fact that the students are very often inattentive, lack patience and try to accomplish the test quickly without bothering to double check it.

Further, regarding to Heaton (1990:13), who states that the test could be unreliable if the two different markers mark it, we can add that this factor should be accepted, as well. For example, one representative of marking team could be rather lenient and have different demands and requirements, but the other one could appear to be too strict and would pay attention to any detail. Thus, we can come to another important factor influencing the reliability that is marker’s comparison of examinees’ answers. Moreover, we have to admit a rather sad fact but not the exceptional one that the maker’s personal attitude towards the testee could impact his/her evaluation. No one has to exclude various home or health problems the marker can encounter at that moment, as well.

To summarize, we can say that for a good test possessing validity and reliability is not enough. The test should be practical, or in other words, efficient. It should be easily understood by the examinee, ease scored and administered, and, certainly, rather cheap. It should not last for eternity, for both examiner and examinee could become tired during five hours non-stop testing process. Moreover, testing the students the teachers should be aware of the fact that together with checking their knowledge the test can influence the students negatively. Therefore, the teachers ought to design such a test that it could encourage the students, but not to make them reassure in their own abilities. The test should be a friend, not an enemy. Thus, the issue of validity and reliability is very essential in creating a good test. The test should measure what it is supposed to measure, but not the knowledge beyond the students’ abilities. Moreover, the test will be a true indicator whether the learning process and the teacher’s work is effective.

***Chapter 3***

***Types of tests***

Different scholars (Alderson, 1996; Heaton, 1990; Underhill, 1991) in their researches ask the similar question – why test, do the teachers really need them and for what purpose. Further, they all agree that test is not the teacher’s desire to catch the students unprepared with what they are not acquainted; it is also not the motivating factor for the students to study. In fact, the test is a request for information and possibility to learn what the teachers did not know about their students before. We can add here that the test is important for the students, too, though they are unaware of that. The test is supposed to display not only the students’ weak points, but also their strong sides. It could act as an indicator of progress the student is gradually making learning the language. Moreover, we can cite the idea of Hughes (1989:5) who emphasises that we can check the progress, general or specific knowledge of the students, etc. This claim will directly lead us to the statement that for each of these purposes there is a special type of testing. According to some scholars (Thompson, 2001; Hughes, 1989; Alderson, 1996; Heaton, 1990; Underhill, 1991), there are four traditional categories or types of tests: proficiency tests, achievement tests, diagnostic tests, and placement tests. The author of the paper, once being a teacher, can claim that she is acquainted with three of them and has frequently used them in her teaching practice.

In the following sub-chapters we are determined to discuss different types of tests and if possible to apply our own experience in using them.

***3.1. Diagnostic tests***

It is wise to start our discussion with that type of testing, for it is typically the first step each teacher, even non-language teacher, takes at the beginning of a new school year. In the establishment the author of the paper was working it was one of the main rules to start a new study year giving the students a diagnostic test. Every year the administration of the school had stemmed a special plan where every teacher was supposed to write when and how they were going to test their students. Moreover, the teachers were supposed to analyse the diagnostic tests, complete special documents and provide diagrams with the results of each class or group if a class was divided. Then, at the end of the study year the teachers were demanded to compare the results of them with the final, achievement test (see in Appendix 1). The author of the paper has used this type of test for several times, but had never gone deep into details how it is constructed, why and what for. Therefore, the facts listed below were of great value for her.

 Referring to Longman Dictionary of LTAL (106) diagnostic tests is a test that is meant to display what the student knows and what s/he does not know. The dictionary gives an example of testing the learners’ pronunciation of English sounds. Moreover, the test can check the students’ knowledge before starting a particular course. Hughes (1989:6) adds that diagnostic tests are supposed to spot the students’ weak and strong points. Heaton (1990:13) compares such type of test with a diagnosis of a patient, and the teacher with a doctor who states the diagnosis. Underhill (1991:14.) adds that a diagnostic test provides the student with a variety of language elements, which will help the teacher to determine what the student knows or does not know. We believe that the teacher will intentionally include the material that either is presumed to be taught by a syllabus or could be a starting point for a course without the knowledge of which the further work is not possible. Thus, we fully agree with the Heaton’s comparison where he contrasts the test with a patient’s diagnosis. The diagnostic test displays the teacher a situation of the students’ current knowledge. This is very essential especially when the students return from their summer holidays (that produces a rather substantial gap in their knowledge) or if the students start a new course and the teacher is completely unfamiliar with the level of the group. Hence, the teacher has to consider carefully about the items s/he is interested in to teach. This consideration reflects Heaton’s proposal (ibid.), which stipulates that the teachers should be systematic to design the tasks that are supposed to illustrate the students’ abilities, and they should know what exactly they are testing. Moreover, Underhill (ibid.) points out that apart from the above-mentioned the most essential element of the diagnostic test is that the students should not feel depressed when the test is completed. Therefore, very often the teachers do not put any marks for the diagnostic test and sometimes even do not show the test to the learners if the students do not ask the teacher to return it. Nevertheless, regarding our own experience, the learners, especially the young ones, are eager to know their results and even demand marks for their work. Notwithstanding, it is up to the teacher whether to inform his/her students with the results or not; however, the test represents a valuable information mostly for the teacher and his/her plans for designing a syllabus.

 Returning to Hughes (ibid.) we can emphasise his belief that this type of test is very useful for individual check. It means that this test could be applicable for checking a definite item; it is not necessary that it will cover broader topics of the language. However, further Hughes assumes that this test is rather difficult to design and the size of the test can be even impractical. It means that if the teacher wants to check the students’ knowledge of Present simple, s/he will require a great deal of examples for the students to choose from. It will demand a tiresome work from the teacher to compose such type of the test, and may even confuse the learners.

At that point we can allude to our experience in giving a diagnostic test in Form 5. It was the class the teacher had worked before and knew the students and their level rather good. However, new learners had joined the class, and the teacher had not a slightest idea about their abilities. It was obvious that the students worried about how they would accomplish the test and what marks would they receive. The teacher had ensured them that the test would not be evaluated by marks. It was necessary for the teacher to plan her future work. That was done to release the tension in the class and make the students get rid of the stress that might be crucial for the results. The students immediately felt free and set to work. Later when analysing and summarizing the results the teacher realized that the students’ knowledge was purely good. Certainly, there were the place the students required more practice; therefore during the next class the students were offered remedial activities on the points they had encountered any difficulties. Moreover, that was the case when the students were particularly interested in their marks.

To conclude, we can conceive that interpreting the results of diagnostic tests the teachers apart from predicting why the student has done the exercises the way s/he has, but not the other, will receive a significant information about his/her group s/he is going to work with and later use the information as a basis for the forming syllabus.

***3.2 Placement tests***

Another type of test we are intended to discuss is a placement test. Concerning Longman Dictionary of LTAL again (279-280) we can see that a placement test is a test that places the students at an appropriate level in a programme or a course. This term does not refer to the system and construction of the test, but to its usage purpose. According to Hughes (1989:7), this type of test is also used to decide which group or class the learner could be joined to. This statement is entirely supported by another scholar, such as Alderson (1996:216), who declares that this type of test is meant for showing the teacher the students’ level of the language ability. It will assist to put the student exactly in that group that responds his/her true abilities.

Heaton (ibid.) adheres that the following type of testing should be general and should purely focus on a vast range of topics of the language not on just specific one. Therefore, the placement test typically could be represented in the form of dictations, interviews, grammar tests, etc.

Moreover, according to Heaton (ibid.), the placement test should deal exactly with the language skills relevant to those that will be taught during a particular course. If our course includes development of writing skills required for politics, it is not appropriate to study writing required for medical purposes. Thus, Heaton (ibid.) presumes that is fairly important to analyse and study the syllabus beforehand. For the placement test is completely attributed to the future course programme. Furthermore, Hughes (ibid.) stresses that each institution will have its own placement tests meeting its needs. The test suitable for one institution will not suit the needs of another. Likewise, the matter of scoring is particularly significant in the case of placement tests, for the scores gathered serve as a basis for putting the students into different groups appropriate to their level.

At this point we can attempt to compare a placement test and diagnostic one. From the first sight these both types of tests could look similar. They both are given at the beginning of the study year and both are meant for distinguishing the students’ level of the current knowledge. However, if we consider the facts described in sub-chapter 2.1 we will see how they are different. A diagnostic test is meant for displaying a picture of the students’ general knowledge at the beginning of the study year for the teacher to plan further work and design an appropriate syllabus for his/her students. Whereas, a placement test is designed and given in order to use the information of the students’ knowledge for putting the students into groups according to their level of the language. Indeed, they are both used for teacher’s planning of the course their functions differ. A colleague of mine, who works at school, has informed me that they have used a placement test at the beginning of the year and it appeared to be relevant and efficient for her and her colleague’s future teaching. The students were divided according to their English language abilities: the students with better knowledge were put together, whereas the weaker students formed their own group. It does not mean discrimination between the students. The teachers have explained the students the reason for such actions, why it was necessary – they wanted to produce an appropriate teaching for each student taking his/her abilities into account. The teachers have altered their syllabus to meet the demands of the students. The result proved to be satisfying. The students with better knowledge progressed; no one halted them. The weaker students have gradually improved their knowledge, for they received due attention than it would be in a mixed group.

***3.3 Progress test***

Having discussed two types of tests that are usually used at the beginning, we can approach the test typically employed during the study year to check the students’ development. We will speak about a progress test. According to Alderson (1996:217), progress test will show the teacher whether the students have learnt the recently taught material successfully. Basically, the teacher intends to check certain items, not general topics covered during the school or study year. Commonly, it is not very long and is determined to check the recent material. Therefore, the teacher might expect his/her learners to get rather high scores. The following type is supposed to be used after the students have learnt either a set of units on a theme or have covered a definite topic of the language. It will display the teacher whether the material has been successfully acquired or the students need additional practice instead of starting a new material.

A progress test will basically display the activities based on the material the teacher is determined to check. To evaluate it the teacher can work out a certain system of points that later will compose a mark. Typically, such tests do not influence the students’ final mark at the end of the year.

The authorities of school demand the teachers to conduct progress tests, as well. However, the teachers themselves decide on the necessity of applying them. Nevertheless, we can claim that progress test is inevitable part of the learning process. We can even take a responsibility to declare that progress test facilitate the material acquisition in a way. The students preparing for the test look through the material again and there is a chance it can be transferred to their long-term memory.

Further, we can come to Alderson (ibid.) who presumes that such type of testing could function as a motivating fact for the learners, for success will develop the students’ confidence in their own knowledge and motivate them study further more vigorously. In case, there will be two or three students whose scores are rather low, the teacher should encourage them by providing support in future and imply the idea that studying hard will allow them to catch up with the rest of the students sooner or later. The author of the paper basing on her experience agrees with the statement, for she had noticed that weaker students when they had managed to write their test successfully became proud of their achievement and started working better.

However, if the majority of the class scores a rather low grade, the teacher should be cautious. This could be a signal that there is either something wrong with the teaching or the students are low motivated or lazy.

***3.4 Achievement tests***

Apart from a progress test the teachers employ another type – achievement test. According to Longman Dictionary of LTAL (3), an achievement test is a test, which measures a language someone has learned during a specific course, study or program. Here the progress is significant and, therefore, is the main point tested.

Alderson (1996:219) posits that achievement tests are “more formal”, whereas Hughes (1989:8) assumes that this type of tests will fully involve teachers, for they will be responsible for the preparation of such tests and giving them to the learners. He repeats the dictionary defining the notion of achievement tests, adding just that success of the students, groups of students, or the courses.

Furthermore, Alderson (ibid.) conceives that achievement tests are mainly given at definite times of the school year. Moreover, they could be extremely crucial for the students, for they are intended either to make the students pass or fail the test.

At this instant the author of the paper is determined to compare a progress and achievement test. Again if we look at these two types they might seem similar, however, it is not so. Drawing on the facts listed above (see sub-chapter 2.3) we can report that a progress test is typically used during the course to check the acquisition of an excerpted material. An achievement test checks the acquisition of the material, as well. Although, it is far different in its application time. We basically use an achievement test at the end of the course to check the acquisition of the material covered during the study year, not bits of it as it is with a progress test.

Quoting Hughes (ibid.) we can differentiate between two kinds of achievement tests: final and progress tests. Final tests are the tests that are usually given at the end of the course in order to check the students’ achieved results and whether the objectives set at the beginning have been successfully reached. Further Hughes highlights that ministries of education, official examining boards, school administration and even the teachers themselves design these tests. The tests are based on the curriculum and the course that has been studied. We assume, that is a well-known fact that teachers usually are responsible for composing such tests, and it requires a careful work.

Alternatively, Alderson (ibid.) mentions two usage types of achievement tests: *formative* and *summative.* The notion of a formative test denotes the idea that the teacher will be able after evaluating the results of the test reconsider his/her teaching, syllabus design and even slow down the pace of studying to consolidate the material if it is necessary in future. Notwithstanding, these reconsiderations will not affect the present students who have taken the test. They will be applied to the future syllabus design.

Summative usage will deal precisely with the students’ success or failure. The teacher will immediately can take up remedial activities to improve a situation.

Further, Alderson (ibid.) and Heaton (1990:14) stipulate that designing an achievement test is rather time-consuming, for the achievement test is basically devised to cover a broad topic of the material covered during the course. In addition, one and the same achievement test could be given to more than one class at school to check both the students’ progress and the teachers’ work. At that point it is very essential to consider the material covered by different classes or groups. You cannot ask the students what they have not been taught. Heaton (ibid.) emphasises the close cooperative work of the teachers as a crucial element in test design. However, in the school the author of the paper used to work the teachers did not cooperate in designing achievement tests. Each teacher was free to write the test that best suits his/her children.

Developing the topic, we can focus on Hughes’ idea that there is an approach how to design a test; it is called *syllabus-content approach*. The test is based on a syllabus studied or a book taken during the course. This test could be described as a fair test, for it focuses mainly on the detailed material that the students are supposed to have studied. Hughes (ibid.) points out that if the test is inappropriately designed, it could result in unsuccessful accomplishment of it. Sometimes the demands of the test may differ from the objectives of the course. Therefore, the test should be based directly on the objectives of the course. Consequently, it will influence the choice of books appropriate to the syllable and syllable itself. The backwash will be positive not only for the test, but also for the teaching. Furthermore, we should mention that the students have to know the criteria according to which they are going to be evaluated.

To conclude we shall state again that achievement tests are meant to check the mastery of the material covered by the learners. They will be great helpers for the teacher’s future work and will contribute a lot to the students’ progress.

***3.5 Proficiency tests***

The last type of test to be discussed is a proficiency test. Regarding Longman Dictionary of LTAL (292) proficiency test is a test, which measures how much of a language a person knows or has learnt. It is not bound to any curriculum or syllabus, but is intended to check the learners’ language competence. Although, some preparation and administration was done before taking the test, the test’s results are what being focused on. The examples of such tests could be the American Testing of English as Foreign Language test (further in the text TOEFL) that is used to measures the learners’ general knowledge of English in order to allow them to enter any high educational establishments or to take up a job in the USA. Another proficiency test is Cambridge First Certificate test that has almost the same aim as TOEFL.

 Hughes (1989:10) gives the similar definition of proficiency tests stressing that training is not the thing that is emphasised, but the language. He adds that ‘proficient’ in the case of proficiency tests means possessing a certain ability of using the language according to an appropriate purpose. It denotes that the learner’s language ability could be tested in various fields or subjects (art, science, medicine, etc.) in order to check whether the learner could suit the demands of a specific field or not. This could refer to TOEFL tests. Apart from TOEFL we can speak about Cambridge First Certificate test, which is general and does not concern any specific field. The aim of this test is to reveal whether the learners’ language abilities have reached a certain standard set. The test could be taken by anyone who is interested in testing the level of language knowledge. There are special tests levels, which can be chosen by a candidate. If a candidate has passed the exam s/he can take another one of a different level. However, these entire tests are not free of charge, and in order to take it an individual has to pay for them.

Regarding Hughes (ibid.) who supposes that the only similar factor about such tests that they are not based on any courses, but are intended to measure the candidates’ suitability for a certain post or course at the university, we can add that in order to pass these tests a candidate has to attend special preparatory courses.

 Moreover, Hughes (ibid.) believes that the proficiency tests affect learners’ more in negative way, than in positive one.

The author of the paper both agrees and does not agree with the Hughes’ proposed statement. Definitely, this test could make the testee depressed and exhausted by taking a rather long test. Moreover, the proficiency tests are rather impartial; they are not testee-friendly.

However, there is a useful factor amongst the negative ones. It is preparation to proficiency tests, for it involves all language material starting from grammar finishing with listening comprehension. All four skills are being practised during the preparation course; various reading task and activities have been incorporated; writing has been stressed focusing on all possible types of essays, letters, reviews, etc. Speaking has been practiced as well. The whole material has been consolidated for many times.

To summarize we can claim that there are different types of tests that serve for different purposes. Moreover, they all are necessary for the teacher’s work, for them, apart from a proficiency test, could contribute to successful material acquisition by learners.

# Chapter 4

# Ways of testing

 In this chapter we will attempt to discuss various types of testing and if possible compare them. We will start with the most general ones and move to more specific and detailed ways of testing.

***4.1 Direct and indirect testing***

The first types of testing we are intended to discuss are direct and indirect testing. First, we will try to define each of them; secondly, we will endeavour to compare them.

We will commence our discussion with direct testing that according to Hughes (1989:14) means the involvement of a skill that is supposed to be tested. The following view means that when applying the direct testing the teacher will be interested in testing a particular skill, e.g. if the aim of the test is to check listening comprehension, the students will be given a test that will check their listening skills, such as listening to the tape and doing the accompanying tasks. Such type of test will not engage testing of other skills. Hughes (ibid.) emphasises the importance of using authentic materials. Though, we stipulate that the teacher is free to decide him/herself what kind of material the students should be provided with. It the teacher’s aim is to teach the students to comprehend the real, native speech, s/he will apply the authentic material in teaching and later, logically, in tests. Developing the idea we can cite Bynom (2001:8) who assumes that direct testing introduces real-life language through authentic tasks. Consequently, it will lead to the usage of role-plays, summarising the general idea, providing the missing information, etc. Moving further and analysing the statements made by the linguists (Bynom, 2001; Hughes,1989) we can posit the idea that direct testing will be task-oriented, effective and easy to manage if it tests such skills as writing or speaking. It could be explained by the fact that the tasks intended to check the skills mentioned above give us precise information about the learners’ abilities. Moreover, we can maintain that when testing writing the teacher demands the students to write a certain task, such as an essay, a composition or reproduction, and it will be precisely the point the teacher will be intended to check. There will be certain demands imposed on writing test; the teacher might be just interested in the students’ ability to produce the right layout of an essay without taking grammar into account, or, on the contrary, will be more concerned with grammatical and syntactical structures. What concerns testing speaking skills, here the author of the paper does not support the idea promoted by Bynom that it could be treated as direct testing. Definitely, you will have a certain task to involve your speaking skills; however, speaking is not possible without employment of listening skills. This in turn will generate the idea that apart from speaking skills the teacher will test the students’ ability to understand the speech s/he hears, thus involving speaking skills.

It is said that the advantages of direct testing is that it is intended to test some certain abilities, and preparation for that usually involves persistent practice of certain skills. Nevertheless, the skills tested are deprived from the authentic situation that later may cause difficulties for the students in using them.

Now we can shift to another notion - indirect testing. It differs from direct one in the way that it measures a skill through some other skill. It could mean the incorporation of various skills that are connected with each other, e.g. listening and speaking skills.

Indirect testing, regarding to Hughes, tests the usage of the language in real-life situation. Moreover, it suits all situations; whereas direct testing is bound to certain tasks intended to check a certain skill. Hughes (ibid.) assumes that indirect testing is more effective than direct one, for it covers a broader part of the language. It denotes that the learners are not constrained to one particular skill and a relevant exercise. They are free to elaborate all four skills; what is checked is their ability to operate with those skills and apply them in various, even unpredictable situations. This is the true indicator of the learner’s real knowledge of the language.

Indirect testing has more advantages that disadvantages, although the only drawback according to Hughes is that such type of testing is difficult to evaluate. It could be frustrating what to check and how to check; whether grammar should be evaluated higher, than composition structure or vice versa. The author of the paper agrees with that, however, basing on her experience at school again, she must claim that it is not so easy to apply indirect testing. This could be rather time-consuming, for it is a well-known fact that the duration of the class is just forty minutes; moreover, it is rather complicated to construct indirect test – it demands a lot of work, but our teachers are usually overloaded with a variety of other duties. Thus, we can only hope on the course books that supply us with a variety of activities that involve cooperation of all four skills.

***4.2 Discrete point and integrative testing***

Having discussed the kinds of testing that deal with general aspects, such as certain skills and variety of skills in cooperation, we can come to the more detailed types as discrete point and integrative testing. According to Longman Dictionary of LTAL (112), discrete point test is a language test that is meant to test a particular language item, e.g. tenses. The basis of that type of tests is that we can test components of the language (grammar, vocabulary, pronunciation, and spelling) and language skills (listening, reading, speaking, and writing) separately. We can declare that discrete point test is a common test used by the teachers in our schools. Having studied a grammar topic or new vocabulary, having practiced it a great deal, the teacher basically gives a test based on the covered material. This test usually includes the items that were studied and will never display anything else from a far different field. The same will concern the language skills; if the teacher’ aim is to check reading skills; the other skills will be neglected. The author of the paper had used such types of tests herself, especially after a definite grammar topic was studied. She had to construct the tests herself basing on the examples displayed in various grammar books. It was usually gap-filling exercises, multiple choice items or cloze tests. Sometimes a creative work was offered, where the students had to write a story involving a certain grammar theme that was being checked. According to her observance, the students who studied hard were able to complete them successfully, though there were the cases when the students failed. Now having discussed the theory on validity, reliability and types of testing, it is even more difficult to realize who was really to blame for the test failures: either the tests were wrongly designed or there was a problem in teaching. Notwithstanding, this type was and still remains to be the most general and acceptable type in schools of our country, for it is easy to design, it concerns a certain aspect of the language and is easy to score. If we speak about types of tests we can say that this way of testing refers more to a progress test (You can see the examples of such type of test in Appendix 2).

Nevertheless, according to Bynom (2001:8) there is a certain drawback of discrete point testing, for it tests only separated parts, but does not show us the whole language. It is true, if our aim is to incorporate the whole language. Though, if we are to check the exact material the students were supposed to learn, then why not use it.

Discussing further, we have come to integrative tests. According to Longman Dictionary of LTAL, the integrative test intends to check several language skills and language components together or simultaneously. Hughes (1989:15) stipulates that the integrative tests display the learners’ knowledge of grammar, vocabulary, spelling together, but not as separate skills or items.

Alderson (1996:219) poses that, by and large, most teachers prefer using integrative testing to discrete point type. He explains the fact that basically the teachers either have no enough of spare time to check a certain split item being tested or the purpose of the test is only considered to view the whole material. Moreover, some language skills such as reading do not require the precise investigation of the students’ abilities whether they can cope with definite fragments of the text or not. We can render the prior statements as the idea that the teachers are mostly concerned with general language knowledge, but not with bits and pieces of it. The separate items usually are not capable of showing the real state of the students’ knowledge. What concerns the author of the paper, she finds integrative testing very useful, though more habitual one she believes to be discrete point test. She assumes that the teacher should incorporate both types of testing for effective evaluation of the students’ true language abilities.

***4.3 Criterion-referenced and norm referenced testing***

The next types of testing to be discussed are criterion-referenced and norm referenced testing. They are not focused directly on the language items, but on the scores the students can get. Again we should concern Longman Dictionary of LTAL (17) that states that criterion-referenced test measures the knowledge of the students according to set standards or criteria. This means that there will be certain criteria according to which the students will be assessed. There will be various criteria for different levels of the students’ language knowledge. Here the aim of testing is not to compare the results of the students. It is connected with the learners’ knowledge of the subject. As Hughes (1989:16) puts it the criterion-referenced tests check the actual language abilities of the students. They distinguish the weak and strong points of the students. The students either manage to pass the test or fail it. However, they never feel better or worse than their classmates, for the progress is focused and checked. At this point we can speak about the centralized exams at the end of the twelfth and ninth form. As far as the author of the paper is concerned, the results of the exams are confident, and the learners after passing the exams are conferred with various levels relevant to their language ability. Apart from that, once a year in Latvian schools the students are given tests designed by the officials of the Ministry of Education to check the level of the students and, what is most important, the work of the teacher. They call them diagnostic tests, though according to the material discussed above it is rather arguable. Nevertheless, we can accept the fact that criterion-referenced testing could be used in the form of diagnostic tests.

Advancing further, we have come to norm-referenced test that measures the knowledge of the learner and compares it with the knowledge of another member of his/her group. The learner’s score is compared with the scores of the other students. According to Hughes (ibid.), this type of test does not show us what exactly the student knows. Therefore, we presume that the best test format for the following type of testing could be a placement test, for it concerns the students’ placement and division according to their knowledge of the foreign language. There the score is vital, as well.

***4.4 Objective and subjective testing***

It worth mentioning that apart from scoring and testing the learners’ abilities another essential role could be devoted to indirect factors that influence evaluating. These are objective and subjective issues in testing. According to Hughes (1989:19), the difference between these two types is the way of scoring and presence or absence of the examiner’s judgement. If there is not any judgement, the test is objective. On the contrary, the subjective test involves personal judgement of the examiner. The author of the paper sees it as when testing the students objectively, the teacher usually checks just the knowledge of the topic. Whereas, testing subjectively could imply the teacher’s ideas and judgements. This could be encountered during speaking test where the student can produce either positive or negative impression on the teacher. Moreover, the teacher’s impression and his/her knowledge of the students’ true abilities can seriously influence assessing process. For example, the student has failed the test; however, the teacher knows the true abilities of the student and, therefore, s/he will assess the work of that student differently taking all the factors into account.

***4.5 Communicative language testing***

 Referring to Bynom (ibid.), this type of testing has become popular since 1970-80s. It involves the knowledge of grammar and how it could be applied in written and oral language; the knowledge when to speak and what to say in an appropriate situation; knowledge of verbal and non-verbal communication. All these types of knowledge should be successfully used in a situation. It bases on the functional use of the language. Moreover, communicative language testing helps the learners feel themselves in real-life situation and acquire the relevant language.

 Weir (1990:7) stipulates that the current type of testing tests exactly the “performance” of communication. Further, he develops the idea of “competence” due to the fact that an individual usually acts in a variety of situations. Afterwards, reconsidering Bachman’s idea he comes with another notion – ‘communicative language ability’.

 Weir (1990:10-11) assumes that in order to work out a good communicative language test we have to bear in mind the issue of precision: both the skills and performance should be accurate. Besides, their collaboration is vital for the students’ placement in the so-called ‘real life situation’. However, without a context the communicative language test would not function. The context should be as closer to the real life as possible. It is required in order to help the student feel him/herself in the natural environment. Furthermore, Weir (ibid.) stresses that language ‘fades’ if deprived of the context.

 Weir (ibid., p.11) says: “*to measure language proficiency adequately in each situation, account must be taken of: where, when, how, with whom, and why the language is to be used, and on what topics, and with what effect.”* Moreover, Weirs (ibid.) emphasises the crucial role of the schemata (prior knowledge) in the communicative language tests.

 The tasks used in the communicative language testing should be authentic and ‘direct’ in order the student will be able to perform as it is done in everyday life.

 According to Weir (ibid.), the students have to be ready to speak in any situation; they have to be ready to discuss some topics in groups and be able to overcome difficulties met in the natural environment. Therefore, the tests of this type are never simplified, but are given as they could be encountered in the surroundings of the native speaker. Moreover, the student has to possess some communicative skills, that is how to behave in a certain situation, how to apply body language, etc.

Finally, we can repeat that communicative language testing involves the learner’s ability to operate with the language s/he knows and apply it in a certain situation s/he is placed in. S/he should be capable of behaving in real-life situation with confidence and be ready to supply the information required by a certain situation. Thereof, we can speak about communicative language testing as a testing of the student’s ability to behave him/herself, as he or she would do in everyday life. We evaluate their performance.

To conclude we will repeat that there are different types testing used in the language teaching: discreet point and integrative testing, direct and indirect testing, etc. All of them are vital for testing the students.

***Chapter 5***

***Testing the Language Skills***

 In this chapter we will attempt to examine the various elements or formats of tests that could be applied for testing of four language skills: reading, listening, writing and speaking. First, we will look at multiple-choice tests, after that we will come to cloze tests and gap filling, then to dictations and so on. Ultimately, we will attempt to draw a parallel between them and the skills they could be used for.

***5.1 Multiple choice tests***

 It is not surprising why we have started exactly with multiple-choice tests (MCQs, further in the text). To the author’s concern these tests are widely used by teachers in their teaching practice, and, moreover, are favoured by the students (Here the author has been supported by the equivalent idea of Alderson (1996:222)). Heaton (1990:79) believes that multiple-choice questions are basically employed to test vocabulary. However, we can argue with the statement, for the multiple choice tests could be successfully used for testing grammar, as well as for testing listening or reading skills.

It is a well-known fact how a multiple-choice test looks like:

* 1. **---- not until the invention of the camera that artists correctly painted horses racing.**
1. **There was**
2. **It was**
3. **There**
4. **It**

**“Cambridge Preparation for the TOEFL Test”:**

 A task basically is represented by a number of sentences, which should be provided with the right variant, that, in its turn, is usually given below. Furthermore, apart from the right variant the students are offered a set of distractors, which are normally introduced in order to “deceive” the learner. If the student knows the material that is being tested, s/he will spot the right variant, supply it and successfully accomplish the task. The distractors, or wrong words, basically slightly differ from the correct variant and sometimes are even funny. Nevertheless, very often they could be represented by the synonyms of the correct answer whose differences are known to those who encounter the language more frequently as their job or study field. In that case they could be hardly differentiated, and the students are frustrated. Certainly, the following cases could be implied when teaching vocabulary, and, consequently, will demand the students’ ability to use the right synonym. The author of the paper had given the multiple-choice tests to her students and must confess that despite difficulties in preparing them, the students found them easier to do. They motivated their favour for them as it was rather convenient for them to find the right variant, definitely if they knew what to look for. We presume that such test format as if motivated the learners and supplied them additional support that they were deprived during the test where nobody could hope for the teacher’s help.

Everything mentioned above has raised the author’s interest in the theory on multiple-choice test format and, therefore, she finds extremely useful the following list of advantages and disadvantages generated by Weir. He (1990:43) lists four advantages and six disadvantages of the multiple-choice questions test. Let us look at the advantages first:

* According to Weir, the multiple-choice questions are structured in such a form that there is no possibility for the teacher or as he places “*marker*” to apply his/her personal attitude to the marking process.

The author of the paper finds it to be very significant, for employing the test of this format we see only what the student knows or does not know; the teacher cannot raise or lower the marker basing on the students’ additional ideas displayed in the work. Furthermore, the teacher, though knowing the strong and weak points of his/her students, cannot apply this information as well to influence the mark. What s/he gets are the pure facts of the students’ knowledge.

Another advantage is:

* The usage of pre-test that could be helpful for stating the level of difficulty of the examples and the test in the whole. That will reduce the probability of the test being inadequate or too complicated both for completing and marking.

This could mean that the teacher can ensure his/her students and him/herself against failures. For this purposes s/he just has to test the multiple-choice test to avoid troubles connected with its inadequacy that later can lead to the disaster for the students receiving bad marks due to the fact that the test’s examples were too complicated or too ambiguous.

The next advantage concerns the format of the test that clearly implies the idea of what the learner should do. The instructions are clear, unambiguous. The students know what they are expected to do and do not waste their precious time on trying to figure out what they are supposed to do.

The last advantage displayed by Weir is that the MCQs in a certain context are better than open-ended or short-answer questions, for the learners are not required to produce their writing skills. This eliminates the students’ fear of mistakes they can make while writing; moreover, the task does not demand any creative activity, but only checks the exact knowledge of the material.

Having considered the advantages of MCQs, it is worth speaking about its disadvantages. We will not present all of them only what we find of the utmost interest and value for us.

The first disadvantage concerns the students’ guessing the answers; therefore, we cannot objectively judge his/her true knowledge of the topic. We are not able to see whether the student knows the material or have just luckily ticked or circled the right variant. Therefore, it could be connected with another shortcoming of the following test format that while scoring the teacher will not get the right and true picture of what the students really know.

Another interesting point that could be mentioned it that multiple-choice differ from the real-life situation by the choice of alternatives. Usually, in our everyday life we have to choose between two alternatives, whereas the multiple-choice testing might confuse the learner by the examples s/he even has not thought about. That will definitely lead to frustration, and, consequently, to the student’s failure to accomplish the task successfully.

Besides, regarding Weir (ibid.) who quotes Heaton (1975) we can stipulate that in some cases multiple-choice tests are not adequate and it is better to use open-ended questions to avoid the pro-long lists of multiple-choice items. This probably will concern the subject, which will require a more precise description and explanation from the students’ side.

To finish up with the drawbacks of MCQs we can declare that they are relatively costly and time-consuming to prepare. The test designer should carefully select and analyse each item to be included in the test to avoid ambiguity and imprecision. Furthermore, s/he should check all possible grammar, spelling and punctuation mistakes, evaluate the quality of information offered for the learners’ tasks and choose the correct and relevant distractors for the students not to confuse them during the test.

To conclude we can cite Heaton (1990:17) who stipulates that designing a multiple-choice items test is not so fearful and hard as many teachers think. The only thing you need is practice accompanied by a bit of theory. He suggests for an inexperienced teacher to use not more than three options if the teacher encounters certain difficulties in supplying more examples for the distractors. The options should be grammatically correct and of equal length. Moreover, the context should be appropriate to illustrate an example and make the student guess right.

***5.2 Short answer tests***

 A further format that is worth mentioning is short answer test format. According to Alderson (1996:223) short answer tests could be substitutes to multiple-choice tests. The only difference is that apart from the optional answers the students will have to provide short answers. The author of the paper had not used this test format, thus, she cannot draw on her experience. Therefore, she will just list the ideas produced by other linguists, to be more exact Alderson’s suggestions.

Alderson (ibid.) believes that short answer tests will contribute to the students’ results, for they will be able to support their answers and, if necessary, clarify why they responded in that way but not the other. It could be explained that the students will have an opportunity to prove their answers and support them if necessary.

 Nevertheless, the short answer tests are relatively complicated for the teacher to be designed. The teacher has to consider a variety of ideas and thoughts to create a fairy relevant test with fairly relevant items. May be that could explain the fact why this test format is not such a common occasion as MCQs are.

 At this point we have come to advantages and drawbacks of short answer tests. Weir (1990:44) says that this type of testing differs from MCQs by the absence of the answers. The students have to provide the answer themselves. That will give the marker the clear idea whether the students know what they write about or not. Certainly, the teacher will be definite about the students’ knowledge, whereas in MCQs s/he can doubt whether the students know or have just guessed the correct answer. Moreover, short answer test could make the students apply their various language skills techniques they use while dealing with any reading, listening or speaking activity.

Finally, Weir (ibid.) stipulates that if the questions are well formulated, there is a high chance the student will supply short, well-formulated answer. Therefore, a variety of questions could be included in the test to cover a broader field of the student’s knowledge, and certainly it will require a great work from the teacher.

Nevertheless, there are certain drawbacks displayed by the following test format. One of the major disadvantages could be the students’ involvement in writing. For if we are determined to check the students reading abilities, it is not appropriate to give the students writing tasks due to the high possibility of the spelling and grammar mistakes that may occur during the process. Therefore, we have to decide upon our priorities – what do we want to test. Furthermore, the students while writing can produce far different answers than expected. It will be rather complicated to decide whether to consider them as mistakes or not.

***5.3 The cloze test and gap-filling tests***

Before coming to the theory on cloze tests we assume that it is necessary for us to speak about a term “cloze”. Weir (1990:46) informs that it was coined by W.L. Taylor (1953) from the word ‘closure’ and meant the individual’s ability to complete a model.

However, to follow the model one has to posses certain skills to do so. Hence, we can speak about introduction of such skill that Weir calls deduction. Deduction is an important aspect for dealing with anything that is unknown and unfamiliar. Thus, before giving a cloze test the teacher has to be certain whether his/her students are familiar with the deduction technique.

 Alderson (1996:224) assumes that there are two cloze test techniques: pseudo-random and rational cloze technique. In the pseudo-random test the test designer deletes words at a definite rate, or as Heaton (1990:19) places it, systematically, for example every 7th word should be deleted occasionally with the initiate letter of the omitting word left as a prompt:

 *Although you may think of Britain as England ,i...is really four countries in one. There a.. …..four very distinct nations within the British I………: England, Scotland, Wales and Ireland, each with their o…..unique culture, history, cuisine, literature a…..even languages.*

*(Discovering Britain, Pavlockij B. M., 2000)*

 However, the task could be more demanding if the teacher will not assist the learners’ guesses and will not provide any hints:

 Scotland is in the north and Wales in the west were………separate countries. They have different customs,……………….., language and, in Scotland’ s case, different legal and educational……………….

(ibid.)

 The examples shown above do not yield to be ideal examples at all. Without doubt, the material used in the task should more or less provide the students with the appropriate clues to form correct guessing. Notwithstanding, the author of the paper has used such tests in her practice and according to her observations; she can conclude that the tasks with the first letter left are highly motivating for the students and supply a lot of help for them. Moreover, having discussed the following test format the teacher has revealed that the students like it and receive a real pleasure if they are able to confirm their guess and find the right variant.

However, according to Alderson (ibid.), the teacher commonly does not intend to check a certain material by the cloze test. The main point here is the independence of the student and his/her ability to apply all the necessary techniques to fill in the blank spaces. Concerning the mentioned-above scholars, we have to agree that the following type of test is actually relatively challenging, for it demands vast language knowledge from the student. Heaton (ibid.) believes that each third or fourth deleted word can turn into the handicap for the learner due to the lack of prompting devices, such as collocations, prepositions, etc. Whereas, the removal of each ninth word may even lead to the exhausting reading process.

 On the contrary, the rational cloze technique, or as it is usually called gap-filling, is based on the deletion of words connected with the topic the teacher wants or intends to check. At this time the teacher controls the procedure more than it is in the pseudo-random test discussed above. Moreover, s/he tries to delete every fifth or sixth word, but does it rather carefully not to distort the meaning and mislead the learner. Besides, a significant factor in this type of testing is that the teacher removes exactly the main words that are supposed to be checked, i.e.:

 *Britain…….a deceptively large island and ……surrounded by some very beautiful coastline. The south of England has popular sandy beaches, especially in the west. But the coast in the south west Wales…..a unique coastal National Park. Its beaches…… great for sunbathing and the rock pools and cliffs ……..havens for wildlife. Up in Scotland, the striking white beaches of the west coast and islands……excellent places for explorative walks.*

*(Discovering Britain, Pavlockij B. M., 2000)*

It is evident that the teacher’s aim by the help of the rational cloze test is to check the students’ knowledge of the Present simple of the verb *“to be”.* Thereof, the cloze tests could be successfully used for testing grammar, as well.

We have come again to the point when we are going to mention the advantages and disadvantages of cloze and gap-filling testing coined by Weir. Regarding Weir, there are more disadvantages than advantages in applying the cloze tests. He says that to design a cloze test is fairly easy, and they are easy to evaluate, and it is the best means to check reading comprehension. Concerning the drawbacks, we can emphasise that randomly removed words usually will act as distractors and will not be of true importance for the students to comprehend a message if, for example, it is a reading task.

 Compared to the cloze test, gap filling is more material based, for it checks the students’ knowledge of a particular topic. Therefore, we can speak about the first advantage that is the learners will know exactly what they should insert. Moreover, the selectively deleted items allow focusing exactly on them and do not confuse the student.

The last what could be said about gap filling tests is that this technique limits us to check only a certain language skill, e.g. a vocabulary on different topics.

***5.4 C-Tests***

 It is worth mentioning that in the 80s German school introduced an alternative to cloze test another type of testing – C-Tests. This test was based on the cloze test system; however, every second word there was deleted. It could seem quite a complicated type, though it is not. According to Weir (1990:47) in this type every deleted word is partially preserved. Thus, the students, if they possess a fairly good knowledge of the language and can activate their schemata, or background knowledge of a topic or the world, they will succeed in completing the test. Such test format could look as follows:

 *Cats ha…. always been surro………by superstitions. In anc……Egypt ca….were cons……. sacred, but in medi…..Europe ma….. people beli…… cats we…. witches in disgu…… A popular supers……... about ca…. is that a blac…cat, cros… your pa… from left to rig…., will bri… you bad lu…. However, in some cult….. a black ca… is thought to be a go… omen rat… than a ba… one.*

*(First certificate Star, Luke Prodromou, p.134)*

 Definitely there are advantages and disadvantages of the following test format. According to Weir, due to the frequency of the deleted items there is a great possibility to include more tested items in the test. Moreover, this test is economical. However, despite all the advantages, the test can mislead the students as it is fragmented. The examples are deprived from the context that could be very helpful for the students’ guessing of the missing parts.

***5.5 True/False items***

 This test format is familiar for all the teachers and students. Each reading task will always be followed with true/false activities that will intend to check the students’ comprehension of a text. The students will be offered a set of statements some of which are true and some are wrong, e.g.:

* + 1. *People went to see ‘Cats’ because of the story. T F*
		2. *Lloyd Webber’s father helped his career. T F*
		3. *Lloyd Webber comes from a musical family. T F*

*( Famous Britons, Michael Dean)*

 They usually should be ticked, and in order to tick the correct variants the students have to be able to employ various guessing strategies.

 According to Weir (1990:48), the advantage of such test is found in its applicability and suitability. One can write more true/false statements for a test and use them to check the students’ progress or achievement. Furthermore, the current sort of testing could be more motivating for the students than a multiple-choice test. It will not make the students confused offering just one possibility than a multiple-choice test, which typically proposes more than one option to choose from. Moreover, it is easy to answer for the students and check for the teachers.

***5.6 Dictation***

Another test format that could be applied in the language classroom is dictation. We commonly use dictations to check spelling; nevertheless, it could be applied to test listening comprehension, as well. It is obvious that to dictate something we have either to speak or read. It means that while writing a dictation the student has to be able to perceive the spoken language efficiently enough to produce in on paper. For this purpose the student will require a variety of techniques such as schemata and its application, predictions, guessing and context clues, etc. Further, it also is constrained that dictation help the students develop their abilities to distinguish between phonemes, separate words and intonation. Besides, dictations function in spoken language; thereof the students have an opportunity to learn to understand the language through listening. To conclude what has been mentioned above we can agree with Weir (1990:49) that dictations will force the students to use the variety of skills: listening, reading, speaking and writing skills.

Heaton (1990:28) advises that to enable the students comprehend successfully, the teacher need to read carefully and clearly, however avoiding slow, word for word reading. Moreover, to allow the students to check what they have written the repetition will be required. The author of the paper when giving dictations to her students had encountered the need for repetition for a number of times. The following could be explained by many factors, such as the students are not able to perceive spoken speech through listening; they are not able to elaborate various guessing, inferring of the meaning techniques or their pace of writing is simply rather slow. Thus, we entirely support the next statement claimed by Heaton that it is wise after the first reading of a dictation to ask a set of comprehension questions to make the students aware of the general idea of a text. It will simplify the process of the understanding.

Notwithstanding, even an ideal variant will definitely contain some drawbacks. The same could be applied to dictations. First, to write a dictation, the student requires a good memory. S/he has to retain information they have heard in order to display it later; moreover, the information should be identical to the original. Therefore, we can claim that the student has to recognize at least seventy-eighty per cent of what has been dictated. In that case we short-term memory should be well developed.

Apart from memory, scoring could be problematic, as well. Weir (1990:50) believes that is difficult to decide what to pay attention to: whether to evaluate spelling and grammar, or just perceived information. Thus, the teacher has to work out a certain set of criteria, as we have already mentioned that in Chapter 1, the criteria s/he will be operating with. Besides, the students should be acquainted with it, as well.

In addition, Weir (ibid.) says that dictating is more efficient if it is recorded on the tape and is delivered by a native speaker. It could mean that the students will have a chance to fell themselves in the real-life situation; for this is the actual purpose they learn the language for. The following has been expanded by Heaton (ibid.) that speaking face to face with a speaker is even more beneficial, for we can compensate the lack of understanding by his/her facial expression, gestures and movements. Listening to a cassette does not provide us with such a chance, and therefore, it is more challenging and requires more developed skills to understand a recorded message.

***5.7 Listening Recall***

 This test format is specifically applied to testing listening skills. It differs from a dictation that it supplies the students with a printed text. However, the text is given not as the complete script of the tape. Certain words that carry the meaning load are deleted from a passage, and the students after listening to the tape are supposed to insert them. Hence, it could be related to a gap-filling test. Here the cassette is usually played for two times; first, the students listen for information and attempt to insert the missing details. The second time allows them to add what they had failed to understand at the beginning. The author of the paper had not used that as a direct test format but as a while-listening activity during her classes. According to her scrutiny the students with more advanced language abilities were able to comprehend the texts immediately, whereas the weaker students sometimes could not manage to understand the message even listening for the tape for the third time. That again proves the significance of usage of pre-, while and post-listening activities in the language classroom. Weir (ibid.) states that such type of testing involves the students’ short-time memory, which they need to switch while listening to the tape.

 According to Weir (ibid.), one of the advantages of listening recall is uncomplicated construction, administration and marking.

 Nevertheless, there are several disadvantages, as well. There is a danger, that the students will read the passage before listening to the tape, thus we will not be able to evaluate exactly their listening skills. The author of the current paper had encountered the similar situation, where the teacher warns the students not to read but just listen. However, they start reading immediately after receiving the text, even though the tape record being still turned off.

***5.8 Testing Grammar Through Error-recognition Items and Word Formation Tasks***

 One of the test formats for testing grammar is error-recognition items. Here the teacher writes sentences underlining various words. One of the words is obligatory wrong, and the students have to identify what word is wrong and should be corrected. Heaton (ibid.) introduces a variation of that type, saying that the teacher can supply the students with incorrect sentences asking the students to provide the right variant. This again demands a fairly good knowledge of the subject from the students to differentiate between the right and wrong variants. In that case the error-recognition format could be compared with multiple-choice format and even called a branch of it. Below you can find the example of error-recognition items format:

* 1. *I can’t come to the phone – I have / I’m having a shower!*
	2. *I watched/ I was watching TV when suddenly the telephone rang.*
	3. *I had been waiting/ I had waited in the rain for ages when she finally turned up.*

*(First certificate Star, Luke Prodromou, p.12)*

 Further, for testing grammar and language structures we often use word-formation tasks, e.g.:

*Making friends and ………people is a gift that some influence*

*………….people seem to be born with, while for others it luck*

*is a skill that has to be ……..through practice and acquire*

*hard work. It is, however, …….to know that most skills, comfort*

*particularly ………….skill, can be learnt and that it is never society*

*too late to start improving.*

*(First certificate Star, Luke Prodromou, p.41)*

or

|  |  |  |  |
| --- | --- | --- | --- |
| ***verb*** | ***noun*** | ***person*** | ***Adjective*** |
| Invent |  |  |  |
|  |  | discoverer | - |
|  | creation |  |  |

 It is frequently used in centralized exams to know the students’ ability to coin new words that displays the students’ advanced level of the language. The students are demanded coining nouns from verbs, adjectives from nouns, etc. This requires certain knowledge of prefixes, suffixes and roots in order to create a necessary word. Word coinage is an inevitable skill for recognizing new word items either.

***5.9 Controlled writing***

 In order to check the students grammar and writing ability the teacher can use different test formats: transformation, broken sentences, sentence and paragraph completion, form filling, notes and diaries.

 According to Heaton (1990:32), transformation deals with re-writing sentences. For example, the students are asked to change a sentence in Active voice into a sentence in Passive voice. To differ the task the teacher can put the required word in brackets at the end of each sentence. The students will need to transform a sentence to fit the word in brackets. Or another example of transformation could be changing the focus of the sentence, e.g.:

1. *Berlin is not an easy city to move about in.*

*Difficult*

*It………………………in Berlin.*

1. *I wonder if you could open the window.*

*Could*

*You couldn’t ………………….*

1. *When did you start to learn English?*

*Been*

*How…………………….English?*

*(First certificate Star, Luke Prodromou, and p.40)*

 Further, he discusses the sentences that are divided into fragments (he calls them broken sentences), and the student’s task is to arrange the words in order to produce correct examples. Thus, the students have to know grammar and syntaxes to make a right sentence with the correct word order. Sometimes the students are asked to alter the words to make grammatically correct sentences, e.g.:

1. *a German/hunting/huge/black dog*
2. *a 25-year-old/Opera/tall singer*
3. *a brand-new/plastic/shopping/green bag*
4. *an English/young/interesting teacher*

*(First certificate Star, Luke Prodromou, and p.80)*

 Afterwards, the students can be asked to complete the whole paragraphs, finish dialogues, write diaries using the given information, and fill the form, for example hotel check-in. The author of the paper had used writing a diary in her 8th form, when the learners had to write the diary of captain’s wife whose husband disappeared in the sea. They also had to write the diary of the captain himself before the catastrophe. The students liked the task immensely.

***5.10 Free writing***

 Heaton (ibid.) believes that the most suitable way to check the students’ writing skills is asking them to write a composition. The teacher can include a variety of testing criteria there depending on what is really being tested. The topics for a composition should be appropriate to the age of the students and respond to their interest. However, the teacher has to establish clearly what s/he is going to check (the material studied: e.g. grammar) and what could be neglected. The students have to know whether the teacher is interested in the context or may be s/he is concerned with grammar and spelling, as well.

***5.11 Test Formats Used in Testing Speaking Skills***

 We are not going to deep into details of test formats used for testing speaking skills. Heaton (ibid.) displays that one of the most essential elements of testing speaking is pronunciation. To check how the students pronounce certain testing items the teacher may ask his/her students to read aloud and retell stories. Moreover, the teacher will receive the impression how well his/her students can operate with the spoken language.

 Afterwards, the teachers can use pictures to test the students’ speaking skills. This is widely used task, and a lot of teachers use it to check the students speaking skills and the knowledge of the vocabulary. Moreover, while describing the picture the student will have to imply the correct grammar and knowledge of the English sentence structure. The description could be done on the spot and does not require a lot of time for preparation, though Heaton (ibid.) stipulates that the teacher should ensure his/her students with a time during which they can formulate their ideas before presentation.

Apart from the pictures the students could be offered to describe a person if their topic is people’s appearance or jobs, stay the sequence of events basing on the provided information or pictures accompanying the task, spot differences between two pictures and compare them. Further, Heaton (ibid.) displays a rather interesting task. The students receive a picture with speech bubbles. They are asked to write what they think people are saying. This in turn involves creativity from the students and could be assessed as an additional element and contribute to the students’ marks. Definitely, each teacher will develop and give the students various tasks regarding the criteria and demands to be tested.

 In conclusion we can say that the teacher can use a variety of test formats, such as multiple-choice questions, transfer of information; reordering the words, describe a picture, listening to the instructions to check the language skills of his/her students. Every teacher has to choose him/herself the tasks that will be appropriate to their way of teaching and the needs of the students.

 Below we have attached the table of four language skills and test formats applicable for each skill.

|  |  |
| --- | --- |
| ***Language Skills*** | Test Formats |
| Reading skills | 1. 1 Multiple-choice items
2. Short answers test
3. Cloze test
4. Gap-filling test
5. False/true statement
 |
| Listening skills | 1. Multiple-choice items
2. False/true statements
3. Gap-filling tests
4. Dictations
5. Listening recall
 |
| Writing skills | 1. Dictations
2. Compositions
3. Reproductions
4. Writing stories
5. Writing diaries
6. Filling-in forms
7. Word formation
8. Sentence transformation
 |
| Speaking skills | 1. Retelling stories
2. Describing pictures
3. Describing people
4. Spotting the differences
 |

## Chapter 6

**Analysis of the Test of English as a Foreign Language and Cambridge First Certificate test according to test design criteria.**

The present chapter deals with the practical part of the research. It will be based on the analysis and comparison of two proficiency tests formats TOEFL (The Test of English as a Foreign Language) test and CFC (Cambridge First Certificate) test. We will start with the brief description of their overall features; afterwards we will make an attempt to contrast them and draw relevant conclusions.

The first test to be discussed is Cambridge First Certificate test. It will usually consist of five papers: reading with the duration time 1 hour and 15 minutes, writing -1 hour 30 minutes, use of English -1 hour 15 minutes, listening - 40 minutes and speaking approximately 14 minutes. There is no absolute pass mark, but the candidates need to get about 60% of the total marks to pass with a Grade C (Prodromou, 1998:6-7).

TOEFL test is an examination that intends to evaluate the level of the English language of a foreign speaker (Gear, 1996:3-4). Moreover, it is commonly one of the aspects included into the entrance exams of any university in the USA. The institution the person requires the test for could implement the demanded score here. Nevertheless, the highest score does not differ from that of the CFC. TOEFL test as CFC test consists of four different parts: listening comprehension that occupies approximately 35 minutes and consists of three parts, structure and written expression with time limit 25 minutes composed of two tasks and reading comprehension – 55 minutes, consisting of several passages.

Here we can notice some differences between CFC and TOEFL tests: when TOEFL test consists of just four parts, CFC includes a speaking part more. Moreover, each part of each test will include a various range of tasks, i.e. each part TOEFL test will mainly be composed of two tasks, whereas CFC will classically contain four different activities.

Doing the tests in both cases the students will get special answers sheets where they will have to mark the answers they think are the right ones. The instructions before the taking the test usually warn the participants not to write in the books with questions. Moreover, both tests are checked by the scoring machine, therefore the students should be aware of what type of marking the answers they have to use. In both cases it should be a black lead pencil for the scoring machine to read. The answers should not be circled or lightly marked; in TOEFL test the students are supposed to fill in an oval answer with a letter inside corresponding to the question, whereas in CFC the students will have to fill in a small rectangular under a certain letter. Together the two tests remind the participants to choose only one answer. If the student changes his/her mind and decides to choose another answer, s/he can easily rub out the previous answer.

We can call the both tests valid, for they test what is supposed to be tested and measured and they usually have the same format and length; regarding reliability, we cannot say exactly whether there is reliability or not, for if the student was not lucky for the first time taking the test, s/he can study hard and take the test again for the second time and, thus, score a better result.

Both of the tests involve the four skills: reading, listening, speaking and writing. The difference could be found in the sequence of them, for example if CFC test will start with reading first, TOEFL test will deal with listening. The types of tasks and activities implied in the test differ as well. We will start our analysis with reading part.

***Reading Comprehension Part***

CFC reading paper will test the students’ ability to read in a variety of ways: reading for gist (understanding of the text), reading for details, understanding how a text is organized and deducing the meaning from the context. (Typically, the students could be given four parts to fill) (Prodromou, 1998:8). For that purpose CFC reading paper will offer the students multiple matching. The students will have to match headings or summary sentences to the parts of the text. They will have to show their ability to grasp the overall meaning of the text involving various kinds of knowledge such as morphological, semantic and syntactical one. For example:

*Meet the Flinstones, a modern Stone Age family. From the town of Bedrock, here’s a bit about their history….*

1. *Somewhere in the world, every hour of every day, The Flinstones is being broadcast. An incredible 300 million fans tune in to watch it regularly. Whether you like them or not, Fred, Wilma and their neighbours are impossible to avoid….*

*A) Rocky jokes B) A Stone Age family in skins C) A new idea D) A popular show, etc.*

 *Prodromou, First Certificate Star, 1998*

 Thus, basing on the theory we have discussed in the first part of our paper, we can claim that it is integrative type of test, though being direct, that denotes testing one particular skill directly, but not through other language skills.

Afterwards, CFC may offer the students multiple choice, gapped texts and again multiple matching only connected with information. In multiple-choice activity the students will have to answer four-option multiple-choice questions about a text. For example:

*Mad Cow Disease is a deadly illness of the brain and it is the non-technical term for BSE or Bovine Spongiform Encephalitis. This so difficult to say that journalists and even some doctors prefer the more vivid Mad Cow Disease…*

1. *We use the expression Mad Cow Disease because*
	1. *it is more accurate.*
	2. *It is easier to say.*
	3. *It links cows with people.*
	4. *It sounds less scientific.*

*Prodromou, First Certificate Star, 1998*

It is obvious that only one answer will be the right one, but the others will be distractors that will try to confuse the reader. It will limit the students and make them use a variety of reading strategies, knowledge of vocabulary and syntaxes to discover the right variant. However, the students will not have an opportunity to support their choices and prove why the answer they have chosen is the exact one. Moreover, the students will be checked whether they understand the general meaning of the text, its details, whether they can infer the meaning from the text and understand references (who refers *who*). Thus, we can declare that this type of test is integrative, for it involves the students’ abilities to apply various reading strategies and still direct, for it tests just reading skills.

The same could be said about gapped texts that will check the students’ knowledge of reading strategies, such as organization of the text, reading for gist, etc. (examples available in Appendix p.17) To complete it the students will have to show their knowledge of the certain areas of the language. Multiple matching will require the students to match pieces of information either with a certain text divided into fragments or with several texts joined together with one topic (examples available in Appendix p.8).

CFC will display various types of texts in order to see how well the students can cope with any authentic material when dealing with reading. They will have to show their capability of dealing with advertisements, letters, stories, travelling brochures, guides, manuals, and magazine and newspaper articles. The type of test applicable here will be integrative, including a variety of strategies and direct checking the students reading skills.

TOEFL test’s reading part usually involves the students’ general comprehension of a text. It is regularly a text followed by a number of questions about it typically in the form of multiple choice items format. However, this part of the test requires the students to show their skills in reading for gist, the students have to define the main idea of a text; afterwards, the students will have to display their knowledge of the vocabulary, especially synonyms, ability to infer the meaning, define the words and apply their skills connected with working with references, i.e.:

*….The biggest disadvantage was that the sound and pictures could become unsynchronised if, for example, the gramophone needle jumped or if the speed of the projector changed. This system was only effective for a single song or dialogue sequence…..*

*47. The word “sequence” in line 14 is closest in meaning to*

1. *interpretation*
2. *progression*
3. *distribution*
4. *organization*

*Gear, Cambridge preparation for the TOEFL test, 1996*

The students will be offered to read several passages, usually historical, scientific, medical, etc. facts. They will intend to check the students’ ability to understand specific types of tests taken from specific fields, the skill required at the universities, whereas CFC will offer the students the texts they can encounter in their everyday life being abroad. Each text will be typically accompanied with seven questions.

TOEFL test will chiefly use multiple-choice items; there will be no gap filling or matching implied. Thus, we can call a reading part of TOEFL test as a direct, for it tests the students reading skills, and more discrete point tests than integrative, for it is mainly concerned with checking the students’ knowledge of vocabulary (examples available in Appendix p.391-396).

The above mentioned could be stated as the first difference: TOEFL test is a discreet point test, while CFC is integrative one.

Another difference between CFC and TOEFL reading part could be a variety of tasks given to test the students’ reading skills. CFC will mostly offer a great range of tasks (headings, summary, fragmented texts) and texts types, while TOEFL will not vary a lot.

***Listening Part***

The listening part of CFC test aims to test the students’ ability to listen and understand the gist, the main points, and specific information is to deduce meaning. TOEFL test will check whether the students are able to understand conversations and talks in English.

CFC test will offer the students a variety of activities in order to check whether the students can imply effective listening strategies to comprehend the message. It suggests the idea of the test being integrative, for it will focus on different means that could be used to deal with a listening task. For example, CFC offers multiple choices as a task (examples available in Appendix p.37): the students listen to several short extracts that are taken from different contexts. They could be dialogues or monologues as well. The answer sheet will display the three answer items from which the students will have to choose the correct one. The task could ask the students to guess who the speaker is, where the action takes place, what the conversation is about and even it can include the question about the feelings and emotions of speakers that could be guessed from the contexts.

Afterwards, there will be another task – note taking or blank filling that will check the student’s ability to listen for gist and for details. This type will demand the student’s capability to use his/her writing skills to put down information they will hear. They will have to be able to pick up the necessary information and retain it in their memory in order to fulfil the task (examples available in Appendix p.87).

Subsequently, a further task could engage multiple matching where the students will have to concentrate on a particular kind of information. This task could be displayed in the form of a dialogue or a monologue. The students will be given several answers with letters that should be inserted into the right box. However, there will always be one option that does not suit any question, the so-called distractor. Moreover, asking the students to complete a grid, i.e. advantages and disadvantages of anything, could expand the task, i.e. advantages and disadvantages of keeping a certain pet:

|  |  |  |
| --- | --- | --- |
|  | ***Advantages*** | ***Disadvantages*** |
| ***dog*** |  |  |
| ***cat*** |  |  |
| ***fish*** |  |  |

 *Prodromou, First Certificate Star, 1998*

Moreover, the listening task could involve True/false activities where the students will have to listen to a dialogue or a monologue and react to it (examples available in Appendix). The students will have to display how well they have comprehended the message ticking the statements whether they are true or false. In spite of that, Yes/No questions could take place. We have been discussing them already in our theoretical part and mentioned that the so-called open-ended questions allow the students support their answers. Answering them, the students are having a chance to prove why they have chosen a certain answer, but not the other. Usually, if the students are aware of such a possibility, they fill more secure and motivated, for they can be certain that the examiner will be able to realize the students’ point. However, it is not a very appropriate type for such a test as CFC, for checking such tests will be rather time-consuming.

Listening part of the TOEFL test differs a lot from that of CFC, for it is fully based on the multiple-choice items that focus mainly on the understanding of the main idea of a message (examples available in Appendix p.379-384) The participants are exposed to a set of short dialogues that are accompanied with four answers, where three are usually distractors and the rest one is correct, i.e.:

*(man) I think, I’ll have the curtains changed.*

*(woman) They are a bit worn.*

*(narrator) What does the woman mean?*

1. *She thinks every bit of change is important.*
2. *She wants to wear them.*
3. *She thinks they’ve been worn enough.*
4. *She thinks they’re in bad condition.*

*Gear, Cambridge preparation for the TOEFL test, 1996*

 The test implies the idea that to do it the students have to use a variety of listening strategies, but it is not directly aimed at it. Whereas, the listening part of CFC test is structured so that the students would be able to display their listening skills and strategies, that are so useful for them to comprehend the real message in the real-life situation dealing with a native speaker.

Thus, we can distinguish certain similarities and differences, which we can encounter comparing them. They are both direct aiming at checking one exact skill; however, CFC is integrative, but TOEFL is discreet point test. Moreover, the test formats differ as well. CFC is richer in activities, than TOEFL test, which offer the students just multiple-choice items test.

The author of the paper presumes that CFC listening part is more testee-friendly, while TOEFL listening part is more “reserved” and does not allow the students fill free, but alarmed.

***Writing Part***

Writing part of CFC test tests the students’ ability to write different types of writing texts. These could be transactional letters, simple letters, compositions, descriptions, reports, etc. Moreover, the students could be asked to write an opinion composition and even an article (examples available in Appendix p.38).

Transactional letters are aimed at making somebody do something. Writing them, the students have to keep in mind that they are supposed to get a relevant answer.

There are different types of transactional letters, such as a letter of complaint, a letter of invitation, a letter asking for information and a letter describing something. The task requiring the students ability to write these letters will supply the students with necessary information, may be even pictures, and usually will ask for the students’ personal opinion. Moreover, the students have to be aware of the style that should be used depending on the requirements. Furthermore, the students will have to know how the letters are structured, for it will be the factor that will be evaluated as well.

Another writing task such as writing articles for a magazine will require the students to display their writing abilities, the knowledge of the vocabulary, the style and letter organization knowledge (examples available in Appendix 38).

Writing a report will be based on the students’ capability to gather facts and analyse them. It could involve a kind of a research work and knowledge how to express and link the ideas together (examples available in Appendix 30).

Writing a narrative story will ask the creativity from the students to make it interesting and original. Again the students will have to be able to express and link their ideas to produce a meaningful text.

Opinion composition will involve the students’ abilities to state advantages and disadvantages of the topic being discussed, expressing own opinion, stating the problem and possible solutions of it and expansion on the topic analysing various aspects of a topic.

Another writing task could be a book review. The students will have to know how to plan and organize the review, giving brief information about an author and some essential details about a book. Moreover, personal opinion of the students will be required as well.

Thus, looking at the facts stated above we can declare that the writing part of CFC is purely integrative type of test, for it involves all possible written tasks and strategies that should be used to accomplish the tasks effectively. Furthermore, it will be a direct testing aimed at testing the students’ writing skills. The tasks and activities presented in this part of CFC reflect the students’ needs they may meet in a real-life situation, for every possible writing piece is given.

The writing part of TOEFL test will generally involve essay writing. There will not be any letters or book reviews. The students will be given a topic that is typically a statement and they will have to expand it and write about it giving the facts, ideas and sometimes even a personal opinion, i.e.: ***“ If the earth to be saved from environmental catastrophe, we shall all have to make major changes in our lifestyles” (****Gear, Cambridge preparation for the TOEFL test, 1996)*. This type of writing will focus on expressing ideas and their linking as well. To write a good essay the students will require the knowledge of the topic, or schemata, the knowledge of a relevant vocabulary, appropriate style and organization of the written text, i.e. thesis sentence, paragraphs, etc (examples available in Appendix p. 377 – 378).

Therefore, we can conclude that the writing part of TOEFL test could be called also an integrative type of test involving the range of strategies. Moreover, it could be defined as direct testing, for it implies testing exactly the writing skill. Furthermore, it is totally based on the knowledge how to organize an essay with all necessary paragraphs, introductions and conclusions.

***Use of English or Structure and Written Expressions***

An import role in both tests is occupied by use of English or as it is called in TOEFL Structure and Written Expressions part. It aims at testing the students’ knowledge of grammar and vocabulary used in the English language.

CFC offers the students a range of various activities and task to be done during the testing time. They are multiple choice cloze, open cloze, key word transformations, error correction and word formation. Whereas, the usual procedure of the same part in TOEFL test will mostly include multiple-choice cloze and error correction.

The multiple choice cloze in CFC will usually be in the form of a gapped text followed by fifteen multiple questions with four options, as always the only one will be the correct. It will mostly be concerned with vocabulary items or grammar issues (examples available in Appendix p.44). For example:

*Robin Williams was creative and gifted from an early age. He was a/an (1)\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_child and at school was always a (2)\_\_\_\_\_\_\_\_\_\_\_\_\_pupil: he wrestled, ran cross-country and worked (3)\_\_\_\_\_\_\_\_\_\_\_\_\_at his studies.*

*1. A imaginary B imaginative C fantastic D mythical*

*2. A classic B model C superior D spoilt*

*3. A quickly B easily C hard D fast*

*Prodromou, First Certificate Star, 1998*

Open cloze will mostly be presented in the form of a text with several spaces, which the students will have to complete with an appropriate word. It will imply the students’ knowledge of grammar and vocabulary and will involve the students’ ability to predict and guess from the context (examples available in Appendix p.94). The task will be rather complicated, for it will not be a C-test type where the words to be inserted preserve the initial letter or letters to make the guessing process easier. In our case the students will have to know how the words and phrases are connected together, how the sentences are linked, and they will have to know the grammar forms and structures, so, for example, if they see have/has, they should immediately know that Present perfect is used. For example:

*When you join the International Bird Society, your membership (1)\_\_\_\_\_\_\_\_\_\_\_\_\_ make a positive difference to birds everywhere – even if the only ones you see are the blue tits…..*

*Prodromou, First Certificate Star, 1998*

Key word transformations will make the students alter the sentences structures, however preserving the entire meaning of them. They will have to complete a sentence with a given word; here the vocabulary and grammar will be of major interest again (examples available in Appendix p.86). The usual change will occur with phrasal verbs, active and passive voice, verbs and prepositions that go together, etc.:

*1. I didn’t like the story and I didn’t like the actors.* ***neither***

*I \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_the actors.*

*Prodromou, First Certificate Star, 1998*

Error correction will implement the students’ knowledge of grammar structures. The students will receive a passage in which they will have to find incorrect item and highlight it (examples available in Appendix p.55). Such types of activities will usually include an extra or unnecessary word. These words could be relative pronouns, prepositions, articles, conjunctions, etc. For example:

*\_\_\_\_\_\_\_\_ If you want to find out about someone’s personality, one way of to do it is to*

*\_\_\_\_\_\_\_\_take a sample of their handwriting and analyse it; this is called by*

*\_\_\_\_\_\_\_\_graphology. To do graphology properly, it is important to use fairly typical…..*

 *Prodromou, First Certificate Star, 1998*

Word formation will based on completing a text by making an appropriate word form from a word stem given, i.e. discover – discovery (examples available in Appendix p.104). This part will focus mainly on vocabulary, especially on word formation rules. Here the knowledge of suffixes and prefixes will be essential for the students. For example:

*Who is mad? Cows or farmers?*

*Bovine Spongiform Encephalitis is a (1)\_\_\_\_\_\_\_\_\_\_\_ brain DEAD*

*Disorder found amongst cows. As this medical term is almost*

*(2) \_\_\_\_\_\_\_\_\_for the majority of ordinary people to say, the illness POSSIBLE*

*is (3)\_\_\_\_\_\_\_\_known as Mad Cow Disease. POPULAR*

*Prodromou, First Certificate Star, 1998*

Concerning TOEFL test, we might say that it is similar to CFC use of English; however, it displays just several types of tasks. As we have already mentioned they are error correction and multiple choice cloze. Multiple choice cloze typically consists of a range of statements in which there will be a certain grammar structure missing. It is usually based on grammar, than on vocabulary (examples available in Appendix p. 385 – 386). The students will have to know how the subject and predicate go together, how the words and sentence parts are linked with each other. For example:

1. *--------infinitely large number of undiscovered galaxies.*
	1. *An*
	2. *There are an*
	3. *From an*
	4. *Since there are*

*Gear, Cambridge preparation for the TOEFL test, 1996*

Error correction will differ from that in CFC, for in TOEFL test we will have a statement with the underlined words that are supposed to be wrong. The students will have to choose the correct variant (examples available in Appendix p. 387 – 390). It will usually be based on the students’ knowledge of grammar items and word formation as well. For example:

*Drying food by means of solar energy is ancient process applied wherever food and climate conditions*

 *A B C D*

*make it possible.*

*Gear, Cambridge preparation for the TOEFL test, 1996*

In conclusion we can state that Use of English is both discreet and integrative type of testing, for in some tasks of CFC the knowledge of word formation is demanded, but in some grammar will be included either.

The Use of English of CFC and TOEFL will be a direct testing, for it will test the students’ grammar and vocabulary knowledge.

***Speaking***

Speaking is another part of the test that is present in CFC and is not included into TOEFL test. It could be explained by the fact that if the student passes TOEFL test successfully, s/he will be interviewed directly at the place s/he needed the test for.

Therefore, will briefly look at CFC speaking part and discuss it. It aims at the students’ ability to use spoken language effectively in different types of interaction. The students could be asked to give personal information, talk about pictures and photographs, be involved in pair work task or even in discussion.

In personal information part the students could be asked to supply the personal details about themselves: i.e. their job, family position, studies, etc.( examples available in Appendix 10 – 11).

In describing pictures or photographs they will have to share their opinion about them speaking with an examiner. There will be a time limit set for the talk.

In pair work task and discussion the students will be supplied either by pictures or photos or by charts and diagrams. They will be joined in pairs and will have to carry out the task together. It could be either the solving the problem, planning something, putting something in order or discussing a certain topic. Discussion will certainly require the students’ personal opinion and analysis of a topic (examples available in Appendix 63).

In CFC the students will have to cooperate with another interlocutor: either the examiner or another participant.

The author of the paper assume that this part is both integrative and indirect testing. It is integrative, for it will involve the students’ knowledge of the whole aspects of the language: grammar, sentence structure, vocabulary, listening skills and may be even reading skills if the task will be written. To communicate successfully the students will require listening and comprehending the other speaker’s message to respond. Grammar should be accurate to produce a good and correct dialogue or a monologue, for accuracy is an important factor there. The rich word stock will be inevitable element as well.

Indirect testing means that the whole material will be included while testing speaking skills.

To conclude we can declare that CFC and TOEFL tests are both integrative and discreet point tests. They are also direct, however, speaking part in CFC could be defined as indirect one involving all four skills to be used.

**Conclusions**

The present research attempted to investigate the essence of two types of tests, such as TOEFL and CFC tests. The research has achieved the initially set goals and objectives. It dealt with the basic data about testing, where the author had displayed the ideas what was the essence of tests, why the students should be tested, what consequences could tests produce and whom they would mostly influence. Afterwards, the reasons for testing were discussed, where the author of the paper had gradually showed why tests were significant in the process of learning and the role of testing in the teaching process. After the basic data had been discussed, the author came directly to types of testing. At that point the author of the research made an attempt to review various sources on the topic she was able to find. She had presented the definitions of the types of tests offered in Longman dictionary of LTAL and then had compared them with the definitions given by various authors. Later, the author of the research displayed the ways of their applications and reasons for that. She had also presented several examples of tests types in the Appendix. The author of the paper had also discussed ways of testing, such as discrete point test and integrative test, objective and subjective tests, direct and indirect tests, etc. The attention was drawn to the significance of their usage and the purpose for it. Furthermore, the discussion had changed the focus on another important issue, such as tests formats and approaches for testing four language skills. Here the author had broadly and explicitly discussed and analysed the tests formats, such as MCQs, false/true items, cloze tests, gap-filling tests, etc. She had focused on their application and skills for which they are used. Moreover, she had displayed various examples to exemplify each test format, offering several of them in Appendix of the paper. Likewise, a table with the language skills and test formats applicable for them was attached to the work as well. Further, a practical part in the form of the tests’ analysis was presented.

The author of the paper had also dealt with the main issues that are very vital and essential in analysis of the tests. She had focused on the reliability and validity of the tests and tried to trace them in TOEFL and CFC tests. She had thoroughly discussed the tasks and activities composing the tests designed to test the students’ language skills. Moreover, she had attempted to compare the two tests and find out any similarities and differences between them. She had methodically studied each part of the tests, starting from reading skills finishing with speaking. She had presented a detailed investigation into the matter together with the examples that could be observed in Appendix, as well.

Eventually, she had gained her aim having checked the theory into practice and had proved that it really functioned in the real world. Moreover, she had revealed that though being sometimes different in their purpose, design and structure, the TOEFL test and CFC test are constructed according to the universally accepted pattern.

Thus, the hypothesis of the present research has been confirmed.

**Theses**

1. The role of tests is very useful and important, especially in language learning, for they indicates how much the learners have learnt during a course, as well as display the strength and weaknesses of the teaching process and help the teacher improve it.
2. The tests can facilitate the students’ acquisition process and function as a tool to increase their motivation; however, too much of testing could be disastrous changing entirely the students’ attitude towards learning the language, especially if the results are usually dissatisfying.
3. Assessment and evaluation are important aspects for the teacher and the students and should be correlated in order to make evaluation and assessment *“go hand in hand”*.
4. The test should be valid and reliable. They should test what was taught, taking the learner’s individual pace into account. Moreover, the instructions of the test should be unambiguous.
5. Validity deals with what is tested and degree to which a test measures what is supposed to measure.
6. Reliability shows that the test’s results will be similar and will not change if one and the same test will be given on various days.
7. There are four traditional categories or types of tests: proficiency tests measuring how much of a language a person knows or has learnt; achievement tests measuring a language someone has learned during a specific course, study or program; diagnostic tests displaying the knowledge of the students or lack of it, and placement tests placing the students at an appropriate level in a programme or a course.
8. There are two important aspect direct and indirect testing. Direct testing means the involvement of a skill that is supposed to be tested, whereas indirect testing tests the usage of the language in real-life situation and is assumed to be more effective.
9. Discrete point test is a language test that is meant to test a particular language item, whereas the integrative test intends to check several language skills and language components together or simultaneously.
10. There are various tests formats, such as multiple-choice tasks, gap-filling tests, cloze tests, true/false statements, etc. used to check four language skills.
11. To enter any foreign university the students are supposed to take the TOEFL or CFC tests. Besides, they can be taken to reveal the student’s level of the English language.
12. Serving for almost similar purpose, however being sometimes different in their design and structure, the TOEFL and CFC tests are usually constructed according to the accepted universal pattern.

#### Bibliography

1. Bynom, A. 2001. *Testing terms*.English Teaching *professional.* Forum. July. Issue Twenty
2. Gear, 1996. *Cambridge Preparation for the TOEFL Test.* Cambridge University Press.
3. Grellet. 1981. *Developing Reading skills*. Cambridge University Press
4. Heaton, J. 1990. *Classroom Testing.* Longman
5. Hedge. T. 2000. *Teaching and Learning in the Language Classroom*. Oxford University Press
6. Hughes, A. 1989. *Testing for Language Teachers.* Cambridge University Press
7. Hicks, D. Littlejohn, A. 1998. *Cambridge English for Schools* (CES). Teacher’s Book. Level Two. Cambridge University Press.
8. Hicks, D. Littlejohn, A. 1997. *Cambridge English for Schools* (CES). Student’s Book. Level Two. Cambridge University Press.
9. Kruse, A. 1987. *Vocabulary in Context in Vocabulary Learning in Long*. M. (ed.), Methodology in TESOL. New York. Newbury House Publishers.
10. Kramiņa, I. 2002. *Lingua-Didactic Theories Underlying Multi-purpose Language Acquisition*. LU
11. Prodromou, L. 1998. *First Certificate Star,* McMillean
12. Richards, J. 1992. *Language Teaching and Applied Linguistics. Longman Dictionary*. Longman
13. Thompson, M. 2001. *Putting students to the test*. Issue Twenty. Forum. July
14. Wallace, K. 1992. *Reading*. Oxford University Press
15. Weir, C.1990. *Communicative Language Testing*. Prentice Hall
16. Underhill, N. 1987. *Testing Spoken Language*. Cambridge University Press

17. Forum for Teachers

file://A:\Forum\Vol

1. www.ets.org.
2. www.ets.org./TOEFL/
3. www.ielts.org.
4. www.cambridge-efl.org.
5. www.britishcouncil.org.

**Appendix**