

To my beloved father Mahmut Duran,  
whose character has always illuminated my life...

"Deniz babasıyla vedalaşırken ona bir söz verdi. Artık büyüyecek ve hiç ağlamayacaktı. Çünkü onun göklerde uçmasa da süper kahraman bir babası vardı. Deniz süper kahramanın gitmesine izin verdi. Çünkü yeni maceralar onu bekliyordu. Bu mutlu sonla biten hikayeler de hep başkalarına anlatılmalıydı."

(Çağın Irmak, 2005, Babam ve Oğlum)

Teachers' and Students' Perceptions about Classroom-Based Speaking Tests and  
Their Washback

The Graduate School of Education  
of  
Bilkent University

by  
Özlem Duran

In Partial Fulfilment of the Requirements for the Degree of  
Master of Arts  
in  
The Program of  
Teaching English as a Foreign Language  
Bilkent University  
Ankara

July 2011

BILKENT UNIVERSITY  
THE GRADUATE SCHOOL OF EDUCATION  
MA THESIS EXAMINATION RESULT FORM

July 29, 2011

The examining committee appointed by the Graduate School of Education for the  
thesis examination of the MA TEFL student

Özlem Duran

has read the thesis of the student.

The committee has decided that the thesis of the student is satisfactory.

Thesis Title:	Teachers' and Students' Perceptions about Classroom-based Speaking Tests and Their Washback
Thesis Advisor:	Asst. Prof. Dr. Philip Durrant Bilkent University, MA TEFL Program
Committee Members:	Asst. Prof. Dr. Robin Martin Bilkent University, Faculty of Education Res. Asst. Büşra Delen Gazi University, Gazi Faculty of Education, ELT Program

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Teaching English as a Foreign Language.

---

(Assist. Prof. Dr. Philip Durrant)

Supervisor

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Teaching English as a Foreign Language.

---

(Asst. Prof. Dr. Robin Martin)

Examining Committee Member

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Teaching English as a Foreign Language.

---

(Res. Asst. Büşra Delen)

Examining Committee Member

Approval of the Graduate School of Education

---

(Visiting Prof. Dr. Margaret Sands)

Director

ABSTRACT

TEACHERS' AND STUDENTS' PERCEPTIONS ABOUT CLASSROOM-BASED  
SPEAKING TESTS AND THEIR WASHBACK

Özlem Duran

M.A. Department of Teaching English as a Foreign Language

Supervisor: Asst. Prof. Dr. Philip Durrant

July 2011

Testing is an indispensable part of the teaching and learning processes. Since testing, teaching, and learning are closely related, it is inevitable for them to have an influence on each other. Tests are thought to affect teaching and learning positively or negatively. Direct testing is seen to have greater effect on productive and receptive skills than other tests do. Speaking skills is one of the English language skills which is tested through direct tests. While the washback effect of worldwide or nationwide tests has been studied to a great extent, the washback effect of speaking tests has received little attention from researchers. As for the washback effect of classroom-based speaking tests, the researcher has not been able to find one.

This present study mainly aimed to investigate teachers' and students' perceptions of the washback effects of classroom-based speaking tests. In addition, since the subject is closely related to teachers' and students' attitudes towards and beliefs about teaching and testing speaking, these issues were also addressed.

The study was conducted at Akdeniz University School of Foreign Languages in Antalya, Turkey with 307 preparatory class intermediate level students and 45 instructors of English. The data were collected through teacher and student questionnaires and teacher and student interviews.

The results revealed that teachers stated that they are not influenced by the speaking tests in terms of what they do in classes, but they have positive attitudes towards teaching and testing speaking and they believe that speaking tests have a positive effect on their students' speaking ability. Teachers and students believe that getting ready for speaking tests improves the general speaking skills of students. Students are also quite positive towards teaching and testing speaking and speaking tests' positive effects. The students and instructors think that these speaking tests should remain as a component of all the exams. Moreover, the students think that speaking tests' weight should be increased. The curriculum development department and testing office of Akdeniz University School of Foreign Languages can utilize the results of the current study in order to create more positive washback.

**Key words:** washback effect, speaking tests, testing speaking

ÖZET  
ÖĞRETMENLERİN VE ÖĞRENCİLERİN SINIF TEMELLİ KONUŞMA  
SINAVLARINA VE ONLARIN ÖĞRETİM ÜZERİNE OLAN ETKİSİNE BAKIŞ  
AÇISI

Özlem Duran

Yüksek Lisans, Yabancı Dil Olarak İngilizce Öğretimi Bölümü

Tez danışmanı: Yrd. Doç. Dr. Philip Durrant

Temmuz 2011

Ölçme, öğretim-öğrenim sürecinin vazgeçilmez bir parçasıdır. Ölçme, öğretme ve öğrenme arasında sıkı bir bağ olduğu için, birbirlerini etkilemeleri kaçınılmazdır. Testlerin, öğretim ve öğrenimi olumlu ya da olumsuz şekilde etkilediği düşünülür. Doğrudan ölçme üretime ve algılamaya dayalı beceriler üzerinde diğer sınavlardan daha büyük etkiye sahiptir. Konuşma becerileri, doğrudan ölçme yoluyla ölçülen becerilerden biridir. Dünyaca ünlü ya da ulusal üne sahip sınavların öğrenim-öğretim üzerindeki etkileri büyük oranda çalışılmış olsa da, konuşma sınavlarının öğretime ve öğrenime olan etkisi araştırmacılar tarafından az ilgi görmüştür. Araştırmacı sınıf temelli konuşma sınavlarının eğitim-öğretim üzerindeki etkileri hakkında bir çalışma bulamamıştır.

Bu çalışma öğretmenlerin ve öğrencilerin sınıf temelli konuşma sınavlarının öğretim üzerindeki etkileri hakkındaki görüşlerini araştırmayı hedeflemiştir. Buna ilaveten, konu öğretmenlerin ve öğrencilerin konuşmanın öğretilmesi ve ölçülmesi

hakkındaki tutum ve inançlarıyla da çok yakından alakalı olduğu için bu konulara da gönderme yapılmıştır.

Bu çalışma Antalya’da Akdeniz Üniversitesi Yabancı Diller Yüksekokulu’nda 307 orta seviyeli hazırlık sınıfı öğrencisiyle ve 45 İngilizce okutmanı ile uygulanmıştır. Veriler öğretmen ve öğrenci anketleri ve öğretmen ve öğrenci röportajları yoluyla toplanmıştır.

Sonuçlara göre öğretmenler konuşma sınavlarının sınıfta yaptıklarını etkilemediğini, ama konuşmayı öğretme ve ölçme konusunda olumlu tutumlar içerisinde olduklarını ve konuşma sınavlarının öğrencinin konuşma yeteneği üzerinde olumlu etkileri olduğuna inandıklarını belirtmişlerdir. Öğretmenler ve öğrenciler konuşma sınavlarına hazırlanmanın öğrencinin genel konuşma yeteneğini geliştirdiğine inanmaktadırlar. Öğrenciler de konuşmayı öğretme, ölçme ve konuşma sınavlarının olumlu etkileri konularına karşı oldukça ılımlıdır. Öğrenciler ve öğretmenler, bu konuşma sınavlarının tüm sınavların bir parçası olarak kalması gerektiğini düşünmektedirler. Buna ek olarak, öğrenciler konuşma sınavlarının not ağırlığının artırılması gerektiğini düşünmektedirler. Akdeniz Üniversitesi Yabancı Diller Yüksekokulu Program Geliştirme ve Ölçme Değerlendirme Birimleri bu çalışmanın sonuçlarından, sınavlar yoluyla öğrenim öğretim üzerinde olumlu etki yaratmak için yararlanabilirler.

Anahtar kelimeler: sınavların öğrenim öğretim sürecine etkileri, konuşma sınavları, konuşma becerisini ölçme



## ACKNOWLEDGEMENT

If it was not for Fulbright Commission, it would be just a dream for me to do my MA degree at Bilkent University TEFL program. That is why I would firstly like to thank the Fulbright Commission and Bilkent University for offering such an invaluable opportunity, and thank Julie Mathews-Aydınlı, the head of the MA TEFL, for directing such a qualified program.

I owe special thanks to my supervisor Asst. Prof. Dr. Philip Durrant without whose support and patience it would be impossible for me to finish this thesis.

I would also like to thank my other MA TEFL lecturers Asst. Prof. Dr. JoDee Walters and Dr. Maria Angelova. Mrs. Walters introduced us to the corpus, which was a totally new phenomenon for me, and she helped me determine the direction my thesis would take by reminding me of the term ‘washback’. Dr. Angelova, who was humanistic, empathetic, and understanding, meant more than a lecturer to me.

I would also like to thank my defence jury members Asst. Prof. Dr. Robin Martin and Res. Asst. Büşra Delen for their invaluable support and feedback.

Although all these special people helped me within the process if Prof. Dr. Mustafa KİNSİZ, the head of Akdeniz University Foreign Languages School, had not given me the support to do my MA degree at Bilkent University and conduct my study at Akdeniz University Foreign Languages School, it would not have been possible to study this program. I am also thankful to my volunteer participants.

It was especially the good impression that my dear colleague Gülnihal Şakrak had left on the Bilkent people on behalf of Akdeniz University in 2008-2009, that made things easier for me. Gülnihal provided great help and her thesis did not leave me alone in Ankara by always being by my side.

I am grateful to Ahmet Kütük, Alperen Güçlü, Başak Ercan, Demet Çimen, Emre Başaran, F. Birgül Erdemir, Kevser Demirtaş, Mustafa Başaran, and Kağan Ataalp for being just a phone call away, supporting and helping me in many different stages of this long story.

I owe special thanks to my brother Volkan Duran for giving me a hand with transcribing the longest and the most difficult interviews.

When I decided to apply for the program, my mom and my siblings -Canan and Volkan Duran- were quite sure that I would be accepted since they always believed in my success in any subject. Moreover, it was my mother Sakine Duran and my father Mahmut Duran, who always respected and supported my decisions and let me study ELT.

Inst. Hülya Uzun, one of the former vice principals of Foreign Languages School, a very special person, has done a lot for me such as writing a reference letter, motivating me to do my very first piloting to make up my mind on which way to go, and checking my translations of the questionnaires. I have always felt that her eyes have been on me wherever I am, which make me feel that I am safe.

In Ankara I never felt lonely thanks to my grandmother Hatice Yiğit, my aunts Halise Buzsun and Fatma Yiğit, and my cousins Ebru Buzsun, Onur Buzsun, Melda Karakuş, and Yiğit Mahir Karakuş who listened to my MA TEFL days in his mother's womb but passed away when he was just a fourteen-day-old baby. To know that I had a family in Yenimahalle made me feel relieved. I want to thank my grandmother for providing a hot stove in cold Ankara days and my cousin Onur Buzsun for being with me and getting excited while I was doing the reliability analysis for the scales of my questionnaires. Though my beloved grandfather Mahir

Yiğit passed away in 2003, I am pretty sure that he is in peace since I became a teacher as he had always wished and did my MA degree in Ankara.

I thank all my MA TEFL classmates for their support whenever I thought that the MA year would never end. Though all of them have a different place deep in my heart, it was for Demet Kulaç, who is my biggest gain in that program, I got used to live in a dark and formal city without a sea. She has been far more than a sister to me as she will always be.

Without my friends, I would not be able to finish my thesis in hard summer school conditions and without having a permanent accommodation. In this respect, my special and lifelong thanks go to Burcu Göçmen, Ege Özince, Judehan Yağcıoğlu, Şebnem Kurt, and Tuğba Yalçın since they warmly welcomed me to their homes and accommodated me. Though I met Mehmet Galip Zorba this summer, it is as if I knew him as long as I have known my best friends. With the others, he did not let me feel that I would never finish the thesis in July.

I would also like to thank engineer Gülten Güller for taking care of me in my last BA year when I suffered quite difficult times after my father's death and for helping me finish my school.

Finally I thank my lifelong friends Burcu Kılıç, Havva Özay Hıdıroğlu, and Hamdiye Avcı for always being with me and being the half of my heart and soul.

Last but not least I owe thanks to Orhan and Canan Şahin. While Mr. Şahin, my primary school teacher, planted all the good things in me when I was a little girl, Mrs. Şahin, my English teacher in high school, was the reason why I became an English teacher.

Thank you

Your Özlem

## TABLE OF CONTENTS

ABSTRACT .....	iv
ÖZET .....	vi
ACKNOWLEDGEMENT .....	viii
TABLE OF CONTENTS .....	xi
LIST OF TABLES .....	xiii
LIST OF FIGURES .....	xiv
CHAPTER 1: INTRODUCTION .....	1
Introduction .....	1
Background of the Study .....	2
Statement of the Problem .....	5
Significance of the Study .....	6
Conclusion .....	7
CHAPTER II: LITERATURE REVIEW .....	8
Washback and Test Impact .....	8
Measurement-driven Instruction and Curriculum Alignment .....	9
Washback, Systematic, and Consequential Validity .....	10
Types of Washback .....	12
Washback Hypotheses .....	16
Studies Investigating Washback Effects .....	17
Conclusion .....	22
CHAPTER III: METHODOLOGY .....	23
Introduction .....	23
Research Questions .....	23
Setting .....	23
Speaking Tests Administered at Akdeniz University Prep Classes .....	25
Participants .....	26
Instruments .....	30
The Piloting Procedure .....	32
Data Collection Procedures .....	34
Data Analysis .....	37
Conclusion .....	37
CHAPTER IV: DATA ANALYSIS .....	38
Overview of the Study .....	38
Forming the Scales of the Questionnaires .....	39
Analysis of the Questionnaires .....	45
Teachers' attitudes towards and beliefs about teaching and testing speaking ...	45
Students' attitudes towards and beliefs about teaching and testing speaking ...	51
Compared and contrasted items on teachers' and students' attitudes towards and beliefs about teaching and testing speaking .....	54
Teachers' perceptions of the washback effects of classroom-based speaking tests .....	55
Students' perceptions of the washback effects of classroom-based speaking tests .....	61

Compared and contrasted items on teachers' and students' perceptions of the washback effects of classroom-based speaking tests .....	67
Conclusion .....	72
CHAPTER V: CONCLUSION .....	74
Introduction .....	74
General Results and Discussion .....	75
Teachers' and students' attitudes towards and beliefs about teaching and testing speaking.....	75
Teachers' and students' perceptions of the washback effects of classroom-based speaking tests .....	77
Pedagogical Implications .....	81
Limitations .....	84
Suggestions for Further Research .....	85
Conclusion .....	85
REFERENCES.....	87
APPENDIX A: INFORMED CONSENT FORM .....	89
APPENDIX B: BİLGİ VE KABUL FORMU .....	90
APPENDIX C: INFORMED CONSENT FORM .....	91
APPENDIX D: BİLGİ VE KABUL FORMU .....	92
APPENDIX E: TEACHER QUESTIONNAIRE.....	93
APPENDIX F: ÖĞRETMEN ANKETİ.....	97
APPENDIX G: STUDENT QUESTIONNAIRE .....	101
APPENDIX H: ÖĞRENCİ ANKETİ .....	105
APPENDIX I: CATEGORIZATION OF THE TEACHER QUESTIONNAIRE ITEMS ON TEACHERS' ATTITUDES TOWARDS AND BELIEFS ABOUT TEACHING AND TESTING SPEAKING .....	109
APPENDIX J: CATEGORIZATION OF TEACHER QUESTIONNAIRE ITEMS ON THE TEACHERS' PERCEPTIONS OF THE WASHBACK EFFECT OF CLASSROOM-BASED SPEAKING TESTS .....	111
APPENDIX K: CATEGORIZATION OF THE STUDENT QUESTIONNAIRE ITEMS ON STUDENTS' ATTITUDES TOWARDS AND BELIEFS ABOUT TEACHING AND TESTING SPEAKING .....	113
APPENDIX L: CATEGORIZATION OF STUDENT QUESTIONNAIRE ITEMS ON THE STUDENTS' PERCEPTIONS OF THE WASHBACK EFFECT OF CLASSROOM-BASED SPEAKING TESTS .....	115
APPENDIX M: THE STUDENT QUESTIONNAIRE OF CAINE (2005) .....	117
APPENDIX N: 2 NUMARALI ÖĞRENCİ RÖPORTAJINDAN BİR KESİT .....	119
APPENDIX O: A PART OF THE STUDENT 2 INTERVIEW.....	120
APPENDIX P: 1 NUMARALI ÖĞRETMEN RÖPORTAJINDAN BİR KESİT ...	121
APPENDIX Q: A PART OF THE TEACHER 1 INTERVIEW .....	122

## LIST OF TABLES

Table 1- Descriptive statistics for age, experience in teaching at Akdeniz University, and in administering speaking tests of the instructors .....	27
Table 2 - BA and MA majors of the instructors.....	27
Table 3 - The type of high school the students graduated from.....	28
Table 4 - The future faculties of the students.....	29
Table 5 - The ranges of the speaking scores of midterm .....	29
Table 6 - The ranges of speaking quiz scores of the students.....	30
Table 7 - Descriptive statistics for teachers' attitudes towards and beliefs about teaching and testing speaking .....	46
Table 8 - Descriptive statistics for students' attitudes towards and beliefs about teaching and testing speaking .....	51
Table 9 - Descriptive statistics for teachers' and students' attitudes towards and beliefs about teaching and testing speaking .....	55
Table 10 - Descriptive statistics of the teachers' perceptions of the washback effects of classroom-based speaking tests .....	56
Table 11 - Descriptive statistics of students' perceptions of the washback effects of classroom-based speaking tests.....	62
Table 12 - Descriptive statistics for teachers' and students' perceptions of the washback effects of classroom-based speaking tests.....	67

## LIST OF FIGURES

Figure 1 - A basic model of washback.....	16
Figure 2 - Getting ready for speaking tests improves students' speaking skills.....	68
Figure 3 - The usage of the things that the students have studied for the speaking test, in lessons after the test .....	69
Figure 4 - If speaking skills were not tested, I would not spend so much time on improving speaking skills .....	70
Figure 5 - Even if speaking skills were not tested, they should have a place in classes .....	71
Figure 6 - Students easily forget the things that they have studied for the speaking test, after the test .....	71

## CHAPTER 1: INTRODUCTION

### Introduction

We, as teachers, all want what we teach to be learnt by our students. We have been looking for ways to make our classes more important for students. One way which has been heavily used is to test what you teach to make students learn. If there is a test at the end of a period of instruction and students are graded accordingly, they have a good reason to study. Teaching and testing go hand in hand. Thus, testing has an important place in the field of education.

It is the same case in language teaching. Testing is an indispensable part of second language teaching. Although testing itself has been studied to a great extent, ‘the influences of tests on teaching and learning’ (Bailey, 1996, p.259), which is known as the washback effect, has not been studied adequately. The reason for this can stem from the fact that it is a complex phenomenon (Alderson and Wall, 1993; Bailey, 1996; Cheng, 2000; Watanabe, 2004).

Researchers have largely studied the washback effect of high stake tests, such as the effect of English Tests on Spanish University Entrance Examination (Amengual-Pizarro, 2009), Hong Kong Certificate of Education Examination in English in Hong Kong secondary schools (Cheng, 1997), the Graduation Threshold (GT) on English proficiency among graduating students in Taiwan (Hsu, 2009), the Foreign Language Test in a new competence-based State Examination for the Admission into Higher Education in Colombia (Manjarres, 2005), the use of TOEIC (Newsfields, 2005), ‘washback to the learners’ from the TOEFL (Reynolds, 2010), and the General English Proficiency Test (GEPT) on English teaching in Taiwan



(Shih, 2009). However, there has been little investigation on the washback effect of classroom-based tests. The fact that classroom-based tests have not received much attention may stem from the fact that research interests and perhaps more importantly, research funding, which is often provided by the company that produce the large tests, tends to be directed mostly towards the single big tests since these, individually, have high impact. But of course classroom tests, because they are far more common, are likely to have just as big an impact cumulatively.

With regard to the testing of speaking ability, while a number of studies have looked at ways of improving the reliability and validity of tests (Hughes, 2003; Messick, 1996) and at the tasks used in testing speaking (Elder; 2002; Fulcher and Marquez Reiter, 2003; Hyun, 2003; Taguchi, 2007; Tavakoli, 2009), little attention has been given to the influences of these speaking tests on teaching and learning.

This study mainly aims to reflect teachers' and students' perceptions of the washback effect of classroom-based speaking tests on the teaching and learning process.

### Background of the Study

As Taylor (2005) states, tests have long been believed to have a variety of direct influences on educational processes. It is commonly assumed that 'teachers will be influenced by the knowledge that their students are planning to take a certain test and will adapt their teaching methodology and lesson content to reflect the test's demands' (Taylor, 2005, p.154). Similarly, McEwen (as cited in Cheng, 2000, p.1) summarizes this situation by claiming 'what is assessed becomes what is valued, which becomes what is taught'. Madaus (as cited in Spratt, 2005, p.5) states that 'it is

testing, not the 'official' stated curriculum, that is increasingly determining what is taught, how it is taught, what is learned, and how it is learned'. The term 'washback' or 'backwash' refers to this 'influence of testing on teaching and learning' (Cheng, 2000, p.2).

The washback effect can either be 'positive' or 'negative' (Alderson and Wall, 1993). Vernon (as cited in Cheng, 2000) believes that if subjects or activities in a curriculum cannot directly contribute to passing the exam, they will most probably be ignored by teachers. Davies et al. (as cited in Taylor, 2005) provide a good example to illustrate this situation. They state that if the writing skill is tested through multiple choice tests, in-class practice will be more multiple choice-oriented rather than focused on writing itself.

Studies investigating the washback effect so far have mainly focused on high stake tests which are thought to influence learners' lives to a great extent. Some of these studies investigated the washback effects of worldwide tests such as TOEFL, IELTS, and TOEIC (Alderson & Hamp-Lyons, 1996; Newsfields, 2005; Reynolds, 2010; Rhami and Nazland, 2010). Others examined the washback effects of nationwide tests (Amengual-Pizarro, 2009; Cheng, 1997; Hsu, 2009; Manjarrés, 2004; Mohammadi, 2007; Shih, 2009; Yıldırım, 2010). However, the investigation of the washback effect of speaking tests has been neglected. Andrews et al. (2002), Caine (2005), Ferman (2004), and Munoz and Alvarez (2010) are among the studies which investigated washback effect of speaking tests.

Andrews et al. (2002) did an experimental study and they tried to measure and compare the spoken performances of the students who had to take the 'Use of English' (UE) test as a component of the Hong Kong Advanced Supplementary Test

with that of those who did not have to. The results suggest that the existence of UE might have had a positive influence on students' spoken English performance.

Caine (2005) examined the effects of the communicative curriculum of Japanese Ministry of Education into existing tests in Japan and proposed a direct test of speaking and investigated the washback effect of the direct tests. The results reveal that positive washback may be obtained by changing the testing tools to communicative testing but in-service training is necessary for teachers to teach and test communicative skills better.

Ferman (2004) aimed to investigate the washback of an EFL National Oral Matriculation Test administered in Israel. According to the results, there is a strong washback effect of EFL National Oral Matriculation Test on teaching and learning.

Munoz and Alvarez (2010) studied the washback of an Oral Assessment System (OAS). The results suggest that washback may be increased when students are informed on some things such as assessment procedures, scoring scales, and self-assessment mechanisms.

Although the studies mentioned here have contributed to the field of English Language Teaching, they have not investigated the effects of classroom-based speaking tests from the teachers' and students' point of view. Not only the Turkish EFL context but also the ELT world is in need of a broader mirror that reflects the influences of classroom-based speaking tests on the teaching and learning process for teachers and students.

To fulfil this need, the study aims to get the perspectives of EFL learners and instructors on the washback effect of speaking tests. The issues such as what is learned and taught in speaking classes, the time allocated for the in-class speaking

activities, and their reliability will be examined from teachers' and students' point of view.

### Statement of the Problem

The term washback effect has been a popular subject matter in educational contexts for a long time (Alderson & Wall, 1993; Cheng, 2000; Pan, 2009; Shawcross, 2007; Taylor, 2005; Vernon, as cited in Alderson, 1993). The washback effect of high-stake tests such as TOEFL and IELTS has been examined by many researchers (Alderson & Hamp-Lyons, 1996; Reynolds, 2010; Rhami & Nazland, 2010). In addition to these worldwide-known tests, a great deal of research has been conducted on the washback effect of some other high stake tests which are known nationwide (Amengual-Pizarro, 2009; Cheng, 1997; Hsu, 2009; Manjarrés, 2004; Mohammadi, 2007; Shih, 2009; Yıldırım, 2010). However, as Munoz and Alvarez (2010) state, there are not many studies focusing on the effects of classroom-based assessment. Moreover, no research has investigated the washback effect of classroom-based speaking tests from the teachers' and students' points of view. The purpose of this study is to reflect the perceptions of instructors and students on the washback effect of speaking tests in EFL context. Moreover, owing to the fact that the subject is closely related to teachers' and students' attitudes towards and beliefs about teaching and testing speaking, these issues will also be addressed.

Akdeniz University School of Foreign Languages has been applying speaking tests as a part of mid-term and final examinations for a long time. The English speaking ability of students has been tested more for the last two years by giving more weight to speaking quizzes and by allocating one out of five points of the

exams to the speaking parts. However, what teachers and students really think about the washback effects of these speaking tests is unknown. This study aims to reveal whether these speaking tests applied as a part of the evaluation process have any effects on the teaching and learning process from the instructors' and students' points of view.

#### Research Questions

This study will investigate the following research questions:

1. What are teachers' and students' attitudes towards and beliefs about teaching and testing speaking?
2. What are the washback effects of classroom-based speaking tests, as reflected in teachers' and students' perceptions?

#### Significance of the Study

Ever since the Communicative Approach was adopted, many modifications have been observed in testing 'speaking'. Institutions or national ministries of education such as the Japanese Ministry of Education (Caine, 2005) have started testing communication skills more. Although there have been many studies on speaking exam tasks, there is little research reflecting what teachers and students really think about the washback effects of speaking tests. Thus, this study may contribute to the literature by reflecting the perceptions of the instructors and students on the washback effect of classroom-based speaking tests.

At the local level, this study may contribute to Akdeniz University School of Foreign Languages by revealing the effects of the speaking exams done as a part of evaluation on the instructors and students. The results of this study may help the

testing department review the speaking tests they prepare and if necessary work more cooperatively with the curriculum development department to make necessary changes taking the ideas of the instructors who teach speaking and administer the exams, and the students who sit them, into account. The results may show whether these tests have an effect on teaching and learning speaking skills from the teachers' and students' points of view. If these exams have a washback effect, the instructors who work for the curriculum development department can use speaking tests to achieve their objectives to improve the speaking ability of the students. Making these revisions to test speaking ability may provide positive washback for teaching and practicing speaking skills.

### Conclusion

This chapter presented the background of the study, the statement and significance of the problem, and the research questions. The following chapter will provide the literature review for the theoretical background for the study. The third chapter will present detailed information on the methodology of the study which includes the participants, the data collection tools, data collection and analysis procedure. In the fourth chapter the collected data will be analysed and the findings will be presented. In the final chapter, which is the fifth one, general results and discussion, pedagogical implications, limitations of the study, and suggestions for further research will be presented.

## CHAPTER II: LITERATURE REVIEW

### Washback and Test Impact

The term washback has an important place in language testing. A number of researchers have proposed definitions for the term ‘washback’ for many years (Alderson and Hamp-Lyons, 1996; Alderson and Wall, 1993; Bachman and Palmer, 1996; Bailey, 1996; Hughes, 2003; McNamara, 2000; Messick, 1996; Pan, 2009; Shohamy et al., 1996; Watanabe, 1996). The point that has been emphasized in all these definitions is that testing affects teaching and learning. Although in the washback definitions of Bailey (1996), Bigg (as cited in Cheng and Curtis, 2004), Hughes (2003), McNamara (2000), Pan, (2009); Shohamy et al. (1996), and Watanabe (1996) both teaching and learning are stated to be influenced by tests, the washback definitions of Alderson and Hamp-Lyons (1996) and Alderson and Wall (1993) include the influences of tests just on teaching.

In this present study, the definition of Bailey (1996, p.259) stating that ‘the influences of tests on teaching and learning’ will be used to carry the meaning of washback. The main reason for this choice is the fact that it focuses on the effects of tests on both teachers and students.

Some researchers have used the term ‘backwash’ instead of washback to mean the same thing (Bigg, as cited in Cheng, 2000; Hughes, 2003; Spolsky, as cited in Pan, 2009; Tsegari, 2007). However, Alderson and Wall (1993) do not see any pragmatic or semantic difference between the terms.

Another common related term is **test impact**. Hamp-Lyons (1997) states that the term ‘impact’ is preferred in the general education and educational measurement

literature instead of washback. Actually, the main difference between ‘impact’ and washback is that impact can affect wider educational contexts. In general, if the effects of tests are mainly class-based and related to teaching and learning, it especially affects the curriculum, methodology, and students’ learning and these effects are related to washback. However, if the tests influence individuals, policies or practices, education system, society, and publishing, it is accepted as test impact (Bahman and Palmer, 1996; McNamara, 2000; Taylor, 2005; Wall, as cited Cheng et al., 2004).

### Measurement-driven Instruction and Curriculum Alignment

There are some different terms related to the ‘relationship between testing and teaching / learning’ other than washback.

Cheng and Curtis (2004, p.4) state that ‘tests or examinations can or should drive teaching, and hence learning’ and that this result is described as ‘**measurement-driven instruction**’ by Popham (as cited in Cheng and Curtis, 2004). Measurement-driven instruction brings positive connotations to mind, claiming that ‘testing should drive curriculum and thereby teaching and learning’ (Hamp-Lyons, 1997, p.295). Cheng and Curtis (2004) state that if driving teaching is the target, there should be a parallelism between the test format and content / curriculum. Shepherd (as cited in Cheng and Curtis, 2004) refers to this as ‘**curriculum alignment**’. Since this fact narrows the curriculum (Madaus; Cooley both cited in Alderson and Hamp-Lyons, 1996) and teachers’ training practices, it brings negative connotations to mind (Hamp-Lyons, 1997). According to Cheng and Curtis (2004) this alignment, including the situation in which a new examination is added to the



education system with the purpose of having a beneficial effect on teaching and learning process, has been labelled differently by different researchers. While Frederiksen and Collins (as cited in Cheng and Curtis, 2004) refer to this alignment as systematic validity, Messick (1996) sees it as consequential aspect of construct validity and Bahman and Palmer (1996) and Baker (as cited in Cheng and Curtis, 2004) call it test impact.

### Washback, Systematic, and Consequential Validity

Alderson and Wall (1993) claim that some writers have tended to relate the validity of a test to the extent to its good effects on teaching processes. In other words, the more beneficial effects a test has, the more valid a test is and vice versa. Morrow (as cited in Alderson and Wall, 1993) has used the term ‘washback validity’ to emphasize the degree of this relationship between a test and associated teaching. Morrow (as cited in Alderson and Wall, 1993, p.116) states that ‘[t]he first validity criterion that I would...put forward for [these examinations] would be a measure of how far the intended washback effect was actually being met in practice’. Pan (2009) explains Morrow’s idea by stating that the extent to which the needs of students, educators, researchers, administrators of tests, and anyone who uses the test are met is an issue that is directly related to washback validity.

In a similar way, Frederiksen and Collins (as cited in Alderson and Wall, 1993, p.116) have used the term systematic validity, which they define as ‘one that induces in the education system curricular and instructional changes that foster the development of the cognitive skills that the test is designed to measure’. They count improvement in skills after the test has taken place as a proof of systematic validity.

Though the term ‘consequential validity’ has been used in some studies (Alderson and Hamp-Lyons, 1996; Brown, 2004; Cheng and Curtis, 2004; McNamara, 2000; Pan, 2009), Messick (1996, p.251) does not view it as a separate type of validity, but instead he views it as one aspect of construct validity, which ‘includes evidence and rationales for evaluating the intended and unintended consequences of score interpretation and use in both short- and long-term, ... , with unfairness in test use, and with positive or negative washback effects on teaching and learning’.

McNamara (2000) provides an example which he sees as consequential validity of tests. He states that in an assessment reform which turns out to be based on ongoing projects rather than tests, the discrimination of the students’ skills can be more difficult in that rich families can hire a teacher to help their children in order to help them get good enough marks to be able to be accepted by good universities. This example situation brings to the researcher’ mind the situation in her own country. In Turkey, students have to take high-stake tests in order to be accepted by good schools for their secondary and higher education. However, the type of these tests does not fit the education system provided in their own schools, which forces them to take private courses in order to be successful in these tests. If the parents are able to afford the cost of the private course, they let their children go to these courses. The test developer should try to see the unexpected or unintended results of the tests because it is not fair to discriminate between students in the way of the example provided above. In this system the tests have less to do with the skills of the students than the wealth of their families. The consequences of tests should be

carefully considered by test developers before implementing anything new to the existing system.

### Types of Washback

Hughes (2003, p.1) and Buck (as cited in Bailey, 1996, p.258) agree on the issue that washback of testing on teaching and learning can either be 'harmful or beneficial'. If the influences of a test harm the teaching, and hence learning process, then it is considered to be negative washback. Hughes (2003) provides an example of the negative effect of washback. He gives the example of a student who is getting ready to study in an English-speaking medium and trying to gain all the language skills. If the exam which can determine the student's ability to study in that English-speaking environment does not address language skills at all, but focuses on a multiple-choice test, it will most probably cause the student to study for this multiple-choice tests rather than learning the necessary language skills. This effect is seen as undesirable.

Morris, Swain, and Alderson (as cited in Alderson and Wall, 1993) as well as Andrews et al. (2002) and Bailey (1996) think that washback can affect the teaching and learning process positively. One way in which tests can have beneficial effects is to use them as teaching and learning activities. For example Pearson (as cited in Hsu, 2009) considers good tests to be usable class activities.

In order to put the curriculum into practice effectively, Morris (as cited in Alderson and Wall, 1993) believes having examinations is fundamental. Andrews et al. (2002) similarly suggest that bringing changes into testing will let innovations occur in the language curriculum. In this respect, direct performance tests are

expected to promote the performance skills. Hughes (2003, p.18) believes that ‘a helpful washback effect’ can be achieved as a result of direct testing. Davies (as cited in Cheng, 2000, p.11) holds the belief that good tests are both ‘obedient servants of teaching’ and ‘leaders [of teaching]’.

Hughes (2003, p.53-56) lists seven ways to achieve the positive washback, as follows:

- ‘1. Test the abilities whose development you want to encourage.
2. Sample widely and unpredictably.
3. Use direct testing.
4. Make testing criterion-referenced.
5. Base achievement tests on objectives.
6. Ensure the test is known and understood by students and teachers.
7. Where necessary provide assistance to teachers.’

A question which can be asked here is whether only good tests bring positive effects to teaching and learning. Alderson and Wall (1993, p.117) hypothesize that not only good tests have a beneficial effect on teaching and learning process. They state that ‘poor’ tests can also be beneficial if they can make students and teachers ‘do good things they would not otherwise do’ by motivating students to do their homework, take the subject being tested more seriously, pay more attention to the lesson and hence be more successful and motivate teachers to prepare lessons more thoroughly no matter how valid the tests are.

Cheng and Curtis (2004) think that the educational context in which it appears can have a role on determining the type of washback. They summarise this educational context with four *wh* questions and a *how* question. While four *wh*

questions stand for the people *who* teach and manage the program, the school *where* the teaching and testing take place, *when* the program takes place (including the length of the program and a particular testing tool), the reason *why* the tool is adopted, and *how* stands for different teaching and testing methods are applied by the people in that context. The teachers and the administration of a school who make important decisions on the methods of teaching, length of teaching, and the rationale behind the methods and length can lead to washback to be positive or negative. For example in a school, the administration may want teaching to be communicative in all classes and set communicative exams but two different teachers in the same school can apply different methodologies in their classes. One of them may have more grammar-based classes while the other may have more communication-based classes. In the class of the latter, more positive washback is expected since the direct tests can foster practising the skills to be tested. In sum, educational context can have great impact on positive or negative washback. Spratt (2005) also sees the teacher as having a key role on the type and intensity of washback, which constitutes a part of *who* question of Cheng and Curtis (2004).

Bailey (1996, p.263-264) categorizes the effects of washback, into two main headings: 'washback to the learners' and 'washback to the programme'. According to her, while 'washback to the learners' is about supplying test-derived information to the test-takers which leads to direct impact of the tests on the test takers; 'washback to the programme' refers to supplying the test-derived information to the 'teachers, administrators, curriculum developers, counsellors, etc.'. It can be concluded that washback to the program affects the test-takers indirectly.

Washback is not only a relationship between testing and teaching / learning, but also a relationship between tests and curriculum, program and materials. Hughes (as cited in Bailey, 1996, p.262) believes that a test can affect three components, namely participants, process, and product. By participants he means all of those 'whose perceptions and attitudes towards their work may be affected by a test'. He defines the process as the 'actions taken by the participants which may contribute to the process of learning'. He sees the product as 'what is learned and the quality of learning'. He thinks that in the first place participants' attitudes and perceptions are affected by the tests, then the participants get in a process doing something according to the test, and finally this process lead to product, which is beneficial washback. Building on Hughes's model (as cited in Bailey, 1996) and Alderson and Wall's (1993) washback hypotheses, Bailey (1996) created a figure (Figure 1) to investigate how washback works. This figure clearly shows that a test can have a direct impact on the participants who are involved in the process of learning, and this involvement leads the products peculiar to participants.

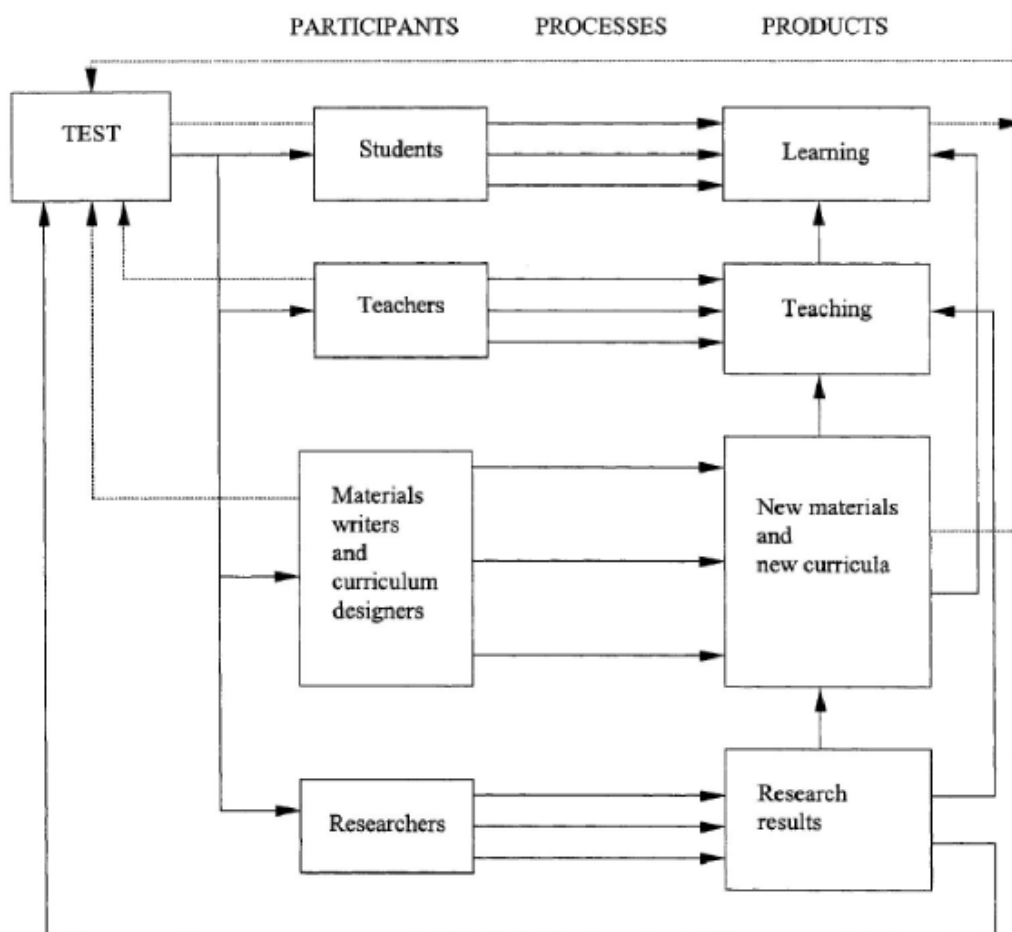


Figure 1 - A basic model of washback

### Washback Hypotheses

Alderson and Wall (1993, p.120-121) list a number of washback hypotheses, which have been referred in nearly all the washback studies. The hypotheses reveal how complex washback is. This present study was inspired by them and took all these hypotheses into account while doing the research.

1. A test will influence teaching.
2. A test will influence learning.
3. A test will influence what teachers teach.
4. A test will influence how teachers teach.

5. A test will influence what learners learn.
6. A test will influence how learners learn.
7. A test will influence the rate and the sequence of teaching.
8. A test will influence the rate and the sequence of learning.
9. A test will influence the degree and depth of teaching.
10. A test will influence the degree and depth of learning.
11. A test will influence attitudes to the content, method, etc. of teaching and learning.
12. Tests that have important consequences will have washback.
13. Tests that do not have important consequences will have no washback.
14. Tests will have washback on all learners and teachers.
15. Tests will have washback effects for some learners and some teachers, but not for others.

#### Studies Investigating Washback Effects

Disappointingly, there has been little empirical research done on washback effects in educational contexts. When language education is considered, it is much more disappointing.

The washback effect of tests has generally been associated with high-stake tests, i.e. those which are used for making important educational and professional decisions, such as admissions, graduation, employment, or promotions, and therefore affect people's futures' (Munoz and Alvarez, 2010, p.33). Studies done so far have investigated the washback effect of some worldwide-known high-stake tests such as TOEFL, TOEIC, and IELTS (Alderson and Hamp-Lyons, 1996; Newsfields, 2005;



Reynolds, 2010; Rhami and Nazland, 2010) and some nationwide-known high-stake tests (Amengual-Pizarro, 2009; Cheng, 1997; Hsu, 2009; Manjarres, 2005; Mohammadi, n.d.; Shih, 2009; Watanabe, 1996; Yıldırım, 2010). Since it is very difficult to measure the effect of washback, these kinds of studies have generally been qualitative ones. Though there are many studies investigating the washback of high-stake tests, there is little research on washback of speaking tests.

Rahimi and Nazland's study (2010) is one of the studies which focused on the washback effect of one of the world-wide known speaking tests. They conducted a study of the washback effect of IELTS preparation courses to learn students' perceptions of their speaking instruction. 60 Iranian students studying via e-learning IELTS courses or through non IELTS e-learning courses took part in the study. They had a six week-course. Students expressed what they thought about the speaking instruction through questionnaires conducted at the beginning and end of the courses. The same questionnaires were used for both groups. There were a number of differences in the perceptions of the two groups. In particular, the learners reported that they had learned different things; teachers on the different course types had different goals; differences in course contents were driven by differences in learner expectations; and the group getting ready for IELTS had more test-related content. In general, the IELTS exam washback was judged to have had a negative effect on learners and the programmes since the speaking skill was not given enough importance during the course just because it is not tested in IELTS.

As for nation-wide tests, one empirical study, conducted by Andrews et al. (2002), targeted washback of the Hong Kong Advanced Supplementary (AS) Use of English (UE) oral examination. The researchers investigated if the addition of an oral

component to the UE exam influences students' spoken English performance. Students take the UE in their final year of schooling to gain acceptance to university. This oral component (UE) had a two-part design. The first part included an individual oral presentation based on a text and in the second part a group discussion was held. The UE accounted for 18% of the total score of AS. Questionnaires were used to compare the views of the 'innovators' (the members of the Working Party who designed UE oral component) and implementers (the teachers of Secondary 6 and 7 classes) (Andrews et al., 2002). There were 31 students in each group from 1993-94-95 secondary 7 cohorts. The 1993 group did not get prepared for the oral examination since the UE started to be applied in 1994. Thus, the 1994 and 1995 groups were the first groups which took UE. The oral performances were video-taped and rated by eight assessors. According to the compared ratings, no significant differences can be seen between the mean performances of the 1993 and 1994 groups. However, the 1995 group had higher scores than the other two groups. Thus, it can be said that tests influence what students learn and UE Oral Component might have had a positive influence on students' spoken English performance at the end of Secondary 7.

Caine (2005) focused on the mismatch between the levels of curriculum planning adopted by the Japanese Ministry of Education and actual classroom implementation. He examined the effects of existing English tests in Japan, which are used to test speaking and writing indirectly. He also proposed an original direct test of speaking and investigated the washback effects of the new and trialled speaking test. Classroom observations and teacher and student questionnaires were used to collect the data. Seven Japanese high school teachers of English and two

groups of students, in total 46, were the subjects of the study. The results suggest that despite the official communicative changes made in the syllabus of Japanese Ministry of Education in order to solve the problem that Japanese students have in using the language for purposeful communication, teachers were still seen applying a more grammar-based methodology in their classes. The results suggest that changing the examination may have an effect on the methodology of the teacher in that more communicative assessment may lead to communicative approach to teaching. However, positive washback can occur on that issue when comprehensive in-service teacher training programs are combined with the changes in order to train teachers on communicative teaching and testing.

Ferman (2004) also conducted a study on the washback of an EFL National Oral Matriculation Test, which is held in Israel, to teaching and learning. The EFL National Oral Matriculation Test is taken by high school grade 12 students in order to enter university. It has been a component of the National Matriculation Examination since 1986 and has a 20% weight for English subject in total. It is administered just after the national matriculation exam and has four sections, which are extended interview, modified role-play, an extensive reading part, and a literature component. The study was conducted in three different types of high school and three different levels of classes. The subjects of the study were 18 EFL teachers, 120 students, and 4 EFL inspectors. Structured questionnaires, structured interviews, open interviews, and document analyses were used in order to collect the data. The results suggest that there was a strong washback effect of EFL National Oral Matriculation Test on the educational processes. The attention paid by the teachers, students, and parents, content, allocation of time for developing speaking skills, and

anxiety levels of the teachers and students were all affected by the test. It was seen that while the existence of the test promoted learning oral skills, it narrowed the scope and content of teaching and learning processes, led teachers to feel more pressure to cover the material, and increased the anxiety levels of the teachers' and students'.

Munoz and Alvarez (2010) aimed to explore the possibility of creating positive washback by focusing on some of the principles underlying the Oral Assessment System (OAS). The OAS was developed in 2001 at the language centre of a small private university where the researchers worked in Colombia, South America. The participants were 14 EFL teachers and 110 college students. A comparison and an experimental group were formed. Although the OAS was used in both groups, the experimental group was trained on the use of the OAS and how to teach their students to use the rubrics in the OAS. In this way, students could assess themselves. The experimental group had periodical meetings. However, in the comparison group it was the teachers' own decision what to assess, when and how to assess the students. All teacher and student surveys, class observations, and external evaluations of students' oral performance were used to collect the data. The study has three main conclusions. First, washback may be fostered by informing students of assessment procedures and scoring scales, specifying objectives, and structuring assessment tasks. Second, positive washback will be promoted when both teachers and students clearly establish the connection between educational goals and assessment. Third, assessment and the use of self-assessment mechanisms foster washback to the learners as they can take control of the assessment.

Although Poonpon (2010) did not aim to investigate the washback effect of speaking tests but rather examined how oral language assessment could be integrated into an English language class, one part of her findings is directly related to washback. Through her study she aimed to get students' opinions about the integration of speaking tests into their English class and their speaking ability before and after taking the speaking tests through a questionnaire. The students stated that their level of English improved after they started to take the speaking exam. It is the result of the direct speaking tests which has positive washback on the students' improvement in English.

#### Conclusion

Most of the washback studies in the literature have investigated the effects of worldwide-known or nationwide-known high stake tests. When it comes to the washback effect of speaking tests, there are very few studies. Existing studies have not attempted to examine the washback effect of existing classroom-based speaking tests. This present study intends to fill this gap by exploring students' and teachers' perceptions of the washback effect of classroom-based speaking tests.

## CHAPTER III: METHODOLOGY

### Introduction

This present study seeks teachers' and students' perceptions of washback effect of classroom-based speaking tests and teachers' and students' attitudes towards and beliefs about teaching and testing speaking at a Turkish university prep school. This chapter gives the methodological details of the study. The chapter starts with the research questions, answers of which are sought to be found. In the first section, the information about the setting and participants are provided. The upcoming sections present the instruments used to collect the data and data collection procedure. The final section focuses on the data analysis part.

### Research Questions

This study will investigate the following research questions:

1. What are teachers' and students' attitudes towards and beliefs about teaching and testing speaking?
2. What are the washback effects of classroom-based speaking tests, as reflected in teachers' and students' perceptions?

### Setting

This study was conducted at Akdeniz University, which is a state university, School of Foreign Languages. The school has been providing English preparatory class programs since 1998. The School of Foreign Languages also offers prep classes in German and French. There are 52 students enrolled in prep class in German and

26 students enrolled in prep class in French this year. As for English, while 459 students are studying compulsory prep class, 352 students are studying elective prep class. In total, 889 students are getting the advantage of studying at Akdeniz University prep classes in 2010-2011 academic year. 83 language instructors are doing their best to fulfil the needs of the students under the roof of Foreign Languages School.

At the very beginning of each academic year, generally in September, the school offers a proficiency test for the students who have failed the prep class before and for the freshmen of that year. According to the results of this test, the students who have 70 and above out of 100 pass the prep class. The students who have studied prep class before and fail in September again try their chance to pass in the next proficiency exam, which has been started to be held also in January. The freshmen who fail the test have a placement test, according to the grades of which they are placed in their classes. Students of English Language Teaching, Medicine and Civil Aviation (evening classes) are grouped with students of their own departments, again graded according to their scores on the placement tests.

Since the 2008-2009 academic year some of the prep classes have been compulsory, though, some others have been offered as electives. Since Akdeniz University is a Turkish medium university, just a few departments, namely Medicine, English Language Teaching, Economics, and Management provide at least 30% English in their departments. The students of these departments have to study prep class unless they have passed it at the beginning of the year in the proficiency test. In all the other departments, studying prep class is offered as an elective.

In order to evaluate the total grade of the students, first of all the cumulative average scores are calculated. This includes 6 midterms, 19 quizzes, writing portfolios and teachers' opinion marks. In all of the tests, all the language skills, including speaking, are tested. The midterms have 70%, the quizzes have 20%, and both the writing portfolio and teachers' opinions have 5% weight each. Students whose cumulative average scores are at least 50 can take the final exam. Lastly, a final course grade is given, 60% of which is determined by the cumulative average and 40% by the final exam score. Students need a final score at least 70 in order to pass the prep class.

After the placement test, the students take their places in A1, A2, and B1 levels. While A1 level students have 25 hours of English classes a week, A2 and B1 level students have 20 hours a week. In each level, four hours of writing and four hours of reading classes are included in the program. This year a new system has been adopted by the administration. Each midterm is used as an indicator of the students' success and after each midterm the students who get 70 and above are permitted to go on studying the next units. For those who get below 70, new classes are formed to repeat the units with new teachers and new textbooks. However, each level can be repeated once. Even if the students are unsuccessful after they repeat the class again, they are allowed to go on studying the next units. Moreover, in the repeat classes, the students do not study writing and reading classes again.

#### Speaking Tests Administered at Akdeniz University Prep Classes

The prep class students of Akdeniz University Foreign Languages School have six midterms, 19 quizzes, and a final exam during the academic year. All the



midterms and the final exam have a speaking component which has a 20% weight on scoring. Each of the quizzes measures the four main skills respectively. There are therefore at least four speaking quizzes administered during the year. The speaking tests are based on Preliminary English Test (PET) and Key English Test (KET). The main types of the exams are making dialogues, photo description, and answering the questions. Some announcements about the type of the speaking task, which will be asked in the speaking test, are made for teachers and students when it is thought necessary. In these speaking tests there are two interlocutors, one of whom is one of the class teachers of the examinees. The other interlocutor will be another instructor who teaches also at that level. Before the speaking tests no training is provided for the teachers. In the speaking document envelopes there is guidance for teachers about the task/s. There are two evaluation sheets for the interlocutors but how to use them is up to the interlocutors. Some partners evaluate the performances together and use one sheet while some others use individual sheets and then add up their marks and divide them into two. Although they have these evaluation sheets as criteria, they are not informed about what 0, 1, or 2 mean in the criteria.

### Participants

The teacher participants of his study were 45 instructors of English at Akdeniz University Foreign Languages School. While 40 of them (89.9%) of them were female, five of them (11.1%) were male.

	Age	Experience in teaching (years)	Experience at Akdeniz University	Experience in administering speaking tests
Mean	33.84	11.20	7.11	7.09
Minimum	22	1	1	1
Maximum	62	40	23	18

Table 1- Descriptive statistics for age, experience in teaching at Akdeniz University, and in administering speaking tests of the instructors

While the youngest instructor was 22 years old, the oldest instructor participated in the study was 62 years old. The average age was 34. Experience in teaching English was eleven years on average. The instructors had been working at Akdeniz University for an average of seven years when the data were collected. Experience in administering speaking tests varied between one to 18 years. The average of experience in administering speaking tests was seven years.

	BA	MA
English Language Teaching	29	14
English Language and Literature / American Culture and Literature	13	3
Educational Sciences	0	5
Translation and Interpretation	2	0
English Linguistics	1	0
Total	45	22

Table 2 - BA and MA majors of the instructors

As table 2 shows, a great number of the instructors, not surprisingly, studied ELT as their BA. Just under half of the teachers had done, or were in the process of doing their masters degrees. Again, the most popular major was ELT. One participant had also completed a PhD (in Educational Sciences).

The students who participated in this study were intermediate level preparatory class students at Akdeniz University School of Foreign Languages. All the participants, 307 in total, took part voluntarily. While 129 (42%) of them were male students, 178 (58%) of them were female students.

266 students answered the question asking their ages in the questionnaire.

The youngest participants were 17 years old and the oldest participants were 25 years old. The average age was 19.53.

	Frequency	Percentages
High School / Vocational High School Graduates	110	35.8%
Anatolian Technical / Anatolian Vocational High School Graduates	13	4.2%
College Anatolian / Science High School Graduates	104	33.9%
Total	227	73.9%

Table 3 - The type of high school the students graduated from

While 80 students did not reply to the question asking the type of high school they graduated from, 227 students answered it. Analysing the type of the high schools the students graduated from, the general density of English classes they had up to this prep class year can be concluded. While the 35.8% of the students who graduated from high schools or vocational high schools must have had English classes four hours a week for at least three years, the 38.1% of the students who graduated from other types of school should have taken relatively denser English classes in their high school years. The average number of English classes offered in the first year of Anatolian, science, Anatolian technical / vocational high schools, and colleges is 10 hours a week. Although Anatolian and science high schools and colleges give more emphasis to English, Anatolian Technical and Anatolian Vocational High Schools also offer ten hours of English a week.

While seven students stated that they had never had any English classes before they were enrolled in the preparatory class this year, five of the students stated

that they had studied preparatory class before either when they were in high school or at another university.

FACULTY	Frequency	Percent
Faculty of Agriculture	8	2.6
Faculty of Literature	25	8.1
Faculty of Economics and Administrative Sciences	170	55.4
Faculty of Engineering	52	16.9
Vocational School of Social Sciences	25	8.1
School of Physical Education and Sports	2	.7
Faculty of Communication	25	8.1
Total (N)	307	100

Table 4 - The future faculties of the students

As table 4 reveals, while the future students of Akdeniz University Faculty of Economics and Administrative Sciences make up the majority of the participants, at 55.4%, the future students of School of Physical Education and Sports were the smallest group, at .7%.

		Frequency	Percent
Valid	16-20	106	34.5
	12-16	133	43.3
	8-12	27	8.8
	4-8	2	.7
	0-4	1	.3
	Total	269	87.6
Missing	System	38	12.4
Total		307	100.0

Table 5 - The ranges of the speaking scores of midterm

		Frequency	Percent
Valid	85-100	73	23.8
	70-85	108	35.2
	55-70	74	24.1
	40-55	11	3.6
	25-40	1	.3
	0-25	1	.3
	Total	268	87.3
Missing	System	39	12.7
Total		307	100

Table 6 - The ranges of speaking quiz scores of the students

Table 5 and 6 show the average speaking scores of the students from the midterms and quizzes. In both of the tests, the mode score range is the second option ranges. As table 5 shows 43.3% of the students got 12-16 points in the midterms and 34.5% of the students got 16-20 points. Just 3% of the students got 0-4. As can be seen in table 6, in the speaking quizzes the most frequent range is 70-85 point option, at 35.2%. 85-100 range follows it with 23.8. In total, the students who got over 70 make up 59% of the students. Just 6% of students were in the 40-0 range. Taking these percentages into account it is possible to conclude that the students are successful at the speaking sections of the midterms and speaking quizzes.

### Instruments

In order to get teachers' and students' perceptions of washback effect of classroom-based speaking tests and teachers' and students' attitudes towards and beliefs about teaching and testing speaking; teacher and student questionnaires and teacher and student interviews were used. The questionnaires comprised 5-point Likert-scale items (1=strongly disagree; 5=strongly agree). In the first parts of the

questionnaires, some basic information about themselves was asked to the participants (the results of which were given in the previous section).

One of the questionnaires was applied to teachers to get their perceptions of the washback effect of speaking tests and their attitudes towards and beliefs about teaching and testing speaking. As well as answering 32 items, the teachers also gave the personal information reported in the previous section.

The other questionnaire investigated students' perceptions of the washback effect of speaking tests and their attitudes towards and beliefs about teaching and testing speaking, with 34 items. In student questionnaires their names, classes, age, sex, the type of high school they graduated from, their average midterm and quiz grades in speaking, and their departments were required to be filled in.

Interviews, as second data collection tool in this study, were held with 6 teachers and 7 students, who had already answered the questionnaires. While most of the participants were chosen randomly, just two of them (a teacher and a student) were chosen on purpose since their questionnaire responses were a bit different from the other participants. To analyze the answers, interviews were recorded and transcribed.

The questionnaires and interviews were held in participants' L1, which is Turkish, in order to prevent any communication breakdown. Since measuring students' comprehension ability was not aimed in this study by the questionnaires or interviews, it was hoped that giving the tools in participants' L1 would increase the reliability of the collected information. It was also thought that holding the interviews in Turkish would increase the sincerity between the researcher and the participants in face to face communication.

For the questionnaires, all the items were originally written in English owing to the fact that the researcher believed that she expressed what she had in her mind better in English. This may stem from the fact that the literature she had reviewed in English led some concepts settled down in her mind in English.

A back translation method was adopted to ensure accuracy of translation of the questionnaire into Turkish. The items written in English were translated into Turkish by the researcher. Then, a proficient non-native speaker of English translated the Turkish versions into English again. The original version and the version translated by the proficient non-native speaker were given to the native speaker of English to compare the two translations. The English and Turkish versions of the questionnaires can be seen in Appendices E, F, G, and H.

### The Piloting Procedure

In November 2010 a first trial for student questionnaires, which were compiled using the questionnaire items from Caine's (2005) and Poonpon's (2010) studies, was done at Akdeniz University in one of the Medicine prep classes. This trial, which was done before the first piloting, just aimed to help the researcher to prepare better items, organize the scales in a better way, and most importantly to identify the direction of the study clearly. After the whole literature was reviewed on the topic, the researcher was able to prepare her own scales and to the point questionnaire items. However, Q.6 and Q.31 in the student questionnaire clearly show the heavy influence of Caine's (2005) student questionnaire items which are 2 and 3 (Appendix M). In the second week of March 2011, the latest teacher and student questionnaires were piloted for the first time to determine if there were any

unclear items because of the translation or word ordering and to see the reliability scores of the scales.

In the randomly chosen classes, 51 students were asked to complete the student questionnaires. Before they started answering the questionnaires, they were asked to mark the items which they had difficulty in understanding or found problematic for any other reasons. They were also asked to explain their reasons why they found unclear or problematic very briefly in the space left for further comments in the questionnaires. According to the results, items which were found unclear or problematic were revised. Another contribution of this piloting procedure was that it helped to determine the necessary time to administer the students' questionnaire. It was observed that the average time to fill out the questionnaire was between 15-20 minutes. The results were analyzed quantitatively and reliability scores of the scales were questioned. The scales which had low reliability scores were revised and the scales were re-organized. Since there were many changes in the new form of the questionnaire it was piloted again with 35 other students again at an Akdeniz university prep class. It was analyzed quantitatively and some scales which had low reliability scores were re-organized and the necessary modifications were made accordingly for the main study.

The teacher questionnaires were piloted with 17 instructors of English, who were doing their MA degree at TEFL program at Bilkent University and working in different parts of Turkey. The changes that the instructor participants offered were taken into consideration and necessary changes were done. In addition, the questionnaire results were analyzed quantitatively and the scales which had low reliability scores were revised.



The Cronbach's alpha coefficient was calculated to examine the internal consistency of the data collection tool by using the collected data. The Cronbach's alpha coefficient can be seen in Appendices I, J, K, and L for the scales of the two questionnaires.

### Data Collection Procedures

In the middle of February 2011, Akdeniz University School of Foreign Languages vice principals were informed about the study. In the first week of March, the head of the school was informed about the content and aim of the study, the necessary number of participants, and classes. With the consent of the administration of Akdeniz University Foreign Languages School, the study was piloted for the first time in a B1 level Medicine prep class at the second week of March and necessary changes were made accordingly. The next piloting took place a week after the first piloting. Another group of students studying prep class at Akdeniz University were asked to answer the questionnaire items. After analyzing the results, the questionnaire was revised and necessary modifications were made.

In the second week of April, last 190 A2 level students just started to study in B level classes. There were also an existing 621 B1 level students. In order to determine the participants of the main study, the researcher herself explained the aim of the study by visiting the classes in their class time, except Medicine and English Language Teaching classes. The students were told about the study, and volunteer students that were willing to take part in this study were given a consent form (the English and Turkish versions of consent forms are in the Appendices A, B, C, and D). They were also given the questionnaires in the same lesson hour. The researcher

herself stayed in the class in which the questionnaire was being administered in order to prevent any misunderstanding in the questionnaires and to provide uniformity in administering the questionnaires. The participants' questions were answered in detail in all the classes. Visiting all the classes and administering the student questionnaires to all the classes by the researcher took 2 days in total. 307 students took part voluntarily in the study.

The teacher questionnaires were administered in the first week of May. 45 instructors answered the questionnaire. While 41 of them handed the printed versions of the questionnaires in, four of them answered the questionnaires online and sent them via e-mail.

After the questionnaires were analyzed, the interviews started to take place. In the first instance the students who had stated that they would like to take part in the interviews, though they had not been asked in the questionnaire, were chosen as the student interviewees. However, since some of these students were absent on the day of the interviews, some other students were also asked to be interviewed by the researcher herself by visiting the classes and some of the students kindly accepted that offer and became volunteers. There were two participants whose answers for the questionnaire items were different than the other students' in that they were more on the negative side of the continuum. They were specially asked to be interviewed. Although both of them agreed to be interviewed at first, one of them did not take part and the other student wanted his/her classmate to be with him/her during the interview. This request was not rejected since the student was needed because of his/her different points of view. Student interviews were held in lesson hours of the students in some available classes. Five of them individually and two of them in

pairs, totally seven students were interviewed. Before the interviews took place, the students were reminded of the study, were asked to sincerely state what they really thought about the issues, and their permission was taken to record the interviews. The students were informed on the topics they were going to face during the interviews. The length of the interviews varied between nine and 23 minutes. Appendices N and O present the beginning of the interview done with student 2 for illustrational purposes.

Teacher interviewees were the colleagues of the researcher. Two out of six teachers were chosen on purpose since their answers were different from the others in that the answers of one stood on the negative side of the continuum and those of the other were different than the others' in that s/he both agreed with Q.4 and Q. 21 in the questionnaire (See Appendix E). They accepted the interview offer just like the other four interviewees, who were the volunteer instructors of English who wanted to help the researcher to carry out her study easily. Before the interviews, the interviewees were reminded of the study and the subscales which formed the topics of the questions that were to going to be asked. The teacher interviews took minimum four and maximum 19 minutes.

The interview questions clearly reflected the subscales of the questionnaires about teachers' and students' attitudes towards and beliefs about teaching and testing speaking (See Appendices I and K for the subscales) and teachers' and students' perceptions of the washback effect of classroom-based speaking tests (See Appendices J and L for the subscales). Though similar, not all the interviews went through the same sequence of questions since some of the interviewees had already

mentioned the issues before being asked. Appendices P and Q present the beginning of the interview done with teacher 1 for illustration purposes.

During the interviews, not much was written down by the interviewer in order not to lead communication breakdowns by losing eye contact.

### Data Analysis

In this study, while quantitative data was collected through questionnaires, qualitative data was collected through interviews. The quantitative data was analyzed using Statistical Package for Social Sciences (SPSS).

### Conclusion

This chapter provided information about the methodology of this study in terms of setting and participants, instruments, and data collection procedures. The data analysis will be explained in detail in the upcoming chapter.

## CHAPTER IV: DATA ANALYSIS

### Overview of the Study

The main purpose of this study was to learn teachers' and students' perceptions of the washback effect of classroom-based speaking tests, which were implemented as quizzes, as a part of midterms and the final exam at the preparatory class program of Akdeniz University School of Foreign Languages. Moreover, teachers' and students' attitudes towards and beliefs about teaching and testing speaking were also addressed. This chapter presents the results by providing the analysis of the collected data.

The research questions posed for the study were:

1. What are teachers' and students' attitudes towards and beliefs about teaching and testing speaking?
2. What are the washback effects of classroom-based speaking tests, as reflected in teachers' and students' perceptions?

Before presenting the results, some information on the scales of the questionnaires will be provided. The results will be presented in four phases. The first phase will present the analysis of the Likert-scale questionnaires completed by the teachers and students and the interviews done with the teachers and the students in order to get information about the teachers' and students' attitudes towards and beliefs about teaching and testing speaking. Then, the same and similar questions about the teachers' and students' attitudes towards and beliefs about teaching and testing speaking which were asked both to the teachers and the students will be compared and contrasted. In the third phase analysis of the Likert-scale

questionnaires and the interviews done with the teachers and the students will be carried out with the aim of providing information on the teachers' and students' perceptions of the washback effects of classroom-based speaking tests. Finally, the same and similar questions about the teachers' and students' perceptions of the washback effects of classroom-based speaking tests which were asked both to the teachers and the students will be compared and contrasted.

### Forming the Scales of the Questionnaires

In the teacher questionnaire (See Appendices E and F), there are 32 items, which form nine scales all together (See Appendices I and J). As well as having scales on teachers' perceptions of washback effects of classroom-based speaking tests, it also has scales on teachers' attitudes towards and beliefs about teaching and testing speaking. The aim of collecting data about general attitudes towards and beliefs about teaching and testing speaking is to provide background data to the main question of washback. The teachers' attitudes towards and beliefs about teaching and testing speaking were investigated in six scales (See Appendix I) which are listed below together with their subscales where appropriate.

- a. Do teachers have positive attitudes towards the possibility of teaching speaking?
- b. Do teachers have positive attitudes towards the importance of teaching speaking?
- c. Do teachers believe that speaking should be tested?
- d. Do teachers believe that speaking can be tested?

- Do teachers believe that testing speaking skills is difficult?
  - Do teachers believe that speaking ability can be measured accurately?
  - Do teachers believe that speaking tests' scores are an accurate reflection of speaking ability?
- e. Do teachers think that speaking skills should be tested through speaking?
- Do teachers believe that speaking ability can be tested effectively through writing?
- f. Do teachers believe that the results of speaking tests can be used as a reliable diagnostic tool?

For the teacher questionnaire, one piloting procedure was applied. The scales which had low reliability scores were examined. The reasons why they had low reliability scores were investigated. Some scales which including both behaviour and attitude questions were revised in that these two different type of questions were divided into different scales. In addition, much more straightforward questions were asked instead of some questions. In some scales some questions were revisited by reverse-coding.

Once the final version of the questionnaire had been administered and an initial analysis of the data conducted, it was decided to treat some of the questions which had been written as a part of a scale individually as separate items since their reliability scores showed that they did not match with the scales they had been written for.

The major changes will be presented here in turn. The scale on teachers' attitudes towards the importance of teaching speaking (scale a) originally had three

items, but the reliability analysis revealed that the main question of the scale, which is Q.12, did not fit well with the scale. Due to the fact that question 12 is the main item, it was analysed individually under this scale:

Q. 12. I think that teaching speaking skills is important.

As for the scale about whether speaking can be tested (scale d), the Alpha .27 showed that the three items needed to be treated individually. The sub-questions listed below were therefore analysed individually:

Q. 20. I think that testing speaking skills is difficult.

Q. 3. Speaking skills can be measured accurately.

Q. 14. Speaking tests do not reflect speaking skills accurately.

With regard to speaking as the medium of testing speaking (scale e), questions 21, 25, and 4 were designed to go together as a single scale. However, Q.4 was seen to decrease the reliability of the scale, so questions 21 and 25 were analysed together as a scale, but Q.4 was analysed individually. The reason why Q.4 did not fit well with the scale may stem from the fact that the teachers may think that speaking skills can also be tested in other ways as an alternative but not through writing:

Q. 21. Speaking skills should be tested through speaking.

Q. 25. Speaking skills of students can be effectively measured without requiring them to speak.

Q. 4. Speaking skills can be measured effectively through written tests.

As for the student questionnaire, it was piloted twice with different students. It has 34 items (See Appendices G and H), which aim to investigate the students' attitudes towards and beliefs about teaching and testing speaking and their



perceptions of washback effect of classroom-based speaking tests. Four scales aim to reveal the students' attitudes towards and beliefs about teaching and testing speaking. The main questions of the scales are listed below. Since the reliability analysis of the main questionnaire did not give important breakdowns or coherence problems, none of the questions were analysed individually.

- a. Do students have positive attitudes towards learning speaking?
- b. Do students believe in the importance of testing speaking?
- c. Do students enjoy speaking tests?
- d. Do students think that speaking tests can really show their speaking ability?

With regard to the washback effect scales, the teacher questionnaire has three scales.

- a. Do teachers tailor their speaking classes according to speaking tests in terms of content and allocation of time for speaking activities?
- b. Do teachers have positive attitudes towards organising extra speaking activities in classes, which are not alike with the activity that will be asked in the speaking test?
- c. Do teachers believe that speaking tests have a positive effect on their students' speaking ability?

As for tailoring their speaking classes according to speaking tests in terms of content and allocation of time for speaking activities (scale a), while five questions were analysed as a scale, one of the questions (Q. 5) was treated individually since it

correlated weakly with the rest of the scale. This weak correlation may stem from the fact that Q.5 focuses more on the days leading up to the speaking test while the other questions are more general questions:

Q. 5. I spend more time on speaking parts in the days leading up to the speaking test.

Teachers' beliefs about speaking tests' positive effects on their students' speaking ability was written as a scale (scale c). However, the main study results showed that while question 16 and 24 made a good pair, all the other four questions needed to be analysed individually. 16 and 24 form a good pair since they both focus on the fact that speaking tests help students to notice their strengths or weaknesses in their speaking performances. However, no other questions worked well together, so each was analysed individually:

Q. 16. Speaking tests help students to notice the weaknesses in their speaking performances.

Q. 24. Speaking tests help students to notice the strengths in their speaking performances.

Q. 7. Getting ready for the speaking test improves the general speaking skills of students.

Q. 10. Speaking tests encourage students to speak more in lessons.

Q. 27. Students can also use many of the things that they have studied for the test, in lessons after the test.

Q. 30. Students tend to forget lots of the things they have studied for the speaking test, after the test.

As for the students' perceptions of washback effect of classroom-based speaking tests, the questionnaire again has three scales.

- a. Do students believe that speaking tests have positive effects on their speaking ability?
- b. Do speaking tests have an influence on what and how students learn?
- c. Do speaking tests have any effects on what students really do?

Although they all aimed to elicit speaking tests' influences on what and how students learn, items 5, 31, and 20 did not fit well with this scale and it was decided to treat them individually. What they share in common is that they are all about what students really do. The items in speaking tests' influences on what and how students learn scale were mostly about imaginary situations such as '*if speaking sections were given 10 points in total in midterms, I would participate less in speaking activities in lessons*'. Though they are all about what students really do, they are not related to each other so they were treated individually:

Q. 5. I try to practice the speaking activities which will be asked in speaking tests in daily life.

Q. 31. I participate more in speaking parts in the days leading up to the speaking tests.

Q. 20. I give more importance to the speaking parts which will be asked in the test.

There was only one question (Q.17) which does not belong to any of the scales. Just like 5, 31, and 20, it was also written for speaking tests' influences on what and how students learn, but not only did it fit the scale but it also did not fit

with the new scale (scale c) formed for 5, 31, and 20. The reason why it did not fit the scale c may be that it states an opinion rather than providing information on what the students really do.

Q. 17. Even if they were not tested, speaking skills should have a place in lessons.

### Analysis of the Questionnaires

1. What are teachers' and students' attitudes towards and beliefs about teaching and testing speaking?

#### *Teachers' attitudes towards and beliefs about teaching and testing speaking*

Table 7 provides mean scores and standard deviations of the scales and sub-questions to reflect the teachers' attitudes towards and beliefs about teaching and testing speaking. The administered questionnaire is a Likert-scale questionnaire. For all the questions, while 5 represents 'totally agree' option, 1 represents 'totally disagree'. Number 3 stands for 'undecided'. The responses to each scale and sub-question will be discussed in turn.

	Mean	SD	Percentages of the participants who		
			TA/A	UD	D/TD
a. Do teachers have positive attitudes towards the possibility of teaching speaking?	4.46	.53	96.6	1.1	2.2
b. Do teachers have positive attitudes towards the importance of teaching speaking?	4.73	.44	100	0	0
c. Do teachers believe that speaking should be tested?	4.46	.60	91.1	6.65	2.2
d. Do teachers believe that speaking can be tested?					
• I think that testing speaking skills is difficult.	3.51	1.34	62.2	4.4	33.4
• Speaking skills can be measured accurately.	3.98	.69	80	17.8	2.2
• Speaking tests do not reflect students' speaking skills accurately.	2.73	.88	24.4	28.9	46.6
e. Do teachers think that speaking skills should be tested through speaking?	4.40	.78	87.75	7.75	4.4
• Speaking skills can be measured effectively through written tests.	1.93	1	8.9	13.3	77.8
f. Do teachers believe that the results of speaking tests can be used as a reliable diagnostic tool?	2.87	.62	32.42	27.1	40.42

Note: TA: Totally agree, A: Agree, UD: Undecided, D: Disagree, TD: Totally disagree

Table 7 - Descriptive statistics for teachers' attitudes towards and beliefs about teaching and testing speaking

- a. Do teachers have positive attitudes towards the possibility of teaching speaking?

The attitudes of the instructors of English at Akdeniz University towards the teachability of speaking skills in class are quite positive. The mean score for this scale, which is 4.46, shows that nearly all of the instructors believe that teaching speaking skills in classes is possible. The total percentage of the participants who totally agree and agree, 96.6%, also supports the mean score.

- b. Do teachers have positive attitudes towards the importance of teaching speaking?

The mean score for this question, 4.73, clearly reveals that nearly all the participants strongly agree with the importance of teaching speaking skills. Taking the mean scores of the questions 'a' and 'b' into account, it can be concluded that the instructors think that teaching speaking skills is possible and important. The percentage of the participants who totally agree and agree is also 100%.

Participant 5 sums up the reason for these feelings as follow:

(Participant 5)...Teaching speaking skills **is important**. The purpose of learning a language is already to be able to speak [and] express yourself.

- c. Do teachers believe that speaking should be tested?

The instructors again form a huge group who do not see the time spent for testing speaking as a loss of time. Moreover, the mean score, 4.46, displays the teachers' strong beliefs in the necessity of testing speaking. The percentage of the participants who totally agree and agree, 91.1%, also supports the idea.

Five of the six interviewees, representing nearly all the participants, also support testing speaking. Participant 2 thinks that testing their speaking is motivating for students.

(Participant 2)...It (speaking skills) should be tested because it is a motivation...The students want to see **the evaluation of their (speaking) performances in a concrete way...**

Participant 5 thinks that every skill which is taught should be tested.

(Participant 5)...**If we teach** it (speaking skills), **we should test it.**

d. Do teachers believe that speaking can be tested?

- Do teachers believe that testing speaking skills is difficult?

The mean score of the instructors' responses for the difficulty of testing speaking ability is 3.51. While 62.2% of the participants think that testing speaking skills is difficult, 33.4% of the participants disagree with this idea. Since more than half of the participants are on the 'agree' side of the continuum, it can be concluded that the instructors on the whole think that testing speaking skills is difficult.

- Do teachers believe that speaking ability can be measured accurately?

The mean score of the possibility of measuring speaking skills accurately is 3.98. While 80% of the participants think that speaking skills can be measured accurately, 2.2% of them do not think so. 17.8% are neutral. 80% of the participants constitute a good number to conclude that the teachers believe that speaking ability can be measured accurately.

The interview results for this scale are a bit different from the questionnaire results in that the numbers of the participants who think that speaking ability can be

measured accurately do not form the majority. However, participants 2, 3 and 5 raised the issue of importance of clear-cut and reliable criteria.

(Participant 2)...That is to say, we have **specific criteria**, or we prepare a certain criteria taking the ability we will measure into account...Of course it (speaking skills) can be measured.

(Participant 3)...It (speaking skills) **can be measured accurately if the criterion is determined in a clear-cut way...**

(Participant 5)...It (speaking skills) can be tested but **it seems to me controversial** how objective it is [and] to what extent it reflects the truth...**Criteria** is very important.

- Do teachers believe that speaking tests' scores are an accurate reflection of speaking ability?

The mean score for this scale is 2.73. While 46.6% of the instructors disagree with this statement, 24.4% of them agree with it. 28.9% of the participants are undecided about the issue. Based on the percentages, it can be concluded that teachers are closer to disagreeing with the item.

e. Do teachers think that speaking skills should be tested through speaking?

- Do teachers believe that speaking ability can be tested effectively through writing?

4.40 mean score for this scale reveals that the instructors who participated in this study strongly believe that the medium to test speaking skills should be speaking itself. The percentage of the participants who totally agree and agree is 87.75. As for testing speaking ability effectively through writing, 1.93 mean score shows that the teachers do not see writing as an effective alternative to test speaking. While just 8.9% of the participants see writing as an alternative, 77.8% of them do not.



In general, the interviewees reported that listening and reading can be used in a speaking exam but that they should not be evaluated. Two of the instructors (3 and 6) believe that while practising speaking ability, writing tasks can be used but not to test speaking ability.

(Participant 3)...When teaching speaking skills, **it can be taught through writing**, the structures they have to use etc. **but it should be tested through speaking**, I think.

One of the teachers (1) stated that writing can also be used to test speaking as an alternative but it is not preferred.

(Participant 1)... It **(writing) is not preferred** but when you have no other options it is an **alternative...**because it is **production**, the child (the students) can produce the same thing through **writing**.

- f. Do teachers believe that the results of speaking tests can be used as a reliable diagnostic tool?

The mean score, 2.87 is very close to being neutral. While 32.42% of the participants believe that the results of the speaking tests can be used as a reliable diagnostic tool, 40.42% of them disagree with this statement. 27.1% of the participants are undecided. As it can be seen in the percentages, the participants who agree and disagree on this issue are close to each other. The interviews show that nearly all the interviewees raised the issue of students giving unexpected performances in the speaking tests, but the percentages of the students who unexpectedly outperform or have poor performances in the speaking tests they have in mind are different. While some of the teachers think that it happens a lot, some of them state that it happens rarely. While the teachers who think that it happens rarely believe that the results of speaking tests can be used as a reliable diagnostic tool, the

ones who think that it happens a lot do not believe so though they experience the same thing.

(Participant 3)...A student who is very good at speaking in classes cannot be able to show the same performance in a speaking test or the students who do not participate [to the speaking parts] in classes will be shining.

Participant 6 is more interested in the task itself and how the task is administered when the subject is related to reliability.

(Participant 6)...I had students who surprised me a lot, **but it depends on how it (a speaking test) was put into practice.** The English of a student was very bad but since the dialogues were given beforehand (before the exam), he learned it by heart [and] studied it. He got a very high mark.

#### *Students' attitudes towards and beliefs about teaching and testing speaking*

The students' responses to the questionnaire about attitudes towards and beliefs about teaching and testing speaking are summarised in Table 8. As before, 5 represents 'totally agree' option, 1 represents 'totally disagree'. 3 stands for 'undecided'.

	Mean	SD	Percentages of the participants who		
			TA/A	UD	D/TD
a. Do students have positive attitudes towards learning speaking?	4.17	.52	83.16	11.6	5.18
b. Do students believe the importance of testing speaking?	4.18	.58	83.81	10.75	5.53
c. Do students enjoy speaking tests?	2.95	1.02	36.65	26.2	37.15
d. Do students think that speaking tests can really show their speaking ability?	2.82	.80	40.85	31.45	27.1

Note: TA: Totally agree, A: Agree, UD: Undecided, D: Disagree, TD: Totally disagree

Table 8 - Descriptive statistics for students' attitudes towards and beliefs about teaching and testing speaking

a. Do students have positive attitudes towards learning speaking?

The mean score 4.17 for the scale about the students' attitudes towards learning speaking clearly shows that the students think that learning speaking is important and it can be learnt in classes. The percentage of the participants who totally agree and agree, 83.16%, also supports the mean score.

One of the interviewees stated that learning speaking in classes depends on the willingness of the students to learn it.

(Participant 3)... **Of course it (speaking) is something to be learnt in classes...** However, the most of the students need to be willing to do (to learn) this.

Being one of the representatives of 5.18% of the students, participant 6 does not think that speaking skills are learnt in classes. S/he stated that teachers give more importance to grammar points in speaking performances and this hinders improving speaking skills.

(Participant 6)...I think speaking ability is not learnt in classes because (grammar) rules are stuck to too much... Sometimes details are felt over (by the teacher) too much. For example a classmate can do it (speak English in class) but makes a mistake. Then, (the teacher) says that it should not be (said) in that way, it should be...

b. Do students believe in the importance of testing speaking?

Again with a high mean score, 4.18, the participants strongly believe that testing speaking skills is important. 83.81% of the participants totally agree and agree with the scale related to the importance of testing speaking.

On the importance of testing speaking, the interviewees stated that speaking tests' weight should be increased.

(Participant 1)...Speaking exams are the ones to which I give importance most...I think by **giving more weight its (speaking skills') importance should be adjusted...**

Although participant (7) thinks that speaking should be tested, s/he is not in favour of grading.

(Participant 7)...I think it (speaking) should be tested but it should be tested in order to see whether everything is going well in class or to see the students' performances. I mean, not to give grades....

c. Do students enjoy speaking tests?

The mean score, 2.95, shows that the students taking the speaking tests are neutral about enjoying or being nervous in speaking tests. The students who do and who do not enjoy speaking tests are nearly the same. While 36.65% of the participants stated that they enjoy speaking tests, 37.15% of them stated that they do not.

According to Participant 1, some students can be nervous in speaking tests but they can also enjoy the same speaking test.

(Participant 1)...I get a bit nervous before I start to take (the speaking test). However, when I see the questions or the pictures **my tension disappears...There are some cases (in speaking tests) which let us enjoy...**

d. Do students think that speaking tests can really show their speaking ability?

The mean score, 2.82, is again very close to being neutral about the statement whether speaking tests can really show their speaking ability. The percentages of the students who agree and disagree with the idea whether speaking tests can really show their speaking ability are 40.85% versus 27.1%. The percentage of the students who are neutral on the issue is 31.45%. While a bit more students agree with the idea, the diversity in the answers show that the students have different opinions about whether speaking tests can really show their speaking ability.

Participant 1 stated that s/he can show more than his/her real speaking ability in the exams.

(Participant 1)...Maybe brain works better under that excitement and stress...

Participant 2 thinks that s/he cannot show his/her speaking ability in speaking tests.

(Participant 2)...It is because of me since all the words I know disappear suddenly since I get very nervous. However, when I talk to a friend outside (the class) though little bit, you get relaxed but I get a grade in front of the teacher. As a result, it gets worse.

Participant 5 raised the issue of topic areas that are asked in speaking tests.

(Participant 5)...I cannot show my real speaking ability thoroughly...You cannot say whatever you know or you do not remember the sentence at that moment...It is only about not having enough information (on the asked topic)...

*Compared and contrasted items on teachers' and students' attitudes towards and beliefs about teaching and testing speaking*

The following analysis aims to investigate the differences between teachers' and students' attitudes towards and beliefs about teaching and testing speaking.

Three questions were the same on the teachers' and students' version of the questionnaire. Kolmogorov-Smirnov and Shapiro-Wilk tests showed the data from these questions not to be normally distributed, so the non-parametric Mann-Whitney tests were used.

	Teachers		Students		<i>U</i>	<i>r</i>	<i>p</i> (two-tailed)
	<i>Mdn</i>	<i>IQR</i>	<i>Mdn</i>	<i>IQR</i>			
1. Importance of teaching / learning speaking skills	5	1	5	1	6145.50	-.07	.18
2. Speaking skills should be tested	4.5	1	4.5	1	6022.50	-.08	.14
3. Speaking exams show students' real level of speaking ability	3	1.5	3	2	5768	-.10	.06

Table 9 - Descriptive statistics for teachers' and students' attitudes towards and beliefs about teaching and testing speaking

As can be seen in table 9, there were no significant differences between teachers and students in the importance given to teaching and learning speaking, whether to test speaking skills, and whether speaking tests are an accurate reflection of students' speaking skills. The interviews also did not reveal any differences.

R.Q.2. What are the washback effects of classroom-based speaking tests, as reflected in teachers' and students' perceptions?

#### *Teachers' perceptions of the washback effects of classroom-based speaking tests*

Table 10 shows the mean scores and standard deviations of the scales and individual questions for the teachers' perceptions of the washback effect of classroom-based speaking tests. Just like the previous Likert-scale items while 5 represents 'totally agree' option, 1 represents 'totally disagree'. Number 3 stands for 'undecided'. The questions will be discussed in turn.

	Mean	SD	Percentages of the participants who		
			TA/A	UD	D/TD
a. Do teachers tailor their speaking classes according to speaking tests in terms of content and allocation of time for speaking activities?	3.03	.52	44.44	11.54	44
<ul style="list-style-type: none"> <li>I spend more time on speaking parts in the days leading up to the speaking test.</li> </ul>	3	.98	40.9	18.2	40.9
b. Do teachers have positive attitudes towards organising extra speaking activities in classes, which are not alike with the activity that will be asked in the speaking test?	2.77	.44	84.45	7.75	7.8
c. Do teachers believe that speaking tests have a positive effect on their students' speaking ability?	3.93	.75	82.2	11.1	6.6
<ul style="list-style-type: none"> <li>Getting ready for the speaking test improves the general speaking skills of students.</li> </ul>	3.93	.78	80	13.3	6.7
<ul style="list-style-type: none"> <li>Students can also use many things that they have studied for the test, in lessons after the test.</li> </ul>	4.11	.64	88.8	8.9	2.2
<ul style="list-style-type: none"> <li>Students tend to forget lots of the things which they have studied for the speaking test, after the test.</li> </ul>	3.04	1.08	37.8	17.8	44.4
<ul style="list-style-type: none"> <li>Speaking tests encourage students to speak more in lessons.</li> </ul>	3.53	.91	60	22.2	17.8

Note: TA: Totally agree, A: Agree, UD: Undecided, D: Disagree, TD: Totally disagree

Table 10 - Descriptive statistics of the teachers' perceptions of the washback effects of classroom-based speaking tests

- a. Do teachers tailor their speaking classes according to speaking tests in terms of content and allocation of time for speaking activities?

As for washback effect of speaking tests in terms of content and allocation of time for speaking activities, the mean score is 3.03. The percentages of the participants who state that they tailor and do not tailor their speaking classes according to speaking tests in terms of content and allocation of time for speaking activities is nearly the same, at 44.4% versus 44%. In the interviews, some teachers stated that they tailor their speaking classes in terms of content and allocation of time for speaking activities but not according to the speaking tests, according to the things that they find useful teaching and the flow of the activity in a class.

(Participant 2)...Sometimes students have lots of things to say, sometimes they have nothing to say (on a topic in a speaking activity)...If they are not interested much, I do not force them to speak (on that subject) since I think that I can cover it in another way.

(Participant 3)...I do not like some speaking parts of the course book (we use) or I do not think that they are appropriate... They (some speaking activities) may not be asked in the exam but if I think that the students need to know them, I teach them.

- I spend more time on speaking parts in the days leading up to the speaking test.

With regard to the statement about spending more time on speaking parts in the days leading up to the speaking test, the mean score, which is 3, suggests that teachers cannot come together in an agree or disagree group but rather there is a diversity. The same percentage, 40.9, show that equal number of the teachers spend



and do not spend more time on speaking parts in the days leading up to the speaking test.

Surprisingly the interviews revealed that most of the interviewees spent more time on speaking activities.

(Participant 1)...(in the days leading up to the speaking test) I continuously repeat the words [and] structures which I want them (the students) to use [in the speaking test].

One of the teachers thinks that getting prepared for the speaking test in class together is something that students like.

(Participant 2)... the students get happier when we say let's practice the task that will be asked in the (speaking) exam.

- b. Do teachers have positive attitudes towards organising extra speaking activities in classes, which are not alike with the activity that will be asked in the speaking test?

Although the mean score for this item is only 2.77, the percentage of the instructors who organize extra speaking activities which are not similar to the activity that will be asked in the speaking test in classes is quite high (84.45%). 7.8% of the instructors' responses show that they do not organize extra speaking activities which are not similar to the activity that will be asked in the speaking test in classes and 7.75% of them are neutral. Based on the percentages it is possible to say that teachers have positive attitudes towards organising extra speaking activities in classes, which are not alike with the activity that will be asked in the speaking test.

As for extra speaking activities, the interviewees generally stated that they organize extra speaking activities in their classes not taking the speaking exam task into account. They also believe that the students participate more in these activities.

(Participant 6)...**They (the students) are always more interested in them** (extra speaking activities)...I did not see them (the students) being more interested with anything just because it will be asked in the exam.

- c. Do teachers believe that speaking tests have a positive effect on their students' speaking ability?

The mean score of 3.93 shows that the teachers believe that speaking tests help students to notice the weaknesses and strengths in their speaking performances under the umbrella title of positive effects of speaking tests' on their students' speaking ability. The percentage of the participants who believe that speaking tests have a positive effect on their students' speaking ability is 82.2.

The other individual questions also investigate the positive effects of speaking tests on students' speaking ability, but since each one of them investigates a different point, they are analysed individually.

- Getting ready for the speaking test improves the general speaking skills of students.

The same mean score which was also given for the previous item, 3.93, suggests that teachers believe that getting ready for the speaking test improves the general speaking skills of the students. The percentage of the participants who think in that way is 80.

The interviews also support the mean score. All the interviewees think just like participant 1.

(Participant 1)...The preparation process before the exam definitely has a (positive) effect on (the students' speaking ability).

- Students can also use many things that they have studied for the test, in lessons after the test.

4.11 mean score reveals that the teachers strongly believe that students can also use many things that they have studied for the test, in lessons after the test. 88.8% of the teachers support this idea.

Participant 4 thinks that students can also use many things that they have studied for the test in lessons after the test because when they speak they create their own sentences in English and this helps them remember and use the structures they studied for the tests.

(Participant 4)...The students already listen to, read, [and] write something. However, I believe in that they (the students) comprehend it in a better way when they form their own sentences. **They try to use these structures, phrases, [and] sentences in other classes, as well.**

- Students tend to forget lots of the things which they have studied for the speaking test, after the test.

The teachers' mean score is 3.04 for the item 'students tend to forget lots of the things which they have studied for the speaking test, after the test'. The percentage of the participants who agree with that statement is 37.8%, who disagree with it is 44.4%, and who are neutral on it is 17.8%. The percentages show that there is diversity of opinion on that issue.

- Speaking tests encourage students to speak more in lessons.

Finally, for the statement that speaking tests encourage students to speak more in classes, the mean score, 3.53, reveals that teachers are between being undecided and agreeing. The percentages show that 60% of the teachers think that speaking tests encourage students to speak more in classes. While 22.2% of the teachers are neutral, 17.8% of them do not think that speaking tests encourage students to speak more in classes. Taking the 60% of the teachers' ideas into account,

it can be said that teachers believe that there is a beneficial washback owing to the fact that the existence of the speaking test increases motivation.

While participants 2 and 4 were clear about the encouraging effect of the speaking tests on speaking, participant 5 had some doubts on generalising this idea.

(Participant 4)...The students give more importance at least to the speaking parts in classes when they know that they will be also tested on speaking. They try to be more participatory... If we did not test it (speaking ability), they would not take the speaking parts seriously.

(Participant 5)...I think the tests do not encourage them but yesterday a student of mine asked me what kind of things s/he could do in order to improve his / her speaking ability by stating that the next exam would also have a speaking part. I cannot figure out whether s/he asked it [just] for the tests or to improve his / her speaking ability. **May be it (speaking tests) encourages the students (to speak more in classes) but I cannot say that for all of them.**

*Students' perceptions of the washback effects of classroom-based speaking tests*

Table 11 shows the mean scores and standard deviations of the scales and individual questions for the students' perceptions of the washback effect of classroom-based speaking tests. Just like all the previous Likert-scale items, while 5 represents 'totally agree' option, 1 represents 'totally disagree'. Number 3 stands for 'undecided'.

	Mean	SD	Percentages of the participants who		
			TA/A	UD	D/T D
a. Do students believe that speaking tests have positive effects on their speaking ability?	3.84	.57	71.91	17.2	10.91
b. Do speaking tests have an influence on what and how students learn?	2.28	.77	14.65	19.17	66.2
c. Do speaking tests have any effects on what students really do?					
<ul style="list-style-type: none"> <li>• I try to practice the speaking activities which will be asked in speaking tests in daily life.</li> </ul>	3.44	1.05	54.7	23.5	21.8
<ul style="list-style-type: none"> <li>• I participate more in speaking parts in the days leading up to the speaking tests.</li> </ul>	3.05	1.05	38.9	24.9	36.2
<ul style="list-style-type: none"> <li>• I give more importance to the speaking parts which will be asked in the test.</li> </ul>	3.26	1.06	46	27.5	26.5
<ul style="list-style-type: none"> <li>• Even if they were not tested, speaking skills should have a place in lessons.</li> </ul>	4.30	.87	88	7.8	4.3

Note: TA: Totally agree, A: Agree, UD: Undecided, D: Disagree, TD: Totally disagree

Table 11 - Descriptive statistics of students' perceptions of the washback effects of classroom-based speaking tests

- a. Do students believe that speaking tests have positive effects on their speaking ability?

The mean score of 3.84 and the percentage of the students who totally agree and agree, which is 71.91%, reveal that students believe that speaking tests have positive effects on their speaking ability. Supporting this, most of the interviewees also think that speaking tests have positive effects on their speaking ability.

(Participant 1)... I have not forgotten any of them [structures or words learnt for the speaking test] owing to the fact that I will need them in the upcoming speaking tests...

(Participant 4)...I measure how much I can speak [and] how much English I have learned in the speaking tests... They (speaking tests) affect it (speaking ability) positively because at least I study. I learn something new there (in speaking test environment).

Participants 1 and 5 think that speaking tests are a good practice chance for real life. Especially participant 5 thinks that the speaking tests and real life English have something in common.

(Participant 5)... in the speaking tests the questions are asked spontaneously. The questions will be asked in the same way on abroad or wherever it is [in real life]...

- b. Do speaking tests have an influence on what and how students learn?

Expected washback effect is not seen at this point owing to the fact that the students' responses show what and how students learn are not influenced by speaking tests, with a 2.28-mean score. Just 14.65% of the students' responses show that speaking tests have an influence on what and how they learn.

Participant 2 does not seem to be much influenced by the speaking test itself in terms of what and how s/he learns. However, s/he stated that the more points an exam brings, the more s/he studies for it.

(Participant 2)... If speaking had more weight in grading, I would focus more on it. I study less for the things which I will get less points.

Participant 3's statement shows that s/he is an autonomous student who does not work just for the test.

(Participant 3)...Even though there were not any (speaking) tests, I would still pay a special attention to speaking. According to me, it (speaking) is something that I should learn.

Participants 2 and 5 state that they do not give equal importance to every speaking activity in the class. They give more importance to the ones which they think that they will use more in the future.

(Participant 2)...Some of them (speaking activities) are very important since I think about working in summer. I mean, how to ask for something [or] request. There are these kinds of nice things but I consider some of them as unnecessary... I study for the things which are important for me.

c. Do speaking tests have any effects on what students really do?

The mean score for practicing the speaking activities which will be asked in speaking tests in daily life is 3.44. While 54.7% of the students state that they practice the speaking activities which will be asked in speaking tests in daily life, 21.8% of them do not try to practice them in their daily life. 23.5% of them are neutral on the issue.

Being one of representatives of the disagree group, participant 2 stated that s/he does not have much chance to practice speaking in his/her daily life. S/he tried to speak with some friends whose English were good, in his/her daily life. However, s/he said that it was not to practice for the speaking test.

The mean score about whether they participate more in speaking parts in the days leading up to the speaking tests or not is 3.05. While the percentages of the participants who state that they participate more is 38.9, the ones who do not think in

that way constitute 36.2%. The percentages of the participants who are neutral is 24.9. The participants who agree and do not agree are very close to each other.

Participant 2 stated that s/he does not participate more in speaking parts in the days leading up to the speaking tests because since s/he knows that the type of the speaking task will be the same.

(Participant 2)...It (having speaking exams) does not have an effect because the style (in the speaking tests) is the same...I mean it does not change. As a result, we do not take the (speaking) exams seriously.

The quotation of participant 2 suggests that when students get used to do something in the same way, it gets easier for them. If they do not find the style or the task challenging and get high grades, they discredit the thing they are doing.

The mean score 3.05 shows that the students were not able to form a big group stating that they participate more in speaking parts in the days leading up to the speaking tests or not. However, the percentages of the students who state that they participate more in speaking parts in the days leading up to the speaking tests, 54.7% , is twice more than the percentages of the students who state that they do not participate more in speaking parts in the days leading up to the speaking tests, which is 21.8%. 23.5% of the participants are undecided.

None of the interviewees stated that they participate more in speaking parts in the days leading up to the speaking tests. Instead, they all stated that they have a quick look at the questions that will be asked in the test before the test.

The students tend to agree (the mean score is 3.26) with the statement that they give more importance to the speaking parts which will be asked in the test. While 46% of the participants think that they give more importance to the speaking parts which will be asked in the test, 26.5% of them disagree with this idea. In



addition, 27.5% of the participants are neutral. This result suggests that a [speaking] test can influence what learners learn (Alderson and Wall, 1993).

Participant 1's statement shows that the students give special importance to the speaking parts which will be asked in the test. It can be inferred in his/her speech that whether they participate well or not in speaking parts in classes, they all try to do something for the test by giving more importance to the speaking parts which will be asked in the test.

(Participant 1)...Before the speaking tests, everyone detects the questions [and] writes them on papers [to study]...

One of the questions was analysed individually without having a scale. It aims to investigate the fact that even if they were not tested, speaking skills should have a place in classes. The mean score, 4.30, indicates that the students are strongly in favour of this idea. The percentage of the students who totally agree and agree with this idea is 88. This result suggests that the students are aware of the fact that language is for communication.

Like his/her prep school friends, participant 4 stated that speaking skills should have a place in classes even if they were not tested because communicative skills have been ignored for a long time in the education system of Turkey.

(Participant 4)... (Even if they were not tested) Speaking skills should have a place (in classes) because English has been taught as grammar so far, but speaking has been ignored.

*Compared and contrasted items on teachers' and students' perceptions of the washback effects of classroom-based speaking tests*

	Teachers		Students		<i>U</i>	<i>r</i>	<i>p</i> (two-tailed)
	<i>Mdn</i>	<i>IQR</i>	<i>Mdn</i>	<i>IQR</i>			
1. Getting ready for speaking tests improves students' speaking skills	4	0	4	1	5749.50	-.10	.05
2. Spending more time on speaking skills in the days leading up to the test	3	2	3	2	6459	-.01	.78
3. Giving more importance to the speaking parts which will be asked in the test	3	2	3	2	6257	-.05	.30
4. The usage of the things that the students have studied for the speaking test, in lessons after the test	4	.50	4	1	5570.50	-.12	.02
5. Students easily forget the things that they have studied for the speaking test, after the test	3	2	4	1	5110.50	-.15	.005
6. Students notice their weaknesses in speaking after speaking tests	4	.00	4	1	6714.50	-.02	.73
7. Even if speaking skills were not tested, they should have a place in classes	5	1	4	1	5265	-.15	.004
8. If speaking skills were not tested, I would not spend so much time on improving speaking skills	2	1	2	2	5320.50	-.14	.01

Table 12 - Descriptive statistics for teachers' and students' perceptions of the washback effects of classroom-based speaking tests

In the following paragraphs, the very same questions, asked both to the teachers and students about their perceptions of washback effect of classroom-based speaking tests, are compared. Kolmogorov-Smirnov and Shapiro-Wilk tests showed the data not to be normally distributed, so non-parametric Mann-Whitney tests were used.

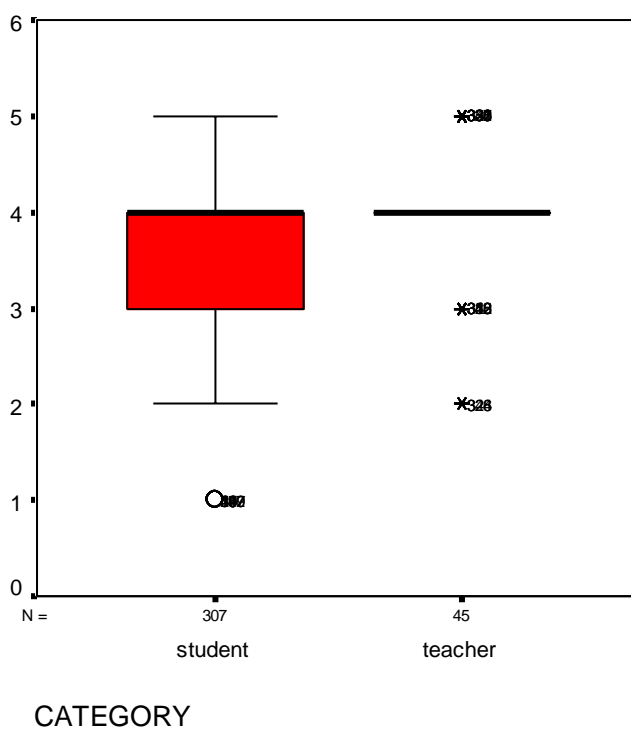


Figure 2 - Getting ready for speaking tests improves students' speaking skills

As it can be seen in figure 2, there is a small but significant difference between teachers' and students' thoughts on students' improvement on speaking skills by getting ready for speaking tests. However, this significant difference is very small and does not show up in the medians. While 80% of the teachers agree with that statement, the percentage of the students who think the same is 63.6%.

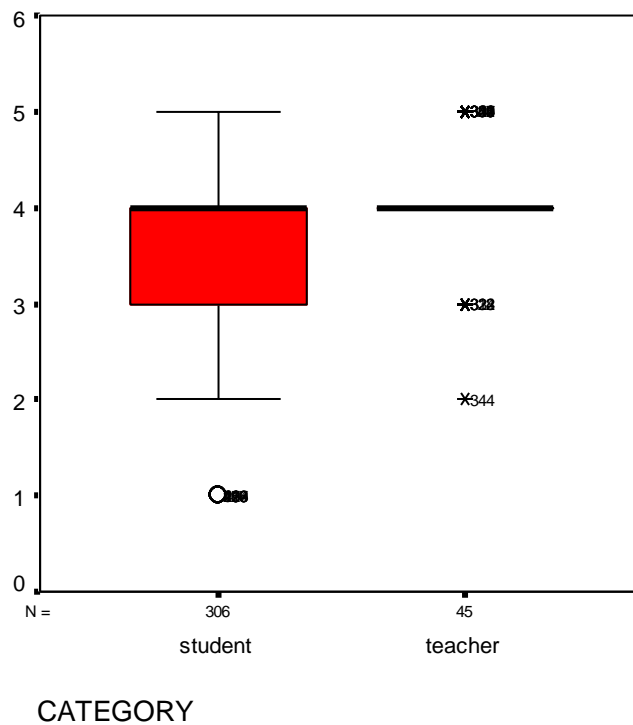


Figure 3 - The usage of the things that the students have studied for the speaking test, in lessons after the test

Again small but significant differences were found between teachers and students on the usage of the items that the students have studied for the speaking test in classes after the test (See figure 3). However, the significance does not show up in the medians. While the percentage of the teachers who think that students can use many of the things, in lessons after the test is 88.8%, the percentage for the students is 68.6%.

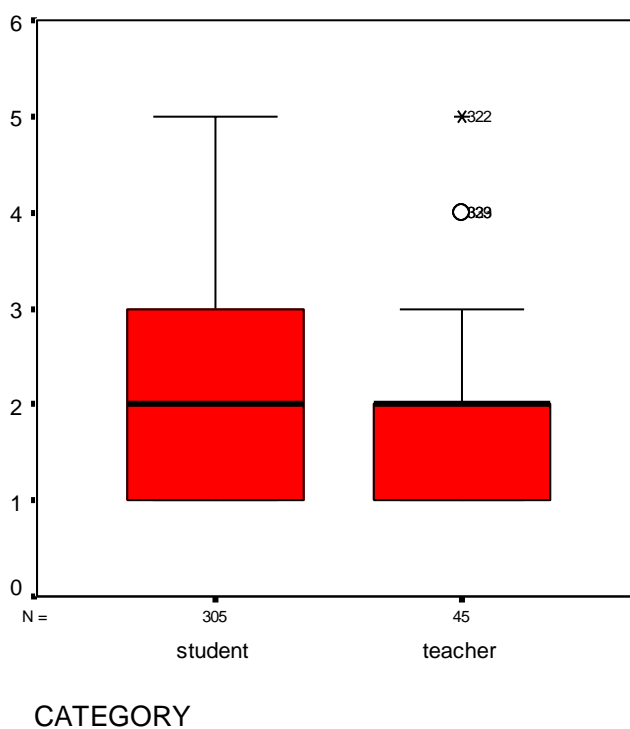


Figure 4 - If speaking skills were not tested, I would not spend so much time on improving speaking skills

In regard to the statement that if speaking skills were not tested, I would not spend so much time on improving speaking skills, there is a significant difference between the teachers and students as it can be seen in figure 4. However, the significant difference is very small and it does not show up in the medians. 84.4% of the teachers and 68.6% of the students disagree with that statement.

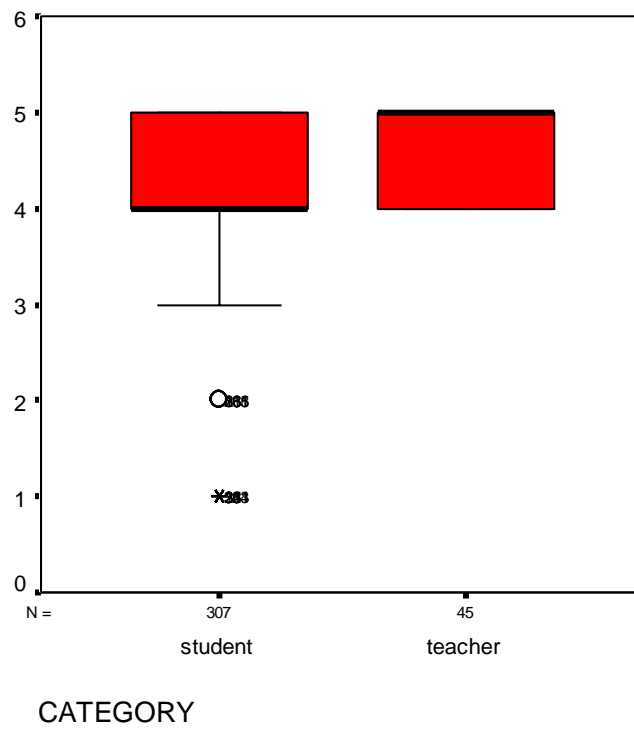


Figure 5 - Even if speaking skills were not tested, they should have a place in classes

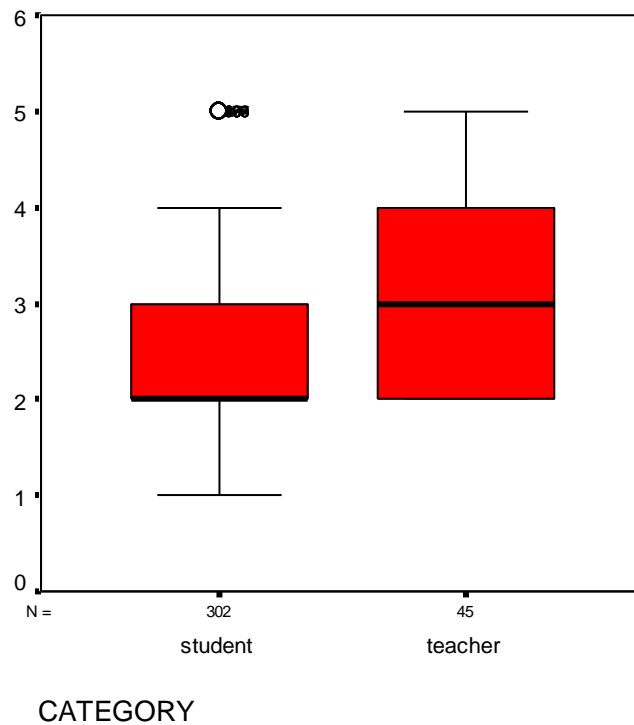


Figure 6 - Students easily forget the things that they have studied for the speaking test, after the test

In related to speaking skills' place in classes even if they were not tested (See figure 5) and on students' easily forgetting the items that they have studied for the speaking test after the test (See figure 6) there is a small but significant difference between the teachers and students.

As for what teachers and students do in classes because of the speaking tests, there are no significant differences between teachers and students on giving more importance to the speaking parts which will be asked in the test and spending more time on speaking skills in the days leading up to the test. It is the same case with students' noticing their weaknesses in speaking after speaking tests.

### Conclusion

In this chapter, the data gathered from the teacher and student questionnaires and interviews were analysed and reported in four phases. The first phase presented the analysis of the Likert-scale questionnaires of teachers and students and the interviews done with the teachers and the students in order to get information about the teachers' and students' attitudes towards and beliefs about teaching and testing speaking. Then, the same and similar questions about the teachers' and students' attitudes towards and beliefs about teaching and testing speaking were compared and contrasted. In the third phase analysis of the Likert-scale questionnaires and the interviews of the teachers and the students were carried out with the aim of providing information on the teachers' and students' perceptions of the washback effects of classroom-based speaking tests. Finally, the same and similar questions about the teachers' and students' perceptions of the washback effects of classroom-based speaking tests were compared and contrasted.

In the next chapter the results of the study will be discussed, pedagogical implications, limitations and suggestions for further research will be given respectively.



## CHAPTER V: CONCLUSION

### Introduction

The main purpose of this study was to learn teachers' and students' perceptions of the washback effect of classroom-based speaking tests. Moreover, teachers' and students' attitudes towards and beliefs about teaching and testing speaking were also addressed.

The study was conducted at Akdeniz University, School of Foreign Languages in the 2010-2011 academic year. The participants who took part in this study were 45 instructors of English and 307 preparatory class students. These 307 participants were the future students of various departments at Akdeniz University but they were all intermediate level students. The data were collected through teacher and student questionnaires and teacher and student interviews. The teacher questionnaire had 32 items in order to investigate teachers' perceptions of the washback effect of classroom-based speaking tests and their attitudes towards and beliefs about teaching and testing speaking. The student questionnaire had 34 items in order to reveal students' perceptions of the washback effect of classroom-based speaking tests and their attitudes towards and beliefs about teaching and testing speaking.

In this chapter, the findings of the study will be presented and discussed. Then, pedagogical implications will be provided. Finally, limitations of the study will be given and suggestions for further research will be made.

## General Results and Discussion

### *Teachers' and students' attitudes towards and beliefs about teaching and testing speaking*

The first research question aimed to reveal teachers' and students' attitudes towards and beliefs about teaching and testing speaking.

The questionnaires and interviews show that teachers think that it is quite possible to teach speaking skills and none of the teachers have any doubts on the importance of teaching speaking. All the teachers think that teaching speaking skills is crucial. Similarly, the study conducted by Caine (2005), who tried to reveal the extent and nature of washback resulting from the direct speaking test he proposed, revealed that teachers thought that it was very important for their learners to develop communicative ability. As for whether speaking skills should be tested or not, just like Caine's (2005) findings, the teachers believe that it is important to test speaking. Most of the teachers believe that, though testing speaking skills is difficult, they can be measured accurately. Given this, it is perhaps surprising that teachers did not generally agree with the item stating that speaking tests reflect students' speaking skills accurately. However, this apparent conflict may stem from the fact that the teachers, just like participant 6, may not trust the speaking tests that they administer in their program but they may believe that there are some ways to measure speaking skills accurately.

(Participant 6)...I think that there is a problem in administering these (speaking) tests...I do not know how reliable information it provides to give the subjects before the test, learning the given dialogues by heart (before the test) or answering the questions. I believe that speaking tests should be more simultaneous, more natural, and more creative. I think that it should not be tested through question-answer way.

As for the medium of testing speaking skills, teachers believe that the best medium to test speaking skills is speaking itself. Based on that belief it can be suggested that speaking tests should be continued to be measured through speaking.

In their beliefs about the usage of the results of the speaking tests as a reliable diagnostic tool, the teachers were not able to create a big agree or disagree group, but instead there is diversity. This may stem from the fact that some of the teachers face with unexpected performances of the students which are quite different from their in-class performances. This result can stem from the type of the speaking tests. The teachers who experience this situation a lot do not think that the results of the speaking tests can be used as a reliable diagnostic tool. It can be suggested that the tasks of the speaking tests should be chosen accordingly in order to increase the reliability.

As regards students, the collected data clearly show that students, like teachers', have quite positive attitudes towards learning speaking skills. In addition, students' attitudes towards teaching speaking and the importance of speaking are very positive. They believe that it is very important to test speaking skills. Both teachers and students totally agree with the statement that speaking skills should be tested. Similarly Poonpon (2010) found that for students speaking tests were necessary and they should remain in the English course they took. The questionnaires and interviews show that students' opinions about whether they enjoy or are nervous in speaking tests change. Although the issue of different examiners was not addressed in the questionnaires, some of the students mentioned it in their interviews. (Participant 2)... Up to now there have been eight or ten speaking quizzes, midterms etc. Two or three of them were fun since I was relaxed because I had these exams with an acquaintance teacher. I mean, I felt under stress in the rest [of them].

The students' perceptions may be limited to the examiners' attitudes whom they have met. Every other examiner, among nearly 20 examiners, can leave different impressions about testing speaking on the examinees.

Just like the teachers, students have different ideas about whether speaking tests can really show their speaking ability. Taking this diversity into account, the tasks of the speaking tests should be modified.

*Teachers' and students' perceptions of the washback effects of classroom-based speaking tests*

The second research question aimed to reveal teachers' and students' perceptions of the washback effects of speaking tests.

There is diversity in teachers' opinions on questions of whether they tailor their speaking classes according to speaking tests in terms of content and allocation of time for speaking activities and spending more time on speaking parts in the days leading up to the speaking test. As it can be concluded here, the speaking tests can have some influences on the content and allocation of time for speaking activities and intensity of the activities for some teachers.

They have positive attitudes towards organizing extra speaking activities in classes which are not similar with the activity that will be asked in the speaking test. While some teachers may organize activities both alike and not alike with the activity that will be asked in the speaking test, some others may just organize extra speaking activities which are not similar with the activity that will be asked in the speaking test. For the former group there may be washback effect to some extent while in the latter no direct washback effect is seen. These facts suggest one of the hypotheses of

Alderson and Wall (1993, p.121) which says 'Tests will have washback effects for some learners and some teachers, but not for others'.

Although not all the teachers claim to be influenced by the speaking tests in terms of what they do in classes, they believe that speaking tests have a positive effect on their students' speaking ability. They believe that getting ready for the speaking tests improves the general speaking skills of students. In their experimental study, Andrews et al. (2002) also stated that adding an oral component called Use of English to the existing Hong Kong Advanced Supplementary exam had an influence on what students learned and commented that the Use of English Oral Component might have had a positive influence on students' spoken English performance. In this current study, teachers also believe that students can also use many of the things that they have studied for the speaking test in lessons after the test. Another fact which teachers believe about the positive effects of a speaking test is that speaking tests encourage students to speak more in classes. As Hughes (2003, p.18) states, this is 'a helpful washback effect' achieved as a result of direct testing. However, teachers have different ideas on the question of whether students tend to forget lots of the things that they have studied for the speaking. In spite of believing that students can use many of the things that they have studied in lessons after the test, it is interesting that teachers have different opinions on this point. This may stem from the fact that though teachers can witness students' usage of the studied points in lessons, they are not able to say whether students tend to forget the target things studied for the speaking tests since students may not have forgotten them but prefer not to use them or are not participatory in classes.

The questionnaires and interviews show that students also believe that speaking tests have positive effects on their speaking ability. Caine's (2005) study also revealed that students thought that studying for the speaking test improved their English. Similarly, in a study conducted by Poonpon (2010) it was also seen that students' self-perceptions of their own speaking ability increased after taking speaking tests. Poonpon (2010) interpreted this self-perception as positive washback. In one of the interviews done with a student for the current study, the interviewee stated that s/he learned something new in the testing speaking environment and two of them stated that speaking tests were a good chance to practice. This perception of the interviewees is similar to that of Pearson (as cited in Hsu, 2009), who considers good tests to be usable class activities. The students' belief in the positive effects of speaking tests on their speaking ability fits one of the washback hypothesis of Alderson and Wall (1993) in that [classroom-based speaking] tests influence learning. However, students do not believe that they are influenced by speaking tests in regard to what and how they learn. As regards speaking tests' effects on what students really do such as whether they spend more time on speaking skills in the days leading up to the test and whether they give more importance to the speaking parts which will be asked in the test, there is a diversity in both teachers' and students' responses. This contrasts with the findings of Caine (2005), whose students stated that they made a greater effort to speak English in the weeks leading up to the test. The scales of the current study reveal that students do not do something special because of the speaking tests. However, they are closer to agreeing with the item that they try to practice the speaking activities which will be asked in speaking tests in daily life, rather than being neutral. This may stem from the fact that if students, like

participants 2 and 5, find the speaking activities which will be asked in the speaking tests necessary for their general English and future life, they may try to practice them in their daily life.

(Participant 2)...Some of them (speaking activities) are very important since I think about working in summer. I mean, how to ask for something [or] request. There are these kinds of nice things but I consider some of them (speaking activities) as unnecessary... I study for the things which are important for me.

(Participant 5)...I say I will use these somewhere, [I give more importance them] not according to the subject but, well, the things which will be useful for me.

Both teachers and students also agree with the items which state that students use the things which they have studied for the speaking test in lessons after the test and that they notice their weaknesses in speaking after the test. There is diversity in the teachers' replies to the question which states that students easily forget the things that they have studied for the speaking test, after the test. However, students agree that they easily forget the things that they have studied for the speaking test, after the test.

Teachers and students do not believe that, if speaking skills were not tested, they would spend less time on improving speaking skills. Their claim is that they do not have a place for speaking in their classes just because it is tested. While teachers *strongly* agree with the statement that even if speaking skills were not tested, they should have a place in classes, students agree with that statement. This contrasts with Ferman's (2004) finding that teachers stated that they would stop teaching oral proficiency right after the oral test. This difference in perceptions may stem from the fact that in Ferman's study (2004) the oral exam was used as a component of a high stake exam which had a strong washback effect on teaching and learning. In her

study the teachers felt that they were under pressure because of the speaking component and it narrowed the curriculum. However, the current study is based on classroom-based speaking tests in which there are many speaking tests, unlike a high stake one, which will affect the grades cumulatively. That is to say, the students and teachers have a chance to cover the low grades in the upcoming speaking tests.

The results all reveal the general consensus of the scales of the teachers' and students' attitudes towards and belief in teaching and testing speaking and their perceptions of washback effect of speaking tests. However, in most of the scales about the washback effect teachers and students have different ideas. In other words while some participants agreed, some others disagreed and were neutral. Based on these findings it can be concluded that, as Alderson and Wall (1993, p.121) suggest, 'tests will have washback effects for some learners and some teachers, but not for others'. Similar to the current study, Watanabe's research (1996) revealed that washback happened to some teachers, not to others.

### Pedagogical Implications

The collected data revealed that both teachers and students have positive attitudes towards the possibility and necessity of teaching and testing speaking skills. Based on these findings, it can be suggested that speaking exams administered at the preparatory class program of Akdeniz University Foreign Languages School should remain as a part of the testing system.

While teachers believe that speaking should be tested, they have some concerns about the reliability of these tests. Most of the interviewees stated that these exams should be more detailed than they are now. In addition, they raised the issue



of the necessity of detailed and reliable criteria to be used in these speaking tests. In order to be more objective and provide unity, there should be some training for teachers on how to evaluate the speaking performances. As for the medium to test speaking, participants are in favour of using speaking tests. They believe that other skills can be made use of while testing speaking, but that they should not be evaluated. Some students stated that if they do not see something to read in any part of the speaking test, they get more anxious. The testing committee can consider this fact while preparing the speaking exams.

About the test itself, another issue is the topic area of the questions asked. Some students stated that sometimes they cannot answer the questions because they do not know anything about the topic asked. They said that if they cannot give an answer to a question in Turkish, it is impossible to do it in a foreign language. The testers state that they choose the topics which are mentioned or are similar to the topics in the units of the course book used. However, choosing the same or similar topics of the course book for the speaking test does not always ensure that they are suitable topics. If this is what is wanted by the testers, the topics which draw more attention than the others during the lessons can be written down and these popular ones can be used.

Another issue is the weight of scoring in the criteria. Teachers stated that generally all the components of speaking skills such as fluency and accuracy are given nearly the same weight without taking the task into account. One of the students also complained about grammar's weight in speaking performances. S/he stated that teachers check the grammar knowledge of the students who try to speak, to a great extent. New criteria including the proper weights for the speaking

components can be prepared taking the task that will be carried out by the students in the test into account. One of the students also stated that the tasks are always the same. This belief and testing system can decrease the value of the speaking tests for students.

One of the teachers (6) stated that s/he does not believe that students can evaluate their actual performances after taking speaking tests. This statement raises the feedback issue, which is always important. They have to be informed about their performances in detail not just by grades. Notes about students' speaking performances should be taken and the recordings should be kept. The students should be provided with feedback systematically and this system should not be left to the teachers' individual preference.

Most of the students stated that they give more importance to some dialogues or structures than the other dialogues or structures in classes. However, it is not a result of the speaking tests but the result of the students' being conscious. That is to say, they give more importance or participate more in the activities which they think they will be beneficial for their future jobs. To increase class participation in speaking parts, the topics related to the students' departments can be given place in speaking activities.

Finally, a lot of students wrote in the further comments part in the questionnaires and most of the interviewees stated that they are looking for a real reason to speak English in lessons. They would like a foreign student, for example Erasmus exchange students, in their speaking classes. A system can be developed to have a foreign student in each class in some classes to motivate the students to speak English.

## Limitations

First of all owing to the fact that the questionnaires were created by the researcher herself, more than one piloting procedure was required in order to increase the reliability. Although the student questionnaire was piloted twice, the teacher questionnaire had only one opportunity to be piloted because of the time constraints. As a result, there came out many individual questions which were needed to be treated individually.

Another limitation is that the study was conducted in only one institution, which is Akdeniz University Foreign Languages Department. Owing to the fact that the testing of speaking skills can be quite different at other universities, this study may not be generalizable to other testing speaking settings.

One other limitation is that the questionnaires and the interviews mainly address positive washback. This thesis has not addressed the issue of whether there are harmful effects –negative washback- of these classroom-based speaking tests on speaking or on the curriculum such as increasing the stress level and anxiety of the students or distorting the curriculum. In other words, the possible negative consequences and influences of speaking tests on many educational aspects are neglected.

The final and most important limitation is that this study is based only on stakeholders' perceptions. This fact raises two important issues. First, it is not clear whether the teachers or the students are sufficiently aware or have enough experience to know whether it would be better not to have exams. The experience of teaching cannot always be a good indicator on that issue since if people do not improve

themselves in their field, just spending more time in their job does not make a good difference. The second issue is that participants may have been reluctant to admit what they really do or believe to look like good teachers and students even if they do not believe in the way they answered the questions. Since this is not an empirical study this fact should be kept in mind while evaluating the results.

### Suggestions for Further Research

Since it is very difficult to measure the washback effect of speaking tests, there are few empirical studies on this subject. Due to the fact that each university has a different assessment system on speaking, this study could be usefully replicated studying another setting.

It would be particularly interesting to do research at a university where speaking tests are not administered as a part of the exams. Having two groups of students, one of which do not take speaking exams, it is possible to compare the two groups' speaking performances to see whether there is a washback effect of having speaking tests.

### Conclusion

The current study aimed to reveal teachers' and students' attitudes towards and beliefs about teaching and testing speaking and their perceptions of the washback effects of classroom-based speaking tests. The study revealed teachers' and students' positive attitudes towards the importance of teaching and testing speaking. Although teachers believe that it is difficult to test speaking, they also believe that it can be measured accurately. They are in favour of testing speaking through speaking performances. The answers of the teachers include diversity on the

usage of the speaking tests as reliable diagnostic tools. There is also diversity in students' answers on whether they enjoy speaking tests or be nervous during the tests.

As for washback effect, there is again diversity between the teachers and students on what they teach / learn and do in classes because of speaking tests. It can be concluded that no washback can be seen on what teachers teach, what students learn, and what they do in classes. However, they think that these speaking tests have positive effects on the students in many respects. Both teachers and students think that they would go on teaching / learning speaking skills in class even if they were not tested.

## REFERENCES

- Alderson, C., & Wall, D. (1993). Does Washback Exist? *Applied Linguistics*, 14 (2), 115-129.
- Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL Preparation Courses: a study of washback. *Language Testing*, 13 (3), 280-297.
- Amengual-Pizarro, M. (2009). Does the English Test in the Spanish University Entrance Examination Influence the Teaching of English? *English Studies*, 90 (5), 582-598.
- Andrews, S., Fullilove, J., & Wong, Y. (2002). Targeting washback—a case-study. *System*, 30 (2), 207–223. Elsevier.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in Practice*. Oxford: Oxford University Press.
- Bailey, K. M. (1996). Working for washback: a review of the washback concept in language testing. *Language Testing*, 13 (3), 257-279.
- Brown, H. D. (2004). *Language Assessment Principles and Classroom Practices*. New York: Pearson Education.
- Caine, N. A. (2005). *EFL Examination Washback in Japan: Investigating the Effects of Oral Assessment on Teaching and Learning*.
- Cheng, L. (1997). How Does Washback Influence Teaching? Implications for Hong Kong. *Language and Education*, 11 (1), 38-54.
- Cheng, L. (2000). Washback or backwash: a review of the impact of testing on teaching and learning.
- Cheng, L., Watanabe, Y., & Curtis, A. (2004). *Washback in language testing: Research contexts and methods*. New Jersey: Lawrence Erlbaum Associates.
- Ferman, I. (2004). The Washback of an EFL National Oral Matriculation Test to Teaching and Learning. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in Language Testing* (pp. 191-210). New Jersey: Lawrence Erlbaum Associates, Inc., Publishers.
- Hamp-Lyons, L. (1997). Washback, impact and validity: ethical concerns. *Language Testing*, 14 (3), 295-303.
- Hsu, H.-F. (2009). The impact of implementing English proficiency tests as a graduation requirement at Taiwanese universities of technology University of York. *Educational Studies*.

- Hughes, A. (2003). *Testing for Language Teachers*. Cambridge: Cambridge University Press.
- Manjarrés, N. B. (2005). Washback of the foreign language tests of the state examination in Colombia: A case study. *University of Arizona, 12*, 1-19.
- McNamara, T. (2000). *Language Testing*. (H. G. Widdowson, Ed.). Hong Kong: Oxford University Press.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing, 13* (3), 241-256.
- Mohammadi, M. (n.d.). High-Stakes Testing Washback: A Survey on the Effect of Iranian MA Entrance Examination on Teaching. faculty.ksu.edu.sa.
- Munoz, A. P., & Alvarez, M. E. (2010). Washback of an oral assessment system in the EFL classroom. *Language Testing, 27* (1), 33-49.
- Newfields, T. (2005). TOEIC ® Washback Effects on Teachers : A Pilot Study at One University Faculty. *English, 31* (1), 83-105.
- Nguyen, P. N., Griffin, P., & Izard, J. (n.d.). Toward Beneficial Washback Effects of a Test. edfac.unimelb.edu.au, 1-27.
- Pan, Y. (2009). A review of washback and its pedagogical implications. *Evaluation, 25*(February), 257-263.
- Rahimi, Z., & Nazhand, N. (2010). Perspectives on IELTS Preparation Courses. *Test*.
- Reynolds, J. (2010). An Exploratory Study of TOEFL Students as Evaluators of “ Washback to the Learners ”
- Shohamy, E., Donitsa-Schmidt, S., & Ferman, I. (1996). Test impact revisited: washback effect over time. *Language Testing, 13* (3), 298-317.
- Taylor, L. (2005). Washback and impact. *ELT Journal, 59*(April), 2000-2001.
- Tsagari, D. (2007). Review of Washback in Language Testing: What Has Been Done? What More Needs Doing? Online Submission, 69.
- Watanabe, Y. (1996). Does grammar translation come from the entrance examination? Preliminary findings from classroom-based research. *Language Testing, 13* (3), 318-333.

## APPENDIX A: INFORMED CONSENT FORM

My Dear Colleague;

I am Özlem Duran, one of the instructors of English at Akdeniz University Foreign Languages School. I have been doing MA degree in the department of Teaching English as a Foreign Language at Bilkent University. The main purpose of my thesis subject is to investigate the washback effect of English speaking tests administered in our prep class program on the teaching and learning of speaking ability from the teachers' and students' perspective. In this study, the ideas of our valuable instructors and dear students' ideas will be acquired through teacher and student questionnaires and teacher and student interviews. After the analysis of the questionnaires a short interview will take place with some of my colleagues who will be volunteers. The information about your identification will not be published in any reports at the end of the research.

The responses that you will give to the questionnaire items will contribute to the study to a great extent. If you accept taking part in the study, fill in the related blanks at the bottom of the page and sign.

Özlem Duran

Supervisor: Asst. Prof. Dr. Philip Durrant

MA TEFL Bilkent University / ANKARA

I have read the information in this form and I accept participating in the study.

Name:

Signature:

Date:



## APPENDIX B: BİLGİ VE KABUL FORMU

Değerli Meslektaşım;

Ben Akdeniz Üniversitesi Yabancı Diller Yüksek Okulu İngilizce okutmanlarından Özlem Duran. Bilkent Üniversitesi Yabancı Dil Olarak İngilizce Öğretimi programında yüksek lisans yapmaktayım. Tez konumun temel amacı, yüksekokulumuz hazırlık programında uygulanan İngilizce konuşma sınavlarının konuşma becerilerini öğretmeye ve öğrenmeye olan etkisini, öğretmenlerin ve öğrencilerin bakış açısı doğrultusunda araştırmaktır. Bu çalışma kapsamında, siz değerli okutmanlarımızın ve sevgili öğrencilerimizin fikirleri öğretmen ve öğrenci anketleri ve röportajları yoluyla edinilecektir. Anketlerin analizi sonrasında, gönüllülük esasına uygun olarak, bazı meslektaşarımla bu çalışmayla ilgili kısa bir röportaj yapılacaktır.

Kimliğinizle ilgili bilgiler bu araştırma sonucu herhangi bir raporda yayınlanmayacaktır.

Anket sorularına vereceğiniz cevaplar araştırmaya çok büyük katkı sağlayacaktır. Araştırmaya katılmayı kabul ediyorsanız, sayfanın altındaki ilgili yerleri doldurarak imzalayınız.

Özlem Duran

Tez danışmanı: Yrd. Doç. Dr. Philip Durrant

MA TEFL Programı Bilkent Üniversitesi / ANKARA

Bu formdaki bilgileri okudum ve araştırmaya katılmayı kabul ediyorum.

İsim:

İmza:

Tarih:

## APPENDIX C: INFORMED CONSENT FORM

Dear Student;

I am Özlem Duran, one of the instructors of English at Akdeniz University Foreign Languages School. I have been conducting a study to investigate the washback effect of English speaking tests administered in our prep class program on the teaching and learning of speaking ability from the teachers' and students' perspective. In one of the class hours of your English classes, a questionnaire will be administered. After the analysis of the questionnaire, short interviews will take place with some of you, who are volunteer again, in a class hour

The information about your identification will not be published in any reports at the end of the research. The responses you will give with your name will not be known by anybody.

The responses that you will give to the questionnaire items will contribute to the study to a great extent. If you want to take part in the study, fill in the related blanks at the bottom of the page and sign.

Instructor of English Özlem DURAN

MA TEFL

Bilkent / ANKARA

I have read the information in this form and I accept participating in the study.

Name & surname:

Department:

Class:

Signature:

## APPENDIX D: BİLGİ VE KABUL FORMU

Sevgili Öğrenci,

Ben Akdeniz Üniversitesi Yabancı Diller Yüksek Okulu İngilizce okutmanlarından Özlem Duran. Yüksekokulumuz hazırlık programında uygulanan İngilizce konuşma sınavlarının konuşma becerilerini öğretmeye ve öğrenmeye olan etkisini, öğretmenlerin ve öğrencilerin bakış açısını değerlendirerek edinmek amacıyla bir araştırma yürütüyorum. İngilizce derslerinizin herhangi bir saatinde, bu araştırmayla ilgili bir anket uygulanacaktır. Anket analizi sonrası, gönüllülük usulüne bağlı kalınarak aranızdan bazı arkadaşlarımızla yine İngilizce derslerinizin herhangi bir saatinde kısa bir röportaj yapılacaktır.

Kimliğinizle ilgili bilgiler bu araştırma sonucu herhangi bir raporda yayınlanmayacaktır. Adınızla beraber verdiğiniz cevaplar kimse tarafından bilinmeyecektir.

Anket sorularına vereceğiniz cevaplar araştırmaya çok büyük katkı sağlayacaktır. Araştırmaya katılmak istiyorsanız, sayfanın altındaki ilgili yerleri doldurarak imzalayınız.

İngilizce Okutmanı Özlem DURAN

MA TEFL Programı

Bilkent / ANKARA

Bu formdaki bilgileri okudum ve araştırmaya katılmayı kabul ediyorum.

Adım & Soyadım:

Bölümüm:

Sınıfım:

İmza:

Tarih:

## APPENDIX E: TEACHER QUESTIONNAIRE

My Dear Colleague;

This questionnaire was designed to get the perceptions of the teachers on the washback effect of speaking tests, which is a part of my thesis namely teachers' and students' perceptions of washback effect of speaking tests in MA TEFL program at Bilkent University. The responses that you will give sincerely are very important for the sake of the study's validity and the reliability. The answers you give will be analysed taking your privacy into account. The questionnaire has two parts. I request you not to leave any of the questions empty and to fill in the optic form.

**Name and Surname:**.....

**Signature:**.....

**Date:** .....

### PART 1

1. **Age:** .....

2. **Graduated BA program:**

A. English Language Teaching

B. English Language and  
Literature / American Culture and  
Literature

C. Translation and Interpretation

D. Comparative Literature

Other

.....

3. **MA degree:**

.....

A. I did not do.

B. I did in ELT.

C. I did in Educational  
Sciences.

D. I have been doing in ELT.

E. I have been doing in  
Educational Sciences.

Other .....

4. **PhD:** .....

5. **Experience in teaching:** ..... years

6. **The length of time spent at your current institution:** ..... years

7. **How long have been administering speaking tests?:**..... years

**PART 2**

The statements below are equal to the letters in the boxes. Mark the letters which are next to the statements and fill in the optic form. Mark the best statement which reflects your idea.

Example:	Totally agree	Agree	Undecided	Disagree	Totally disagree
1. Speaking skills should be tested in our school.	A	B	C	D	E

A- Totally **agree**

B- Agree

C- Undecided

D- Disagree

E- Totally disagree

	Totally agree	Agree	Undecided	Disagree	Totally disagree
1. I believe that speaking skills can be taught in lessons.	A	B	C	D	E
2. Speaking skills may not be taught.	A	B	C	D	E
3. Speaking skills can be measured accurately.	A	B	C	D	E
4. Speaking skills can be measured effectively through written tests.	A	B	C	D	E
5. I spend more time on speaking parts in the days leading up to the speaking test.	A	B	C	D	E
6. There is no point in including extra speaking materials in the lessons.	A	B	C	D	E
7. Getting ready for the speaking test improves the general speaking skills of students.	A	B	C	D	E

8. Some of my students, though being unsuccessful in speaking tasks, perform well in speaking tests.	A	B	C	D	E
9. If speaking skills weren't tested, I wouldn't spend so much time on improving speaking skills.	A	B	C	D	E
10. Speaking tests encourage students to speak more in lessons.	A	B	C	D	E
11. The scores obtained from speaking tests reflect students' speaking levels in English.	A	B	C	D	E
12. I think that teaching speaking skills is important.	A	B	C	D	E
13. Speaking skills should be tested.	A	B	C	D	E
14. Speaking tests do not reflect students' speaking skills accurately.	A	B	C	D	E
15. The content of the speaking tests has an effect on my decision of the subjects which I will put emphasis on in lessons.	A	B	C	D	E
16. Speaking tests help students to notice the weaknesses in their speaking performances.	A	B	C	D	E
17. Speaking tests provide reliable information about students' speaking ability.	A	B	C	D	E
18. Speaking skills should be emphasized in lessons.	A	B	C	D	E
19. Trying to test speaking skills is a waste of time.	A	B	C	D	E
20. I think that testing speaking skills is difficult.	A	B	C	D	E
21. Speaking skills should be tested through speaking.	A	B	C	D	E
22. I give more importance to the parts which will be asked in the speaking test.	A	B	C	D	E
23. I organize extra speaking activities in classes, which are not alike with the activity that will be asked in the speaking test.	A	B	C	D	E

24. Speaking tests help students to notice the strengths in their speaking performances.	A	B	C	D	E
25. Speaking skills of students can be effectively measured without requiring them to speak.	A	B	C	D	E
26. Even if speaking skills were not tested, it would take part in my classes.	A	B	C	D	E
27. Students can also use many of the things that they have studied for the test, in lessons after the test.	A	B	C	D	E
28. Some of my students, though being successful in class activities, cannot perform well in speaking tests.	A	B	C	D	E
29. I give equal importance to all the speaking activities I do.	A	B	C	D	E
30. Students tend to forget lots of the things they have studied for the speaking test, after the test.	A	B	C	D	E
31. Relying on the speaking test scores, important decisions can be taken about students.	A	B	C	D	E
32. I do not believe that speaking skills can be taught in lessons.	A	B	C	D	E

Thank you for your participation

MA TEFL Student Özlem Duran

[ozlemduran@akdeniz.edu.tr](mailto:ozlemduran@akdeniz.edu.tr)

Supervisor: Asst. Prof. Dr. Philip Durrant

## APPENDIX F: ÖĞRETMEN ANKETİ

Değerli Meslektaşım;

Bu anket Bilkent Üniversitesi Yabancı Dil Olarak İngilizce Öğretimi bölümünde yaptığım ‘Öğretmenlerin ve Öğrencilerin, Konuşma Sınavlarının Öğretime Olan Etkisine Bakış Açısı’ adlı yüksek lisans tez çalışmamın öğretmenlerin bakış açısını öğrenmeye yönelik kısmı için düzenlenmiştir. Samimiyetle vereceğiniz cevaplar çalışmamın geçerliliği ve güvenilirliği açısından çok önemlidir. Vereceğiniz cevaplar gizlilik ilkelerine sadık kalınarak ele alınacaktır. Anket iki bölümden oluşmaktadır. Hiçbir soruyu boş bırakmamanızı ve **optik formu** doldurmanızı rica ederim.

**Adınız Soyadınız:**.....

**İmza:**.....

**Tarih:** .....

### 1.BÖLÜM

1. **Yaşınız:** .....

2. **Mezun olunan lisans programı:**

- A. İngilizce Öğretmenliği
- B. İngiliz Dili Edebiyatı / Amerikan Kültürü ve Edebiyatı
- C. Mütercim-Tercümanlık
- D. Karşılaştırmalı Edebiyat
- E. Diğer  
.....

3. **Yüksek lisans:**

.....

- A. Yapmadım.
- B. ELT alanında yaptım.
- C. Eğitim Bilimleri alanında yaptım.
- D. ELT alanında yapmaktayım.
- E. Eğitim Bilimleri alanında yapmaktayım.
- Diğer  
.....

4. **Doktora:** .....

5. **Öğretmenlikteki tecrübeniz:** ..... yıl

6. **Şu anki kurumunuzdaki toplam hizmet süreniz:** ..... yıl

7. **İngilizce konuşma sınavlarınızı kaç yıldır uygulamaktasınız?:** ..... yıl



## 2.BÖLÜM

Aşağıdaki ifadeler kutuların içindeki harflerle eşdeğerdir. Her ifade için fikrinizi en iyi yansıtan şıkkı işaretleyiniz ve optik forma kodlayınız.

Örnek:	Kesinlikle katılıyorum	Katılıyorum	Kararsızım	Katılmıyorum	Kesinlikle katılmıyorum
1. Okulumuzda konuşma becerileri test edilmelidir.	A	B	C	D	E

A- Kesinlikle **katılıyorum**

B- Katılıyorum

C- Kararsızım

D- Katılmıyorum

E- Kesinlikle katılmıyorum

	Kesinlikle katılıyorum	Katılıyorum	Kararsızım	Katılmıyorum	Kesinlikle katılmıyorum
1. Derslerde konuşma becerilerinin öğretilebileceğine inanırım.	A	B	C	D	E
2. Konuşma becerileri öğretilmese de olur.	A	B	C	D	E
3. Konuşma becerileri doğru bir şekilde ölçülebilir.	A	B	C	D	E
4. Konuşma becerileri yazılı sınavlar kullanılarak etkili bir şekilde ölçülebilir.	A	B	C	D	E
5. Konuşma sınavına yaklaşan günlerde konuşma bölümlerine daha fazla zaman harcarım.	A	B	C	D	E

	Kesinlikle katılıyorum	Katılıyorum	Kararsızım	Katılmıyorum	Kesinlikle katılmıyorum
6. Derslere ekstra konuşma materyali dahil etmenin anlamı yoktur.	A	B	C	D	E
7. Konuşma sınavı için hazırlanmak öğrencilerin genel konuşma becerilerini geliştirir.	A	B	C	D	E
8. Bazı öğrencilerim sınıftaki konuşma aktivitelerinde başarılı <u>olmamalarına</u> rağmen, konuşma sınavlarında iyi bir performans göstermektedirler.	A	B	C	D	E
9. Konuşma yeteneği test edilmeseydi, derslerde konuşma yeteneğini geliştirme üzerine bu kadar vakit harcamazdım.	A	B	C	D	E
10. Konuşma sınavları öğrencileri derslerde daha fazla konuşmaya teşvik eder.	A	B	C	D	E
11. Öğrencilerin konuşma sınavlarından aldıkları notlar İngilizce konuşma seviyelerini gösterir.	A	B	C	D	E
12. Konuşma becerilerinin öğretilmesinin önemli olduğunu düşünürüm.	A	B	C	D	E
13. Konuşma becerileri test edilmelidir.	A	B	C	D	E
14. Konuşma sınavları öğrencilerin konuşma yeteneklerini doğru bir şekilde yansıtmaz.	A	B	C	D	E
15. Konuşma sınavlarının içeriği derslerde üzerinde duracağım konulara karar vermemde etkilidir.	A	B	C	D	E
16. Konuşma sınavları öğrencilerin konuşma performanslarındaki zayıf yönlerini fark etmelerine yardım eder.	A	B	C	D	E
17. Konuşma sınavları öğrencilerin konuşma yetenekleri ile ilgili güvenilir bilgiler sağlar.	A	B	C	D	E
18. Derslerde konuşma becerilerinin üzerinde durulması gerekir.	A	B	C	D	E
19. Konuşma becerilerinin test edilmeye çalışılması vakit kaybıdır.	A	B	C	D	E

	Kesinlikle katılıyorum	Katılıyorum	Kararsızım	Katılmıyorum	Kesinlikle katılmıyorum
20. Konuşma becerilerinin test edilmesinin zor olduğunu düşünürüm.	A	B	C	D	E
21. Konuşma becerileri, konuşma yoluyla test edilmelidir.	A	B	C	D	E
22. Konuşma sınavında sorulacak olan konuşma bölümlerine daha fazla önem veririm.	A	B	C	D	E
23. Derslerde konuşma sınavında sorulacak olan aktiviteye <u>benzemeyen</u> ekstra konuşma aktiviteleri düzenlerim.	A	B	C	D	E
24. Konuşma sınavları öğrencilerin konuşma performanslarındaki güçlü yönlerini fark etmelerine yardım eder.	A	B	C	D	E
25. Öğrencilerin konuşma becerileri, konuşmalarına gerek olmadan da etkili bir şekilde ölçülebilir.	A	B	C	D	E
26. Konuşma becerileri, test edilmese bile derslerimde yer alır.	A	B	C	D	E
27. Öğrenciler konuşma sınavı için çalıştıkları birçok şeyi sınavdan sonra derslerde de kullanabilirler.	A	B	C	D	E
28. Bazı öğrencilerim sınıftaki konuşma aktivitelerinde başarılı olmalarına rağmen, konuşma sınavlarında iyi bir performans <u>gösterememektedirler</u> .	A	B	C	D	E
29. Derslerimde tüm konuşma aktivitelerini hepsine eşit önem vererek işlerim.	A	B	C	D	E
30. Öğrenciler konuşma sınavı için çalıştıkları birçok şeyi, sınavdan sonra kolayca unutma eğilimindedirler.	A	B	C	D	E
31. Konuşma sınavı notlarına güvenilerek öğrencilerle ilgili önemli kararlar verilebilir.	A	B	C	D	E
32. Konuşma becerilerinin derslerde öğretilbileceğini <u>düşünmem</u> .	A	B	C	D	E

Katılımınız için teşekkür ederim.  
ozlemduran@akdeniz.edu.tr

MA TEFL Öğrencisi Özlem Duran  
Tez Danışmanı: Yrd. Doç. Dr. Philip Durrant

## APPENDIX G: STUDENT QUESTIONNAIRE

**THE QUESTIONNAIRE OF PREP CLASS STUDENTS TO GET THEIR PERCEPTIONS OF WASHBACK EFFECT OF SPEAKING TESTS**

This questionnaire, which constitutes a part of a study aimed to survey the washback effect of English speaking tests that of Akdeniz University prep class students. The statements do not have only one correct answer, thus it is very important to understand the statements and mark the box which reflects your idea best for the sake of the validity and reliability of the questionnaire. Optic form also needs to be filled in without leaving missing parts. The responses you will give will be kept as a secret.

- **Name-Surname:**  
.....
  - **Signature:**  
.....  
...
  - **Date:**.....
  - **Age:** .....
1. I have not taken any English classes before.
  2. I am a graduate of high / vocational high school.
  3. I am a graduate of Anatolian technical / vocational high school.
  4. I am a graduate of a college / Anatolian high school.
  5. I have studied prep class before.
- **Circle the number(s) which fits your situation.**
  - **The general ranges of scores I get form the speaking parts of the midterms out of 20:**  
a. 20-16      b. 12-16      c. 12-8      d. 8-4      e. 0-4
  - **The ranges of scores I generally get from the speaking quizzes out of 100:**  
f. 100-85      g. 85-70      h. 70-55      i. 55-40      j.40-25      k. 25-0

The statements below are equal to the letters in the boxes. Mark the letters from A-E which are next to the statements and fill in the optic form. Mark the best statement which reflects your idea.

A- Totally **agree**

D- Disagree

B- Agree

E- Totally disagree

C- Undecided

<b>EXAMPLE</b>	<b>Totally agree</b>	<b>Agree</b>	<b>Undecided</b>	<b>Disagree</b>	<b>Totally disagree</b>
<i>Example statement:</i> It is necessary to test speaking skills in our school.	<input checked="" type="radio"/> A	<input type="radio"/> B	<input type="radio"/> C	<input type="radio"/> D	<input type="radio"/> E

	<b>Totally agree</b>	<b>Agree</b>	<b>Undecided</b>	<b>Disagree</b>	<b>Totally disagree</b>
1. I believe that English speaking skills can be learned.	<input type="radio"/> A	<input type="radio"/> B	<input type="radio"/> C	<input type="radio"/> D	<input type="radio"/> E
2. Speaking skills should be tested.	<input type="radio"/> A	<input type="radio"/> B	<input type="radio"/> C	<input type="radio"/> D	<input type="radio"/> E
3. It is fun to have a speaking test.	<input type="radio"/> A	<input type="radio"/> B	<input type="radio"/> C	<input type="radio"/> D	<input type="radio"/> E
4. I can also use many of the things which I have studied for the speaking test in lessons.	<input type="radio"/> A	<input type="radio"/> B	<input type="radio"/> C	<input type="radio"/> D	<input type="radio"/> E
5. I try to practice the speaking activities which will be asked in speaking tests in daily life.	<input type="radio"/> A	<input type="radio"/> B	<input type="radio"/> C	<input type="radio"/> D	<input type="radio"/> E
6. Getting ready for speaking tests has improved my English.	<input type="radio"/> A	<input type="radio"/> B	<input type="radio"/> C	<input type="radio"/> D	<input type="radio"/> E
7. I believe that learning speaking skills is important.	<input type="radio"/> A	<input type="radio"/> B	<input type="radio"/> C	<input type="radio"/> D	<input type="radio"/> E
8. The grade I get from speaking exams correctly reflects my speaking ability.	<input type="radio"/> A	<input type="radio"/> B	<input type="radio"/> C	<input type="radio"/> D	<input type="radio"/> E
9. Speaking tests make me nervous.	<input type="radio"/> A	<input type="radio"/> B	<input type="radio"/> C	<input type="radio"/> D	<input type="radio"/> E
10. If I knew that a speaking activity would not be asked, I would not spend time practicing it in class.	<input type="radio"/> A	<input type="radio"/> B	<input type="radio"/> C	<input type="radio"/> D	<input type="radio"/> E

11. I notice my weaknesses in speaking after speaking tests.	A	B	C	D	E
12. Speaking tests do not show my real speaking level.	A	B	C	D	E
13. The administration of speaking tests should be continued in the preparatory school program.	A	B	C	D	E
14. I take the opportunities to improve my English speaking skills in lessons.	A	B	C	D	E
15. Getting ready for speaking tests has improved my speaking skills.	A	B	C	D	E
16. It is important to be successful in speaking test.	A	B	C	D	E
17. Even if they were not tested, speaking skills should have a place in lessons.	A	B	C	D	E
18. Speaking skills can be improved in lessons.	A	B	C	D	E
19. My real speaking ability is reflected in my scores in the speaking exams.	A	B	C	D	E
20. I give more importance to the speaking parts which will be asked in the test.	A	B	C	D	E
21. It is important to test speaking skills.	A	B	C	D	E
22. Speaking tests have decreased my speaking skills.	A	B	C	D	E
23. I participate in all speaking activities in the book in classes by giving equal importance to all of them.	A	B	C	D	E
24. I think that the time spared for speaking skills is unnecessary.	A	B	C	D	E
25. If speaking sections were given 10 points in total in midterms, I would participate less in speaking activities in lessons.	A	B	C	D	E

26. The speaking activities which are not alike with the speaking activities that will be asked in the test do not take my attention.	A	B	C	D	E
27. 20 points that are spared for the speaking sections in midterms are too much.	A	B	C	D	E
28. If it were not tested, I would not try to practice speaking skills in lessons.	A	B	C	D	E
29. Speaking tests are necessary in order to learn to speak English.	A	B	C	D	E
30. Speaking activities are necessary in class.	A	B	C	D	E
31. I participate more in speaking parts in the days leading up to the speaking tests.	A	B	C	D	E
32. I easily forget many of the things I have studied for the test after the test.	A	B	C	D	E
33. It is a loss of time to test speaking skills.	A	B	C	D	E
34. Speaking exams provide an accurate picture my speaking ability.	A	B	C	D	E

You can write your questions and points of view here. (The participants who want to get answers to their questions are requested to leave their e-mail addresses in order to be get in touch with them.)

.....

.....

.....

.....

.....

Thanks for your participation  
 Bilkent University MA TEFL Student Özlem Duran  
 Supervisor: Asst. Prof. Dr. Philip Durrant  
 e-mail: [ozlemduran@akdeniz.edu.tr](mailto:ozlemduran@akdeniz.edu.tr)

## APPENDIX H: ÖĞRENCİ ANKETİ

**HAZIRLIK ÖĞRENCİLERİNİN İNGİLİZCE KONUŞMA SINAVLARININ  
ÖĞRETİMLERİNE OLAN ETKİSİNİ ÖLÇME ANKETİ**

Bu anket Akdeniz Üniversitesi hazırlık sınıfı öğrencilerinin İngilizce konuşma sınavlarının öğretimlerine olan etkisini ölçmek amacıyla yapılan araştırmanın bir parçasını oluşturmaktadır. İfadelerin tek bir doğru yanıtı yoktur, bu yüzden maddeleri anlamamız ve fikrinizi en iyi yansıtan kutuyu işaretlemeniz anketin geçerliliği ve güvenilirliği açısından oldukça önemlidir. Optik formun da eksiksiz olarak doldurulması gerekmektedir. Adınızla vereceğiniz yanıtlar gizli tutulacaktır.

- **Adı – Soyadı:** .....
- **İmza:** .....
- **Tarih:** .....
- **Yaş:** .....
- **Size uygun olan rakam(lar)ı yuvarlak içine alınız.**
  - 6. Daha önce hiç İngilizce dersi almamıştım.
  - 7. Düz lise / Meslek lisesi mezunuyum.
  - 8. Anadolu Teknik / Anadolu Meslek Lisesi mezunuyum.
  - 9. Kolej / Anadolu Lisesi mezunuyum.
  - 10. Daha önce de hazırlık okumuştum.
- **Vizelerin konuşma bölümlerinden 20 üzerinden genellikle aldığım not aralığı:**
  - a. 20-16      b. 12-16      c. 12-8      d. 8-4      e. 0-4
- **Konuşma quizlerinden 100 üzerinden genellikle aldığım not aralığı:**
  - f. 100-85      g. 85-70      h. 70-55      i. 55-40      j.40-25      k. 25-0

Aşağıdaki ifadeler kutuların içindeki harflerle eşdeğerdir. Yazılı metnin yanındaki A' dan E' ye kadar olan harfleri işaretleyiniz ve optik forma kodlayınız. Her soru için fikrinizi en iyi yansıtan ifadeyi işaretleyiniz.

A- Kesinlikle **katılıyorum**

D- Katılmıyorum

B- Katılıyorum

E- Kesinlikle katılmıyorum

C- Kararsızım



ÖRNEK:	Kesinlikle katılıyorum	Katılıyorum	Kararsızım	Katılmıyorum	Kesinlikle katılmıyorum
<i>Örnek ifade:</i> Okulumuzda konuşma becerilerinin test edilmesi gereklidir.	A	B	C	D	E

	Kesinlikle katılıyorum	Katılıyorum	Kararsızım	Katılmıyorum	Kesinlikle katılmıyorum
1. İngilizce konuşma becerilerinin öğrenilebileceğine inanırım.	A	B	C	D	E
2. Konuşma becerileri test edilmelidir.	A	B	C	D	E
3. Konuşma sınavı olmak eğlencelidir.	A	B	C	D	E
4. Konuşma sınavı için çalıştığım birçok şeyi sınavdan sonra derslerde de kullanabilirim.	A	B	C	D	E
5. Konuşma sınavlarında sorulacak konuşma aktivitelerini günlük hayatta pratik etmeye çalışırım.	A	B	C	D	E
6. Konuşma sınavlarına hazırlanmak İngilizcemi geliştirdi.	A	B	C	D	E
7. Konuşma becerilerinin öğrenilmesinin önemli olduğunu düşünürüm.	A	B	C	D	E
8. Konuşma sınavlarından aldığım puan konuşma yeteneğimi doğru bir şekilde yansıtır.	A	B	C	D	E
9. Konuşma sınavları beni gerer.	A	B	C	D	E
10. Sınavlarda herhangi bir konuşma aktivitesinin sorulmayacağını bilsem, derste onu pratik etmek için vakit harcamam.	A	B	C	D	E
11. Konuşma sınavlarından sonra konuşmadaki zayıflıklarımı fark ederim.	A	B	C	D	E

	Kesinlikle katılıyorum	Katılıyorum	Kararsızım	Katılmıyorum	Kesinlikle katılmıyorum
12. Konuşma sınavları gerçek konuşma seviyemi göstermez.	A	B	C	D	E
13. Konuşma sınavları hazırlık programında uygulanmaya devam edilmelidir.	A	B	C	D	E
14. Derslerde İngilizce konuşma becerilerini ilerletmek için fırsatları değerlendiririm.	A	B	C	D	E
15. Konuşma sınavlarına hazırlanmak konuşma yeteneğimi geliştirdi.	A	B	C	D	E
16. Konuşma sınavlarında başarılı olmak önemlidir.	A	B	C	D	E
17. Konuşma becerileri test edilmese bile derslerde yer almalıdır.	A	B	C	D	E
18. Konuşma becerileri derslerde geliştirilebilir.	A	B	C	D	E
19. Gerçek konuşma yeteneğim konuşma sınavından aldığım notlara yansır.	A	B	C	D	E
20. Konuşma sınavında sorulacak olan konuşma bölümlerine derslerde daha fazla önem veririm.	A	B	C	D	E
21. Konuşma yeteneğinin test edilmesi önemlidir.	A	B	C	D	E
22. Konuşma sınavları konuşma yeteneğimi azalttı.	A	B	C	D	E
23. Derslerde kitaptaki tüm konuşma aktivitelerine, hepsine eşit önem vererek katılıyorum.	A	B	C	D	E
24. Derslerde konuşma becerilerine ayrılan vaktin gereksiz olduğunu düşünürüm.	A	B	C	D	E
25. Vizelerdeki konuşma bölümlerine toplamda 10 puan ayrılrsa, derslerdeki konuşma aktivitelerine daha az katılıyorum.	A	B	C	D	E
26. Konuşma sınavında sorulacak olan aktiviteye <u>benzemeyen</u> konuşma aktiviteleri derslerde ilgimi çekmez.	A	B	C	D	E

	Kesinlikle katılıyorum	Katılıyorum	Kararsızım	Katılmıyorum	Kesinlikle katılmıyorum
27. Vizelerde konuşma bölümleri için ayrılan 20 puan çok fazladır.	A	B	C	D	E
28. Test edilmeseydi, derslerde konuşma becerilerini pratik etmek için çaba harcamazdım.	A	B	C	D	E
29. İngilizce konuşmayı öğrenmek için konuşma sınavları gereklidir.	A	B	C	D	E
30. Sınıfta konuşma aktiviteleri gereklidir.	A	B	C	D	E
31. Konuşma sınavına yaklaşan günlerde konuşma bölümlerine daha fazla katılırım.	A	B	C	D	E
32. Konuşma sınavı için çalıştığım birçok şeyi, sınavdan sonra kolayca unuturum.	A	B	C	D	E
33. Konuşma becerilerinin test edilmesi vakit kaybıdır.	A	B	C	D	E
34. Konuşma sınavları konuşma yeteneğimin sınırları hakkında doğru bilgiler sağlar.	A	B	C	D	E

Bu çalışmayla ilgili soru ve görüşlerinizi buraya yazabilirsiniz. (Sorularına yanıt almak isteyen katılımcıların kendilerine ulaşılabilmesi için e-posta adreslerini not etmeleri rica olunur.)

.....

.....

.....

.....

.....

.....

.....

Katılımınız için teşekkür ederim.

Bilkent Üniversitesi MA TEFL Öğrencisi Özlem Duran  
Tez Danışmanı: Yrd. Doç. Dr. Philip Durrant  
e-posta: [ozlemduran@akdeniz.edu.tr](mailto:ozlemduran@akdeniz.edu.tr)

APPENDIX I: CATEGORIZATION OF THE TEACHER QUESTIONNAIRE  
ITEMS ON TEACHERS' ATTITUDES TOWARDS AND BELIEFS ABOUT  
TEACHING AND TESTING SPEAKING

*a. Do teachers have positive attitudes towards the possibility of teaching speaking? (Cronbach's Alpha= .3937)*

1. I believe that speaking skills can be taught in lessons.

32. I do not believe that speaking skills can be taught in lessons. (R)

*b. Do teachers have positive attitudes towards the importance of teaching speaking?*

12. I think that teaching speaking skills is important.

*c. Do teachers believe that speaking should be tested? (Cronbach's Alpha = .5802)*

13. Speaking skills should be tested.

19. Trying to test speaking skills is a waste of time. (R)

*d. Do teachers believe that speaking can be tested?*

3. Speaking skills can be measured accurately. (Individually treated)

14. Speaking tests do not reflect students' speaking skills accurately. (R)  
(Individually treated)

20. I think that testing speaking skills is difficult. (R) (Individually treated)

*e. Do teachers think that speaking skills should be tested through speaking? [Cronbach's Alpha= .6716 (Q.21 and 25)]*

21. Speaking skills should be tested through speaking.

25. Speaking skills of students can be effectively measured without requiring them to speak. (R)

4. Speaking skills can be measured effectively through written tests.

(Individually treated)

*f. Do teachers believe that the results of speaking tests can be used as a reliable diagnostic tool? (Cronbach's Alpha= .6626)*

8. Some of my students, though being unsuccessful in speaking tasks, perform well in speaking tests. (R)

11. The scores obtained from speaking tests reflect students' speaking levels in English.

17. Speaking tests provide reliable information about students' speaking ability.

28. Some of my students, though being successful in class activities, cannot perform well in speaking tests. (R)

31. Relying on the speaking test scores, important decisions can be taken about students.

APPENDIX J: CATEGORIZATION OF TEACHER QUESTIONNAIRE ITEMS  
ON THE TEACHERS' PERCEPTIONS OF THE WASHBACK EFFECT OF  
CLASSROOM-BASED SPEAKING TESTS

*a. Do teachers tailor their speaking classes according to speaking tests in terms of content and allocation of time for speaking activities?*

*[Cronbach's Alpha= .5948 (Q. 9, 15, 22, 26, and 29)]*

5. I spend more time on speaking parts in the days leading up to the speaking test. (Individually treated)

9. If speaking skills weren't tested, I wouldn't spend so much time on improving speaking skills.

15. The content of the speaking tests has an effect on my decision of the subjects which I will put emphasis on in lessons.

22. I give more importance to the parts which will be asked in the speaking test.

29. I give equal importance to all the speaking activities I do. (R)

*b. Do teachers have positive attitudes towards organising extra speaking activities in classes, which are not alike with the activity that will be asked in the speaking test? (Cronbach's Alpha= .6046)*

6. There is no point in including extra speaking materials in the lessons. R

23. I organize extra speaking activities in classes, which are not alike with the activity that will be asked in the speaking test.

*c. Do teachers believe that speaking tests have a positive effect on their students' speaking ability? [Cronbach's Alpha= .7590 (Q.16 and 24)]*

16. Speaking tests help students to notice the weaknesses in their speaking performances.

24. Speaking tests help students to notice the strengths in their speaking performances.

7. Getting ready for the speaking test improves the general speaking skills of students. (Individually treated)

27. Students can also use many of the things that they have studied for the test, in lessons after the test. (Individually treated)

30. Students tend to forget lots of the things they have studied for the speaking test, after the test. (R) (Individually treated)

10. Speaking tests encourage students to speak more in lessons. (Individually treated)

*d. An independent question*

26. Even if speaking skills were not tested, it would take part in my classes.

APPENDIX K: CATEGORIZATION OF THE STUDENT QUESTIONNAIRE ITEMS ON  
STUDENTS' ATTITUDES TOWARDS AND BELIEFS ABOUT TEACHING AND  
TESTING SPEAKING

*a. Do students have positive attitudes towards learning speaking? (Cronbach's  
alpha= .6779)*

- 1. I believe that English speaking skills can be learned.
- 7. I believe that learning speaking skills is important.
- 14. I take the opportunities to improve my English speaking skills in lessons.
- 18. Speaking skills can be improved in lessons.
- 24. I think that the time spared for speaking skills is unnecessary. (R)
- 30. Speaking activities are necessary in class.

*b. Do students believe the importance of testing speaking? (Cronbach's alpha=  
.7389)*

- 2. Speaking skills should be tested.
- 13. The administration of speaking tests should be continued in the preparatory school program.
- 16. It is important to be successful in speaking test.
- 21. It is important to test speaking skills.
- 27. 20 points that are spared for the speaking sections in midterms are too much. (R)
- 33. It is a loss of time to test speaking skills. (R)

*c. Do students enjoy speaking tests? (Cronbach's Alpha= .5697)*

- 3. It is fun to have a speaking test.
- 9. Speaking tests make me nervous. (R)



*d. Do students think that speaking tests can really show their speaking ability?*

*(Cronbach's Alpha= .7778)*

8. The grade I get from speaking exams correctly reflects my speaking ability.

12. Speaking tests do not show my real speaking level. (R)

19. My real speaking ability is reflected in my scores in the speaking exams.

34. Speaking exams provide an accurate picture of my speaking ability.

APPENDIX L: CATEGORIZATION OF STUDENT QUESTIONNAIRE ITEMS ON THE STUDENTS' PERCEPTIONS OF THE WASHBACK EFFECT OF CLASSROOM-BASED SPEAKING TESTS

*a. Do students believe that speaking tests have positive effects on their speaking ability? (Cronbach's Alpha= .6809)*

4. I can also use many of the things which I have studied for the speaking test in lessons.

6. Getting ready for speaking tests has improved my English.

11. I notice my weaknesses in speaking after speaking tests.

15. Getting ready for speaking tests has improved my speaking skills.

32. I easily forget many of the things I have studied for the test after the test. (R)

*b. Do speaking tests have an influence on what and how students learn? (Cronbach's Alpha= .7145)*

10. If I knew that a speaking activity would not be asked, I would not spend time practicing it in class.

25. If speaking sections were given 10 points in total in midterms, I would participate less in speaking activities in lessons.

26. The speaking activities which are not alike with the speaking activities that will be asked in the test do not take my attention.

28. If it were not tested, I would not try to practice speaking skills in lessons.

*c. Do speaking tests have any effects on what students really do?*

5. I try to practice the speaking activities which will be asked in speaking tests in daily life. (Individually treated)

31. I participate more in speaking parts in the days leading up to the speaking tests. (Individually treated)

20. I give more importance to the speaking parts which will be asked in the test.

(Individually treated)

*d. An independent question*

17. Even if they were not tested, speaking skills should have a place in lessons.

## APPENDIX M: THE STUDENT QUESTIONNAIRE OF CAINE (2005)

## Student questionnaire #2

The following statements refer to the end-of-term speaking test you have just done in Oral Communication. Grade each one on a 4-point scale, where:

1 = strongly disagree

2 = disagree

3 = agree

4 = strongly agree

Write your answers in the brackets

- (1) [ ] I enjoyed doing the speaking test with a partner.
- (2) [ ] Studying for the speaking test improved my English.
- (3) [ ] I made a greater effort to speak English in the weeks leading up to the test.
- (4) [ ] It was important to do well on the test.
- (5) [ ] It was possible to do well on the test without much preparation.
- (6) [ ] If the test was included as part of the Seibu Moshi or Center Test I would do more preparation.
- (7) [ ] I enjoyed practising for the test during class.
- (8) [ ] It was important to practise for the test during class.
- (9) [ ] There wasn't enough time to practise for the test.
- (10) [ ] It is more important to practice other language skills during class.
- (11) [ ] When not in class, it is better to study for a speaking test on your own.
- (12) [ ] When not in class, it is better to study for a speaking test with friends.

(13) [ ] I would have liked to have had access to more materials for practice out of class.

(14) [ ] It is difficult to study for a speaking test in your own time without a teacher.

(15) [ ] It wasn't necessary to study for the speaking test in my own time.

(16) [ ] In order to do well on the test it is necessary to memorize key phrases/vocabulary.

(17) [ ] It is only necessary to speak English during Oral Communication classes.

(18) [ ] I want to improve my English speaking skills.

(19) [ ] It isn't important for me to speak English, so speaking shouldn't be tested.

Comments:

– That is the end of the questionnaire –

Thank you very much for your help

## APPENDIX N: 2 NUMARALI ÖĞRENCİ RÖPORTAJINDAN BİR KESİT

Araştırmacı: Derslerde konuşma becerilerini öğrenebileceğine /öğrenilebileceğine inanıyor musun?

Öğrenci katılımcı 2: Evet inanıyorum. Bu gerçek hayata göre tecrübe oluyor. İnanıyorum yani.

A: Peki... Sence önemli mi?

ÖK2: Bence çok çok önemli. Şu andaki sistemden konuşmaya daha çok önem vermeliyiz derslerde.

A: Bu şekilde tamam... Konuşma sınavları ile ilgili ne düşünüyorsun?

ÖK2: Bence konuşma sınavları daha da çok olmalı.

A: Bildiğim kadarıyla 6 midtermin hepsinde oluyor.

ÖK2: Evet 6 midtermin hepsinde... Bir de 4 tane quiz çeşidi var. İşte 6 haftada bir ancak (üzgün☹)geliyor. Biraz az oluyor. Biraz...

A: Hımmmm...

ÖK2: İşte az oluyor.

A: Bir buçuk ayda bir speaking quizi oluyor denk gelirse... Önemli olduğunu düşünüyorsun?

ÖK2: Evet.

A: Önemli olduğunu düşünüyorsun. Peki, sence 20 puan nasıl? Çok mu?

ÖK2: Bence hepsinin 20 puan olması biraz kötü oldu.

A: Hangi açıdan? Hangi konularda daha iyisin?...

## APPENDIX O: A PART OF THE STUDENT 2 INTERVIEW

Researcher: Do you believe that you can learn or speaking skills can be learned in classes?

Student interviewee 2: Yes, I do. It is an experience for the real life. I mean I believe.

R: Ok, do you think that it is important?

SI2: I think it is very important. We should give more importance to speaking in classes rather than the system we are in.

R: All right... What do you think about speaking tests?

SI2: I think speaking tests should be a lot more.

R: As far as I know it is included in all the six midterms.

SI2: Yes, in all six midterms... In addition there are four types of quiz types. So, it is speaking's turn once six weeks (sad). It is a bit few. Few...

R: Himm...mm...

SI2: As a result it is few.

R: There is a speaking quiz once in every one and a half month. You think that it is important?

SI2: Yes.

R: You think that it is important. Ok, what about 20 points? Is it a lot?

SI2: I think it turned out to be bad that all have 20 points.

R: In what aspects? On which subjects are you better?...

APPENDIX P: 1 NUMARALI ÖĞRETMEN RÖPORTAJINDAN BİR KESİT

Araştırmacı: Evet, merhaba .....(özel isim) hocam.

Öğretmen katılımcı 1: Merhaba.

A: Derslerde konuşma becerilerinin öğretilebileceğine inanıyor musunuz?

ÖK1: Tabi ki.

A: Önemli olduğunu düşünüyorsunuz o zaman?

ÖK1: Evet, çok.

A: Sizce peki test edilebilir bir yetenek midir konuşma?

ÖK1: Evet (gülüşmeler)

A: Konuşma yeteneği doğru bir şekilde ölçülebilir mi?

ÖK1: Diğer beceriler ne kadar doğru bir şekilde ölçülebilirse o da, evet, o da ölçülebilir.

A: Yani diğerlerinin, burada bir kinaye var mı? Yani diğerlerinin doğru ölçülebileceğine inanıyor musunuz?

ÖK1: Yani productiona dayalı olan şeyler ne kadar ölçülebilirse, tabi ki speaking de o ölçüde...

A: Ölçülebilir. Mesela anketinizde de şunu analiz ettiğimde çok enteresan gelmişti ve ne düşündüğünüzü gerçekten çok merak ediyorum. Konuşma sınavları konuşma yoluyla mı test edilmelidir yoksa alternatif çözümler gerçekten işe yarıyor mu? Writing yoluyla ölçülmesi...

ÖK1: Tabi ki konuşma yoluyla test edilmeli ama hani olmadığı durumlarda çünkü Türkiye koşulları çok kalabalık. Farklı yöntemlerle de test edilmesi olabilir. Yani tercih edilmez ama mecbur kalınca bir alternatif...



## APPENDIX Q: A PART OF THE TEACHER 1 INTERVIEW

Researcher: Well, hi ..... (name) teacher.

Teacher interviewee 1: Hi.

R: Do you believe that speaking skills can be taught?

T11: Of course.

A: So you think it is important?

T11: Yes, a lot.

A: Do you think that speaking skills can be tested?

T11: Yes (laughing).

R: Can speaking skills be measured accurately?

S11: It can be measured to the extent the other skills can be measured accurately.

R: Well, is there a sarcasm here? Do you believe that other skills can be measured accurately?

S11: I mean speaking can be [measured] to the extent the other skills which are based on production can be measured...

R: Measured. While I was analysing your questionnaire, I found something which was very interesting for me and I am really curious about what you think. Should speaking ability be tested through speaking or do alternative solutions work? Testing through writing...