



**T. C.  
PAMUKKALE ÜNİVERSİTESİ  
EĞİTİM BİLİMLERİ ENSTİTÜSÜ  
YABANCI DİLLER EĞİTİMİ ANABİLİM DALI  
İNGİLİZ DİLİ EĞİTİMİ BİLİM DALI  
YÜKSEK LİSANS TEZİ**

**PERCEPTIONS AND ATTITUDES OF EFL LEARNERS  
AND THEIR INSTRUCTORS TOWARDS  
ONLINE SPEAKING ASSESSMENT**

**Gizem KORKMAZ**

**DENİZLİ-2019**

**TR  
PAMUKKALE UNIVERSITY  
INSTITUTE OF EDUCATIONAL SCIENCES  
DEPARTMENT OF FOREIGN LANGUAGES EDUCATION  
ENGLISH LANGUAGE TEACHING PROGRAM  
MASTER'S THESIS**

**PERCEPTIONS AND ATTITUDES OF EFL LEARNERS AND THEIR  
INSTRUCTORS TOWARDS  
ONLINE SPEAKING ASSESSMENT**

**Gizem KORKMAZ**

**Supervisor**

**Prof. Dr. Turan PAKER**

## YÜKSEK LİSANS TEZİ ONAY FORMU

Bu çalışma, Yabancı Diller Eğitimi Anabilim Dalı, İngiliz Dili Bilim Dalı'nda jürimiz tarafından Yüksek Lisans Tezi olarak kabul edilmiştir.

**Başkan:** Prof. Dr. Şevki KÖMÜR

imza



**Üye:** Prof. Dr. Turan PAKER (Danışman)



**Üye:** Dr. Öğretim Üyesi Eda DURUK



Pamukkale Üniversitesi Eğitim Bilimleri Enstitüsü Yönetim Kurulu'nun **19/07/2019** tarih ve **29/15** sayılı kararı ile onaylanmıştır.



**Prof. Dr. Mustafa BULUŞ**

**Enstitü Müdür**

## ETİK BEYANNAMESİ

Pamukkale Üniversitesi Eğitim Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada;

Tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,

- Görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- Başkalarının eserlerinden yararlanması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- Atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- Kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- Bu tezin herhangi bir bölümünü bu üniversitede veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı beyan ederim.

  
Gizem KORKMAZ

## ACKNOWLEDGEMENTS

First of all, I owe my deepest gratitude to my thesis supervisor, Prof. Dr. Turan PAKER whose profound knowledge and experience lighted my way throughout this study. His continuous support, encouragement, and invaluable assistance as well as his expertise in the field have made this thesis possible. I am also thankful to the other jury members Prof. Dr. Şevki K m r and Assistant Professor Dr. Eda Duruk for their helpful suggestions and comments.

I would particularly like to thank Prof. Dr. Demet YAYLI, Assoc. Prof. Dr. Recep Şahin ARSLAN, and Asst. Prof. Dr. Selami OK for sharing their extensive knowledge with us during the MA lessons.

In addition, I also owe special thanks to all students and my colleagues taking part in this study, they helped me a lot to collect the data. My heartfelt appreciation goes to my beloved friend Elif GENÇ for her great help and encouragement. I am also grateful to all my friends for their wholehearted support for making this journey unforgettable.

Last but not least, I would like to show my greatest appreciation to my mother G lsen YEŐIL for her endless love, support, and patience, and to my father Ali YEŐIL for his encouragement and for my beloved brother Gencay YEŐIL for his support. I am also deeply grateful to my precious husband  zg r KORKMAZ and son Tuna KORKMAZ for being always there for me.



**To my family,  
and my dearest son, Tuna**

## ÖZET

### **YDİ Öğrenenlerinin ve Okutmanlarının Çevrimiçi Konuşma Sınavlarına Karşı Algı ve Tutumları**

KORKMAZ, Gizem

Yüksek Lisans Tezi, Eğitim Bilimleri ABD,  
Yabancı Diller Eğitimi Ana Bilim Dalı İngilizce Öğretmenliği

Tez Dnişmanı: Prof. Dr. Turan PAKER

Haziran 2019, 95 sayfa

Bu çalışmada öğrencilere gerçek konuşma ortamlarında iletişim kurma ve dili etkili bir şekilde kullanma fırsatı yaratan yeni bir konuşma değerlendirme sistemi bütün taraflarıyla incelenmiş ve kullanıcıların tutum ve algıları derinlemesine tartışılmıştır. Buna ek olarak çalışmada bu sistemin öğrencilerin konuşma ve dinleme becerilerine etkisinin de keşfedilmesi amaçlanmıştır. Aynı zamanda çevrimiçi konuşma sınavı sisteminin öğrenciler üzerindeki etkilerini araştırmak için bu sistemi uygulayan sınav gözetmenleri, okutmanlar ve idari kadronun da fikirleri analiz edilmiştir.

Katılımcılar İzmir’de bir vakıf üniversitesinde eğitim gören B2 seviyesindeki hazırlık sınıfı öğrencilerinden oluşmuştur. Bunun yanında, veri toplamada iki ana yöntem kullanılmıştır: anket ve görüşme. Görüşme verileri daha ekonomik olması açısından öğrencilerle yarı yapılandırılmış yani küçük gruplar halinde toplanırken, akademisyenlerle ikili görüşmeler şeklinde düzenlenmiştir.

Elde edilen bulgulara göre katılımcıların birçoğu çevrimiçi sınav esnasında kendilerini özgüvenli, rahat ve pozitif hissetmiştir. Ayrıca sonuçlar öğrencilerin konuşma sınavına karşı tutum ve davranışları ve sistemin konuşma ve dinleme becerilerine katkısı açısından değerlendirildiğinde katılımcıların çoğu sistemin pratik ve tamamlanması kolay olarak değerlendirdiği tespit edilmiştir. Aynı zamanda, katılımcıların yarısından fazlası sistemi konuşmadaki genel ve detaylı mesajı anlayabilme yeteneklerini geliştirmesi açısından faydalı bulmuştur.

Sonuçlar akademisyenlerin, idari kadronun ve sınavı uygulayanların da bakış açılarıyla değerlendirildiğinde sistemin öğrenciler ve onların geliřtirmesi hedeflenen konuşma ve dinleme becerileri üzerine hem olumlu hem de olumsuz şekillerde etki ettiđi tespit edilmiştir. Hepsi, çevrimiçi konuşma sınavının olumlu sonuçları üzerinde hemfikir olmuş ama diđer taraftan geliřtirilmesi gereken yönlerinin de bulunduđunu vurgulamışlardır.

**Anahtar Kelimeler:** Çevrimiçi Konuşma, Konuşma Sınavları, Alternatif Deđerlendirme,

Hazırlık Okulu, Dinleme Yetisi, Konuşma Yetisi





## **ABSTRACT**

### **Perceptions and Attitudes of EFL Learners and Their Instructors towards Online Speaking Assessment**

KORKMAZ, Gizem

Master's Degree Dissertation in Educational Sciences,  
Department of Foreign Language Education, ELT Department

Supervisor: Prof. Dr. Turan PAKER

June 2019, 95 pages

In this study, a newly adopted system of online speaking assessment which enables learners to communicate and use the language in a real atmosphere were thoroughly analyzed and discussed by looking at the test takers' perceptions and attitudes. In addition, this study aimed to provide information about test takers' perceptions and attitudes to the online speaking assessment. Another ultimate goal of this study was to discover the effects of this system on test takers' speaking and listening competencies. The study also examined the opinions of test makers', school administrators' and instructors' about the effects of online speaking assessment system on students.

The participants in this study were B2 level preparatory school students in a foundation university in İzmir. In order to collect data, two main instruments were used; a questionnaire and an interview. A questionnaire was administered to the all B2 level students who participated in the speaking system process from the beginning of the year. In addition, interview data were obtained through semi-structured interviews conducted with both students and instructors to get more detailed information.

The results showed that the participants in this study mostly felt themselves confident, relaxed and positive during the assessment. When the results were examined in terms of students' attitudes towards the system and its effect on their speaking and listening competencies, it could be seen that many of the participants thought that it was practical and easy to complete it. In addition, as for the students' ideas about the effect of this kind of assessment on their speaking and listening competencies, On the other hand, more than half of the test takers found the system helpful for them to understand general messages or details during conversation.

When the results were examined from the eyes of the examiners and the instructors, it could be seen that the instructors and the examiners had both negative and positive thoughts about the effects of this system on the test takers and on the competencies to be developed. All of the participants were actually sure about the system and its good results. However, there were also some issues to be improved related to the system such as the topics chosen for the assessment, the testing competencies of the test makers and time allocated for the assessment.

**Key words:** Online Speaking, Speaking Assessment, Alternative Assessments, Washback Effect, Preparatory School, Listening Competency, Speaking Competency



## TABLE OF CONTENTS

YÜKSEK LİSANS TEZİ ONAY FORMU.....	iii
ETİK BEYANNAMESİ .....	<b>Hata! Yer işareti tanımlanmamış.</b>
ACKNOWLEDGEMENTS.....	v
DEDICATION.....	vi
ÖZET .....	vii
ABSTRACT.....	ix
TABLE OF CONTENTS.....	xi
LIST OF TABLES .....	xiv
TABLE OF FIGURES .....	xivi
LIST OF ABBREVIATIONS.....	xivii
CHAPTER 1 : INTRODUCTION.....	1
1.1. Background to the Study.....	1
1.2. Statement of the Problem.....	2
1.3. Research Questions.....	4
1.4. Significance of the Study .....	4
1.5. Assumptions and Limitations of the Study.....	5
CHAPTER 2 : LITERATURE REVIEW.....	6
2.1. Introduction .....	6
2.2. Speaking as a skill .....	6
2.2.1. Aims .....	7
2.2.2. Resources.....	8
2.2.3. Needs.....	9
2.3. Constructivist Approach in Language Teaching.....	9
2.4. Communicative Language Teaching and Testing.....	10
2.5. Communicative Competence .....	10
2.6. Task-Based Language Instruction and Testing .....	13
2.7. Components of Speaking.....	13
2.7.1. Grammar.....	14
2.7.2. Vocabulary .....	15
2.7.3. Pronunciation .....	15
2.7.4. Fluency .....	15

2.8. Testing Speaking .....	16
2.8.1. Reliability .....	16
2.8.2. Validity .....	16
2.8.3. Washback Effect.....	18
2.8.4. Practicality.....	19
2.8.5. Test Administration.....	20
2.9. Speaking Scales .....	21
2.9.1. Scoring and Marking Systems.....	21
2.9.2. Holistic Scoring.....	21
2.9.3. Analytic Scoring.....	22
2.9.4. CEFR.....	23
2.10. Tasks in Speaking Tests .....	24
2.10.1. Advantages of Tasks in Speaking Instruction .....	25
2.10.2. Designing Tasks .....	26
2.11. Alternatives in Language Assessment.....	27
2.11.1. Computer Based Testing .....	28
2.11.2. Computer Adaptive Testing: CAT.....	28
2.11.3. Internet Based Testing.....	29
2.12. The Significance of Feedback in Speaking Assessments .....	30
<b>CHAPTER 3 : METHODOLOGY .....</b>	<b>32</b>
3.1. Introduction.....	32
3.2. Nature of the Study .....	32
3.3. Methodology of the Study .....	33
3.3.1. Setting .....	33
3.3.2. Participants.....	34
3.3.3. Data Collection Instruments and Procedures .....	36
3.3.3.1. Questionnaire .....	37
3.3.3.2. Interview with Students.....	39
3.3.3.3. Interview with Instructors, Test makers and Administrators.....	39
3.3.4. Data Analysis .....	39
<b>CHAPTER 4 : RESULTS.....</b>	<b>41</b>
4.1. Introduction.....	41
4.2. What Are The Test Takers' Perceptions And Attitudes Towards Online Speaking Assessment System? .....	43

4.2.1 Feelings.....	44
4.2.2. System.....	50
4.2.3. Skill.....	56
4.3. What are the test takers’ perceptions and attitudes towards the effect of this system on their speaking competency?.....	59
4.4. What are the test takers’ perceptions and attitudes towards the effect of this system on their listening competency? .....	61
4.5. What are the Perceptions and Attitudes of the Examiners, the School Administrators and the Teachers towards Online Speaking Assessments?.....	63
CHAPTER 5 : DISCUSSION, CONCLUSION AND SUGGESTIONS.....	67
5.1. Discussion.....	67
5.1.1. What Are The Test Takers’ Perceptions And Attitudes Towards Online Speaking Assessment System? .....	67
5.1.2. What are the test takers’ perceptions and attitudes towards the effect of this system on their speaking and listening competencies?.....	70
5.1.3. What are the Perceptions and Attitudes of the Examiners, the School Administrators and the Teachers towards Online Speaking Assessments?.....	71
5.2. Conclusion .....	72
5.3. Implications of the Study.....	73
5.4. Limitations of the Study .....	76
5.5. Suggestions .....	76
REFERENCES .....	78
APPENDICES .....	87
APPENDIX A: Survey .....	88
APPENDIX B: Test Makers’, Instructors’ and Administrators’ Interview Questions.....	93
APPENDIX C: Test-takers’ Interview Questions.....	94
APPENDIX D: CV.....	95

## LIST OF TABLES

Table 2.1. <i>Interview Assessment Scale</i> .....	22
Table 2.2. <i>Analytic Rating Scale</i> .....	23
Table 2.3. <i>Qualitative aspects of spoken language use</i> .....	24
Table 3.1. <i>The distribution of participants according to their gender and age</i> .....	35
Table 3.2. <i>The item numbers for each part</i> .....	37
Table 3.3. <i>Categories and Items in the Questionnaire</i> .....	38
Table 3.4. <i>The Sample Responses of the Test-takers</i> .....	40
Table 4.1. <i>Computer Familiarity</i> .....	42
Table 4.2. <i>Computer Familiarity (gender differences)</i> .....	42
Table 4.3. <i>Prior Experiences of the Test-takers</i> .....	43
Table 4.4. <i>System Use (gender differences)</i> .....	43
Table 4.5. <i>The Easiness of the System (gender differences)</i> .....	43
Table 4.6. <i>The Item Numbers for Each Part</i> .....	44
Table 4.7. <i>Test-takers' Feelings about Online Speaking Assessment</i> .....	44
Table 4.8. <i>The Relationship between Chatting Online and Gaining Self-confidence</i> .....	45
Table 4.9. <i>The Relationship between Chatting Online and Feeling Nervous</i> .....	46
Table 4.10. <i>The Relationship between the Easiness of Taking the exam and Feeling Nervous</i> .....	47
Table 4.11. <i>Test-takers' Opinions about Feeling Nervous (gender differences)</i> .....	47
Table 4.12. <i>Test-takers' Opinions about Feeling Nervous (prior experiences)</i> .....	47
Table 4.13. <i>Test-takers' Opinions about Online Speaking Assessment (motivation)</i> .....	48
Table 4.14. <i>The Relationship between the Comfortable Use of the System and Recommending to Others</i> .....	48
Table 4.15. <i>The Relationship between Computer Ability and Recommending to Others</i> .....	49
Table 4.16. <i>Test-takers' Opinions about Participating Future Online Assessments (prior experience)</i> .....	49
Table 4.17. <i>Test-takers' Opinions about the System of Online Speaking Assessment</i> .....	50
Table 4.18. <i>The Relationship between Following the Instructions and Being Contemporary</i> .....	50

Table 4.19. <i>The Relationship between the Easiness of the System and Being Contemporary</i> .....	51
Table 4.20. <i>The Relationship between the Easiness of the System and Being More Practical than Face-to-face Exams</i> .....	52
Table 4.21. <i>The Relationship between the Comfortable Use of the System and Being More Practical than Face-to-face Exams</i> .....	52
Table 4.22. <i>The Relationship between Reading a Computer Screen and Being More Practical than Face-to-face Exams</i> .....	53
Table 4.23. <i>Test-takers' Perceptions about the Effects of Online Speaking Assessment to Their Skills</i> .....	56
Table 4.24. <i>Test-takers Opinions about Speaking in Class (prior experiences)</i> .....	57
Table 4.25. <i>Relationship between Attending Online Course Before and Improving Speaking Skills</i> .....	58
Table 4.26. <i>The Relationship between Understanding Computer and Improving Speaking Skills</i> .....	59
Table 4.27. <i>Test-takers Opinions about the Effects of Online Speaking Assessment to Their Speaking Competency</i> .....	59
Table 4.28. <i>Test-takers Opinions about the Effects of Online Speaking Assessment to Their Listening Competency</i> .....	61

## LIST OF FIGURES

<i>Figure 2.1.</i> Canale & Swain's (1980) Components of Communicative Competence.....	12
<i>Figure 2.2.</i> Components of Oral Language.....	14
<i>Figure 2.3.</i> A Model of Feedback to Enhance Learning.....	30





## LIST OF ABBREVIATIONS

<b>CALL</b>	Computer Assisted Language Learning
<b>CEFR</b>	Common European Framework of Reference
<b>EFL</b>	English as a Foreign Language
<b>ELT</b>	English Language Teaching
<b>LGS</b>	Liselere Geçiř Sınavı
<b>L2</b>	Second Language
<b>TEOG</b>	Temel Eđitimden Ortaöđretime Geçiř
<b>TL</b>	Target Language
<b>TOEFL</b>	Test of English as a Foreign Language
<b>YKS</b>	Yüksek Öđretim Kurumları Sınavı

## CHAPTER I

### INTRODUCTION

This chapter consists of five sections, which provide information about background to the study related to speaking using the target language and an overview of teaching and assessing speaking English language in Turkey, statement of the problem, research questions, significance of the study, and assumptions and limitations for the study.

#### 1.1. Background to the Study

The ability to speak the target language fluently can motivate and affect the learner positively as it enables the speaker to communicate, express feelings. Although learners are sometimes proficient in using the structures properly, they may not be sufficient enough to use the language communicatively. The reason for this is the lack of real communication environments that help the learners to use the language in a real English spoken atmosphere. As Underhill (1987) stated, in a real oral tests the most important priorities are different. Unlike the written tests, real people communicate face to face without the existence of the idea of 'test'. Instead of providing fake environments while teaching and testing speaking, newly adopted methods such as computer mediated language testing can be made use of. With the help of computers, learners can communicate with real people on real topics without being aware of the fact that they are actually being tested and evaluated.

For students, oral exams are the most stressful part of the exams that cause anxiety. Traditional testing methods for speaking that requires the students sit at the table and talk about on a randomly chosen topic that they may never encounter again throughout their lives, which makes the testing stressful for them. The limitations such as lack of confidence, concentration, limited period of time, and the necessity to comment on the topic make the testing more and more challenging for them.

As Madsen (1983) points out "The testing of speaking has always been seen as the most challenging skill as it is regarded hard to administer and score" (p.147). One reason why testing speaking is seen challenging is that it requires fair, valid and reliable tools and assessment. There are numerous elements (content, organization, appropriate vocabulary, fluency and accuracy; grammar, pronunciation, etc.) that should be taken into account in testing speaking, and it can be hard to identify all of them at once as it is conducted in a very

short and limited period of time. From time to time, even preparing the materials for testing speaking can be a problematic area. As Underhill (1987) stated because of the difficulty of assessing speaking tests in the same way as other traditional tests, little space and time is allocated to oral testing. For this reason, this skill can be sometimes neglected or not given enough attention although it is one of the most important skills to be tested.

The advantages of computers for testing are apparent and well known. Douglas (2014) pointed out that the point is to ask whether to use technologies for language testing, instead the effects of the use of different technologies in language testing and its effect on the test takers' performance and attitudes should be deeply analyzed.

With these things in mind, a newly adopted system of online speaking assessment which enables learners to communicate and use the language in a real atmosphere will be thoroughly analyzed and discussed by looking at the test takers' perceptions and attitudes. In addition, the effect of this system on the test takers' listening and speaking competencies will also be examined. In order to make the research applicable in a broader way, test makers', administrators' and instructors' opinions about the effectiveness of the system will also be analyzed.

## **1.2. Statement of the Problem**

Examinations play a crucial role in Turkish educational system. Students are supposed to take numerous examinations such as TEOG, YKS, and LGS throughout their school lives. One of the main problems is language testing, and assessment in Turkish educational system is not to prepare learners to be able to use the language communicatively. Learners do not know how to communicate with foreigners, or do not know how to express their feelings in the target language. They feel anxious or nervous while using the language even in the class environment. As little attention is given to the area of teaching speaking, testing speaking becomes a challenging area for test developers. Traditionally, one to one, face to face exam types are tried but provides no advantage to the test takers in their real lives.

In order to arrive at the most appropriate testing system, it is necessary to understand the principles of testing and discover new ways to test the knowledge. The idea of using technology has always been with us as it is developing really fast. However, it should be taken into account in language testing that different technologies might be appropriate to test

the test takers' language skills. With the help of technology, test takers could be given opportunity to widen and improve their productivity in using the language that they are struggling to learn.

As stated before, testing speaking is regarded as the hardest skill to prepare, administer and score. During the test, the learners feel less confident and more anxious when they are assessed in a strict and less communicative environment. Therefore, the test design must be in the way that enables learners to use the language communicatively as much as possible. Fulcher (2000) points out that “communicative tests would involve performance (speaking), and the performance would be judged subjectively, qualitatively and impressionistically, by a sympathetic interlocutor/assessor” (p.484).

One way that we can make the speaking assessment environments real and comfortable is to create relaxed atmosphere where the test takers are given the chance to speak on real life tasks. In their article Canale and Swain (1980) stress that teaching and assessment instruments should be designed to develop both communicative competence and communicative performance, so the performance that is to be shown should be in a *real* environment for *authentic* purposes. According to Lado (1961), “speaking ability is described as the ability to express oneself in life situations, or the ability to report acts or situations in precise words, or the ability to converse, or to express a sequence of ideas fluently” (p. 240). Therefore, the idea of using technology for delivering and organizing the tests is getting more and more popular nowadays as it provides more authentic way of assessing the speaking ability. However, it has got both advantages and disadvantages. Spolsky and Hult (2008) in their book *Chapelle*, point out that reading texts or watching videos online are some of the facilities that Computer Assisted Language Learning offers to learners. Communication happens whenever and wherever the learners want, and this communication helps them acquire the knowledge and response quickly, and these types of evaluations give them the chance of getting quick feedback which is extremely important in a learning process. Other than these advantages, there are also some deficits of using technology in language testing such as misinterpretations or misunderstandings of the words that cause extra burden and stress on test takers.

When Lee (2007) analyzed students' reflections, the outcomes of the study showed differences between outgoing and shy students in terms of their perceptions. While the outgoing ones believed that it worked a lot to practice their speaking skills with a native

speaker it was useful to get feedback, the other group who felt shy found it very stressful to speak someone stranger.

This study aims to provide information about test takers' perceptions and attitudes to the online speaking assessment. Another ultimate goal of this study is to discover the effects of this system on test takers' speaking and listening competencies. The study also examines what test makers', school administrators' and instructors' thoughts are about the effects of online speaking assessment system on students.

### **1.3. Research Questions**

This study attempts to find answers to the following questions:

- 1) What are the test takers' perceptions and attitudes towards online speaking assessment system?
  - a) What are the test takers' perceptions and attitudes towards the effect of this system on their speaking competency?
  - b) What are the test takers' perceptions and attitudes towards the effect of this system on their listening competency?
- 2) What do the test makers, instructors and the school administrators think about the system and its effect on the test takers' development?

### **1.4. Significance of the Study**

As Yule and Brown (1988) stated, testing English has always been a headache for the teachers. Even in well-established tests, the focus is on grammar or vocabulary, and the speaking parts are mostly neglected. It is suggested that instead of testing grammar and vocabulary in isolation, they should be evaluated as a part of spoken language. In one research, Wall and Alderson (1993) found out that the content of the teaching is affected by the exam, thus teachers are more interested in covering the parts of the textbook that they feel are most likely to be tested. This means, listening and speaking do not receive the attention they should deserve, because of the attention teachers feel they must pay to reading. That is, test takers need to be able to integrate all the language skills and communicate effectively.

Although more traditional ways of assessing speaking have been used for many years, it is now a common fact that technology is an inevitable part of modern language testing since it provides more authentic, efficient and effective test delivery. Related to this belief, Douglas (2014) stated that “the use of technology in language assessment really is not an issue we can reasonably reject – technology is being used and will continue to be used” (p.139)

This study may provide general information about the perceptions and attitudes of the test takers towards an alternative testing system- Online Speaking Assessment. Their previous experience, their opinions about the system use will help the study to gain an insight about the effectiveness of the system on their speaking and listening ability. The study will also contribute to the field as it is going to give us information about the importance of alternative speaking assessment system. In the light of this study, the new trends in language testing can be taken into account instead of the traditional ones. It will also show us the effectiveness of the online assessment from the point of administrators, instructors and test makers.

### **1.5. Assumptions and Limitations of the Study**

Speaking is seen as one of the most challenging skills among the students. The reason for this may be the inadequate hours spent for speaking and listening skills during the class time. Or, it is just because of the prejudices learnt by the learners throughout their language learning process. In this study, we assume that because the test takers will experience an alternative way of assessment instead of being evaluated face to face, they may foster their abilities in terms of speaking and listening. Another assumption is that the test takers may feel themselves inexperienced and incompetent while communicating a foreigner teacher although they have practiced it for almost a year, so this situation may affect their performances and perceptions in a negative way.

The limitation of the study is that the whole process is based on the technology. Therefore, even a tiny breakdown caused by the system may affect the whole process negatively. The malfunction of the interface of the system or even a speaker may cause big problems.

## **CHAPTER II**

### **LITERATURE REVIEW**

In this chapter, theoretical framework, definitions and importance of language tests, communicative language teaching and testing, task based language instruction and testing, components of speaking, speaking scales, tasks in speaking tests are included. In addition, advantages of tasks in speaking instruction, alternatives in language assessment, computer and internet based testing are reviewed through the analysis of relevant studies. Finally, the significance of feedback in speaking assessments is discussed.

#### **2.1. Introduction**

Speaking can be seen as the most important skill when compared to other skills. (Listening, reading, and writing). As the process of learning and applying the skills of oral English are so closely related, classroom is a place where the use of spoken language is supported. In a classroom environment students are supposed to make connections between what they know and what they are learning.

In practice, many of the learners find it difficult to communicate or interact using the target language. They think that communicating in a foreign language perfectly is something, a very tough task to complete or achieve. This study aims to put a light on what learners think about online and alternative ways of assessing speaking. It covers some issues or points needed to be taken into account while assessing the learners' speaking and listening competencies.

#### **2.2. Speaking as a Skill**

According to Bygate (1998: 23), speaking is a skill which includes certain strategies such as making decisions and interaction during the conversation. This is thought a top-down view of speaking. In the light of this view, Eckard & Kearny (1981), Florez (1999) and Howarth (2001) identify speaking as a process involving a true communication of ideas, information or feelings. This top-down view considers the verbally expressed ideas as the result of participation between at least two interactants in shared time, and a common physical setting. Hence, as a proponent of this view Nunan (1989) suggests that learners

should be encouraged to acquire the language in small units and participate in speaking discourse rather than being taught to make well-formed sentences.

### **2.2.1. Aims**

According to Underhill, ‘Giving a test is like asking a question- it is a request for information. In order to develop a test, you have to know what kind of information you want’ (1987, p.12). Consequently, you get the result of your question. As Underhill stated, ‘tests can be used to ask four basic kinds of questions’ (1987, p.12-13).

#### *1) What is the learner’s general level of language ability?*

This question can be achieved by a proficiency test. As Hughes stated, ‘proficiency tests are designed to measure people’s ability in a language, regardless of any training they may have had in that language’ (1989, p.11). Therefore, proficiency test is not based on the content or objectives of the language courses, rather it is based on what the learners are able to do to be considered proficient. In order to get a quick and quite accurate measure of general proficiency, a well–designed oral test that contain a number of different test techniques is needed.

#### *2) Where does this learner fit in our teaching program?*

The information that is needed here is an accurate placement. According to Underhill (1987), a placement test identifies the right class for learner; there is no bad or good score, there is only a recommendation for the most suitable class. As Hughes stated, ‘placement tests are intended to provide information that will help to place students at the stage of the teaching program most appropriate to their abilities.’ (1989, p. 16)

#### *3) What are the learner’s particular strengths and weaknesses?*

Diagnostic tests identifies the learners’ strengths and weaknesses. Instead of giving information about the form of a score, it shows in what areas the learners are strong or need further practice or do follow-up work. According to Hughes (1989), a diagnostic test enables the learners to see the gaps in their command of the language. Therefore, as Underhill stated (1987), ‘the interactive nature of an oral test allows the interviewer to probe for individual strengths and weaknesses, to ask for repetition or clarification, perhaps to elicit the learner’s own opinion of his ability’ (p.13).



4) *How much has the learner learnt from a particular course?*

Underhill (1987) defines the achievement test as ‘the sample of the language elements or skills that have been covered on the course and aims to test how well the learner has mastered those elements’ (p.13). It aims to see how successful individual students, groups of students, or the courses themselves are in achieving objectives (Hughes, 1989).

### **2.2.2. Resources**

In order to be successful in a testing program, it is inevitable to use available resources. According to Underhill (1987), there are three different types of resources; people, time and physical facilities (p.15).

#### *People*

As Underhill (1987) stated people are the biggest resource in testing and should be used effectively. People are needed to mark the tests, to try out the test in the first place, to check the test and to improve the test if it is necessary. In testing, people are also important to carry out administrative duties, then to present the results as quickly as possible. Bachman and Palmer (2010) described the several roles of human in testing as: test developer, task writer, assessment administrators, individuals who produce assessment records and clerical support.

#### *Time*

Time is a really critical element in language testing. As Bachman and Plamer (2010) suggested it is important to allocate and estimate the correct amount of time needed for personnel and equipments. According to Underhill (1987), the very first stage of the test development start with that question: ‘how much time do you have to develop the test?’ (p. 16). Sharing the work saves time and more people means more time. Second, the operation stage is related to the amount of time to be spent on testing, and the numbers of tests to be conducted, the maximum amount of time for each test or task. Finally, the test improvement stage is the period when the adjustments are done, and the marking systems are made.

### *Equipment and facilities*

First of all, Underhill mentions (1987) the choice of room, testing atmosphere and the arrangement of furniture affects the testing process. Sound and video recording systems are also important. They should be reliable and easy to operate, and give good quality of reproduction. The room should be quiet and free from interruption. A learner taking the test should get the feeling that his test is given a very high priority. Otherwise, he may lose his train of thought or flow of speech and lose confidence. As Hughes (1989) stated, unless a test is not well-administered, it is inevitable for the test to give invalid and unreliable results.

#### **2.2.3. Needs**

According to Underhill (1987), aims are institutional while needs are personal. Ideally, the aims of the program should match the needs of the learner. If there is a mismatch between institutional aims and personal needs, the results may be demotivating for the learner. Therefore, the resources and the materials should be designed and adapted according to the needs of the program and the learner. In order to reach that goal, a needs analysis should be done and then the learners can be classified as clearly distinct group by certain criteria.

### **2.3. Constructivist Approach in Language Teaching**

Constructivists believe that if an individual gives meaning to an experience or activity then knowledge can be constructed (Merriam et al., 2007; Torre, Daley, Sebastian, Elnicki, (2006). While Piaget (1972) is responsible for the basic formation of personal constructivism, the ideas related to social constructivism is mostly related to Vygotsky (1978).

Splitter (2009) describes constructivism as a psychological theory that we learn actively and consciously bring our past experiences. By looking at the ideas of Piaget and Vygotsky, Rutherford (2012) divides into two categories: personal and social constructivism. While in the perspective of personal constructivism, the learner connects meaningful knowledge using previous knowledge and experience, in social constructivism individuals interpret problems in a social environment.

For many educators constructivism has many meaningful descriptions that learners use the previous knowledge when the new one come. Splitter (2009) agrees that students

should be active participants in the learning process where knowledge has been collected by others and transmitted to them.

Gold (2001) describes constructivism as an alternative epistemology of how people learn and assimilate new knowledge. He continues that people are active, thus they generate cognitive structures that are similar to the past experiences and this helps them to transform and interpret experience. According to Gold (2001), in constructivism, learning happens when specific outcomes are produced. In addition, it is the process by which those outcomes are produced. Products need to be interpreted by keeping the contexts they were produced in mind. In addition, in order to provide multimodal and rich assessments, there should be as many different ways of expressing meaning as possible, such as portfolios.

#### **2.4. Communicative Language Teaching and Testing**

As Canale and Swain described (1980), communicative competence is the ability that focuses on the basis of communicative functions such as apologizing, describing or inviting using the proper functions appropriately. On the other hand, grammatical approaches are mostly organized on the basis of grammatical forms such as phonology, morphology or syntax.

Communicative tests, according to Fulcher (2000) should be authentic with a purpose, authenticity and context.

As Canale and Swain (1980) stated, there is a difference between communicative competence and performance. They stated that “communicative testing must be devoted not only to what extent the learners know about the second language but also to what extent the learner is able to actually demonstrate this knowledge in a meaningful communicative situation” (p.34.).

#### **2.5. Communicative Competence**

Competence and performance are the two terms that are mostly used in the discussions of second language approaches. Chomsky (1965) defines the competence and performance using the statements about the methodological necessity of studying language through idealized abstractions and ignoring what appear to be irrelevant. He also claimed that it is the linguistic system that is internalized by an ideal native speaker, but on the other hand performance mainly focuses on the psychological factors that come to existence in the perception and production of speech. That is, communicative competence mostly focuses on

the ability or capability relating to the rules of language use but grammatical competence focuses to the rules of grammar.

As for Canale and Swain (1980) and Canale (1983), communicative competence is a combination of a fundamental system of knowledge and skill needed for communication. In their point of view communicative competence, conscious or unconscious knowledge refers to the knowledge of someone about language and about other aspects of language use. Unlike Hymes (1972), Canale and Swain (1980) or even Widdowson (1983) Savignon (1972) gave much more importance to the aspect of ability in her concept of communicative competence. That is, she described communicative competence as ‘the ability to function in a truly communicative setting – that is, in a dynamic exchange in which linguistic competence must adapt itself to the total informational input, both linguistic and paralinguistic, of one or more interlocutors’ (Savignon, 1972, p.8).

As Hymes (1972) observes:

a normal child acquires knowledge of sentences not only as grammatical, but also as appropriate. He or she acquires competence as to when to speak, when not, and as to what to talk about with whom, when, where, in what manner. In short, a child becomes able to accomplish a repertoire of speech acts, to take part in speech events, and to evaluate their accomplishment by others (p. 277).

According to Canale and Swain (1980), there are four components of communicative competence: linguistic, sociolinguistic, discourse, and strategic competence. (<https://linguisticator.com/communicative-competence/>),

1- Linguistic competence is the knowledge of the language code, i.e. its grammar and vocabulary, and also of the conventions of its written representation (script and orthography). The grammar component includes the knowledge of the sounds and their pronunciation (i.e. phonetics), the rules that govern sound interactions and patterns (i.e. phonology), the formation of words by means of e.g. inflection and derivation (i.e. morphology), the rules that govern the combination of words and phrases to structure sentences (i.e. syntax), and the way that meaning is conveyed through language (i.e. semantics).

2- Sociolinguistic competence is the knowledge of sociocultural rules of use, i.e. knowing how to use and respond to language appropriately. The appropriateness depends on the setting of the communication, the topic, and the relationships among the people communicating. Moreover, being appropriate depends on knowing what the taboos of the

other culture are, what politeness indices are used in each case, what the politically correct term would be for something, how a specific attitude (authority, friendliness, courtesy, irony etc.) is expressed etc. (Canale and Swain, 1980).

3- Discourse competence is the knowledge of how to produce and comprehend oral or written texts in the modes of speaking/writing and listening/reading respectively. It's knowing how to combine language structures into a cohesive and coherent oral or written text of different types. Thus, discourse competence deals with organizing words, phrases and sentences in order to create conversations, speeches, poetry, email messages, newspaper articles etc. (Canale and Swain, 1980).

4- Strategic competence is the ability to recognize and repair communication breakdowns before, during, or after they occur. For instance, the speaker may not know a certain word, thus will plan to either paraphrase, or ask what that word is in the target language. During the conversation, background noise or other factors may hinder communication; thus the speaker must know how to keep the communication channel open. If the communication was unsuccessful due to external factors (such as interruptions), or due to the message being misunderstood, the speaker must know how to restore communication. (Canale and Swain, 1980).

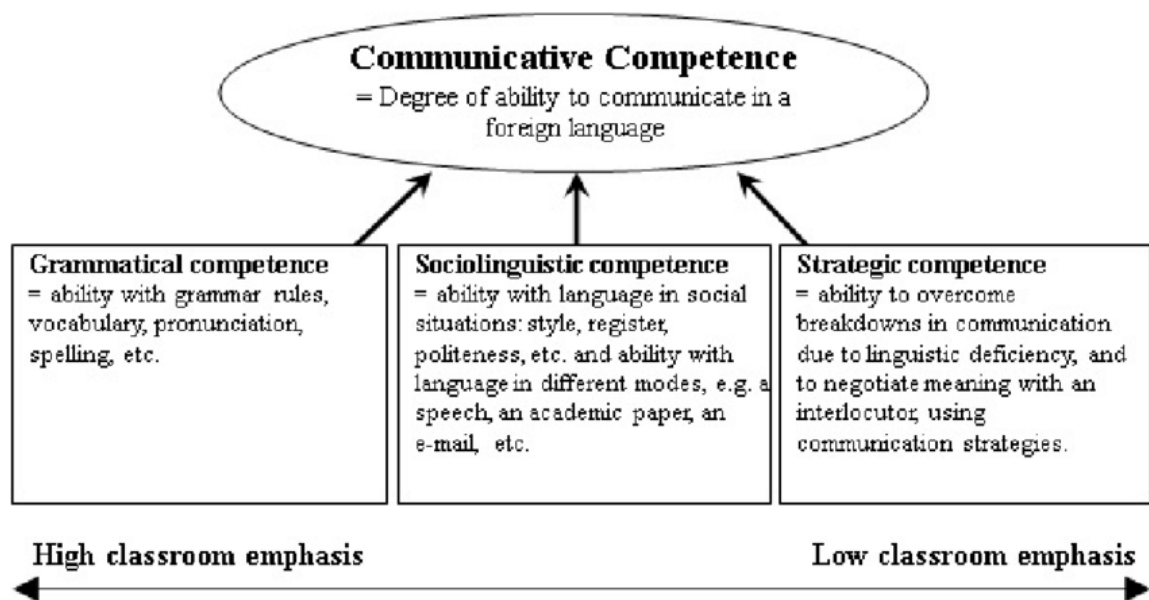


Figure 2.1 Canale & Swain's (1980) components of communicative competence (based on Dörnyei & Thurrell, 1991, p.17).

These strategies may be requests for repetition, clarification, slower speech, or the usage of gestures, taking turns in conversation etc.

## **2.6. Task-Based Language Instruction and Testing**

In most of the EFL concepts there seems to be little or no real communication among the pupils or between teacher and student. Mostly, teacher asks a question and students answer and these unrealistic procedures are sometimes evaluated in terms of speaking competency.

Evidently as Liao (2001) defined, for real communication to occur in the language classroom, teacher-student (and student-student) exchanges must be much more different than just asking and answering questions and should be based on the unknown gap that occurs between speakers when one does not guess or know the next step or what the other is planning to say before.

Much research has proven that achieving the task procedures and goals improve the process of language acquisition and fosters speaking skills. Apparently, tasks engage the learners to take part in different speaking and communicative events. According to Ur (1997), tasks help the learners who are unwilling to talk in front of other people to talk more in the class and thus motivate the learners.

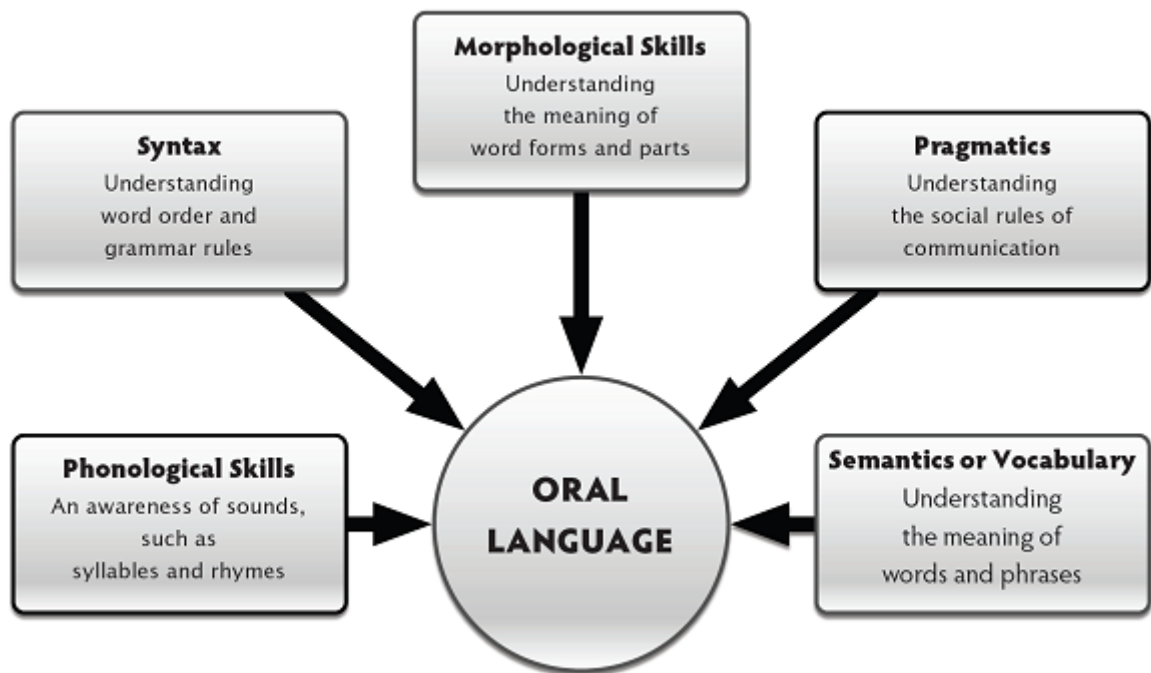
Using tasks is also helpful in detecting the needs of the learning process. As Finch (1997) and Ellis (2003) stated, tasks are helpful for the learners to focus more on the result and outcome, thus determine the needs and help them to self evaluate their own process and communicative competence.

## **2.7. Components of Speaking**

The component is the thing that viewpoint affecting how well individuals communicate in English. "Speaking" is the conveyance of language through the mouth by making sounds with our body. Sometimes it is random or naturally happens but sometimes it is arranged or planned beforehand. It is both formal that happens in scholastic or business circumstances, and it is casual that is utilized with family, companions or individuals you know about.

As proverb suggests, practice makes perfect. Therefore, it is needed for individuals who want to be fluent and accurate in spoken English to practice the language. Accuracy is

the ability to produce grammatically and lexically correct sentences. On the other hand, fluency is the capacity to peruse, talk, or compose effectively, easily, and expressively. In other words, the speaker can read, understand and respond in a language clearly and concisely while relating meaning and context. According to Syakur (1987), speaking is an unpredictable ability since it is not only concerned with grammar and vocabulary but also with pronunciation and fluency.



*Figure 2.2* Components of Oral Language (Lesaux & Harris, 2015)

### 2.7.1. Grammar (Accuracy)

It is inevitable for students to arrange a correct sentence in conversation. It is in line with explanation suggested by Heaton (2003) that student's awareness of the grammatically appropriate and inappropriate one and producing structures properly. The utility of grammar is likewise to get familiar with the right method to pick up ability in a language in oral and written form. Luoma (2004) mentions that students get better from knowing a couple of structures to knowing an ever increasing number of complex ones, and from making errors to making few.

### **2.7.2. Vocabulary**

Vocabulary is the inevitable part of the communication. Without knowing a sufficient amount of vocabulary, one cannot communicate effectively or express their ideas in both oral and written form. Having constrained vocabulary is likewise a barrier that blocks students from learning a language. Language instructors, therefore should process impressive information on the best way to deal with a classroom so the students can pick up an incredible accomplishment in their vocabulary learning. Without grammar very little can be conveyed, without vocabulary nothing can be conveyed. As Harmer asserts (1992), “students should also know that the words they learn have both literal and metaphorical meanings”, which requires an advanced level of knowledge for EFL students (p. 14).

### **2.7.3. Pronunciation**

Pronunciation is a way of delivering clear and understandable language when talking. It touches upon the phonological process that refers to the components of a grammar made up of the elements and principles that determine how sounds vary and pattern in a language. Phonemes and supra-segmental features are the two elements of pronunciation. Unless a speaker who pronounces a range of phonemes correctly, it can be extremely difficult for a speaker from another language community to get the meaning. Pronunciation is the knowledge of how to say a word – that is how to pronounce it. This component comprises three elements which are sounds, stress and intonation (Güllüoğlu, 2004). Pronunciation addresses many features of the speech stream, including individual sounds, pitch, volume, speed, pausing, stress and intonation (Luoma, 2004).

### **2.7.4. Fluency**

As long as what you say is fluent and accurate, it means that speaking is fluent. Speaking fluently in the target language is the ultimate goal of many language learners. If one is capable of speaking at reasonably fast speed and with only a small number of pauses and “ums” or “errs”, it can be said that it is fluent. These indicate that the speaker does not need to invest a ton of energy and time hunting down the language things expected to express the message. These show that the speaker does not need to invest a ton of energy hunting down for the language items to express the message (Brown, 1997: 4). To conclude, “fluency needs practice and time to improve” (Güllüoğlu, 2004, p.15).



## 2.8. Testing Speaking

### 2.8.1. Reliability

Reliability is usually defined as score consistency (AERA, 1999; Brown and Hudson, 2002; Henning, 1987). Reliability plays an important role in assessing as we use it for decision making. Unreliable scores, on the other hand, may lead wrong decision-making procedures, unjustified results.

Underhill (1987) sees reliability as a specific type of general validity; and supports his belief with this sentence “a test cannot be generally valid unless it is reliable” (p.105). He also points out that classical measures for test reliability have nothing to do with oral tests as they are mostly designed for pre-planned tests which include fixed number of individual questions.

On the other hand, oral tests are inseparable from the people who involve in them or they are not like the sets of questions on paper (Underhill, 1987). Therefore, it would be a nice idea for a test designer to design her own system for comparing the results, and producing the most reliable results between markers.

### 2.8.2. Validity

The process of validation is well defined by Underhill (1987) as “Validation is a relative, not an absolute, process; the degree of validity of a test relates only to then particular circumstances in which it was established” (p.104).

As Norris, Brown, Hudson, & Yoshioka (1998) pointed out, “the issues of reliability and validity must be dealt with for alternative assessments just as they are for any other type of assessment—in an open, honest, clear, demonstrable, and convincing way” (p. 5).

Validity answers this question: “Does the test work properly?” According to Underhill (1987), “the process of establishing the general validity of a test procedure is called validation and all aspects of validity are important.” (pp.104-105). It is highly important to understand that validity issue is a little bit vague issue that suffers from lack of definition. Assessing speaking skill precisely is the hardest of all the other language skills since it involves the simultaneous use of wide variety of different abilities that often develop at different rates (Harris, 1969).

The test takers could be questioned to see whether a test is reasonable or not as they are capable of producing informative and objective comments about the tests. To see whether the test is relevant or the items/tasks match what the test is supposed to assess the syllabus and the kind of language generated in the test could be compared. In addition, a test should share the same philosophy as the program of which it is part. Otherwise, the theory behind the test does not match with the test. Concurrent validity has also to do with the same learners' scores on a new test procedure. Finally, according to Underhill (1987) if someone wants to have a predictive validity, it is suggested for you to have clear objectives.

In the literature, in terms of construct, content and concurrent validity language proficiency tests are investigated. Next, face validity is considered - if nothing else can be found, but because it cannot be "measured", sometimes rather slanderously. Some other tests that have specific purposes are assessed and valued in terms of predictive validity.

There are four main types of validity:

#### *Construct validity*

Construct validity is the degree to which the instrument explicitly measures what it is expected to rather than measuring other things. For example, a measure of success should only assess factors relevant to success and not, for instance, whether someone is a hard worker. Underhill (1987) explains that a test should share the same vision and philosophy and have something in common with the program of which it is part. Construct validity is not the simple idea to work with. However, one should keep this question in mind: To what extent does my test match the purposes and basic assumptions and views on language learning?

#### *Content validity*

Content validity describes whether an instrument is explicitly or implicitly representative of the trait it is measuring. For example, a questionnaire aiming to score the level of stress while producing a language should include questions aimed at a broad range of features of stress factors that arouse anxiety while producing a language. According to Underhill (1987), the question one should keep in mind is whether a test matches the aims and needs that are supposed to match. Hughes (1989) also states that a test should measure accurately of what it is supposed to measure and if a test does not represent the specifications thoroughly then it will have harmful backwash effect.

### *Face validity*

As Hughes (1989) stated, face validity is the degree to which a test is subjectively thought to measure what it intends to measure. In other words, does it “look like” it will measure what it should do. For example, when a test aims to measure pronunciation but do not include test elements that are needed to measure pronunciation, then it may be said that it lacks face validity.

### *Criterion validity*

According to Hughes (1989), criterion-related validity can be divided into two kinds: concurrent validity and predictive validity. Hughes (1989) states that we can mention concurrent validity as long as the test and the criterion are administered at about the same time and he also continues that predictive validity is related to the degree to which a test can predict candidates’ future performances.

### **2.8.3. Washback Effect**

Washback, also called backwash in various places in the literature, is the effect of testing and assessment on the language teaching curriculum.

The washback effects of large-scale testing programs on instruction are widely discussed. In the view of instructors and students, such tests contain what students must learn and therefore what must be taught- a reasonable view, given that the tests in many cases represent the language hurdle students must clear before continuing their academic careers. (Chapelle & Douglas, 1993, p.16)

The importance of washback is accepted in language testing. Recently, Alderson and Wall (1993) called into question the existence of washback, and rightly so, given that little if any actual research had ever demonstrated its existence. More recently, a number of studies have further confirmed the existence and complex nature of the washback effect (e.g., Alderson & Hamp-Lyons, 1996; Shohamy, Donitsa-Schmidt, & Ferman, 1996; Wall, 1996; Watanabe, 1992, 1996a, 1996b). These results show that it is not wrong to mention the existence of washback effect in language teaching and testing.

Washback effects can be divided into two categories: negative or positive. When there is a mismatch between the assessment strategies in a curriculum and curriculum’s goals and objectives, then, there will be mismatches between the sets of goals in the curriculum and the implications conducted throughout the teaching and learning process. For example, according to Hughes (1989), if you want to encourage and see the oral ability, then it is needed to test oral ability. If a program sets a series of cooperative and communicative

objectives but at the end of the year evaluates the students' skills with True-False statements, a negative washback effect will be inevitable.

According to Cheng (2005), if the positive washback is supported, then positive changes in teaching and examinations can be achieved. That is, a positive washback effect occurs when the assessment procedures meet the course goals and objectives. For instance, if a program wants to achieve a set of communicative performance objectives, the tests should make students use their communicative skills such as role plays and interviews. Unless it is done, positive washback effect cannot be created in favor of the communicative performance objectives. That is, positive washback occurs when it is tested what it is taught and when the tests measure the same types of materials and skills that are described in the objectives and taught in the courses.

#### **2.8.4. Practicality**

As long as a test is practical, it can be said that it is effective. Rather than being extremely expensive, it should be in accordance with the appropriate time constraints, be relatively easy to administer, and have a scoring/evaluation procedure that is specific and time efficient. Practicality has an effect on test usefulness.

It does not mean that a test is valid as it contains more data to measure, the other thing that is needed to be taken into consideration is to what extent it is practical to apply. When a test arouses the feeling of tiredness in the test taker, it means that test is impractical in terms of the economy of time, effort and money in testing. On the contrary, a test which is assumed as practical should be easy to design, easy to administer, easy to mark, and easy to interpret the results (Bachman and Palmer, 1996). According to Brown (2004), the test that is practical needs to be within the means of financial limitations, appropriate time constraints, easy to administrator, score, and interpret.

Davies (1990, p.6) points out that "It is important to remember that testing is possible only if it is practicable. A good test may in practice be unusable because in the situation for which it is intended it would take up too much time, too much skilled manpower, or it might require expensive or elaborate media systems or scoring arrangements, and so on."

### 2.8.5. Test Administration

Unless the test is organized and administered well, it is inevitable for the test to give unreliable and invalid results. For that reason, there are certain points that should be taken into account before administering a test. As Hughes (1989) stated, the key to successful test organization is to keep the advance preparation in mind. The important factors include materials and equipment, examiners, invigilators or proctors, candidates, rooms and administration.

*Materials and equipment:* arranging and numbering the test materials, organizing the booklets, proofreading the test items to make sure that the items are free of error and checking the tape-recorders, speakers and other materials whether there is something to fix. These allow the test makers much greater validity and security before, during and after test administration.

*Examiners:* in order to create an environment where a test is applied properly, the instructions and the directions should be clear for the examiners and read before the exam. It is important to make sure to create an environment where examiners feel familiar with the exam equipments and test procedures.

*Invigilators/proctors:* to ensure a test which is free from all mistakes and faultless, all of the instructions should be prepared beforehand and supplied for the invigilators.

*Candidates:* all the candidates should be given information such as what to bring, what they are supposed to do before and during the exam if necessary.

*Rooms:* rooms should be large and quiet. Especially for the speaking exams, organizing the equipments needed for speaking and listening activities is important. There should also be enough space in the room to organize the activities comfortably. The layout of the rooms and the seating charts should be prepared before the implementation of the exams.

*Administration:* what is essential about a test is to provide clear and necessary instruction for the examiners before and during a test. If spoken tests instructions are to be given, the examiners should be made to read them aloud. Necessary materials should also be supplied for all the candidates in case they need them before the exam.

## **2.9. Speaking Scales**

### **2.9.1. Scoring and Marking Systems**

Marking system can be said the most important part for oral tests since it has always been a challenge for the testers and markers. As Weir (2005) stated, rating the assessment of spoken language is potentially more challenging than the rating of written scripts.

As for Underhill (1987) unlike the objective tests which requires less time and effort to mark, subjective tests such as oral and writing tasks where there is no pure right or wrong, it cannot be said that they have high reliability. No matter how it is hard to avoid subjectivity in marking the tests, it is highly important to assess the speaking tests as objectively as possible “so the design of the marking system must be consistent with the aims of the program and reflect the answers to all the questions” (Underhill, 1987. p.89). Therefore, to promote a higher standard of reliability and validity, the planning and the selection of the oral test types stage should be given great attention by the test designer and also the marking system should meet the needs accordingly.

Moreover, as it allows the test maker to mark whenever or wherever he/she wants, scoring a test is much easier and practical when it is recorded. As Underhill emphasized (1987), to make the tests transparent, the assessor just records the tasks during the exam but marks them later. He also added that as in a laboratory, you cannot ask the speaker to repeat the sentences.

### **2.9.2. Holistic Scoring**

Holistic scoring can still be defined as an evaluative method that considers the overall quality of the product. Underhill (1989) also named this scoring system as ‘Impression Marking’. According to Madsen (1983), in holistic scoring the focus is on evaluating the overall performance instead of thinking about the errors.

Table 2.1. *Interview Assessment Scale (Carroll, 1982, p.135)*

10	Expert speaker	Speaks with authority on a variety of topics. Can initiate, expand and develop a theme.
9	Very good non-native speaker	Maintains effectively his own part of a discussion. Initiates, maintains and elaborates as necessary. Reveals humor where needed and responds attitudinal tones.
8	Good speaker	Presents case clearly and logically and can develop the dialogue coherently and constructively rather less flexible and fluent than Band 8 performer but can respond to main changes of tone or topic. Some hesitations and repetitions due to a measure of language restriction but interacts effectively.
7	Component speaker	Is able to maintain theme of dialogue, to follow topic switches and to use and appreciate main attitude markers. Stumbles and hesitates at times but is reasonably fluent otherwise. Some errors and inappropriate language but these will not impede exchange of views. Shows some independence in discussion with ability to initiate.
6	Modest speaker	Although gist of dialogue is relevant and can be basically understood, there are noticeable deficiencies in mastery of language patterns and style. Needs to ask for repetition or clarification, and similarly to be asked for them. Lacks flexibility and initiative. The interviewer often has to speak deliberately. Copes but not with great style or interest.
5	Marginal speaker	Can maintain dialogue but in a rather passive manner, rarely taking initiative or guiding the discussion. Has difficulty in following English at normal speed; lacks fluency and probably accuracy in speaking. The dialogue is therefore neither easy nor flowing. Nevertheless gives the impression that he is in touch with the gist of the dialogue even if not wholly master of it. Marked L1 accent.
4	Extremely limited speaker	Dialogue is a drawn-out affair punctuated with hesitations and misunderstandings. Only catches part of normal speech and unable to produce continuous and accurate discourse. Basic merit is just hanging on the discussion gist, without making a major contribution to it.
3	Intermittent speaker	No working facility; occasional, sporadic communication.
2	Non-Understandable	No coherent communication.
1/0	Non- speaker	Not able to understand and/or speak.

### 2.9.3. Analytic Scoring

When it comes to compare analytic scoring with holistic scoring, it can be said that the first is less subjective than the latter and proceeds from a different set of assumptions. In holistic scoring the evaluation is done by looking at the product as a whole but in the analytic scoring skills that form the whole are evaluated. Hunter, Jones and Randhawa (1996) stated that in the analytic scoring, the abilities of students such as organization, content, mechanics,

and diction are judged, then a total score can be calculated. It is “the recording and tabulating of subscores which separates analytic from holistic scoring” (Goulden, 1989, p. 4–5). There are also some parts of analytic scoring to be improved mentioned by Underhill (1987). The first one is that rating scales are built in line with a typical learner, but one should keep this in mind that not all the learners are typical, and assessing speaking is a controversial issue. That is why, there is a need for developing a scale which fits every learner. Second, it is controversial to what extent a rating scale should be detailed.

Table 2.2. *Analytic Rating Scale (Council of Europe (2001) Common European framework of reference for languages: learning, teaching, assessment.)*

Interactive communication
5 Almost wholly effective at communicating both actively and receptively in everyday contexts. Fully sensitive to turn-taking system. Contributes to collaborative topic development and maintenance by asking others to express/expand their opinions and by negotiating meanings both verbally and non-verbally (e.g. ask for clarification, indicate understanding, establish common ground, correct others' utterance and respond to requests for clarification).
4 Communicates effectively by appropriately participating in turn-taking. Responds, comments (e.g. agree/disagree), asks questions, negotiates meanings verbally and non-verbally and develops the interaction in some but not all the occasions.
3 Communicates adequately in most everyday contexts, but could be rather passive with responding and commenting. Asks for clarification (repetition, paraphrasing) verbally or non-verbally, although occasionally it may be unsuccessful. Not effective enough to contribute to develop the interaction.
2 Interaction is ineffective because it is too passive (talks only if required), it lacks coherence or it is monologue only. May show some (verbal or non-verbal) attempts to ask for repetition or paraphrasing, which are frequently unsuccessful.
1 Gives simple responses only when required, but is unable to maintain or develop the interaction. May show a few attempts (mostly non-verbally) to ask for repetition or paraphrasing, which are nearly always unsuccessful.
0 No awareness of other speakers. Understanding and communication is impossible. Neither verbal nor non-verbal attempt to ask for repetition or paraphrasing.

#### 2.9.4. CEFR

The Common European Framework of Reference for Languages, (Council of Europe, 2001) more commonly known as the CEFR, is a principle that defines language learners' language ability. The CEFR is a framework that we use to describe a learner's language ability broadly. In the CEFR, the levels are described in stages. The CEFR identifies the abilities of a language learner with can do statements from the lowest to the highest. It also defines the progress of a learner. It comprises of descriptions called “Can Do” statements that are associated with each of the four skills, Listening, Speaking, Reading and Writing. These descriptions show what a learner is capable of doing by using the language skills. With the help of CEFR, speaking assessment can be more specific and



detailed if it is used in a professional way, the functions and language-use contexts would have to be arranged to suit that test ( Luoma, 2004).

Table 2.3. *Qualitative aspects of spoken language use - Table 3 (CEFR 2018 3.3): Common Reference levels (p.171)*

	RANGE	ACCURACY	FLUENCY	INTERACTION	COHERENCE
B2	Has a sufficient range of language to be able to give clear descriptions, express viewpoints on most general topics, without much conspicuous searching for words, using some complex sentence forms to do so.	Shows a relatively high degree of grammatical control. Does not make errors which cause misunderstanding, and can correct most of his/her mistakes.	Can produce stretches of language with a fairly even tempo; although he/she can be hesitant as he or she searches for patterns and expressions, there are few noticeably long pauses.	Can initiate discourse, take his/her turn when appropriate and end conversation when he / she needs to, though he /she may not always do this elegantly. Can help the discussion along on familiar ground confirming comprehension, inviting others in, etc.	Can use a limited number of cohesive devices to link his/her utterances into clear, coherent discourse, though there may be some "jumpiness" in a long contribution.

## 2.10. Tasks in Speaking Tests

In many EFL classes, the interaction or communication among the pupils or between teacher and student is not real as teachers ask ‘display’ question and students answer. Mostly the answer of the question is already known by the teacher, and the teacher completes the evaluation part and if necessary corrects the answer. This kind of circle has no effect on the development of students’ speaking abilities. “Eventually, this is an unrealistic use of language and these questions have clear limitations in terms of how much genuine communication practice the student receives” (Dinapoli, 2000, p.1). Hence, for a real communication environment in the language classroom, the people who communicate should create an environment where the communication goes beyond display question and the interlocutors should not know what the other is going to say in advance (Liao, 2001).

According to the hypothesis asserted by Swain (*output hypothesis*) (1985) and Long (*interaction hypothesis*) (1996), tasks are good ways for facilitating L2 knowledge. Swain (1985) argued that when learners see the process of producing language, then they will be

able to gain control or internalize knowledge. Therefore, the output produced at the end is the process not a result (Adams, 2003). In addition, according to interaction hypothesis, the interaction between the interlocutors make is possible to use a variety of strategies to negotiate meaning and to achieve mutual comprehension.

Despite of the fact that tasks can be thought to have many purposes ranging from grammatical exercises to complex classroom activities, they are also practical for fostering speaking skills. Nunan (1989) defines a task as a way of facilitating a learner to produce language by interacting using the target language without minding the ‘form’ but rather focusing on the ‘meaning.’ Ellis (1994) also defines the task as an activity that enables the learner to use the language communicatively or reflectively to achieve a goal. Long & Crookes (1992) and Skehan and Foster (1999), widened the definition of tasks and described them in terms of five characteristics. According to their point of view, a task is an activity that gives much of the importance to the meaning and there is a goal needed to be achieved and the whole activity is outcome oriented. In addition, the communication and the relationship is genuine. Finally, there is a time limitation for the interaction and it is problem oriented. All in all, when all the definitions of tasks are considered, it becomes clear that the focus of the tasks is on the achieving a goal for communicative purposes.

Nunan (1989), Finch (1997), Brown (1998) and Nunan (2005) define task elements as follows: *Input* refers to the data for the task. *Activities* are what learners do with the input. *Teachers and learners* are supposed to play in fulfilling the task as well as creating the social and interpersonal relationships between the participants. *Setting* is the arrangements related to where tasks are to be fulfilled.

### **2.10.1. Advantages of Tasks in Speaking Instruction**

As for the advantages of using tasks in language classes, it can be said that learners develop some certain skills such as negotiation and achieving task goals. Basically, with the help of communicative tasks learners learn how to negotiate to establish the meaning of unfamiliar aspects of the task language, achieve a goal and promote their language acquisition skills (Courtney, 1996; Finch, 1999 and Lee, 2000). Therefore, while trying to complete the task goals, one can promote language acquisition. According to Richards and Rodgers (2001), “activities in which language is used for carrying out meaningful tasks promote learning” (p.223). In addition, the tasks are useful in terms of promoting pair or group work as working in pairs or groups help learners to complete a task in given time and

to increase motivation. Therefore, as Richards and Rodgers (2001) pointed out although the focus seems to be on the solo performance first, attention then turn to team task where communication is needed.

Tasks can also be used to give the learners a chance to come up with more coherent and fluent sentences as well as promoting discourse competence. Some tasks which require students to perform the language such as role plays, simulations and drama give some opportunities to the learners to practice the language and engage in different roles. (Kasper, 2001 and Ellis, 2003).

According to Swain's "output hypothesis" (1985) and Long's (1996) "interaction hypothesis", tasks can be made use of for facilitating learners to control the things they have learnt and take the responsibility of their own speaking process. In addition, variety of strategies being used during task completion procedure foster language learning process and increase interaction between interlocutors.

### **2.10.2. Designing Tasks**

The ultimate aim of speaking tasks is to foster and promote communication and interaction as much as possible and there are some factors to be taken into consideration (Torky, 2006).

*Thought:* a speaking task should not only help learners to think out and internalize what they have learnt but also develop speaking abilities. When the definition of thinking in terms of logical processes can be done: generalizations, exemplifications, analysis, synthesis, and evaluation, alternatively (Ur, 1981).

*Result:* in order for a task to achieve its goal there is a need of giving feedback. To see the future needs or what feedback is needed, each task should give a result to teachers. Results are also good for teachers as they help them to understand the sign that the group has done with the task and need further comment or feedback to continue. (Brown, 1998 and Myers, 2000).

*Language practice efficiency:* tasks are helpful for students as they promote not only transactional but also interactional functions of language; long turns and short turns. It should also help students to express themselves in different situations, feelings and relations (Ur, 1981 and Myers, 2000).

*Interaction:* in a task completion process a teacher should make sure that students interact one another and it is impossible for the students to complete the tasks alone much easily (Ur, 1981 and Riggenbach, 1998).

*Interest:* According to Scarcella & Oxford (1992), if tasks aren't interesting and relevant, they won't be able to affect the students' attitudes and motivation in a positive way.

### **2.11. Alternatives in Language Assessment**

As Brown and Hudson (1998) stated 'the various kinds of language assessments are classified into three broad categories: (a) selected-response assessments (including true-false, matching, and multiple-choice assessments); (b) constructed response assessments (including fill-in, short-answer, and performance assessments); and (c) personal-response assessments (including conference, portfolio, and self- or peer assessments)' (p.653).

Nowadays, alternative assessment tools have gained popularity. Huerta-Macías (1995) came up with the alternative assessment procedures such as checklists, journals, logs, videotapes and audiotapes, self-evaluation, and teacher observations. Portfolios, conferences, diaries, self-assessments, and peer assessments can also be added to this list. However, it is necessary to understand the reason why alternative assessments are applicable and contemporary to use, and what makes them different from the traditional assessments. So far, so many answers were given to this question.

Aschbacher (cited from Derakhshan, A., Rezaei, S., & Alemi, M. (2011) listed several common characteristics of alternative assessments, stating that rather than focusing on the product only they also focus on the process, they emphasize higher level thinking skills such as problem solving and analyzing, involve tasks that are worthwhile as instructional activities, use real-world topics and contents, and encourage public disclosure of standards and criteria. Herman, Aschbacher, and Winters (1992, p.6) offer a somewhat different set of characteristics:

Alternative assessments require students to perform, create, produce, or do something; tap into higher level thinking and problem-solving skills; use tasks that represent meaningful instructional activities; approximate real-world applications; ensure that people, not machines, do the scoring, using human judgment; and call upon teachers to perform new instructional and assessment roles.

With all these positive characteristics for alternative assessments in mind, it can be said that they appeal to most language teachers and testers. According to Brown and Hudson (1998), 'Tasks ensure that people, not machines, do the scoring, using human judgment;

encourage open disclosure of standards and rating criteria; and call upon teachers to perform new instructional and assessment roles' (p.654-655).

### **2.11.1. Computer Based Testing**

Basically, although computer-based tests have some demerits, there are also several merits. In a comprehensive study, Alderson (2000) specified a lot of pedagogic and technical advantages of CBTs. He thinks that technically the computer-based tests have the power of eliminating the limitations of test administration such as fixed delivery dates and locations. He also adds that thanks to computer-based tests it is possible to see the results immediately after the test. Furthermore, the test security is greater as a result of examinees' access to a large database of items.

From a pedagogic point of view, Alderson (2000) points to the user friendliness format of computer-based tests. In addition, he emphasizes as the feedback they give to the users is useful, the tests would be more meaningful. They also offer a range of support to learners from help facilities and dictionaries to clear instructions, examples and performance clues. In addition to these, Chalhoub-Deville (2010) also thinks that CBT makes it possible for the test takers to follow their language learning process and allows them to experience new item/task types.

### **2.11.2. Computer Adaptive Testing: CAT**

Chapelle and Douglas (2006) stated that a computer-adaptive test decides and shows the next items according to the test taker's answer to each item. When the first question is answered correctly, a more difficult question is selected from a pool and presented next; if this one is answered correctly, a more difficult one is selected next and it goes in that circle. However, the system selects an easier one if a question cannot be answered by the candidate.

According to Alderson (2000), "Computer adaptive tests are often argued to be more user-friendly, in that they avoid users being presented with frustratingly difficult or easy items. They might thus be argued to be more pedagogically appropriate than fixed-format tests" (p.596).

Like all the other test types, CAT has both advantages and disadvantages. According to Fulcher (2000), CAT has a number of disadvantages. Creating a large number of totally calibrated items for the item bank is time-consuming and costly. In addition, unlike paper-

and-pencil tests, CAT does not allow language examinees to omit items or to review them at the end of the test.

Tung (1986) also highlighted the following advantages: they require fewer items than their paper counterparts, they avoid challenging examinees far beyond their capability by selecting items at the appropriate difficulty level, and they offer improved security by selecting from an item pool to construct individualized tests.

These issues were discussed by Alderson (1990, p. 39–43) who outlined computer capabilities relevant to exploring an innovative agenda for CALT:

1. The computer has the ability to measure time. The time which a learner takes to complete a task, or even the time taken on different parts of a task, can be measured, controlled and recorded by computer.
2. The computer has the ability to record information about the testee's routes through the test.
3. The computer can present information in a variety of ways.
4. The computer can provide quick and easy access to a variety of different types of information.
5. The computer can be linked to other equipment. This can allow different types of input and presentation.
6. The computer can encourage the learner's own strategies for evaluation. In particular the information which a computer can collate and present about test performance could help the learner to feel that his own opinions are of importance.
7. The computer can make use of language rules. At a relatively simple level the computer can do a spelling check on the learner's text. b. Parsers of varying degrees of sophistication can be used not only to check for syntactic errors in the learner's text, but to provide "communicative" tests as well.

### **2.11.3. Internet-based Testing**

When Internet-based tests are thought one question arises in minds: But why do we need to apply Internet-based tests although we already have computer-based assessments? This question have been answered by Fulcher (2000), Alderson (2000), and Roever (2001) by displaying a number of merits of language testing through the Internet. The very first and main advantage of this kind of test is its superiority in terms of time and space. It is possible for a test taker to take the tests whenever or wherever s/he wants as long as there is an Internet access.

When the design issue is taken into consideration, the Internet-based tests are more flexible. It offers the users a great number of texts, images, audio, and video at the same time. In addition, accessing to the online resources, the databases or libraries which are all updated constantly is easy and practical. Moreover, test results can be sent instantly to the score users. As Roever (2001) adds, "Internet based tests are very inexpensive for all parties concerned" (p. 88).

## 2.12. The Significance of Feedback in Speaking Assessments

There may be some mismatches between testing and curriculum in language programs. To correct such a situation, there are issues to be considered: the positive and negative consequences of washback, the significance of feedback, and the importance of using multiple sources of information.

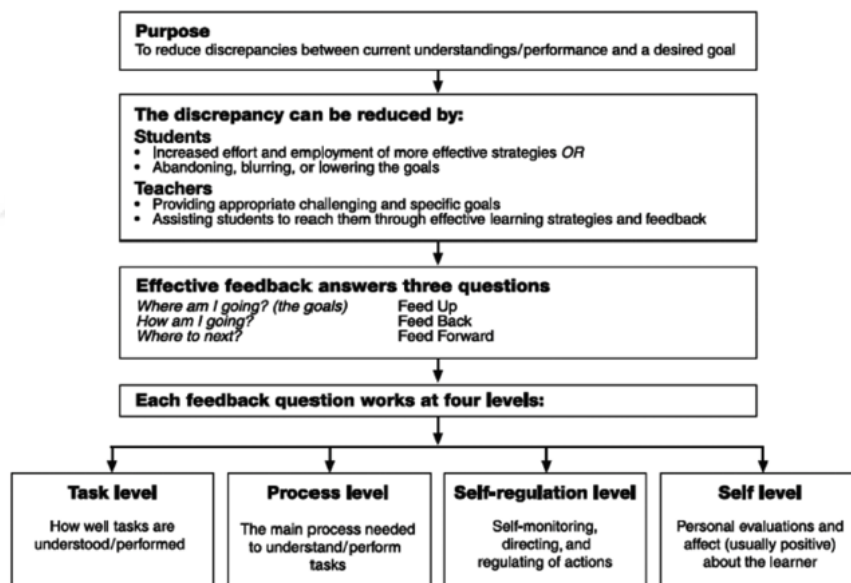


Figure 2.3 A model of feedback to enhance learning (Hattie & Timperley, 2007)

According to Shohamy (1992), the reason why feedback is given vary depending on different situations, yet feedback remains important. For instance, in a diagnostic pretest that is implemented at the beginning of a course the purpose of feedback is to inform students of their strengths and weaknesses. Thus in a diagnostic pretest, the feedback is given in terms of what the students need to do about each course objective. On the other hand, the purpose of the feedback in an achievement test will be quite different. For example, when the focus is on how a student is doing with the objectives of a particular course, the results need to be

interpreted in terms of what the students have learned or learned how to do in the course. Hence, a low score shows the student is lack of the knowledge or skills necessary to master that objective. In that situation, it may be suggested for that students to work hard and realize her/his weaknesses for further studies. Alternatively, if some students have low scores on a number of objectives, it may be possible for the teachers to decide not to promote those students to the next level or that they should fail the course and be required to take it again.

As Heritage (2007) stated, “Formative assessment is a systematic process to continuously gather evidence about learning” (p.2). Formative assessment can be defined as being interactive during the learning process, and the overall progress of the learners can be assessed by using formative assessment strategies. Formative assessment improves the learning environment as it helps learners to understand what they are supposed to learn, how they learn, and what other ways to be applied to learn better.

Providing corrective feedback is a crucial element in teaching and testing processes as it gives a corrective guide to the learner to build up the future actions. According to Irons (2008, p.55), constructive feedback “should be proper to the formative evaluation and to the student learning process”. In giving corrective feedback, the focus should be on the behaviours rather than the personality. In addition, the right balance between the positive and negative statements is needed to be found. Moreover, constructive feedback should specifically define the purpose and offer some suggestions to the learner for the next steps.



## **CHAPTER III**

### **METHODOLOGY**

In this chapter, research design, setting and participants of the study, data collection instruments, data collection process, and data analysis are presented.

#### **3.1. Introduction**

The aim of this study is to find out test takers perceptions and attitudes towards online speaking assessment system. It was also aimed to find their perceptions and attitudes towards the effect of this system on their speaking and listening competencies. It goes one step further by investigating the test makers, the instructors and the school administrators' ideas and their thoughts about the effect of it on the test takers' development.

#### **3.2. Nature of the Study**

According to Nunan (1992) "research is a process which involves (a) defining a problem, (b) stating an objective, and (c) formulating a hypothesis." (p. 2). In other words, it is an activity to find answers to the predetermined questions by asking questions, collecting and analyzing the data, and critically evaluating the results. Creswell (2014) stated that the definition of research can be given as the process that includes some steps for better understanding of a topic or question(s).

As Creswell (2003) stated, combining methods in order to deeply understand research problems requires to make use of the strategies of inquiry that involve collecting data either simultaneously or sequentially. In this study mixed method designs were used to collect data. Therefore, in order to answer the research questions, both quantitative and qualitative research methods were used. While quantitative method was used to reach specific, narrow, measurable, and observable results, qualitative method was used to capture the participants' experiences and implications. In order to achieve this goal, both questionnaires and interviews were adopted. A questionnaire which consists of three parts were used to gather information about the test takers' perceptions and attitudes about computer mediated speaking testing and its effect on their speaking and listening proficiency.

The aim of the questionnaire used in this study was to collect data about the test takers' attitudes and perceptions. In addition, interviews were also carried out both with students and instructors, test makers and administrators to find out more detailed data related to the research questions. The research design was the explanatory sequential design (Cresswell and Plano Clark, 2011). As Creswell (2003) stated, combining methods in order to deeply understand research problems requires to make use of the strategies of inquiry that involve collecting data either simultaneously or sequentially. A questionnaire which consisted of three parts were used to gather information about the test takers' perceptions and attitudes about computer mediated speaking testing and its effect on their speaking and listening proficiency. In addition to the questionnaire, semi-structured interviews were also conducted to strengthen the results in a broader way. As for the second research question, qualitative methods were used to gather details about the test makers,' the instructors' and the school administrators' views and opinions about the effect of online speaking assessment system and its effect on the test takers' overall speaking and listening competency through semi-structured interviews. The reason why mixed methods design was used in this study is defined by Creswell (2003) as:

The researcher bases the inquiry on the assumption that collecting diverse types of data best provides an understanding of a research problem. The study begins with a broad survey in order to generalize results to a population and then focuses, in a second phase, on detailed qualitative, open-ended interviews to collect detailed views from participants (p.21).

Similarly, Mackey & Gass (2005) state that to gather data, combining and mixing different methods are given more and more attention by second language researchers. They also state "When included in a primarily quantitative report, qualitative data or analytic techniques may provide unique insights that would escape both the researcher and the reader if statistical counts and analyses were used in isolation" (p.307).

### **3.3. Methodology of the Study**

The information about the setting, participant, data collections instruments and procedures, and the data analyses parts are discussed in the following parts.

#### **3.3.1. Setting**

The study was conducted in a preparatory school of a private university in İzmir. In the preparatory class, the students are supposed to cover four modules based upon the CEFR

system. The students are placed into the appropriate module according to the exams done at the end of each quarters. In addition to the other skills (reading, writing), speaking and listening skills are highly important for students to pass each level. The school system requires students to take a placement exam which decides the students' level of proficiency at the very beginning of the school year. According to their scores, the students are placed into appropriate levels. The levels are designed in accordance with CEFR: A1, A2, B1, B2. They all go through a process that will last a year and contains four quarters, each contains eight separate weeks. Finally, in order to start their departments, the students are supposed to pass the proficiency exam administered at the end of the academic year.

In addition, online speaking assessment was a program designed to promote learners' speaking skills through systematically prepared activities, the aim of which was to assist learners to speak the target language. Students needed to take 20/25-minute exam twice per week. During each session, they were given some tasks related to the classroom activities and supposed to take active part using the target language. With the help of these tasks and activities, students tried to develop their speaking as well as listening competencies. Some of the tasks used during the assessments were gap filling, pictures/picture story, pictures for describing, pictures for questions and answers. The examiners marked the learners' performances based on a CEFR "can do" criteria that were mentioned earlier in the study. The participants were informed about the criteria before they participated in the activities. Later on, the scores were shared with the institution members along with the participants every week.

### **3.3.2. Participants**

The participants for the questionnaire part in this study were the whole B2 students who were learning English as a foreign language in the Preparatory School at a private university in İzmir. The students had 24 hours of English lessons per week including four skills. Altogether 291 participants took part in the questionnaire part of the study. In addition, while 55% of the students were female, 45% were male. The age ranged from 18 to 23, but 85% of the students were between 18-20. Questionnaires were administered towards the end of the quarter during the class hour.

Table 3.1. *The distribution of participants according to their gender and age.*

	<i>Frequency</i>	<i>Percent (%)</i>
Gender		
<i>Female</i>	157	54.9
<i>Male</i>	129	45.1
Age		
18	91	31.9
19	99	34.7
20	52	18.2
21	24	8.4
22	7	2.5
23	9	4.3

The current study was conducted towards the end of the last quarter with B2 students who completed all the levels successfully. The reason why B2 students were intentionally chosen for the study was that it was considered that those students had already taken lots of online speaking exams since the beginning of the program and were knowledgeable enough about the system to express their perceptions and opinions about the new online speaking system.

The findings which were obtained from the questionnaire were complemented by semi-structured interviews with students ( $N=24$ ) and test makers, administrators and instructors ( $N=5$ ). The students for the interviews were chosen from the volunteer students who wanted to comment on the system and the questions in the interview were mostly on the ideas and perceptions of the students about the speaking tests and its contribution to their speaking and listening skills.

The test makers, administrators and instructors constituted the second group participants of the study as they contributed to the system as the assessor, interviewer, or designer. The ones who conducted the testing process from the beginning to the end of the quarter were kindly asked to participate in the interview process and the volunteers' perceptions and opinions were obtained through interviews. Altogether five participants who were involved in the system from the beginning took part in the interview parts. Two of them were in the administrative group, two of them were the instructors and one of them was the test maker (examiner) who organized the whole process including the tasks from the beginning of the implementation. In addition, while 3 of them were female, the others were male. The age of the participants ranged from 25 to 40. The volunteer administrators, test makers and instructors who had experienced every parts of the assessment from the beginning to the end of the year shared their ideas about online speaking assessment and its

effects on students' development. They also asserted their ideas about the advantages or drawbacks of the system and made comments to make the system more efficient for both the institution and the students.

### **3.3.3. Data Collection Instruments and Procedures**

The study which included one pilot and one main study was based on a questionnaire and semi-structured interviews that were conducted with both students and the instructors, administrators and test makers.

The study had two phases. In the first phase, to eliminate the problems related to data collection tools and provide reliability, a pilot study with 45 B2 students was conducted at the end of the 5<sup>th</sup> week of the 4<sup>th</sup> quarter. After that, the items were revised and we had expert opinion from the staff of ELT Department of Pamukkale University. Later on, the main study was conducted with 291 B2 students who were studying at preparatory school during the spring semester of 2014-2015 academic year. In the second phase, interview data were obtained through semi-structured interviews conducted with both students and instructors, administrators and test makers.. Interviews with students were constructed to cross-check the findings obtained from the questionnaire.

Basically, the study consisted of two parts; first part, a questionnaire containing three main parts, which was designed to obtain information on the students' perceptions and attitudes towards online speaking assessment and its effect on their speaking and listening competency, also demographic information such as computer familiarity, system use and prior experiences was collected through the questionnaire. Likert scale was used in the questionnaire and students indicated their responses on a five-point scale, on which 1 is for "strongly disagree" and 5 is for "strongly agree." Second part, shortly after the week questionnaire was administered, in order to collect qualitative data semi-structured interviews were held in five different sessions with 24 students from different classes. The number of the interviewees varied from four to five in each session. Each interviews lasted about half an hour and nine questions were asked to the interviewees. The interviews were all conducted in Turkish and tape-recorded so that the responses could be transcribed and translated into English. As a follow-up activity, to gather more information about the research questions, and instructors, administrators and test makers ( $N=5$ ) were also interviewed to discuss the effectiveness of the newly adopted speaking test.

### 3.3.3.1. Questionnaire.

A few of the items from some parts (System Use, Prior Experiences, Computer Familiarity) that aim to find out the test takers' perceptions and opinions about Online Speaking Assessment system was adapted from Şanlı's MA (2003). However, most of the parts were added later specifically for this study and applied six weeks after the students started B2 quarter assuming that students were fully familiar with the system.

Likert scale was used in the questionnaire and students indicated their answers on a five-point scale, on which 1 is for "strongly disagree" and 5 is for "strongly agree". The data obtained from the questionnaire were analyzed by using descriptive statistics (mean scores, frequency, percentage).

In the first part of the questionnaire which includes computer familiarity, prior experiences and system use parts, it was aimed to find out demographic information about the test takers and their ideas about the system in general. In the second and third parts on the other hand, finding the test takers' perceptions and attitudes towards the system and its effect on their speaking and listening competencies were tried to be found out. In the process of preparing the items for the questionnaire three experts' (from the field) ideas were taken into account and necessary regulations were done in the light of this feedback. The questionnaire is applied to 291 students but before doing this, piloting of the questionnaire was done with 45 test takers to check the reliability and the validity of it.

The items in the questionnaire can be divided into three parts which are mostly related with feelings of the test takers, system, and skill development.

Table 3.2. *The item numbers for each part*

FEELINGS	1, 7, 9, 19, 20, 21
SYSTEM	2, 5, 10, 12, 14, 17, 18,
SKILL	3, 4, 6, 8, 11, 13, 15, 16

Table 3.3. *Categories and Items in the Questionnaire*

A. FEELINGS	B. SYSTEM	C. SKILL
1. Online speaking assessment system makes me feel less nervous while using the target language.	2. Online speaking assessment system is practical compared to face to face form of speaking assessment.	3. Online speaking assessment system helps to develop my academic speaking skills.
7. Online speaking assessment helps me to gain my self-confidence in speaking.	5. Online speaking assessment is contemporary.	4. Online speaking assessment system helps to develop my daily speaking skills.
9. Online speaking assessment system is motivating.	10. Online speaking assessment system is fair.	6. Online speaking assessment improves my speaking skills through the systematically prepared activities.
19. I get nervous while taking the online speaking assessment system.	12. Cheating is difficult.	8. Online speaking assessment system doesn't help me to improve my speaking skills.
20. I would like to participate in future online assessment activities.	14. The activities we have done in this system are engaging.	11. Online speaking assessment system encourages speak English more in class.
21. I would like to recommend it to the other students.	17. Feedback given by the examiner provides beneficial information about my progress.	13. I do more speaking practice thanks to online speaking assessment system.
	18. The activities I do in online speaking assessment is consistent with the things I learn in class.	15. I have more opportunities to speak English through online speaking assessment system.
		16. I feel that my oral communication skills have improved.

The questionnaire was also designed to get information about the students' perceptions towards the effect of the test on their speaking and listening competency and had three parts. In the first part there were three subcategories (Computer Familiarity, Prior Experiences, and System Use) that aimed to gather demographic information about the students. To get information about the test takers' computer familiarity five items were asked, to check whether they had experienced online assessment before four items were asked and finally to get an insight about what they thought about the system use four items were used. In the second part, 21 items were prepared to find out students perceptions and attitudes towards the test. Some items were similar, some were repetitive. In the third part, two different subcategories, (one for speaking that consists of nine items, the other for listening that consists of six items) were added to find out what students opinions were about the effect of this system on their speaking and listening ability.

The questionnaire that lasted nearly 20 minutes in each class was administered in the students' classrooms during regular class time. The students were informed beforehand and assured that the data gathered from them would be treated as completely confidential, and the results would be used to improve the system.

### **3.3.3.2. Interview with students.**

In addition to the questionnaire that is applied to the test takers, semi-structured interviews were also conducted at the end of the quarter with some of the volunteer students who had something to say about the system. 24 students took part in the interview process and they stated their opinions about the effectiveness of this system. The interviews were audio-recorded and then analyzed.

### **3.3.3.3. Interview with test makers, instructors, and administrators.**

As a follow-up to the interviews conducted with students, interviews were also applied to the test makers', administrators and instructors to find out what they thought about the effectiveness of this system for the students and how they thought this newly adopted speaking assessment system would affect students' speaking and listening competency.

### **3.3.4. Data Analysis**

In this study quantitative data were analyzed with SPSS 20.0. In this study as the questions did not show a normal distribution, non-parametric tests such as Chi-square and Mann Whitney U were applied when examining the relationship between the questions. In order to analyze the data statistically and to see whether there was a relationship between the data Chi-square test was used. Chi-square test was used to see if the comments were valid from a statistical point of view. Mann Whitney U test was also used to compare the data among the groups.

Data were analyzed through pattern-coding. "Pattern-coding is a method of grouping large number of texts into small numbers of sets or themes" (Miles & Huberman, 1994, p. 69). Once all the responses were analyzed and coded, the similar codes were grouped together and some sample responses were given under the related results of the items from the questionnaire. The responses of the test takers were grouped under the feeling, system and skill categories and some of them are as follows:



Table 3.4. *The Sample Responses of the Test-takers*

Feeling	System	Skill
<i>In the lessons, I am a shy student, here I can speak. I speak even more as my teacher corrects my mistakes. (P.6)</i>	<i>Some of the topics are not contemporary. Instead of talking about the same topics in the book we can talk about daily life situations. (P.11)</i>	<i>As we are communicating with native speakers, we can hear the original language. This helped me a lot to improve my language skills in daily usage. ( P.8)</i>
<i>If it had been a face to face speaking exam, I would have been much more stressful. However, online speaking assessment was not a stressful activity for me. (P.16)</i>	<i>The lessons were more to do with Grammar instead of Speaking. (P.13)</i>	<i>It is really hard to speak English outside the school but this system enabled us to communicate using the target language and learn new cultures. It was fun. (P.12)</i>
<i>In the past it was difficult, I was shaking, but as the time went by, I felt more and more confident. (p.17)</i>	<i>Academic English and academic terms can also be taught in this assessments. (P.24)</i>	<i>I began to think my answers in English instead of Turkish and I've also developed my pronunciation. (P.2)</i>

Therefore, the data were grouped according to the responses given by the interviewees and smaller number of subcategories were organized. Subcategories such as “benefits and advantages regarding the assessment”, or “shortcomings or problems attached to this assessment” were generated and the sample responses were also added under each subcategories.

In order to calculate reliability, Cronbach Alpha ( $\alpha$ ) analysis must be done. The  $\alpha$  value for all questions shows the total reliability of the questionnaire. If the  $\alpha$  value is above 0.8, the survey shows that it has high reliability. In this hypothesis, the  $\alpha$  value of the 36 items for 291 students was calculated on R. And the  $\alpha$  value is 0.83. This shows that the reliability of the survey is high.

## CHAPTER IV

### RESULTS

#### 4.1. Introduction

This study aimed to provide information about test takers' perceptions and attitudes to the online speaking assessment. Another aim of the study was to find out how this system had helped the test takers to develop their speaking and listening skills. This study also examined what test makers, school administrators and instructors' thoughts were about online speaking assessment. The findings will be handled according to the research questions.

With this aim in mind, this study attempted to find answers to the following questions:

- 1) What are the test takers' perceptions and attitudes towards online speaking assessment system?
  - a) What are the test takers' perceptions and attitudes towards the effect of this system on their speaking competency?
  - b) What are the test takers' perceptions and attitudes towards the effect of this system on their listening competency?
- 2) What do the test makers, instructors and the school administrators think about the system and its effect on the test takers' development?

As a result of data analysis, each finding was discussed in the light of research questions.

However, first of all, we would like to present some information about the participants regarding the issues as computer familiarity, ability to use computer, ability to read computer screen, ability to solve problems related to computer, prior experiences of the test-takers on online speaking examinations, and system use.

Table 4.1. *Computer Familiarity*

Computer Familiarity N=291	Advanced	Good	Average	Poor	None
	P. %	P. %	P. %	P. %	P. %
1.Your ability to use a computer	24.7	48.1	25.8	1.4	0.0
2.Your ability to read a computer screen	45.0	47.1	7.2	0.7	0.0
3.Your ability to understand the things related to computer	24.4	43.3	28.2	4.1	0.0
4.Your ability to solve the problems if you are stuck while using a computer	15.1	36.1	38.1	9.3	1.4
5.Your ability to chat online	36.1	45.7	16.2	1.7	0.3

When the computer familiarity of the students was examined, most of the students were good at using computer (72.8%), reading a computer screen (92.1%), understanding things related to computer (67.7%), solving the problems when stuck while using a computer (51.2 %) and chatting online (81.8%).

Table 4.2. *Computer Familiarity with Gender Differences*

Computer Familiarity	FEMALE (N=157)		MALE (N=129)	
	Advanced	Average	Advanced	Average
	P. %	P. %	P. %	P. %
1.Your ability to use a computer	35.7	72.6	64.3	27.4
2.Your ability to read a computer screen	45.7	61.9	54.3	38.1
3.Your ability to understand the things related to computer	33.3	67.9	66.7	32.1
4.Your ability to solve the problems if you are stuck while using a computer	27.9	69.4	72.1	30.6

The computer familiarity of male test takers were seen as advanced when compared to the female test takers. When the gender differences were examined in terms of the computer familiarity, it was concluded that male students were more capable of dealing with things related to computer. For instance, while the percentage for understanding things related to computer for male students were 66.7%, it was 33.3% for the female ones. In addition, when the Table 4.2. was examined, it could be seen that the ability of male students solving the problems when stuck was much better than the female students. On the other hand, the results showed that with many of the abilities related to computer familiarity, female students were average.

Table 4.3. *Prior Experiences of the Test Takers*

Prior Experiences N=290	Yes	No
	P. %	P. %
1.I have tests/examinations online before	46.2	53.8
2.I have attended an online course before	13.1	86.9
3.I have taken such kind of online speaking assessments before	4.5	95.5
4.I have used web for instructional purposes	82.3	17.7

When the test takers' prior experiences were examined, it could be seen that 96% of the students didn't take such kind of online speaking assessment at all. However, 82% of them had used web for instructional purposes (Table 4.3).

Table 4.4. *System Use (with gender differences)*

System Use N=291	S.D.	D.	P.A.	A.	S.A.
	P. (%)	P. (%)	P. (%)	P. (%)	P. (%)
1.I follow the instructions without any problem	2.1	7.6	31.7	36.2	22.4
2.It is easy to register the system	3.8	7.2	17.9	41.0	30.0
3.It is easy to take the exam	3.8	11.4	25.5	34.8	24.5
4.It is comfortable to use the system	3.1	8.6	21.7	38.3	28.3

When the data related to the system use was examined, it could be inferred from Table 4.4. that nearly 60% of the test takers thought that it was easy to take the exam and register the system. Similarly, 67% of them pointed that the system was comfortable to use.

Table 4.5. *The Easiness of the System (with gender differences)*

Gender	It is easy to take the exam	
	Strongly Agree	Agree
Female N=156	P. (%)	
	81.8	43.0
Male N=129	P. (%)	
	18.2	57.0
TOTAL	P. (%)	
	100.0	100.0

82% of the female students strongly disagreed with the statement "It is easy to take them exam" while 57% of male students thought it was easy to take the exam (Table 4.5).

#### 4.2. What Are The Test Takers' Perceptions And Attitudes Towards Online Speaking Assessment System?

The questions in the questionnaire are divided into three parts which are mostly related with feelings of the test takers, system, and skill development.

Table 4.6. *The Item Numbers for Each Part*

FEELINGS	1, 7, 9, 19, 20, 21
SYSTEM	2, 5, 10, 12, 14, 17, 18,
SKILL	3, 4, 6, 8, 11, 13, 15, 16

#### 4.2.1. Feelings

Table 4.7. *Test-takers Feelings about Online Speaking Assessment*

Items	Mean	Sd	S.D %	D. %	P.A. %	A. %	S.A %
7. Online speaking assessment helps me to gain my self-confidence in speaking.	3.95	1.04	3.1	7.7	15.3	38.7	35.2
1. Online speaking assessment system makes me feel less nervous while using the target language.	3.63	1.14	5.5	10.3	26.2	31.4	26.6
9. Online speaking assessment system is motivating.	3.49	1.10	4.2	15.3	28.5	31.9	20.1
21. I would like to recommend it to the other students.	3.40	1.29	13.8	6.9	27.2	29.7	22.4
20. I would like to participate in future online assessment activities.	2.97	1.36	20.8	16.3	24.2	22.8	15.8
19. I get nervous while taking the online speaking assessment system.	2.91	1.27	16.3	23.5	27.0	19.4	13.8

These results provided supportive evidence that students had a positive perceptions about the effect of this assessment on their feelings about speaking in target language. In particular, students strongly agreed with the idea that online speaking assessment helped them to gain self-confidence in speaking. Table 4.7 shows that 73.9% of the students thought that they gained self confidence in speaking with this system. When the percentage of the students who attended such kind of online test before is taken, we saw that 96% of them didn't experience online examination. Nevertheless, attending an online assessment helped them to gain self confidence in using the target language as it can be seen in Item 7 ( $M=3.95$ ,  $SD= 1.04$ ). In addition, in item 19 ( $M=2.91$ ,  $SD=1.27$ ), they disagreed that they got nervous while taking the online assessment.

After the semi-structured interviews conducted with 24 students, majority of them stated their positive feelings related to the system. In the light of the responses given during the interviews, several different ideas and categories had emerged. Some of the opinions were related to the feelings of the test takers. Majority of the test takers shared their reflection about how they felt during the implementation of the assessment:

*In the lessons, I am a shy student, here I can speak. I speak even more as my teacher corrects my mistakes. (P.6)*

*If it had been a face to face speaking exam, I would have been much more stressful. However, online speaking assessment was not a stressful activity for me. (P.16)*



Table 4.9. *The Relationship between Chatting Online and Feeling Nervous*

Your ability to chat online		Online speaking assessment system makes me feel less nervous while using the target language					TOTAL
		SD	D	PA	A	SA	
Advanced	Fr.	4	8	19	36	37	104
	P. (%)	3.8%	7.7%	18.3%	34.6%	35.6%	100.0
Good	Fr.	7	14	42	39	31	133
	P. (%)	5.3%	10.5%	31.6%	29.3%	23.3%	100.0
Average	Fr.	5	5	12	16	9	47
	P. (%)	10.6%	10.6%	25.5%	34.0%	19.1%	100.0
Poor	Fr.	0	3	2	0	0	5
	P. (%)	0.0%	60.0%	40.0%	0.0%	0.0%	100.0
None	Fr.	0	0	1	0	0	1
	P. (%)	0.0%	0.0%	100.0%	0.0%	0.0%	100.0
TOTAL	Fr.	16	30	76	91	77	290
	P. (%)	5.5%	10.3%	26.2%	31.4%	26.6%	100.0

$\chi^2=31.053$  sd=16 p= 0.013

Furthermore, when the relationship between feeling less nervous while using the target language and the students ability to chat online was analyzed with Chi-Square tests, it could easily be seen in Table 4.9 that the ones (70.2%) who were advanced in chatting online feel less nervous while the poor ones (60%) in terms of chatting online still felt nervous. This shows that computer familiarity may affect the level of anxiety among the students.

System use was also effective for students in terms of feeling anxiety. That is, after the data relating to the easiness of the system's use and the test takers' anxiety level was examined, it was found that 78.8% of the test takers who strongly agreed and 59% of the test takers who agreed that it was easy to take the exam did not feel nervous (Table 4.10). When asked what they liked about this method of evaluation, more or less similar ideas were presented. In addition to this, the relationship between feeling less nervous and system use showed that the test takers thought that they felt less nervous while using the target language as they thought it was easy to follow the instruction and comfortable to register and use the system. The easiness of the system naturally affected the level of anxiety in using the target language. It means that anxiety level may be affected by the external reasons such as the system use. The harder it gets to use the system, the more the test takers get nervous.

Table 4.10. *The Relationship between the Easiness of Taking the Exam and Feeling Nervous*

It is easy to take the exam		Online speaking assessment system makes me feel less nervous while using the target language					TOTAL
		SD	D	PA	A	SA	
SD	Fr.	2	4	2	1	2	11
	P. (%)	18.2%	36.4%	18.2%	9.1%	18.2%	100.0%
D	Fr.	5	9	8	8	3	33
	P. (%)	15.2%	27.3%	24.2%	24.2%	9.1%	100.0%
PA	Fr.	3	8	25	26	12	74
	P. (%)	4.1%	10.8%	33.8%	35.1%	16.2%	100.0%
A	Fr.	3	6	32	41	18	100
	P. (%)	3.0%	6.0%	32.0%	41.0%	18.0%	100.0%
SA	Fr.	3	3	9	15	41	71
	P. (%)	4.2%	4.2%	12.7%	21.1%	57.7%	100.0%
TOTAL	Fr.	16	30	76	91	76	289
	P. (%)	5.5%	10.4%	26.3%	31.5%	26.3%	100.0%

$\chi^2=83.274$  sd=16 p= 0.000

Table 4.11. *Test-takers' Opinions about Feeling Nervous (gender differences)*

Variable	Gender	Mean Rank	U	Z	P
Online speaking assessment system makes me feel less nervous while using the target language	Female	134.61	8730.00	-1.974	0.048
	Male	153.30			

According to the non-parametric test (Mann-Whitney U) results, it could also be said that male students felt less nervous while using the target language (Table 4.11). The result is statistically significant ( $p < .048$ ).

Table 4.12. *Test-takers' Opinions about Feeling Nervous (prior experiences)*

Variable	Prior Experiences	Mean Rank	U	Z	P
Online speaking assessment system makes me feel less nervous while using the target language	Yes	156.61	8843.000	-2.256	0.024
	No	135.05			

Finally, when the relationship between the Item 1 and prior experiences was analyzed (Table 4.12), it was clear that the ones who took online tests or examinations before felt less nervous with the help of this assessment compared to those who had no experience before, and the result is statistically significant ( $p < .024$ ).

More than half of the test takers (52%) stated that online speaking assessment system was motivating in Item 9 ( $M=3.49$ ,  $SD=1.10$ ) (Table 4.7).



Table 4.13. *Test-takers' Opinions about Online Speaking Assessment (motivation)*

Variable	Prior Experiences	Mean Rank	U	Z	P
Online speaking assessment system is motivating	Yes	155.38	8726.500	-2.250	0.024
	No	134.04			

According to Mann-Whitney U results (Table 4.13), the ones who took online tests or examinations before thought that online speaking assessment was motivating compared to compared to those who had no experience before, and the result is statistically significant ( $p < .024$ ).

Table 4.14. *The Relationship between the Comfortable Use of the System and Recommending to Others*

It is comfortable to use the system		I would like to recommend it to other students					TOTAL
		SD	D	PA	A	SA	
SD	Fr.	5	1	1	1	1	9
	P. (%)	55.6%	11.1%	11.1%	11.1%	11.1%	100.0%
D	Fr.	11	1	8	2	3	25
	P. (%)	44.0%	4.0%	32.0%	8.0%	12.0%	100.0%
PA	Fr.	13	7	25	11	6	62
	P. (%)	21.0%	11.3%	40.3%	17.7%	9.7%	100.0%
A	Fr.	6	8	32	47	18	111
	P. (%)	5.4%	7.2%	28.8%	42.3%	16.2%	100.0%
SA	Fr.	5	3	13	25	36	82
	P. (%)	6.1%	3.7%	15.9%	30.5%	43.9%	100.0%
TOTAL	Fr.	40	20	79	86	64	289
	P. (%)	13.8%	6.9%	27.3%	29.8%	22.1%	100.0%

$\chi^2=90.994$  sd=16 p= 0.000

In addition, in item 21 ( $M=3.40$ ,  $SD=1.29$ ), the test takers stated that they wanted to recommend the system to the others. According to the result obtained from the questionnaire, more than half of the test takers (52.1%) would like to recommend the system to the other students (Table 4.7). When the system use is applicable and reasonable to use, it is almost impossible for the users not to recommend it. This opinion could be supported because a majority of the test takers (74.4%) thought that it was comfortable to use the system, as a result they would like to recommend it to the others as well (Table 4.14). It can be stated that when the test takers have no trouble following the instructions while being tested, they may have positive feelings about attending future online speaking assessment activities.

Table 4.15. *The Relationship between Computer Ability and Recommending to Others*

Your ability to use a computer		I would like to recommend it to other students					TOTAL
		SD	D	PA	A	SA	
Advanced	Fr.	12	5	20	14	20	71
	P. (%)	16.9%	7.0%	28.2%	19.7%	28.2%	100.0
Good	Fr.	23	6	38	43	30	140
	P. (%)	16.4%	4.3%	27.1%	30.7%	21.4%	100.0
Average	Fr.	4	7	21	28	15	75
	P. (%)	5.3%	9.3%	28.0%	37.3%	20.0%	100.0
Poor	Fr.	1	2	0	1	0	4
	P. (%)	25.0%	50.0%	0.0%	25.0%	0.0%	100.0
TOTAL	Fr.	40	20	79	86	65	290
	P. (%)	13.8%	6.9%	27.2%	29.7%	22.4%	100.0

$\chi^2=25.533$  sd=12 p= 0.012

In addition, when the test takers' computer ability was taken into account, the ones who were good (52.1%) or advanced (47.9%) in using a computer stated that they would like to recommend it to the others (Table 4.15). Therefore, it can be understood that computer literacy may affect the students' ideas about recommending the online speaking assessment system to the others positively.

Moreover, when the data was examined to see whether the test takers were willing to participate future online speaking activities, for the item 20, 38.6 % of them stated that they wanted to participate in future online speaking activities while 37.1% disagreed with that idea. (Table 4.7). (M=2.91 SD=1.36)

Table 4.16. *Test-takers' Opinions about Participating Future Online Assessments (prior experiences)*

Variable	Prior Experiences	Mean Rank	U	Z	P
I would like to participate in future online assessment activities	Yes	192.35	1165.500	-2.167	0.030
	No	142.24			

Finally, according to Mann-Whitney U results, there was a relationship between students' ideas about participating future assessment activities and their prior experiences (Table 4.16). That is, the ones who attended such kind of online assessment before would like to participate in future online assessment activities compared to those who had no experience before, and the result is statistically significant ( $p<.030$ ).

#### 4.2.2. System

As a result of data analysis, 76.9% of the test takers indicated that online speaking form of the assessment was contemporary (Table 4.17). In general, the mean score of item 5 (M=4.05, SD=.93) indicated that students found online speaking assessment system contemporary.

Table 4.17. *Test-takers' Opinions about the System of Online Speaking Assessment*

Items	Mean	Sd	S.D %	D. %	P.A. %	A. %	S.A %
5. Online speaking assessment is contemporary.	4.05	.93	2.1	3.5	17.5	40.9	36.0
18.The activities I do in online speaking assessment is consistent with the things I learn in class.	3.98	1.01	3.5	4.2	18.9	37.9	35.4
2. Online speaking assessment system is practical compared to face to face form of speaking assessment.	3.68	1.25	6.9	13.1	18.6	27.9	33.4
17. Feedback given by the examiner provides beneficial information about my progress.	3.44	1.21	9.0	13.1	24.1	32.8	21.0
12.Cheating is difficult.	3.00	1.48	25.3	12.1	22.8	17.3	22.5
14.The activities we have done in this system are engaging.	2.80	1.22	18.5	21.6	31.0	19.2	9.8
10.Online speaking assessment system is fair.	2.56	1.32	30.2	19.1	25.0	16.3	9.4

In order to find the reasons why students thought that it was a contemporary way of assessing speaking, the items and their relationships were deeply analyzed. In the correlation test, we saw that 87.7% of the test takers who followed the instructions without any problem stated that online speaking assessment system was contemporary (Table 4.18).

Table 4.18. *The Relationship between Following the Instructions and Being Contemporary*

I follow the instructions without any problem		Online speaking assessment is contemporary					TOTAL
		SD	D	PA	A	SA	
SD	Fr.	0	0	1	2	3	6
	P. (%)	0.0%	0.0%	16.7%	33.3%	50.0%	100.0%
D	Fr.	1	4	4	11	2	22
	P. (%)	4.5%	18.2%	18.2%	50.0%	9.1%	100.0%
PA	Fr.	1	2	26	39	20	88
	P. (%)	1.1%	2.3%	29.5%	44.3%	22.7%	100.0%
A	Fr.	0	4	15	47	38	104
	P. (%)	0.0%	3.8%	14.4%	45.2%	36.5%	100.0%
SA	Fr.	4	0	4	18	39	65
	P. (%)	6.2%	0.0%	6.2%	27.7%	60.0%	100.0%
TOTAL	Fr.	6	10	50	117	102	285
	P. (%)	2.1%	3.5%	17.5%	41.1%	35.8%	100.0%

$\chi^2=60.998$  sd=16 p= 0.000



success in assessments, the test activities should have a relationship with the classroom activities (Table 4.17). A positive washback effect occurs when the assessment procedures meet the course goals and objectives.

Table 4.20. *The Relationship between the Easiness of the System and Being More Practical than Face-to-face Exams*

It is easy to take the exam		Online speaking assessment system is practical compared to face to face form of speaking assessment					TOTAL
		SD	D	PA	A	SA	
SD	Fr.	1	1	2	3	4	11
	P. (%)	9.1%	9.1%	18.2%	27.3%	36.4%	100.0%
D	Fr.	8	8	5	8	4	33
	P. (%)	24.2%	24.2%	15.2%	24.2%	12.1%	100.0%
PA	Fr.	7	10	18	18	21	74
	P. (%)	9.5%	13.5%	24.3%	24.3%	28.4%	100.0%
A	Fr.	2	14	23	36	25	100
	P. (%)	2.0%	14.0%	23.0%	36.0%	25.0%	100.0%
SA	Fr.	2	4	6	16	43	71
	P. (%)	2.8%	5.6%	8.5%	22.5%	60.6%	100.0%
TOTAL	Fr.	20	37	54	81	97	289
	P. (%)	6.9%	12.8%	18.7%	28.0%	33.6%	100.0%

$\chi^2=59.549$  sd=16 p= 0.000

Table 4.21. *The Relationship between the Comfortable Use of the System and Being More Practical than Face-to-face Exams*

It is comfortable to use the system		Online speaking assessment system is practical compared to face to face form of speaking assessment					TOTAL
		SD	D	PA	A	SA	
SD	Fr.	2	1	2	1	3	9
	P. (%)	22.2%	11.1%	22.2%	11.1%	33.3%	100.0%
D	Fr.	5	7	4	7	2	25
	P. (%)	20.0%	28.0%	16.0%	28.0%	8.0%	100.0%
PA	Fr.	10	10	14	15	13	62
	P. (%)	16.1%	16.1%	22.6%	24.2%	21.0%	100.0%
A	Fr.	1	14	24	37	35	111
	P. (%)	0.9%	12.6%	21.6%	33.3%	31.5%	100.0%
SA	Fr.	2	5	10	21	44	82
	P. (%)	2.4%	6.1%	12.2%	25.6%	53.7%	100.0%
TOTAL	Fr.	20	37	54	81	97	289
	P. (%)	6.9%	12.8%	18.7%	28.0%	33.6%	100.0%

$\chi^2=56.314$  sd=16 p= 0.000

In Table 4.17 for the item 2, 61.3% of the students thought that online form of speaking assessment was more practical than the face to face form of it (M=3.68, SD=1.25). In the correlation test, we saw that many of them (83.1%) who indicated that it was easy to take the exam and 79.3% of the students who said that this system was comfortable to use

said that online form of speaking assessment was more practical (Tables 4.20 and 4.21). It is clear that the effectiveness of the system also affects the test takers' ideas in a positive way. When all the system including the instruction, registration (etc...) are easy and comfortable to use, the system gets practical and easy to use when compared to the old-fashioned form of the assessment strategies.

On the other hand, when students' ideas about the Item 2 (online speaking assessment is practical compared to face to face form of speaking assessment) were compared with their ability to read computer screen, it was seen that nearly half of the students (47.6%) who saw their talents as average in terms of reading a computer screen declared that it was not practical when compared to face to face form of it (Table 4.22). In the light of our data, we can state that when a student is capable of doing the things related to computer or has some talent in using some technology, it is not hard for him/her to get used to the newly adopted assessment systems. The computer literacy level of the test takers may have a bad or good effect on their ideas about the online speaking assessment system.

Table 4.22. *The Relationship between Reading a Computer Screen and Being More Practical than Face-to-face Exams*

Your ability to read a computer screen		Online speaking assessment system is practical compared to face to face form of speaking assessment					TOTAL
		SD	D	PA	A	SA	
Advanced	Fr.	5	14	20	40	51	130
	P. (%)	3.8%	10.8%	15.4%	30.8%	39.2%	100.0%
Good	Fr.	13	16	31	35	42	137
	P. (%)	9.5%	11.7%	22.6%	25.5%	30.7%	100.0%
Average	Fr.	2	8	3	5	3	21
	P. (%)	9.5%	38.1%	14.3%	23.8%	14.3%	100.0%
Poor	Fr.	0	0	0	1	1	2
	P. (%)	0.0%	0.0%	0.0%	50.0%	50.0%	100.0%
TOTAL	Fr.	20	38	54	81	97	290
	P. (%)	6.9%	13.1%	18.6%	27.9%	33.4%	100.0%

$\chi^2=22.177$  sd=12 p= 0.036

More than half of the test takers (54%) stated that the feedback provided by the examiner included beneficial information about the students' progress (Table 4.17). Furthermore, when the results of interviews were examined, a majority of the preparatory school students reported that most of their teachers' feedback contributed a lot to their language skills especially in terms of speaking:

*The thing that I really like about the system is the feedback given by the teachers after the lessons. I love it because it is realistic and not formal. (P. 14)*

Another noteworthy point that was stated by the students was the teachers' helpful, intimate and friendly attitudes. According to students, examiners' attitudes and behaviours were important because they could get the chance to learn some language tricks such as how to respond some utterances, or which vocabulary to use in some situations. Some students also stated that feedback given by the teachers made a great difference in their language development to see their strengths and weaknesses:

*I really like this system because my teachers are so friendly. I've improved my speaking a lot since the beginning of the term. (P.5)*

*I feel comfortable in this system as my teachers are relaxed and it affects my speaking positively. (P.2)*

However, unlike majority of the students who found the examiners really helpful and understanding, some of the students complained about the examiners' attitudes towards them. They stated that some of the examiners' uncaring and indifferent manners caused distrust towards the system. Another thing that the students mentioned as a challenge was the different implementations of different examiners which was something that affected students' motivation negatively:

*Some teachers have no idea about what the others are doing. There is no standardization in the activities. (P.11)*

*Some teachers didn't even ask me how I am. There was no real life communication. (P.13)*

*There should be a kind of in-service assessment system for the examiners to make the things same and standard for everybody. (P.1)*

While 37.4% of the students thought that cheating was easy in this assessment, nearly 39.8% of them stated that it was not easy according to Table 4.17 (M=3.00 SD=1.48). In addition, it could be said that the ones who were really good at solving the problems when stuck while using a computer thought that cheating was easy. On the other hand, cheating was difficult for the students who were poor in solving the problems related to computers.

When the data were examined, it was found out 40.1% of the test takers thought that the topics and the activities were not engaging (M=2.80 SD=1.22). Only 29% stated that the activities were engaging (Table 4.17).

*The teachers always used the same slides. Talking about the same things again and again was boring. (P. 15)*

When the test takers were asked about whether the test was fair, nearly half of them (49.3%) reported that it was not fair (Table 4.17). Only 25.7% were positive about the

fairness issue. The results of the questionnaires were confirmed in the interview session and the ones, especially those who had concerns about the grading system reported that there was a problem related to the fairness issue. Similarly, some students worried that unfairness among the examiners might be a disadvantage because this system was a determining factor that had a big influence on students' pass/fail. Nearly half of them stated that the grades differed from one teacher to another. According to the students, this was something that affected their motivation negatively:

*Grading system differed from one teacher to another. I never understood the reason why. (P. 24)*

*I managed to add only one point to my general score at the end of each quarter. My score was 85 in A2, then it became 86 in B1 and I suppose it is going to be 87 this quarter. I never understood the grading system. (P. 15)*

*The grading system was not fair. (P. 17)*

Finally, when the test takers stated their opinions about the effectiveness of system use, a majority of them frequently referred to the easy and convenient use of the system:

*The interface was easy and convenient to use. (P.6)*

*The system is practical and easy to use. (P.7)*

*Even a person who has no idea about computers can use and understand this system. (P. 2)*

On the other hand, while many students thought that experiencing such an online speaking assessment offered several benefits, there were also some concerns emerged from the most frequent responses given by the test takers related to the challenges of the system. Most B2 students mentioned that the time allocated for the assessment wasn't enough for them to fully cover the things that they are supposed to cover. Students reported that although the time was enough for the beginning levels (A1 A2), it became harder and harder for them to be able to do all the activities in B1 and B1+ because of the limited time. They said that the activities and the tasks became more difficult but the allocated time stayed same:

*There was a problem related to time allocated for the assessment. It should have been 2 hours per week instead of half an hour. (P.8)*

*We start to get used to the activities and tasks after 10 minutes but then the time is up! We do not understand what we did, what I said. (P. 11)*

*Instead of having the assessment once a week, it should have been twice a week. (P. 18)*

Another source of difficulty asserted by some of the test takers was technical and environmental problems. Some of them reported having experienced difficulties in managing the problems related to internet connection, headphones or camera.

*The connection was a big problem. (P.15)*



*Some slides appear on the screen late, then I cannot catch the activity and it distracts my attention.*

*(P. 19)*

*The screen freezes and we start to lose time. (P. 1)*

*I experienced image and sound problems. (P.3)*

### 4.2.3. Skill

Table 4.23. *Test-takers' Perceptions about the Effects of Online Speaking Assessment on their Skills*

Items	Mean	Sd	S.D %	D. %	P.A. %	A. %	S.A %
4. Online speaking assessment system helps to develop my daily speaking skills.	3.96	1.02	2.4	6.9	18.6	36.2	35.9
15. I have more opportunities to speak English through online speaking assessment system.	3.93	1.03	3.5	7.0	15.0	42.0	32.5
13. I do more speaking practice thanks to online speaking assessment system.	3.88	1.07	3.5	8.0	19.1	36.5	33.0
16. I feel that my oral communication skills have improved.	3.79	1.08	3.5	9.4	21.5	35.8	29.9
6. Online speaking assessment improves my speaking skills through the systematically prepared activities.	3.74	1.03	2.1	9.3	28.7	31.8	28.0
3. Online speaking assessment system helps to develop my academic speaking skills.	3.51	1.17	8.0	9.3	28.7	31.5	22.5
11. Online speaking assessment system encourages speak English more in class.	3.44	1.10	5.2	13.5	31.8	30.8	18.7
8. Online speaking assessment system doesn't help me to improve my speaking skills.	2.21	1.30	40.8	24.9	15.2	10.7	8.3

The positive attitudes could be seen in item 4 about the effect of online speaking system on the students' daily speaking skills. (M=3.96 SD= 1.02). Majority of the test takers (72.1%) stated that the system had helped them to develop their daily speaking skills (Table 4.23). Furthermore, in the interviews students stated that thanks to continuous exposure to English every week, they were able to develop their speaking skills, daily spoken English, pronunciation skills and vocabulary knowledge. Most of them were also happy as they were using the language communicatively. In addition, majority of the prep school students stated that thanks to the online speaking assessment system, they got the chance to speak with native speakers, which was something that they found really fruitful to improve their speaking skills. This system also enabled them to learn other cultures by communicating with someone from a different country:

*As we are communicating with native speakers, we can hear the original language. This helped me a lot to improve my language skills in daily usage. ( P.8)*

In Table 4.23, for the item 15 majority of the test takers (74.5%) thought that they had more opportunities to speak English through this system (M=3.93 SD=1.03). In the

interview sessions one of the students' statement about the effect of this system on communicating in the target language was as below:

*It is really hard to speak English outside the school but this system enabled us to communicate using the target language and learn new cultures. It was fun. (P.12)*

When we look at the Table 4.23, we saw that majority of the test takers 69.5% thought that online form of speaking assessment made them practice English more ( $M=3.88$   $SD=1.07$ ). The data obtained from the second part of the questionnaire had similarities with the data obtained from the 'system use' part of the questionnaire. One of the students' statement below from the interview session supports this:

*I began to think my answers in English instead of Turkish and I've also developed my pronunciation. (P.2)*

According to the result obtained from the questionnaire, more than half of the test takers (59.8%) stated that their speaking skills improved through systematically prepared activities with a  $M=3.74$  and  $SD=1.03$ . (Table 4.23). The activities prepared with the help of technology engaged students more in the lesson and gave them a chance of developing the skills.

In addition, as for the item 3 the the results  $M=3.51$  and  $SD=1.17$  showed that more than half of the students found the system helpful for their academic speaking skills (Table 4.23). 54% of them thought that online form of assessment helped them to develop academic skills as well.

According to Table 4.23, nearly half of the test takers 49.5% thought that online form of speaking assessment encouraged them to speak English in class ( $M=3.44$   $SD=1.10$ ). It is clear that the effectiveness of the system also affect the test takers' ideas in a positive way. When all the system including the instruction, registration (etc...) are easy and comfortable to use, the system affects the learners positively and promotes them to use the target language more in class environment.

Table 4.24. *Test-takers' Opinions about Speaking in Class (prior experiences)*

Variable	Prior Experiences	Mean Rank	U	Z	P
Online speaking assessment system encourages students to speak English more in class	Yes	155.96	8783.000	-2.259	0.024
	No	134.53			

Furthermore, the relationship between the item 11 and prior experiences was examined, the ones who took online tests or examinations before agreed that online speaking assessment encouraged to speak English more in class (Table 4.24). That is, the ones who attended such kind of online assessment before thought that online speaking assessment system encouraged them to speak English more in class, and the result is statistically significant ( $p < .024$ ).

When the data were examined, it was found out that 65.7% of the test takers disagreed that online speaking assessment did not help them to develop their speaking skills (Table 4.23). Similarly, again 65.7% of them stated that their oral communication skills improved thanks to this system. Providing authentic materials and giving the students a chance to communicate in a real environment is useful as it promotes their speaking skills.

Table 4.25. *The Relationship between Attending Online Course Before and Improving Speaking Skills*

I have attended an online course before		Online speaking assessment system does not help me improve my speaking skills					TOTAL
		SD	D	PA	A	SA	
Yes	Fr.	14	10	2	9	3	38
	P. (%)	36.8%	26.3%	5.3%	23.7%	7.9%	100.0%
No	Fr.	104	62	42	21	21	250
	P. (%)	41.6%	24.8%	16.8%	8.4%	8.4%	100.0%
TOTAL	Fr.	118	72	44	30	24	288
	P. (%)	41.0%	25.0%	15.3%	10.4%	8.3%	100.0%

$\chi^2=10.494$  sd=4 p= 0.033

It was a striking finding that although more than 66.4% of the participants did not attend an online course or assessment like this before, they still stated that online assessment helped them a lot to develop their speaking (Table 4.25). It can be concluded that attending an online course or assessment before do not have a big influence on the test takers' opinions about the effectiveness of the assessment for their speaking competencies.

Table 4.26. *The Relationship between Understanding Computer and Improving Speaking Skills*

Your ability to understand the things related to computer		Online speaking assessment system does not help me improve my speaking skills					TOTAL
		SD	D	PA	A	SA	
Advanced	Fr.	32	13	6	7	12	70
	P. (%)	45.7%	18.6%	8.6%	10.0%	17.1%	100.0
Good	Fr.	50	27	24	15	9	125
	P. (%)	40.0%	21.6%	19.2%	12.0%	7.2%	100.0
Average	Fr.	32	25	14	8	3	82
	P. (%)	39.0%	30.5%	17.1%	9.8%	3.7%	100.0
Poor	Fr.	4	7	0	1	0	12
	P. (%)	33.3%	58.3%	0.0%	8.3%	0.0%	100.0
TOTAL	Fr.	118	72	44	31	24	289
	P. (%)	40.8%	24.9%	15.2%	10.7%	8.3%	100.0

$\chi^2=24.343$  sd=12 p= 0.018

Finally, another important point about the data showed that no matter what the students' computer literacy level was, many of them disagreed that online speaking assessment did not help them to improve speaking skills (Table 4.26). Both 64.3% of the advanced and 91.6% of the poor students in terms of understanding the thing related to computer stated that online speaking assessment helped them to develop their speaking skills.

### 4.3. What Are the Test Takers' Perceptions and Attitudes towards the Effect of This System on their Speaking Competency?

Table 4.27. *Test-takers' Opinions about the Effects of Online Speaking Assessment on their Speaking Competency*

Items	Mean	Sd	S.D %	D. %	P.A. %	A. %	S.A %
Online speaking assessment system helps me to...							
4.state my opinion on the activity given	3.85	.95	2.5	7.0	19.0	46.5	25.0
7.start the conversation	3.69	1.03	3.5	9.8	23.2	41.1	22.5
5.carry out the activity by checking and confirming	3.63	.96	2.5	8.8	30.3	40.5	18.0
8.end the conversation	3.62	1.05	3.5	10.5	28.4	35.4	22.1
6.fill the gaps while speaking by using proper utterances	3.60	.98	3.5	9.9	25.4	45.8	15.5
1.communicate with some confidence on familiar routine matters	3.60	1.04	4.2	8.8	30.9	35.4	20.7
3.express feelings (surprise, happiness, sadness, etc..)	3.55	1.10	5.3	12.3	24.6	38.2	19.6
9.express my ideas during conversation without having difficulty	3.53	1.05	4.2	10.5	32.6	33.3	19.3
2.communicate with some confidence on familiar non-routine matters	3.22	1.08	8.1	13.7	38.2	28.4	11.6

With this system it was aimed to keep the students active and improve their speaking skills with the activities. It was understood from the data that majority of the students

(71.5%) found the system helpful to state his/her opinion on the activity given (Table 4.27). It is important for a learner to produce an opinion and give appropriate responses in communicative tests.

As for the test takers' perceptions concerning the effect of online speaking assessment system on their speaking competency, majority of the students (83%) in the interviews definitely agreed on the importance of the online speaking assessment system on their speaking skills. Most of them mentioned how developing their speaking skills was important in language learning process. However, with a few students (8%) stated that online speaking assessment system would have no effect on speaking. Another common experience that the test takers had was that many of them improved their self-esteem in speaking. In addition to these, the students reported that their skills such as critical thinking, producing sentences in a short period of time, thinking and responding in English, turn-taking, paraphrasing, understanding different accents had improved a lot. However, some of the students (8%) who had experienced problems related to grading, duration, irrelevant topics chosen in the tasks stated that the system didn't contribute to their speaking competency at all. Only one student out of 24 in the interview was not pretty sure about the effect of this system on speaking. According to CEFR (Council of Europe, 2001), B2 learners can initiate discourse, take his/her turn when appropriate and end conversation when he/she needs, and help the discussion along on familiar ground confirming comprehension. When the test takers' ideas were asked through the interview questions about the effect of this system on their speaking abilities, the results listed below were found;

*Now I'm more confident, I learned not to panic while speaking. ( P.14)*

*I've improved a lot since the beginning of the terms. Now, I can speak without hesitation. (P. 12)*

*I've made a very good progress in terms of pronunciation. (P.4)*

According to the results, 63.6% of the students stated that online speaking assessment helped them to start the conversation while 57.5% of them said it helped them to end the conversation (Table 4.27). That is, more than half of them see themselves capable of starting and ending the conversation.

On the other hand, 58.5% of them pointed out that by checking and confirming they could carry out the activity given (Table 4.27).

Again, 61.3% of them agreed that the system helped them to fill the gaps while speaking by using proper utterances.

As found out in the previous parts of the results, thanks to this assessment student gained self-confidence while using the target language and Table 4.27 showed that 56.1% of the test takers felt confident while communicating on familiar routine matters.

Expressing feelings or ideas and filling the gaps with proper utterances are some of the things that many students find difficult. With the help of this system more than half of the students (57.8%) stated that they started to be able to express their feelings during conversation (Table 4.27).

Finally, 52.6% of the students agreed on the positive effect of this assessment on their ability to express ideas during conversation without having difficulty (Table 4.27). It shows that speaking someone every week on a chosen topic makes them feel active and competent while speaking. In addition, the result showed that the test takers started to achieve the capabilities that were required to sustain the conversation.

*Since time is very tight, I am able to speak and think quickly. (P.17)*

As for communicating with some confidence on non-routine matters, while 40% of the test takers agreed, the other 38.2% of them were not sure. This shows that when the topics are chosen from real-life or routine matters, the students feel more confident.

#### **4.4. What Are the Test Takers' Perceptions and Attitudes towards the Effect of This System on their Listening Competency?**

Table 4.28. *Test-takers' Opinions about the Effects of Online Speaking Assessment on their Listening Competency*

Items	Mean	Sd	S.D %	D. %	P.A. %	A. %	S.A %
Online speaking assessment system helps me to...							
1. understand simple information provided by the examiner if it is suitable to my level	4.00	.93	1.8	5.3	17.2	42.5	33.3
2. follow the talk if the activity is familiar	3.93	.97	1.8	7.4	18.0	42.3	30.6
3. follow the talk if the activity is clearly structured	3.86	.92	1.8	4.9	25.0	42.6	25.7
5. understand general messages	3.84	.96	2.8	5.6	21.8	44.6	25.3
4. follow the main points of the conversation	3.73	.95	2.1	7.4	27.1	41.9	21.5
6. understand details	3.62	1.03	3.9	9.5	27.7	38.2	20.7

According to the results, 75.8% of the students understood the information as long as it was suitable to their level. It shows that when the information about the activity is provided after taking the students' level into consideration, there is no difficulty for the test taker to understand it.

*Now, I am more competent in understanding different accents. (P.19)*

Likewise, if the activity was familiar and clearly structured, the test takers were able to follow the talk. As could be seen from the Table 4.28, 72.9% of the students could follow the talk when the activity was familiar. However, 18% of them also partly agreed this statement. That is, if the activities are taken from non-familiar or off-topic parts, the students find it difficult to carry out the activity.

*With this system I have to catch and understand what the teachers say quickly, so this helped me a lot to improve my listening. (P.2)*

68.3% of the test takers, on the other hand, stated that the system was helpful for them to follow the talk when clearly structured activities were provided.

Finally, according to Table 4.28 more than half of the test takers (69.9%) found the system helpful for them to understand general messages or details during conversation. In addition, majority of the test takers (63.4%) stated that they were able to follow the main parts of the conversation with this system. As for the item 6, 58.9% of the participants stated that they were able to understand the details during the conversations.

When the test takers' perceptions concerning the effect of online speaking assessment system on their listening competency were analyzed in the interview sessions, it was seen that more than half of the students (54.1%) definitely agreed on the importance of the online speaking assessment system on their listening skills with the exception of a few students (33%) who stated that online speaking assessment system would have no effect on listening.

*This system is helpful for listening but not enough. (P. 11)*

*As the assessments focus mostly on the slides, there is no listening. As a result, it does not contribute a lot to the listening. (P.13)*

Most of them mentioned how this system affected their pronunciation skill as they were exposed to hear the pronunciation of the words at first hand. Another common experience was that thanks to online speaking assessment they said that they were able to understand different accents and communicate with someone who had a different accent.

*It improved my pronunciation skills via listening. (P. 5)*

In addition to these, the students reported that their skills such as understanding simple information, following the main points of the conversation had improved a lot. However, some of the students (33.3%) were negative about the effects of this system on their listening. They stated that due to the nature of the activities conducted in the

assessments, they only followed the slides and had no chance to communicate. Some of them also said that there was a mismatch between the accents of the instructors and the examiners. That is, while they were trying to understand the accents of their Turkish teachers in class, they were also supposed to forget everything in speaking assessments and try to understand Philadelphia accent. Finally, according to some of them, the reason why their listening skill did not improve at all was the noise that they were exposed to during the assessments. This noise caused them not to focus on and understand what they hear.

*There is a huge difference between the two pronunciation styles. It is complicated. (P.15)*

#### **4.5. Perceptions and Attitudes of the Test Makers, the Instructors and the School Administrators towards Online Speaking Assessments**

In this part what the test makers, the school administrators and the instructors thought about the results of the system and its effect on the test takers' development were examined.

The test makers' and instructors' ideas who implemented the program ( $N=5$ ) were also analysed, and they were asked to share their views on online speaking assessment in general. From the results it can be seen that the instructors, the test makers and the administrators have both positive and negative thoughts about the effects of this system on the test takers and on the competencies that were being tried to be developed. All of them were actually sure about the system and its good results. On the other hand, there were also some issues to be improved related to the system. In the light of the responses given to the interview questions, the results were categorized as;

- Benefits and advantages regarding the assessment,
- Shortcomings or problems attached to this assessment,
- The match between the content and level of proficiency,
- Components or areas need to be improved and the features or functions of online assessment.

##### *Benefits and advantages regarding the assessment*

First, it was stated that students seem to feel more comfortable when they speak to someone whom they do not know in real life. It was said that the online speaking assessment would, therefore, reflect their actual level of speaking, if administered properly. The system had influenced students positively about how to make use of ICT to improve their language



skills. Both the instructors and administrators shared that because examiners and students can also see each other, it is somehow close to face-to-face communication. Finally, the opportunity to reassess the assessment by listening to the recording again was declared to be one of the advantages of the system. Moreover, it was also stated it was useful that many students were assessed at the same time, on the other hand, injustice would emerge as each assessor was different and their grading was different accordingly.

#### *Shortcomings or problems attached to this assessment*

One of the instructors shared the idea that it was not beneficial for all conditions as it did not represent real life. The system was blamed for being artificial and not as functional as daily life conversations. Assessment of speaking skills was implied to be a controversial issue at the institution for several reasons. The reasons why it was controversial were shared in the following parts. First of all, weekly speaking exams had taken place after school, which placed additional burden on teachers. Second, even though two examiners were present during the exams and both marked students' performances, there were huge gaps identified between the scores given by some of those paired examiners. As a result, students made complaints about the examiners who were labelled as tough graders and they did not want them to be examiners again. They regarded this as the weakness of the system and blamed those examiners for their own failure at the end. Third, there was not a testing and assessment office, so all the assessment materials had to be prepared by level coordinators who already had a lot of other responsibilities. It was quite labour-intensive to develop those materials every week. In addition, the materials they prepared received some criticism from the teachers from time to time, which caused unnecessary conflicts. The administration shared the idea that the online speaking assessment would get rid of all these problems, and therefore, adapted it to the system.

There were also the other weaknesses of the system shared by the instructors. Because students had to take the exam at a pre-scheduled time during their lessons, some of the lessons had to be as short as 20 minutes. It had a negative influence on the effectiveness of those lessons. Furthermore, it damaged the flow of the syllabus and pacing from time to time. From the comments, it was also seen that the assessment had to take place live online, any technical issue such as internet connection breakdown was likely to cause difficulties. For example, some sessions might have to be rescheduled as a result. In addition, it was stated that as the examiners were based in another country, they were not familiar with the

context and culture of the country. Even though they attended an orientation programme at the very beginning and were also informed about the syllabus and materials to address this gap, they were considered not being aware of the objectives and expectations of the institution.

First of all, because the examiners were not familiar with the way of assessing speaking skills, it became a must to train them in the first place. However, further training seemed to be necessary to improve the quality and standardisation. Next, the instructors, the test makers and the examiners were not professionally trained to assess speaking skills. Hence, they were not like the certified examiners of internationally recognised exams such as Cambridge exams. As a consequence, some of them did not take their role seriously enough. The assessors were not seen as qualified enough to conduct this job and grade differences depending on assessors destroyed the validity and reliability of scores.

On the other hand, the company providing the online speaking assessment service was a profit-oriented company. Thus, they saw the students as customers that should always be pleased. For this reason, they tended to avoid giving low scores even if some students performed so badly. Moreover, it took a lot of time and effort to prepare the schedules and run the program smoothly. The assistant director, who already had a lot of other responsibilities, had to devote a lot of time to ensure that nothing goes wrong during the exam period. Despite all, it was not possible to monitor how each assessment was conducted. So, it was not easy to make sure that each was at a standard quality as the institution members could only check the voice recordings of randomly selected sessions and get a general idea.

Finally, because the speaking assessment was fully outsourced, the teachers in the institution could not get an idea about how students were doing in speaking, what sorts of mistakes they usually made, how well they had learnt the topics covered in the previous weeks, which would normally be helpful for formative purposes.

#### *The match between the content and level of proficiency*

When the match between the online assessment content and students' level of proficiency was asked it was understood from one of the interviewees' comments that topics were not in harmony with the curriculum or syllabus and students had difficulty from time to time. Proficiency level and speaking topics did not match each other, and it was one of the greatest issues. On the other hand, majority of the instructors asserted the content of the

online assessment was suitable for the students' level of proficiency in general. It was said to be ensured by checking the content on a regular basis. But, as a deficiency it was found out that some examiners mixed the levels up and assessed some groups of students using wrong content. But in such kind of cases, the institution reported them to the company providing the service.

*Components or areas need to be improved and the features or functions of online assessment*

Topics were stated to be chosen in accordance with curriculum, assessors should be qualified and be the ones who are educated in related field. In addition, the idea of training the examiners on how to test speaking was shared to make sure that they were well aware of how to use criteria for standardisation purposes and to be fair and give more realistic scores in the light of the criteria they were using. As a new idea it was said that a presentation part in which students talk about a predetermined topic for at least three minutes should be added to the test. The instructors who took part in the interviews stated that students should be provided with an orientation at the beginning of an academic year in which they are informed about the online speaking assessment in detail, for example, how it is administered, what they are expected to do etc. Furthermore, it was also advised to provide students a demo video which explains everything in detail and showing how to navigate through the online platform.

To promote positive washback, one of the instructors shared that teachers should employ speaking tasks similar to the ones that the students would come across in the assessment. Furthermore, the content was claimed to be improved in the light of the international standard exams such as KET, PET, FCE, IELTS, and TOEFL. On the other hand, the importance for the examiners to write weekly reports about students' performances and inform the teachers about the common mistakes they were making was also implied by the instructors and administrators. Finally, it was shared that the students could be asked to watch the recording of their performance and reflect on it.

## CHAPTER V

### DISCUSSION, CONCLUSION AND SUGGESTIONS

This chapter starts with the summary of the study which gives information about the participants and their perceptions and attitudes towards the online speaking assessment, research questions, data collection tools and procedures and all the process that had been gone through. In addition, implications of the study, suggestions for the further research and limitation of the study have been presented.

#### 5.1. Discussion

Major findings of the study will be discussed in parallel with research questions.

##### **5.1.1. Research Question 1: What are the Test Takers' Perceptions and Attitudes towards Online Speaking Assessment System?**

From the results it was seen that system use was also effective for students in terms of feeling anxiety. That is, after the data relating to the easiness of the system's use and the test takers' anxiety level was examined, it was found that test takers who strongly agreed that it was easy to take the exam did not feel nervous. In one study, Tarighat & Khodabakhsh (2016) listed the pros of CALL as being able to hear one's own voice, engaging in speaking outside the class, being given chance to do extra practice and speak solo and having no stress.

More than half of the test takers found the online speaking assessment system motivating. According to Anderson (1998), unlike the alternative and technological evaluation methods, traditional evaluation does not help the learners to foster motivation or decrease the anxiety.

In addition, the test takers stated that they wanted to recommend the system to the others. More than half of the test takers would like to recommend the system to the other students. When the system use is applicable and reasonable to use, it is almost impossible for the users not to recommend it. In their study Tarighat & Khodabakhsh (2016), the participants who took CALL assessment stated their opinions positively and wished to attend further activities again except one of them.

In addition, when the test takers' computer ability was taken into account, the ones who were good or advanced in using a computer stated that they would like to recommend

it to the others. Therefore, it can be understood that computer literacy may affect the students' ideas about recommending the online speaking assessment system to the others positively. As Douglas (2014) stated, with the computer technology the difference in test performance between computer familiar and computer unfamiliar people will diminish.

When the data were examined, it was found that the students found the activities done in the assessment consistent with the activities conducted in class time. It indicates in order to achieve the utmost success in assessments, the test activities should have a relationship with the classroom activities. According to Cheng (2005), if the positive washback is supported, then positive changes in teaching and examinations can be achieved. That is, a positive washback effect occurs when the assessment procedures meet the course goals and objectives.

In addition, the students thought that online form of speaking assessment was more practical than the face to face form of it, In Kenyon and Malone's study (2010), it was seen how technology had been used successfully as an alternative to a direct, face-to-face oral proficiency interview assessment.

More than half of the test takers stated that the feedback provided by the examiner included beneficial information about the students' progress. Galaczi (2010) pointed out in her study that as technology offers efficient and transparent results, test takers are provided efficient and instantaneous feedback.

When the data were examined, it was found out some of the test takers thought that the topics and the activities were not engaging. According to Scarcella & Oxford (1992, p.90), "if tasks aren't interesting and relevant, they won't be able to affect the students' attitudes and motivation in a positive way."

When the test takers were asked about whether the test was fair, nearly half of them reported that it was not fair. Only 25.7% were positive about the fairness issue. In the study by Tarighat & Khodabakhsh (2016) about mobile assisted language assessment, one participant noted: "It can't be fair. Speaking should be done in a real situation. At home we are able to think about it but in a real situation we have to improvise" (p. 412). In that study they did not want to be evaluated just to be scored, what they wanted was to be evaluated by their performance during the process. The results of the questionnaires were confirmed in the interview session and the ones, especially those who had concerns about the grading

system reported that there was a problem related to the fairness issue. As Baron and Boschee (1995) listed the drawbacks of alternative assessments, they mentioned the possibility of subjectivity in marking results, or the difficulty of having reliable and valid tests.

Some of the test takers mentioned technical and environmental problems. Some of them reported having experienced difficulties in managing the problems related to internet connection, headphones or camera. In Ounis's study (2017) one of the problems aroused was about the use of technology and it showed that the hardships pertaining to the assessment of the learners' speaking performance though the use of technology comprises the difficulty in handling or in implementing the use of multimedia or digital devices.

The positive attitudes could be seen about the effect of online speaking system on the students' daily speaking skills. Majority of the test takers stated that the system had helped them to develop their daily speaking skills. As Fulcher (2000) stated, language use is unpredictable in real life. That's why, in daily conversations it is important to develop and sustain the dialogue.

Furthermore, majority of the test takers thought that they had more opportunities to speak English through this system. As Alderson (1990) discussed before "the computer can encourage the learner's own strategies for evaluation. In particular the information which a computer can collate and present about test performance could help the learner to feel that his own opinions are of importance" (p. 39-43).

When we look at the results, we saw that majority of the test takers thought that online form of speaking assessment made them practice English more. The study conducted by Tarighat & Khodabakhsh (2016) investigated the feasibility of Mobile Assisted Language Assessment and the EFL learners' attitudes towards it and the results showed that this kind of assessment made it possible for the test takers to participate and practice their speaking skill while being assessed.

According to the result obtained from the questionnaire, more than half of the test takers stated that their speaking skills improved through systematically prepared activities. The activities prepared with the help of technology engaged students more in the lesson and gave them a chance of developing the skills. As Douglas and Chapelle (2006) stated, with the help of rich and authentic input offered to the test-takers by technology, it is possible to attempt the test takers to engage them in productive testing procedure.

According to the results, nearly half of the test takers thought that online form of speaking assessment encouraged them to speak English in class. According to Malabonga et al (2005), in computer-based oral exams test takers can choose the tasks in accordance with their strengths and thus can do their own planning about the level of tasks and this enhances test takers' favourable perceptions of the test.

When the data were examined, it was found out that many of the test takers disagreed that online speaking assessment did not help them to develop their speaking skills and they stated that their oral communication skills improved thanks to this system. Providing authentic materials and giving the students a chance to communicate in a real environment is useful as it promotes their speaking skills. As Brown (2004) suggested teachers need to encourage the use of authentic language and provide students with opportunities to initiate oral communication in the classroom.

#### **5.1.2. Research Questions 1a and 1b: What are the Test Takers' Perceptions and Attitudes towards the Effect of This System on Their Speaking and Listening Competencies?**

With this system, it was aimed to keep the students active and improve their speaking skills with the activities. It was understood from the data that majority of the students found the system helpful to state his/her opinion on the activity given. As it was framed in CEFR (Council of Europe, 2001) B2 students can understand standard spoken language, live or broadcast, on both familiar and unfamiliar topics normally encountered in personal, social, academic or life. It is important for a learner to produce an opinion and give appropriate responses in communicative tests. According to Fulcher (2000) "test taker must be able to recognize communicative purpose and be able to respond appropriately" (p. 490).

In addition, test takers agreed that the system helped them to fill the gaps while speaking by using proper utterances. Fulcher (2000) pointed out that the purpose of communicative tests for the test-taker is to understand the purpose and be able to respond correctly.

As found out in the previous parts of the results, thanks to this assessment student gained self-confidence while using the target language and the test takers felt confident while communicating on familiar routine matters. As Kramersch (2012) pointed out with the

developments in technology (e.g. social media), the communication among humanbeings would revolutionize and be open up new possibilities.

More than half of the test takers, on the other hand, stated that the system was helpful for them to follow the talk when clearly structured activities were provided. According to CEFR (Council of Europe, 2001), B2 learners can understand announcements and messages on concrete and abstract topics spoken in standard dialect at normal speed.

### **5.1.3. Research Question 2: What are the Perceptions and Attitudes of the Test Makers, the Instructors and the School Administrators towards Online Speaking Assessments?**

As discussed before, the instructors, the test makers and the administrators have both positive and negative thoughts about the effects of this system on the test takers and on the competencies that were being tried to be developed. In one study Jamila, Shamic, & Tariq (2012), it was found out that a good majority of teachers expressed that computer based examination system was risky because of system failure or electric failure problems during examinations.

The system was also said to be developed in terms of validity and reliability issues by the instructors and administrators. As Norris, Brown, Hudson, & Yoshioka (1998) pointed out, “the issues of reliability and validity must be dealt with for alternative assessments just as they are for any other type of assessment—in an open, honest, clear, demonstrable, and convincing way” (p. 5). In addition, according to Underhill (1987), “the process of establishing the general validity of a test procedure is called validation and all aspects of validity are important.” (p.104-105).

The instructors and administrators stated that the online speaking assessment would reflect the test takers’ actual level of speaking, if administered properly. They also shared their ideas about how the system had influenced the students positively about how to make use of ICT to improve their language skills. From a pedagogic point of view, Alderson (2000) points to the user friendliness format of computer-based tests. In addition, he emphasizes as the feedback they give to the users is useful, the tests would be more meaningful. In addition, as Fulcher (2000), Alderson (2000), and Roever (2001) stated the importance of language testing through technology by displaying a number of merits of language testing through the Internet. The very first and main advantage of this kind of test is its superiority in terms of time and space.



To promote positive washback, one of the instructors shared that teachers should employ speaking tasks similar to the ones that the students would come across in the assessment. More recently, a number of studies have further confirmed the existence and complex nature of the washback effect (e.g., Alderson & Hamp-Lyons, 1996; Shohamy, Donitsa-Schmidt, & Ferman, 1996; Wall, 1996; Watanabe, 1992, 1996a, 1996b). These results show that it is not wrong to mention the existence of washback effect in language teaching and testing.

## 5.2. Conclusion

The ability to speak the target language fluently can motivate and affect the learner positively as it enables the speaker to communicate and express feelings. Although learners are sometimes proficient in using the structures properly, they may not be sufficient enough to use the language communicatively. The reason for this is the lack of real communication environments that help the learners to use the language in a real English spoken atmosphere. As Underhill (1987) stated in a real oral test, the most important priorities are different. Unlike the written tests, real people communicate face to face without the existence of the idea of 'test'. Instead of providing fake environments while teaching and testing speaking, newly adopted methods such as computer mediated language testing can be made use of. With the help of computers, learners can communicate with real people on real topics without understanding that they are actually being tested and evaluated.

With these things in mind, a newly adopted system of online speaking assessment which enables learners to communicate and use the language in a real atmosphere were thoroughly analyzed and discussed by looking at the test takers' perceptions and attitudes. In addition, this study aimed to provide information about test takers' perceptions and attitudes to the online speaking assessment. Another ultimate goal of this study was to discover the effects of this system on test takers' speaking and listening competencies. The study also examined what test makers', school administrators' and instructors' thoughts were about the effects of online speaking assessment system on students.

The participants in this study were B2 level preparatory school students in a foundation university in İzmir. The total population of the school was more than 600 students and 72 instructors but the target group was limited to 291 (B2) students and 5 instructors, test makers and administrators. The reason why B2 students were intentionally chosen for

the study was that it was considered that those students had already taken lots of online speaking exams since they were A1 level students and were knowledgeable enough about the system to express their perceptions and opinions about the new online speaking system.

In this study, two main instruments were used to collect data; a questionnaire and an interview. Interview data were obtained through semi-structured interviews conducted with both students and instructors to understand the process in detail.

The results were analyzed by dividing the questionnaire questions into three categories: feeling, system and skill. The results showed that the participants in this study felt themselves confident, relaxed and positive during the assessment. When the results were examined in terms of students' attitudes towards the system and its effect on their speaking and listening competencies, it could be seen that many of the participants thought that it was practical and easy to complete it. Nearly 80% percent of them found the system contemporary. Next, they pointed out that the system was applicable and reasonable to use. In addition, as for the students' ideas about the effect of this kind of assessment on their speaking and listening competencies, it could be said that 63.6% of the students stated that online speaking assessment helped them to start the conversation while 57.5% of them pointed out that it helped them to end the conversation. Hence, more than half of them see themselves capable of starting and ending the conversation. Furthermore, more than half of the test takers found the system helpful for them to understand general messages or details during conversation. On the other hand, the participants recommended new ideas related to the grading system, test administrators/designers, technical and environmental issues and the duration of the exam.

When the results were examined from the opinions of the examiners and the instructors, it could be seen that the instructors and the test makers had both positive and negative thoughts about the effects of this system on the test takers and on the competencies that were being tried to be developed. All of them were actually sure about the system and its good results. However, they reported some issues to be improved related to the system such as the fairness problems in grading, time constraints and connection breakdowns.

### **5.3. Implications of the Study**

The ability to speak the target language fluently can motivate and affect the learner positively as it enables the speaker to communicate, express feelings. Although learners are sometimes proficient in using the structures properly, they may not be sufficient enough to

use the language communicatively. The reason for this is the lack of real communication environments that help the learners to use the language in a real English spoken atmosphere. As Underhill (1987) stated, in a real oral tests people communicate face to face without concerning about the idea of being tested and evaluated. Thanks to computer adopted speaking methods, providing real environments for communication is possible.

When analyzed in terms of methodological and pedagogical aspects, the results of the study show important implications. As for the methodological aspects, some critical conclusion can be understood. First, students do not have real environment where they can practice the language that they are learning. They only face the reality of the need of speaking to communicate in the exams. Furthermore, these exams are done in an atmosphere where the students feel the existence of the idea of 'a test'. Consequently, neither the results show the real understanding of the language produced nor help the students develop themselves for further speaking activities in real communication environments.

Second, it is important to analyze the results in terms of anxiety and stress. When the students' attitudes are examined, it can be deduced that when the topics are chosen from real-life or routine matters, the students feel more confident. That is, in order to create a friendly atmosphere for students, the topics should be contemporary and in accordance with the real life communication skills. If the instructors and the test administrators choose the topics accordingly, the students will hopefully want to take more responsibility in speaking activities and in that way will improve their speaking and listening skills.

The third implication of the study is for the nature of the assessment system. As Madsen (1983) mentioned, assessing the speaking skill has always been a controversial and challenging issue in language testing. Therefore, new alternatives should be tried and hopefully good results can be obtained at the end. In addition, there is an increasing need of getting more reliable and valid exams. The participants in this study pointed out that they had concerns about the grading system, topics, tasks, activities chosen for the assessment and the duration of the exam. As an implication it can be said that these issues must be taken into consideration to get better results.

As for the pedagogical aspects, this study aimed to get some insight not only from the students but also from the instructors, test makers and administrators. In order to reach at the utmost success, the attitudes and perceptions of the students related to the effect of

online speaking exams on their speaking and listening abilities were thoroughly analyzed. In addition, the ideas of instructors, test makers and administrators related to the pros and cons of the study were studied. As an implication, this study may enlighten the ways of both the students and the instructors to minimize the effects of speaking tests that had been implemented until then and the study may also help them to improve positive attitudes while preparing more up-to-date speaking assessments.

In consideration of students' perceptions and attitudes, the analysis of interviews and questionnaires provide confirmatory evidence that test takers had positive attitudes towards the effect of this system on their feelings. As for the impact, throughout the interviews students mentioned they felt less and less shy in time. They also stated how this system was useful in terms of gaining self-confidence while speaking in target language. However, while many students thought that experiencing such an online speaking assessment offered several benefits in terms of improving their communication skills and promoting self-confidence, there were also some concerns emerged from the most frequent responses given by the test takers related to the challenges of the system such as the time constraints, technical problems, and unfair grading system.

The results indicate that students benefited from the continuous exposure to English every week, and they were able to develop their speaking skills, daily spoken English, pronunciation skills and vocabulary knowledge. Most of them were also happy as they were using the language communicatively. In addition, majority of the prep school students stated that thanks to the online speaking assessment system, they got the chance to speak with someone else who was living in another country, which was something that they found really helpful to improve their speaking skills.

This system also enabled them to learn other cultures by communicating with someone from a different country. Furthermore, the analysis of the questionnaire presents evidence that students were highly motivated during the implementation of the assessment. They mentioned the decrease in their anxiety.

The results of the study may also provide valuable indications to test makers, instructors and administrators who make decisions about speaking testing system. First of all, to standardize reliability and validity, in-service training may be offered for examiners. In addition, the activities and topics may be chosen in accordance with curriculum so that

the ones who are supposed to carry out the speaking assessment procedure may become well aware of how to use criteria for standardization purposes and to be fair and give more realistic scores. Finally, the institutions that want to promote their students' speaking and listening skills may get some insights and develop their tests in the light of the students' attitudes and perceptions in this study.

#### **5.4. Limitations of the Study**

This study used both quantitative and qualitative data to elicit the feelings and opinions of a group of B2 level preparatory school students in a foundation university. The first limitation is that all data were collected from one particular university and from one particular level of students. Another limitation is the number of the participants. In the quantitative part of the study there were 291 participants, which was actually one third of the whole school population. In the semi-structured interviews done with the students, there should also have been more students who took part in the interviews. The third limitation is the limited amount of time allocated for the study. That is, to get much broader sense of impression about the effect of online assessment, the study could have been done during the whole year instead of restricting it in one module. In sum, the current study had three limitations namely a) the school where the data were collected, b) the number of students, and c) the limited time of the study. It is not possible to generalize the results and understand the whole picture due to these limitations, but the aim is to have a deeper understanding of students' and stakeholders' attitudes and perceptions.

#### **5.5. Suggestions**

This study attempted to investigate the feelings and attitudes of a group of B2 level preparatory school students towards the online speaking assessment. It was aimed to discover the learners' views on the effects of this kind of assessment system on their speaking and listening competencies, so I collected the data from a limited number of the students instead of the whole school. In a further study the number of the participants can be maximized to get much more detailed results.

This study aimed to focus on the learners' views in order to explore their own perceptions about the online speaking assessment. It does not look at the comparison between the other type of assessment system that were implemented before and the new one. Thus, this comparative study may be investigated in a further study.

The participants of the present study were B2 level preparatory school students who had been experiencing online assessment since the beginning of the term. The results would be different with another level of students. For the upcoming research it can be recommended to apply this study with the other levels and in a large sum of time.



## REFERENCES

- Adams, R. (2003). L2 Output, Reformulation and Noticing: Implications for IL Development. *Language Teaching Research*. 7 (3), 347–376.
- Alderson, J. C. (1990). Learner-centered testing through computers: Institutional issues in individual assessment. In J. de Jong & D. K. Stevenson (Eds.) *Individualizing the assessment of language abilities* (pp. 20–7). Clevedon, UK: Multilingual Matters.
- Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
- Alderson, J. C., & Hamp-Lyons, L. 1996. TOEFL preparation courses: a study of washback. *Language Testing* 13(3), 280–97.
- Alderson, J. C., & Wall, D. 1993. Does washback exist? *Applied Linguistics*, 14(2), 155-129.
- Anderson, S. R. (1998). Why talk about different ways to grade? The shift from traditional assessment to alternative assessment. *New Directions for Teaching and Learning*, 74(Summer). doi: 10.1002/tl.7401
- Anivan, S. (1991). *Current Developments in Language Testing*. Anthology Series 25 Published by SEAMEO Regional Language Centre
- Araújo, L. (2010). Computer-based Assessment (CBA) of. *Luxembourg: Publications Office of the European Union*.
- Aschbacher, P. A. (1991). Performance assessment: State activity, interest, and concerns. *Applied Measurement in Education*, 4, 275–288.
- Bachman, L. F. & Palmer, A.S. (1996). *Language Testing in Practice*. Oxford: Oxford University Press.
- Bachman, L. F. & Palmer, A.S. (2010). *Language Assessment in Practice*. Oxford: Oxford University Press.
- Baron, M.A. & Boschee, F. (1995). *Authentic assessment: The key to unlocking student success*. Pennsylvania: Technomic Publishing Company.
- Brown, Gillian. & Yule, George. (1988). *Teaching the Spoken Language: an Approach Based on the Analysis of Conversational English*. Cambridge, New York: Cambridge University Press.
- Brown, S. (October 1998). Current Task Based Language Teaching: Some Issues and Models". Paper presented at the Annual Meeting of the Ohio Teachers of English to Speakers of other Languages. (ERIC Document Reproduction Service No. ED 430389).

- Brown, J. D. & Hudson T. (1998). The Alternatives in Language Assessment. *TESOL Quarterly*, 32 (4), 653-675.
- Brown, J. D. Dean, J., & Hudson, T. (2002). *Criterion-Referenced Language Testing*. Cambridge: CUP.
- Brown, H. D. (2004). *Language assessment: principles and classroom practices*. New York: Pearson Education.
- Bygate, M. (1987). *Speaking*. Oxford: Oxford University Press.
- Bygate, M. (1998). Theoretical Perspectives on Speaking. *Annual Review of Applied Linguistics*. 18 (1) 20-42.
- Bygate, M. (1999). Quality of language and purpose of task: Patterns of learners' language on two oral communication tasks. *Language Teaching Research* 3 (3), 185-214.
- Bygate, M. (2002), Effects of Task Repetition on the Structure and Control of Oral Language. *Researching pedagogic tasks second language learning and testing*. Longman.
- Bygate, M. & P. Skehan & Swain M. (2001) *Researching Pedagogic Tasks: Second Language Learning, Teaching, and Testing*. Applied Linguistics and Language Study. London
- Canale, M. Swain, M. d, I. Assessment, O. & Institute, O. (1980). Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing.1
- Canale, M. & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1, 1-47.
- Canale, M. & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1 (1980)
- Canale, M. & Swain, M. (1981). A Theoretical Framework for Communicative Competence. In Palmer, A. Groot, P. & Trosper, G. (Eds.), *The construct validation of test of communicative competence*, 31-36.
- Canale, M. & Swain, M. (1980). *A Theoretical Framework for Communicative Competence* Retrieved from <https://linguisticator.com/communicative-competence/> on date (12.05.2019)
- Canale, M. (1983). From communicative competence to communicative language pedagogy. In Richards, J. C. & Schmidt, R. W. (Eds.), *Language and Communication*, (pp. 2-27). London: Longman.



- Canale, M. (1984). A communicative approach to language proficiency assessment in a minority setting. In Rivera, C. (Ed.), *Communicative competence approaches to language proficiency assessment: Research and application*, (pp. 107-122). Clevedon: Multilingual Matters.
- Carroll, B. J. (1982). *Testing communicative performance: an interim study*, Pergamon Press, Oxford.
- Chalhoub-Deville, M. (2010). Technology in standardized language assessments. In R. Kaplan (Ed.), *The Oxford handbook of applied linguistics* (2nd ed., pp. 511–26). Oxford, England: Oxford University Press.
- Chapelle, C. (2001). *Computer Applications in Second Language Acquisition: Foundations for Teaching Testing and Research*. Cambridge: Cambridge University Press.
- Chapelle, C. & Douglas, D. (2006). *Assessing Language through Computer Technology*. Cambridge: Cambridge University Press
- Chapelle, C. & Douglas, D. (1993). Foundations and directions for a new decade of language testing. In D. Douglas & C. Chapelle (Eds.), *A new decade of language testing research* (pp. 1-22). Alexandria, VA: TESOL.
- Charles, J. Clapham, C., & Wall, D. (1995). *Language Test Construction and Evaluation*. Cambridge: CUP.
- Cheng, L. (2005). *Changing language teaching through language testing: A Washback study*.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: M.I.T. Press.
- Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Council of Europe (2018). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Courtney, M. (1996). Talking to Learn: Selecting and Using Group Oral Tasks. *ELT Journal*, 5 (4), 318-325.
- Creswell, J. W. (2003). *Research design Qualitative, Quantitative, and Mixed Methods Approaches*. California: SAGE Publications.
- Creswell, J. W. & Plano Clark, V. L. (2006). *Designing and Conducting Mixed Methods Research*. Thousand Oaks, California: SAGE Publications.

- Creswell, J. W. (2014). *Educational research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. Upper Saddle River, N.J: Pearson/Merrill Prentice Hall.
- Creswell, J. W. & Clark, V. L. P. (2015). *Understanding Research: A Consumer's Guide* (2nd ed.) Pearson Education Inc, USA
- Derakhshan, A., Rezaei, S., & Alemi, M. (2011). Alternatives in Assessment or Alternatives to Assessment: A Solution or a Quandary. *International Journal of English Linguistics*, 1(1), 173-178.
- Darryl M. Hunter, Richard M. Jones, Bikkar S. Randhawa (1996) *The Canadian Journal of Program Evaluation* 11 (2), (pp. 61-85) Canadian Evaluation Society.
- Davies. (2000). Practicality of Test. Retrieved from [https://www.academia.edu/20604483/Practicality\\_of\\_Test](https://www.academia.edu/20604483/Practicality_of_Test) 05 28, 2019
- Dinapoli, R. (2000). Promoting Discourse with Task-Based Scenario Interaction. Paper presented at the International Conference on Language for Specific Purposes (Barcelona, Spain, September 7-9, 2000) (ERIC Document Reproduction Service No. ED: 454742).
- Douglas, D. (2014). *Understanding Language Testing*. Routledge.
- Dörnyei, Z. & Thurrell, S. (1991). Strategic competence and how to teach it. *ELT Journal*, 45(1), 16-23.
- Eckard, R. & Kearny, M. (1981). *Teaching Conversational Skills in ESL*. Washington: Center of Applied Linguistics.
- Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.
- Ellis, R. (2003). *Task-Based Language Learning and Teaching*. Oxford: OUP
- Finch, A. (1997). Task based Teaching in a Traditional Setting: Understanding the Students. *Proceedings of the 1997 Korea TESOL Conference*. [http://www.kotesol.org/publications/proceedings/1997/hyun\\_tae-duck.pdf](http://www.kotesol.org/publications/proceedings/1997/hyun_tae-duck.pdf),
- Finch, A. (1999). The Task-Based Classroom in Practice. A Paper presented at PAC2 Conference, Seoul.
- Florez, M. (1998). Improving Adult ESL Learners' Pronunciation Skills. ERIC Digest. *National Clearinghouse for ESL Literacy Education*. (ERIC Document Reproduction Service No. ED: 427553).

- Florez, M. A. (1999). "Improving Adult English Language Learners' Speaking Skills". *ERIC Digest*. (ERIC Document Reproduction Service No. ED: 435204)
- Fulcher, G. (2000). Computers in language testing. In P. Brett and G. Motteram (eds.), *A special interest in computers* (pp. 93–107). Manchester: IATEFL Publications.
- Fulcher, G. (2000, April 11). The 'communicative' legacy in language testing. Elsevier, p. 483-497.
- Galaczi, E. D. (2010) Face-to-face and computer-based assessment of speaking: Challenges and opportunities, in Araujo, L (Ed.) *Computer-based assessment of foreign language speaking skills*, Luxemburg: Publications Office of the European Union, 29-51
- Gold, S. (2001) A Constructivist Approach To Online Training For Online Teachers. *Jaln* 5 (1) p. 35-57.
- Goulden, N.R. (1989). Theoretical and empirical comparisons of holistic and analytic scoring of written and spoken discourse. Paper presented at the annual meeting of the *Speech Communication Association*, San Francisco.
- Güllüoğlu, O. (2004). Attitudes and Perceptions of the Students at Gazi University towards testing speaking. Gazi University, Ankara.
- Harmer, J. (1992). The practice of English language teaching, *Longman*, London.
- Harris, D. P. (1969). *Testing English as a second language*. New York: McGraw Hill.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. doi: 10.3102/003465430298487
- Heaton, J.B. (2003). *Writing English Language Tests*. Longman, London.
- Henning, G. (1987) *A Guide to Language Testing: Development, Evaluation, Research*. Cambridge, MA: Newbury House
- Heritage, M. (2007). Formative assessment: What do teachers need to know and do? *Phi Delta Kappan*, 89 (2), 140-145.
- Herman, J. L. Aschbacher, P. R. & Winters, L. (1992). *A practical guide to alternative assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Howarth, P. (2001). "Process Speaking. Preparing to Repeat Yourself". *MET*. 10(1), 39-44.

- Huerta-Macías, A. (1995). Alternative assessment: Responses to commonly asked questions. *TESOL Journal*, 5 (1), 8–11.
- Hughes, A. (1989). *Testing for Language Teachers*. Cambridge: Cambridge University Press.
- Hymes, Dell H. (1966). Two types of linguistic relativity. In Bright, W. *Sociolinguistics*. The Hague: Mouton. pp.114–158.
- Hymes, D. (1972). On communicative competence. In J. B. Pride and J. Holmes(eds), *Sociolinguistics*. Harmondsworth: Penguin, pp. 269–293.
- Irons, A. (2008). *Enhancing Learning through Formative Assessment and Feedback*. London: Routledge
- Jamila, D. M. Shamic, P. & Tariq, R. (2012, October ). Computer-based vs Paper-based Examinations: Perceptions of University Teachers. *TOJET: The Turkish Online Journal of Educational Technology*, 11.
- Kasper, G. (2001). "Four Perspectives on L2 Pragmatic Development". *Applied Linguistics*. 22 (4), 502-530.
- Kenyon D. and Malone M. (2010) Investigating Examinee Autonomy in a Computerized Test of Oral Proficiency, in Araujo, L (Ed.) *Computer-based assessment of foreign language speaking skills*, Luxemburg: Publications Office of the European Union, 29-51
- Kramsch, C. (2012). Why foreign language teachers need to have a multilingual outlook and what that means for their teaching practice. *Muitas Vozes* 1(1-2). 181-188.
- Lado, R. (1961). *Language Testing: The Construction and Use of Foreign Language Tests*. Longman, London.
- Lee, J. (2000). *Tasks and Communicating in Language Classroom*. New York: McGraw- Hill Higher Education.
- Lee, L. (2007). Fostering second language oral communication through constructivist interaction in desktop videoconferencing. *Foreign Language Annals*, 40 (4), 635-649.
- Lesaux, N., & Harris, J. (2015, October 29). *Heinemann Blog*. Retrieved from <https://blog.heinemann.com/what-is-oral-language>
- Liao, X. (2001). "Information Gap in Communicative Classroom". *English Teaching Forum*. October 2001. 39 (4), pp: 38- 41
- Long, M. (1996). "The Role of Linguistic Environment in Second Language Acquisition". In W. Ritchie and T. Bhatia (Eds.). *Handbook of Second Language Acquisition*, (pp: 413-468), Sand Diego, CA: Academic Press.

- Long, M. and Crookes, G. (Spring 1992). "Three Approaches to Task-Based Syllabus Design". *TESOL Quarterly*. 26 (1), pp: 27-56.
- Luoma, S. (2004). Developing speaking tasks. *In Assessing speaking* (pp.139-169).
- Mackey, A. & Gass, S. (2005). *Second Language Research*. Lawrence Erlbaum Associates Publishers, New Jersey
- Madsen, H. S. (1983). *Techniques in Testing*. Oxford University Press, New York.
- Malabonga, V., Kenyon, D. & Carpenter, H. (2005). Self-assessment, preparation and response time on a computerized oral proficiency test. *Language Testing*, 22, 59-92.
- Merriam, S. Caffarella, R. S. & Baumgartner, L. M. (2007). *Learning in adulthood*. San Francisco, CA: Jossey-Bass
- Miles, M.B. & Huberman, A.M. (1994). *Qualitative Data Analysis (2nd ed.)*. USA: Sage Publications.
- Myers, L. (Fall 2000)." Task Interpretation and Task Effectiveness: A Vygotskian Analysis of a French L2 Classroom Task.". *Texas Papers in Foreign Language Education*, 5 (1), pp: 9-21.
- Norris, J., M. Brown, J. D., Hudson, T., & Yoshioka, J. (1998). *Designing second language performance assessments*. Honolulu: University of Hawai'i Press.
- Nunan, D. (1989). *Designing Tasks for the Communicative Classroom*. New York: Cambridge University Press.
- Nunan, D. (1992). cited in Luoma, 2004, p.30-31 *Research Methods in Language Learning*. Cambridge, Cambridge University Press.
- Nunan, D. (1999). *Second Language Teaching and Learning*. Boston: Heinle & Heinle Publishers.
- Nunan, D. (2005). *Task Based Language Teaching*. Cambridge: Cambridge University Press.
- Ounis, A. (2017, July 15). The Assessment of Speaking Skills at the Tertiary Level. *International Journal of English Linguistics*, 7(4), 95-112.
- Piaget, J. (1972). Intellectual evolution from adolescent to adulthood. *Human Development*, 16, 346-370.
- Richards, J.C. and Rodgers, T.S. (2001). *Approaches and Methods in Language Teaching*. New York: Cambridge University Press.

- Riggenbach, H. (1998). "Evaluating Learner Interactional Skills. Conversation at the Micro Level". In R. Young & A. He (Eds.). *Talking and Testing. Discourse Approaches to the Assessment of Oral Proficiency*, (pp. 53-67). Amsterdam: John Benjamin's Publishing Company.
- Roever, C. (2001). Web-based language testing. *Language Learning & Technology*, 5(2), 84–94.
- Rutherford-Hemming, T. (2012). Simulation Methodology in Nursing Education and Adult Learning Theory. *Adult Learning*, 23(3), 129–137. <https://doi.org/10.1177/1045159512452848>
- Savignon, S. J. (1972). *Communicative Competence: An Experiment in Foreign Language Teaching*. Philadelphia: The Centre for Curriculum Development, Inc.
- Scarcella, R. & Oxford, R. (1992). *The Tapestry of Language Learning*. Boston Heinle & Heinle Publishers.
- Shohamy, E., Donitsa-Schmidt, S., & Ferman, I. 1997. Test impact revisited: washback effect over time. *Language Testing* 13(3), 298–317.
- Skehan, P and Foster, P. (1997b). "The Influence of Task Structure and Processing Conditions on Narrative Retellings". *Language Learning*, 49 (1), 93-120.
- Skehan, P. and Foster, P. (1997). "The Influence of Planning and Post Task Activities on Accuracy and Complexity in Task-Based Learning". *Language Teaching Research*, 1 (3), 16-33.
- Splitter, L. (2009). *Authenticity and Constructivism in Education*. Studies in Philosophy and Education. Springer Science and Business
- Spolsky, B. & Hult, F. M. (2008). *The Handbook of Educational Linguistics*. Malden, MA: Blackwell Pub.
- Syakur, (1987). *Language Testing and Evaluation*. Surakarta: 11 Maret University Press.
- Swain, M. (1985). "Communicative Competence: Some Roles of Comprehensible Input and Comprehensible Output in its Development". In S. Gass and C. Madden (Eds.). *Input in Second Language Acquisition*. (pp: 235-253). Rowley, Mass: Newbury house.
- Şanlı, R. (2003). *Students' perceptions about online assessment: A case study*. Retrieved from <https://tez.yok.gov.tr/UlusalTezMerkezi/> in 2019 published M.A. dissertation, Middle East Technical University.
- Tarighat, S., & Khodabakhsh, S. (2016, July 18). Mobile-Assisted Language Assessment: Assessing speaking. *Computers in Human Behavior*, p. 409-413.

- Torre, D., M. Daley, B. J. Sebastian, J. L. & Elnicki, M. (2006). Overview of current learning theories for medical educators. *American Journal of Medicine*, 119, 903-907 University Press.
- Torky, S. A. (2006) The Effectiveness of a Task-Based Instruction Program in Developing the English Language Speaking Skills of Secondary Stage Students Ain Shams University
- Tung, P. (1986) Computer adaptive testing: Implications for language test developers. In C. Stansfield (ed.), *Technology and language testing* (pp. 13–28). Washington, DC: TESOL Publications.
- Underhill, N. (1987). *Testing Spoken Language, A handbook of Oral Testing Techniques*, Cambridge University Press Edinburgh.
- Ur, P. (1981). *Discussions that Work. Task-Centered Fluency Practice*. Cambridge: Cambridge University Press.
- Ur, P. (1997). *A Course in Language Teaching*. Cambridge: Cambridge University Press
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wall, D., & Alderson, J. (1993, March 1). Examining washback: the Sri Lankan Impact Study. *Language Testing*. 10 (1), 41-69
- Wall, D. 1996. Introducing new tests into traditional systems: insights from general education and from innovation theory. *Language Testing* 13(3), 334–54.
- Watanabe, Y. 1992. Washback effects of college entrance examinations on language learning strategies. *JACET Bulletin*, 23, 175–194.
- Weir, Cyril, J. (1990). *Communicative language testing*. New York: Prentice-Hall.
- Weir, Cyril, J. (1993). *Understanding and Developing Language Tests*. New York: Prentice-Hall.
- Weir, C. (2005). *Language testing and validation: An evidence-based approach*. Springer.
- Widdowson, H. G. (1983). *Learning Purpose and Language Use*. Oxford: Oxford University Press.



**APPENDICES**



## APPENDIX A

### Survey on Test Takers' Perceptions and Attitudes towards Online Speaking Assessment System

#### *Öğrencilerin İnternet Üzerinden Yapılan Konuşma Sınavına Karşı Algı ve Tutumlarına yönelik anket çalışması*

This questionnaire is for the evaluation of online speaking assessment implementation that is currently being used by the Foundation University Preparatory Class students. Your feedback is highly appreciated to make improvements to this assessment system. Your answers will be collected and analyzed by the researcher for the improvement of the assessment system and to see test takers' perceptions and attitudes. Your responses to this questionnaire will be treated as completely confidential. This study is being conducted by Gizem Yeşil Korkmaz (English Instructor & MA student of ELT). Feel free to contact the evaluator, should you have any inquiries about the questionnaire or this study. gzm\_ysl@hotmail.com

Thanks in advance for your invaluable feedback.

*Bu anket şuan bir vakıf üniversitesi Hazırlık Okulu öğrencileri tarafından kullanılan, İnternet üzerinden yapılan konuşma sınavı uygulamasının değerlendirilmesi içindir. Görüşleriniz bu sistemin geliştirilmesi için çok değerlidir. Cevaplarınız bu sistemin geliştirilmesi ve öğrencilerin bu sisteme karşı algı ve tutumlarının bulunması için araştırmacı tarafından toplanıp analiz edilecektir. Bu ankette vereceğiniz cevaplar tamamen gizli tutulacaktır. Bu çalışma halen İngiliz Dili Eğitimi Yüksek Lisans öğrencisi ve İngilizce Okutmanı Gizem Yeşil Korkmaz tarafından yapılmaktadır. Anketle veya bu çalışmayla ilgili sorularınız olursa kendisiyle iletişim kurmaktan çekinmeyiniz. gzm\_ysl@hotmail.com*

*Değerli görüşleriniz için şimdiden teşekkürler.*

#### **Authorization to participate in this study: *Bu çalışmada yer almaya yetkilendirme:***

I have read the information in this consent form and I voluntarily agree to take part in this study.

*Bu onay formundaki bilgileri okudum ve bu çalışmada yer almayı gönüllü olarak kabul ediyorum.*

**Age (Yaş)** : \_\_\_\_\_

**Gender (Cinsiyet)** : \_\_\_\_\_ Female (*Kadın*) \_\_\_\_\_ Male (*Erkek*)

## PART 1

Please check the boxes in the right column for each statement. Check only one box for each.  
(Lütfen her maddeye uygun olan sadece bir kutucuğu işaretleyiniz.)

(Advanced, Good, Average, Poor, None)  
(İleri, İyi, Orta, Zayıf, Hiç)

Computer Familiarity ( <i>Bilgisayara Yatkınlık</i> )	Advanced ( <i>İleri</i> )	Good ( <i>İyi</i> )	Average ( <i>Orta</i> )	Poor ( <i>Zayıf</i> )	None ( <i>Hiç</i> )
1. Your ability to use a computer ( <i>Bilgisayar kullanma yeteneğiniz</i> )					
2. Your ability to read from a computer screen ( <i>Bilgisayar ekranından okuyabilme yeteneğiniz</i> )					
3. Your ability to understand the things related to computer??? ( <i>Bilgisayar ile ilgili meseleleri anlama yeteneğiniz</i> )					
4. Your ability to solve the problems if you are stuck while using a computer ( <i>Bilgisayar kullanırken sorun yaşadığınızda o sorunu çözebilme yeteneğiniz</i> )					
5. Your ability to chat online ( <i>İnternet üzerinden iletişim kurma yeteneğiniz</i> )					

Please check one box in the right column for each statement.  
(Lütfen her maddeye uygun olan sadece bir kutucuğu işaretleyiniz.)

(Yes, No)  
(Evet, Hayır)

Prior Experiences ( <i>Geçmiş Tecrübeler</i> )	Yes ( <i>Evet</i> )	No ( <i>Hayır</i> )
1. I have taken tests/examinations online before. ( <i>Daha önceden internet üzerinden yapılan testlere/sınavlara katıldım.</i> )		
2. I have attended an online course before. ( <i>Daha önceden internet üzerinden yapılan bir kursa katıldım.</i> )		
3. I have taken such kind of online speaking assessments before. ( <i>Daha önceden bu programa benzer bir konuşma sınavına katıldım.</i> )		
4. I have used web for instructional purposes. ( <i>Daha önceden interneti öğrenme amaçlı kullandım.</i> )		

Please check the boxes in the right column for each statement. Check only one box for each.  
(Lütfen her maddeye uygun olan sadece bir kutucuğu işaretleyiniz.)

(1: Strongly disagree 2: Disagree 3: Partly agree 4: Agree 5: Strongly agree)  
(1: Kesinlikle katılmıyorum 2: Katılmıyorum 3: Kararsızım 4: Katılıyorum 5: Kesinlikle katılıyorum)

System Use (Sistem Kullanımı)					
1. I follow the instructions without any problem. (Bu sistemde karşıma çıkan yönergeleri problemsiz takip ederim.)	1	2	3	4	5
2. It is easy to register the system. (Bu sisteme kayıt olmak kolaydır.)	1	2	3	4	5
3. It is easy to take the exam. (Bu sistemde sınava girmek kolaydır.)	1	2	3	4	5
4. It is comfortable to use the system. (Bu sistemi kullanmak rahattır.)	1	2	3	4	5

## PART 2

Please check the boxes in the right column for each statement. Check only one box for each.  
(Lütfen her maddeye uygun olan sadece bir kutucuğu işaretleyiniz.)

(1: Strongly disagree 2: Disagree 3: Partly agree 4: Agree 5: Strongly agree)  
(1: Kesinlikle katılmıyorum 2: Katılmıyorum 3: Kararsızım 4: Katılıyorum 5: Kesinlikle katılıyorum)

<b>Test Takers' Perceptions and Attitudes (Öğrenci Algı ve Tutumları)</b>					
1. Online speaking assessment system makes me feel less nervous while using the target language. (Bu uygulama İngilizce konuşurken daha az gergin hissetmemi sağlar.)	1	2	3	4	5
2. Online speaking assessment system is practical compared to face to face form of speaking assessment. (Yüzyüze yapılan konuşma sınavlarına kıyasla bu uygulama daha pratiktir.)	1	2	3	4	5
3. Online speaking assessment system helps to develop my academic speaking skills. (Bu uygulama akademik konuşma becerilerimi geliştirmemde yardımcı olur.)	1	2	3	4	5
4. Online speaking assessment system helps to develop my daily speaking skills. (Bu uygulama günlük konuşma becerilerimi geliştirmemde yardımcı olur.)	1	2	3	4	5
5. Online speaking assessment is contemporary. (Bu uygulama günceldir.)???	1	2	3	4	5
6. Online speaking assessment improves my speaking skills through the systematically prepared activities. (Bu uygulama sistematik bir şekilde hazırlanan aktivitelerle konuşma becerilerimi geliştirir.)	1	2	3	4	5
7. Online speaking assessment helps me to gain my self-confidence in speaking. (Bu uygulama İngilizce konuşma konusundaki özgüvenimi kazanmama yardımcı olur.)	1	2	3	4	5
8. Online speaking assessment system doesn't help me to improve my speaking skills. (Bu uygulama konuşma becerilerimin gelişmesine katkı sağlamaz.)	1	2	3	4	5
9. Online speaking assessment system is motivating. (Bu uygulama motive edicidir.)	1	2	3	4	5
10. Online speaking assessment system is fair. (Bu uygulama adildir.)	1	2	3	4	5
11. Online speaking assessment system encourages speak English more in class. (Bu uygulama ders içinde daha çok İngilizce konuşmama teşvik eder.)	1	2	3	4	5
12. Cheating is difficult. (Bu uygulamada kopya çekmek zordur.)	1	2	3	4	5
13. I do more speaking practice thanks to online speaking assessment system. (Bu uygulama sayesinde daha fazla konuşma pratiği yapıyorum.)	1	2	3	4	5
14. The activities we have done in this system are engaging. (Bu uygulamada yaptığımız alıştırmalar ilgi çekicidir.)	1	2	3	4	5
15. I have more opportunities to speak English through online speaking assessment system. (Bu uygulamayla daha fazla İngilizce konuşma fırsatı elde ederim.)	1	2	3	4	5
16. I feel that my oral communication skills have improved. (Konuşma becerilerimin ilerlediğini düşünüyorum.)	1	2	3	4	5
17. Feedback given by the examiner provides beneficial information about my progress. (Sınavı uygulayan öğretmenimin verdiği geribildirimler gelişimim hakkında bana faydalı bilgi verir.)	1	2	3	4	5
18. The activities I do in online speaking assessment is consistent with the things I learn in class. (Bu uygulamada yaptığım aktiviteler sınıfta öğrendiklerimle tutarlıdır.)	1	2	3	4	5
19. I get nervous while taking the online speaking assessment system. (Bu uygulamayı kullanırken gergin hissederim.)	1	2	3	4	5
20. I would like to participate in future online assessment activities. (Gelecek online sınav uygulamalarına katılmak isterim.)???	1	2	3	4	5
21. I would like to recommend it to the other students. (Bu uygulamayı diğer öğrencilere tavsiye etmek isterim.)	1	2	3	4	5

### PART 3

Please check the boxes in the right column for each statement. Check only one box for each.

(Lütfen her maddeye uygun olan sadece bir kutucuğu işaretleyiniz.)

(1: Strongly disagree 2: Disagree 3: Partly agree 4: Agree 5: Strongly agree)

(1: Kesinlikle katılmıyorum 2: Katılmıyorum 3: Kararsızım 4: Katılıyorum 5: Kesinlikle katılıyorum)

Speaking Competency (Konuşma Becerisi)					
Online Speaking Assessment System helps me to... (İnternet üzerinden yapılan konuşma sınavı sistemi...)					
1. communicate with some confidence on familiar routine matters. (bildiğim rutin konular üzerine kendime güvenerek iletişim kurmama yardımcı olur.)	1	2	3	4	5
2. communicate with some confidence on familiar non-routine matters. (bildiğim rutin olmayan konular üzerine kendime güvenerek iletişim kurmama yardımcı olur.)	1	2	3	4	5
3. express feelings (surprise, happiness, sadness, interest and indifference) (duygularımı (şaşıрма, mutluluk, üzüntü, ilgi, ilgisizlik) ifade etmeme yardımcı olur.)	1	2	3	4	5
4. state my opinion on the activity given. (verilen aktivitelerde fikir belirtebilmeme yardımcı olur.)	1	2	3	4	5
5. carry out the activity by checking and confirming???. (bilgiyi kontrol ve teyit ederek konuşma aktivitelerini sürdürmeme yardımcı olur.)	1	2	3	4	5
6. fill the gaps while speaking by using proper utterances. (uygun ifadeleri kullanarak konuşmadaki boşlukları doldurmama yardımcı olur.)	1	2	3	4	5
7. start the conversation. (muhabbeti başlatabilmeme yardımcı olur.)	1	2	3	4	5
8. end the conversation. (muhabbeti sonlandırabilmeme yardımcı olur.)	1	2	3	4	5
9. express my ideas during conversation without having difficulty. (muhabbet süresince fikirlerimi zorluk çekmeden ifade edebilmeme yardımcı olur.)	1	2	3	4	5

Please check the boxes in the right column for each statement. Check only one box for each.

(Lütfen her maddeye uygun olan sadece bir kutucuğu işaretleyiniz.)

(1: Strongly disagree 2: Disagree 3: Partly agree 4: Agree 5: Strongly agree)

(1: Kesinlikle katılmıyorum 2: Katılmıyorum 3: Kararsızım 4: Katılıyorum 5: Kesinlikle katılıyorum)

Listening Competency (Dinleme becerisi)					
Online Speaking Assessment System helps me to... (İnternet üzerinden yapılan konuşma sınavı sistemi...)					
1. understand simple information provided by the examiner if it is suitable to my level. (seviyeme uygun olduğu takdirde sınavı yapan hocanın verdiği basit bilgiyi anlayabilmeme yardımcı olur.)	1	2	3	4	5
2. follow the talk if the activity is familiar. (verilen aktivitenin bilindik olması koşuluyla konuşmayı takip edebilmeme yardımcı olur.)	1	2	3	4	5
3. follow the talk if the activity is clearly structured. (aktivitenin açık olması koşuluyla konuşmayı takip edebilmeme yardımcı olur.)	1	2	3	4	5
4. follow the main points of the conversation. (konuşmadaki ana noktaları takip edebilmeme yardımcı olur.)	1	2	3	4	5
5. understand general messages. (konuşmadaki genel mesajları anlayabilmeme yardımcı olur.)	1	2	3	4	5
6. understand details. (konuşmadaki detayları anlayabilmeme yardımcı olur.)	1	2	3	4	5

## **APPENDIX B**

### **TEST MAKERS', INSTRUCTORS' AND ADMINISTRATORS' INTERVIEW QUESTIONS**

1. What are your views on online speaking assessment in general?
2. What do you think of the match between the online assessment content and students' level of proficiency?
3. Are there any shortcomings or problems attached to this assessment? If so, what are the drawbacks of this system?
4. Which component or area needs to be improved most?
5. What can be done to develop features or functions of online assessment?
6. To what degree do you think this online speaking assessment system is appropriate and convenient to use? Why/Why not?
7. What are the other issues or areas that have not been mentioned but need to be improved?

## APPENDIX C

### TEST TAKERS' INTERVIEW QUESTIONS

1. What are your views on online speaking assessment in general?
2. Are there any shortcomings or problems attached to this assessment?
3. To what degree do you think this online speaking assessment system is appropriate and convenient to use? Why/Why not?
4. To what degree do you think this online speaking assessment system has helped you in your learning a new language process? Why/Why not?
5. To what degree do you think this online speaking assessment system improved your speaking skills? Why/Why not?
6. To what degree do you think this online speaking assessment system improved your listening skills? Why/Why not?
7. What are the drawbacks of this system?
8. What are the things that go well in this system?
9. What are the other issues or areas that have not been mentioned but need to be improved?

**APPENDIX D****CV**

<b>Personal Information</b>	
Name	Gizem
Surname	KORKMAZ
Birth place and date	Ceyhan / 01.01.1988
Nationality	T.C
Contact and e-mail address	2380 Sok. No:1 D:8 Güzelbahçe/İzmir gzm_ysl@hotmail.com
<b>Educational Background</b>	
High School	Gürkan (Süper) Lisesi
University	Uludağ University
<b>Foreign Languages</b>	
English – YDS – April, 2017	98,75
<b>Professional Background</b>	
2009-2011	Instructor at Ege University
2011-2016	Instructor at Gediz University
2017-currently employed	Teacher at MEV College