

**AN INVESTIGATION OF KNOWLEDGE, PRACTICES AND
PROFESSIONAL TRAINING NEEDS OF HIGH SCHOOL EFL TEACHERS
FOR ORAL LANGUAGE ASSESSMENT**



Ayten KAPLAN

June 2018

**AN INVESTIGATION OF KNOWLEDGE, PRACTICES AND
PROFESSIONAL TRAINING NEEDS OF HIGH SCHOOL EFL TEACHERS
FOR ORAL LANGUAGE ASSESSMENT**

**A THESIS SUBMITTED TO THE
GRADUATE SCHOOL OF EDUCATIONAL SCIENCES
OF
BAHÇEŞEHİR UNIVERSITY**

Ayten Kaplan

**IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF ARTS
IN THE DEPARTMENT OF ENGLISH LANGUAGE TEACHING**

JUNE 2018

Approval of the Graduate School of Educational Sciences



Assist Prof. Enisa MEDE
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Arts.



Assist Prof. Hatime ÇİFTÇİ
Coordinator

This is to certify that we have read this thesis, and, in our opinion, it is fully adequate, in scope and quality, as a thesis for the degree of Master of Arts.



Assist. Prof. Yeşim Keşli DOLLAR
Supervisor

Examining Committee Members

Assist. Prof. Enisa MEDE

(BAU, ELT)

Assist. Prof. Yeşim KEŞLİ DOLLAR

(BAU, ELT)

Assist Prof. Aynur KESEN MUTLU

(MU, ELT)



I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name : Ayten KAPLAN

Signature :

A handwritten signature in blue ink, appearing to read 'Ayten Kaplan', written over a colon symbol.

ABSTRACT

AN INVESTIGATION OF KNOWLEDGE, PRACTICES AND PROFESSIONAL TRAINING NEEDS OF HIGH SCHOOL EFL TEACHERS FOR ORAL LANGUAGE ASSESSMENT

Kaplan, Ayten

Master's Thesis, Master's Program in English Language Education

Supervisor: Assist. Prof. Yeşim Keşli Dollar

June 2018, 123 pages

Intercultural communication is increasingly getting significant in today's global world; therefore, teaching to speak English as a lingua franca is one of the primary aims of educational systems. Herewith, EFL teachers need to develop their teaching skills for a successful speaking instruction. Accordingly, as assessment is one of the indispensable part of teaching processes (Brown, 2000), EFL teachers also need to assess oral language skills efficiently. The purpose of this thesis was to find out high school EFL teachers' knowledge, current practices, factors influencing their practices and mainly, their training needs to assess oral language skills efficiently. The participants were 50 high school teachers, 42 females and 8 males, who work in Beşiktaş, İstanbul. The data were collected via two surveys and interviews. The results indicated that, although high school teachers try to assess oral language skills in their classrooms, they have various shortcomings in their practices either due to lack of procedural knowledge or challenges in their contexts. Eventually, high school teachers were revealed to have training needs on many aspects of oral language assessment in order to overcome difficulties and carry out an effective assessment.

Keywords: Speaking Assessment, Oral Language Assessment

ÖZ

LİSE İNGİLİZCE ÖĞRETMENLERİNİN KONUŞMA BECERİSİNİN DEĞERLENDİRİLMESİ KONUSUNDAKİ BİLGİLERİNİN, UYGULAMALARININ VE HİZMET İÇİ EĞİTİM İHTİYAÇLARININ İNCELENMESİ

Kaplan, Ayten

Yüksek Lisans, İngiliz Dili Eğitimi Yüksek Lisans Programı

Tez Yöneticisi: Dr. Öğretim Üyesi Yeşim Keşli DOLLAR

Haziran 2018, 123 sayfa

Günümüz küresel dünyasında, kültürler arası iletişimin önemi giderek artmaktadır. Dolayısıyla, ortak dil olan İngilizceyi konuşmayı öğretmek eğitim sistemlerinin de temel amaçlarından biri haline gelmiştir. Bununla birlikte, konuşma becerisini başarılı bir şekilde öğretebilmek için İngilizce öğretmenlerinin de öğretme becerilerini geliştirmeleri gerekmektedir. Bu doğrultuda, değerlendirmenin öğretim süreçlerinin vazgeçilmez bir parçası olduğu göz önüne alındığında (Brown, 2000), İngilizce öğretmenlerin sınıflarında etkili bir konuşma becerisi değerlendirmesi yapması da elzemdir. Bu tezin amacı İngilizce öğretmenlerinin konuşma becerisinin değerlendirilmesi konusundaki bilgilerini, uygulamalarını, bilgi ve uygulamalarına etki eden şartları ve esasen İngilizce konuşma becerisini etkili bir şekilde değerlendirebilmek için hizmet içi eğitim ihtiyaçlarını araştırmaktır. Çalışmaya Beşiktaş'taki liselerde görevli 42'si kadın ve 8'i erkek olmak üzere 50 İngilizce öğretmeni katılmıştır. Veri toplamak için iki anket ve röportajlar kullanılmıştır. Çalışmanın sonuçları, öğretmenlerin sınıflarında konuşma becerisini değerlendirmeye çalışmalarına rağmen gerek yordam bilgisi yetersizliğinden gerekse çalışma koşullarından kaynaklanan çeşitli eksikliklerin olduğunu göstermiştir. Sonuç

olarak, öğretmenlerin zorluklarla başa çıkabilmeleri ve etkili bir konuşma becerisi değerlendirmesi yapabilmeleri için bu kapsamdaki birçok konuda hizmet içi eğitime ihtiyaç duydukları ortaya konmuştur.

Anahtar kelimeler: Konuşma Becerisinin Değerlendirilmesi, Sözlü Dil Becerisinin Değerlendirilmesi





To my beloved fiancé

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my thesis advisor, Assist. Prof. Yeşim Keşli DOLLAR for both her excellent guidance with her valuable feedback that lighted my way throughout my research, and for her invaluable moral support and encouragement in this period.

I also wish to thank to Assist. Prof. Enisa MEDE and Assist. Prof. Aynur KESEN MUTLU for their inspiring comments and feedback that provided me with a better understanding of my research.

I must also express my gratitude to my colleagues at my school who assisted me in various phases of my thesis and to English teachers in Beşiktaş who contributed my research with their sincere answers to my questions in the interviews and questionnaires.

Last but not least, I wish to give my special thanks to my fiancé Tümay Turgay MISIRLIOĞLU, who helped me in each step of my study and gave me a continuous emotional support.

TABLE OF CONTENTS

ETHICAL CONDUCT	iii
ABSTRACT	iv
ÖZ	v
DEDICATION	vii
ACKNOWLEDGEMENTS	viii
TABLE OF CONTENTS	ix
LIST OF TABLES	xiv
LIST OF ABBREVIATIONS	xv
Chapter 1: Introduction	1
1.1 Overview	1
1.2 Theoretical Framework	1
1.2.1 The significance of speaking instruction in language teaching.....	1
1.2.2 The significance of the assessment in language teaching	2
1.2.3 The challenges encountered in the assessment of oral language skills ...	3
1.3 Statement of the Problem	3
1.4 Purpose of the Study.....	4
1.5 Research Questions	5
1.6 Significance of the Study	5
1.7 Definitions	6
Chapter 2: Literature Review	7
2.1 Overview	7
2.2 Fundamental Aspects of Oral Language Assessment	7
2.2.1 Definitions and significance of oral language assessment	7
2.2.2 Considerations in assessment of oral language performance	7
2.2.3 Defining oral language assessment constructs	8
2.2.3.1 Accuracy and fluency	10
2.2.3.2 Vocabulary	10

2.2.3.3 Pronunciation	10
2.2.3.4 Communicative competence	11
2.2.4 Determining criteria and designing rating scales	12
2.2.5 Assessment types	13
2.2.5.1 Formative and summative assessment	13
2.2.5.2 Alternative assessment techniques	14
2.2.5.2.1 Portfolio-assessment	14
2.2.5.2.2 Peer-assessment and Self-assessment	15
2.2.6 Oral language assessment tasks	16
2.2.6.1 Oral language assessment task types	17
2.2.6.2 Determining oral language assessment tasks	18
2.2.7 Analytic and holistic scoring systems	19
2.2.7.1 Pros and cons of analytic and holistic scoring	20
2.2.7.2 Determining which scoring system to use	21
2.2.8 Validity and reliability issues	21
2.2.8.1 Validity	22
2.2.8.2 Reliability	22
2.2.9 Washback effect	23
2.3 Empirical Studies on Teachers Role as Practitioners of Assessment	25
2.3.1 Teachers' knowledge, practices and training needs for language assessment	25
2.3.2 Teachers' knowledge, practices and training needs of oral language assessment	28
Chapter 3: Methodology	34
3.1 Overview	34
3.2 Research Design	34
3.3 Setting and Participants	35
3.4 Procedures	36
3.4.1 Data Collection Instruments	36
3.4.1.1 Questionnaires	37
3.4.1.1.1 Knowledge of speaking assessment questionnaire ...	37
3.4.1.1.2 Training needs analysis for oral language assessment survey	38

3.4.1.2 Semi-structured interviews	39
3.4.2 Data Collection Procedures	40
3.4.3 Data Analysis Procedures	41
Chapter 4: Findings	44
4.1 Overview	44
4.2 Findings	44
4.2.1 Findings for RQ–1: What do high school teachers know about oral language assessment?	44
4.2.1.1 Findings from KOSAQ	44
4.2.1.1.1 Knowledge of speaking assessment constructs	45
4.2.1.1.2 Knowledge of speaking assessment techniques and tasks	47
4.2.1.1.3 Knowledge of speaking assessment criteria	48
4.2.1.2 Findings from semi-structured interviews and open-ended items	50
4.2.1.2.1 Adequacy of high school teachers’ knowledge of oral language assessment	50
4.2.1.2.2 We are required	51
4.2.2 Findings for RQ – 2 What are high school teachers’ current practices of oral language assessment?	53
4.2.2.1 High school teachers try to assess oral language skills	53
4.2.2.2 The frequency of high school teachers’ oral language assessment	55
4.2.2.3 How high school teachers assess oral language skills	56
4.2.2.4 Assessment criteria that high school teachers use Data Analysis Procedures	56
4.2.2.4.1 Intelligibility instead of grammar accuracy	57
4.2.2.4.2 Vocabulary is important	57
4.2.2.4.3 Fluency, accuracy and content	58
4.2.2.5. Oral language assessment tasks that high school teachers use	59
4.2.2.5.1 Oral presentation	59
4.2.2.5.2 Role-Playing	59
4.2.2.5.3 Picture Description	60

4.2.3 Findings for RQ – 3: What are the factors in teachers’ contexts that have influence on speaking assessment?	62
4.2.3.1 Class size	63
4.2.3.2 Curriculum objectives	64
4.2.3.3 Having standardized / prescribed criteria	66
4.2.3.4 Lack of time	67
4.2.3.5 Student level	68
4.2.3.6 Students’ reluctance to speak	69
4.2.3.7 Subjectivity	70
4.2.3.8 Lack of teacher training	70
4.2.4 High school teachers’ training needs for oral language assessment	72
4.2.4.1 I need training	72
4.2.4.2 I need training on all aspects of oral language assessment	75
4.2.4.3 I don’t know what I need to know	77
4.2.4.4 Significance and benefits of assessing oral language skills	78
4.2.4.5 Defining what skills to assess	78
4.2.4.6 Specifying objectives	79
4.2.4.7 Techniques and tasks	79
4.2.4.8 Assessment criteria	81
4.2.4.9 Evaluating / Rating	83
4.2.4.10 Washback in the classroom	84
4.2.4.11 Designing assessment according to student level	84
4.2.4.12 Motivating students	85
4.2.4.13 Pronunciation and fluency	86
Chapter 5: Discussion and Conclusion	88
5.1 Overview	88
5.2 Discussion of Findings for Research Questions	88
5.2.1 Research question 1	88
5.2.2 Research question 2	92
5.2.3 Research question 3	94

5.2.4 Research question 4.....	96
5.3 Limitations.....	99
5.4 Recommendations for Further Studies	100
5.5 Conclusion	101
REFERENCES	104
APPENDICES	115
A. Knowledge of Speaking Assessment Questionnaire	115
B. Training Needs Analysis for Oral Language Assessment Survey	118
C. Semi-Structured Interview	120
D. Permission Correspondence for Knowledge of Speaking Assessment Questionnaire	121
E. Permission Correspondence for Training Needs Analysis for Oral Language Assessment Survey	122
F. Curriculum Vitae	123

LIST OF TABLES

TABLES

Table 1 Demographic Information of Participants	36
Table 2 Reliability Statistics for KOSAQ.....	38
Table 3 Reliability Statistics for TNAFOLAS.....	39
Table 4 Data Collection Tools and Data Analysis Methods	41
Table 5 Knowledge of Speaking Assessment Constructs.....	45
Table 6 Knowledge of Speaking Assessment Techniques and Tasks	47
Table 7 Knowledge of Speaking Assessment Criteria.....	49
Table 8 Participants Who Claimed to Assess Oral Language Skills	53
Table 9 Frequency of The Participants' Oral Language Assessment	55
Table 10 Participants Who Answered How They Assess Oral Language Skills.....	56
Table 11 Sources of Oral Language Assessment Criteria	57
Table 12 Influence of Factors in Participants' Contexts.....	62
Table 13 Participants' Evaluation of Their Previous Training	73
Table 14 High School Teachers' Training Needs	75

LIST OF ABBREVIATIONS

EFL	English as a foreign language
KOSAQ	Knowledge of Speaking Assessment Questionnaire
TNAFOLAS	Training Needs Analysis for Oral Language Assessment Survey



Chapter 1

Introduction

1.1 Overview

This thesis investigated high school EFL teachers' knowledge, current practices, factors influencing their knowledge and practices and their professional training needs for the assessment of oral language skills. In this chapter, the researcher aimed at shedding light on the significance of instruction and assessment of oral language skills and the challenges confronted in the assessment process with a brief theoretical framework, providing an outline of current situation and problems that has urged the researcher to carry out this study in the problem statement section and elucidating purpose and significance of the study in the latter sections with reference to prominent studies which enlightened and contributed to the aim and scope of the thesis.

1.2 Theoretical Framework

1.2.1 The significance of speaking instruction in language teaching.

Integration of four macro skills in language teaching has been accepted and emphasized by several linguists and researches such as Nation and Newton (2009) who suggested an equal rank for each. Many others, on the other hand, highlighted speaking skill as the most important. Ur (1996), for example, stated that most second language learners give priority to learn to 'speak' the language and the people who know a language are referred as 'speakers' of that language as if speaking encloses all other kinds of knowing a language. Richards and Renandya (2002) also pointed out that, a great percentage of language learners in the world is going after developing their proficiency in speaking. Similarly, Genç (2007) claimed that managing a fluent conversation with others is more important than reading, writing or listening. Khamkhien (2010), looking from a different viewpoint, claimed that speaking is the most important skill since it can most clearly indicate the correctness or errors of a learner's language. As confirmed by the language scholar and researchers above, speaking has been accepted as the prominent skill and many EFL learners are interested in learning how to speak English in order to catch up with the global communication with many different purposes. That's why, English teachers are required and constrained to help learners enhance their speaking skill with various strategies and activities in the classroom for an effective teaching.

1.2.2 The significance of assessment in language teaching. In this section, the researcher meant to discuss why a language teacher should feel the need for assessing what have been learned in a teaching /learning process.

Many tend to assume assessment as the final step of a learning process, and all one can do with it is to evaluate and get ultimate scores; conversely, as Brown (1995) claimed, it is the heart of a systematic approach to language curriculum that encloses, links and contributes to all other components (p. 217). Brown (2000) defined assessment as an integral part of a teaching and learning cycle and named the influence of assessment in language learning process as ‘washback effect’. He stated that in each time we teach something we make an assessment, and assessment serves as feedback which students can benefit from for their competence in language learning. Prodromou (1995), taking the language tests as the assessment type into consideration, also stated that they are valuable devices for finding out learners’ errors, defining their interlanguage, fostering extrinsic motivation, enabling discipline as they are dependent on the classroom authority and they provide valuable contribution in assessing learners' proficiency, progress, and achievement. Shepard (2000) claimed that assessment has a constituent role in teaching and learning a language mentioning the significance of improving the content and character of assessment in order that it serves students to learn and teachers to improve the instruction, instead of designing it for only grading students or certificating the end products of the learning. Assessment Reform Group (2002) also stated that both teachers and learners want to know that learning is processing, and assessment provide them and others, who are interested such as parents, administrator or employers, with the knowledge of the progress. In addition, assessment indicates the performance of teachers, schools and educational systems.

The significance of assessment in language teaching / learning has been well-explained in the books and studies above. Speaking has already got the nod from most authorities, scholars and many researchers as the most outstanding skill in language learning. Therefore, assessment of speaking skill is indispensable in language learning / teaching processes so that learners gain a full competence in language they learn.

1.2.3 The challenges encountered in assessment of oral language skills.

Assessment of speaking skills is known as a complex and challenging process comparing to other language skills (Fulcher, 2003); therefore, many researchers have mentioned the omission of speaking skills assessment in language teaching and they dwelled upon the reasons for this negligence. Brown (2000) stated, much of class-time is spared for testing other skills such as reading and listening, but speaking assessment is neglected in schools; therefore, the time allocated for practicing speaking remains limited. Cohen (1980) also threaded the issue stating that throughout the past decades, developing oral language assessment has come into prominence and a great deal of effort has been made on the field; nevertheless, because of practicability issues, it has been disregarded compared to pencil-and-pen tests. Henning (1983) directed our attention to the stationary problems causing the assessment of oral proficiency impracticable or invalid such as the requisite of a long time for large samples, lack of adequate intervals for each person due to the nature of rating scales, subjectivity, lack of validity indicators through various instruments for different language abilities, inadequacy of methods. Ounis (2017) mentioned similar and some other difficulties stating that there are a lot of internal factors that affect either teachers, like fatigue, or learners, like low voice; and external factors such as noise or other people around. Ounis (2017) also pointed out the existence of two distinct approaches, analytic and holistic, to the assessment of speaking skill as another reason that makes it intriguing.

To sum up, there are many drawbacks confronted in the assessment process of oral language performance such as lack of time and equipment, reliability and validity issues, large classes, discordance in determining the criteria, rater subjectivity etc., and teachers are obliged to take all these factors into consideration in order to make a valid and an effective oral language assessment. Teachers also have to be aware of different approaches, methods and techniques, and decide when to use which, and how to use them appropriately in order to carry out a successful assessment and provide a well-qualified washback for classroom learning of speaking (Güllüoğlu, 2004).

1.3 Statement of the Problem

In Turkey, a new curriculum for English instruction in high schools was put into action in the academic year 2016-2017 with the purpose of including all aspects of communicative competence and integration of four language skills in its scope in

order to meet students' needs to learn English and use it productively and communicatively. In line with the purpose and the scope of the new curriculum, to assess student learning in the curricular context effectively, English teachers have been required to implement practice exams to assess four macro skills; speaking, listening, reading and writing for the past two years. The significance of implementing speaking practice exams was emphasized in the guideline of the new curriculum; however, teachers have not been provided with any kind of training for how to carry out oral language assessment. This generates the most important problem, as is also asserted by Underhill (1987), since assessment of spoken language is a toilsome work and requires experience and training for assessors (p. 101). Another problem about the implementation of oral language assessment is large classes. As Morrow (1991) mentioned, it is arduous to administrate speaking tests with large numbers of students in the limited time allocated. Considering a class with 30 students, it will be impossible if at least 5 minutes is given to each student as it totally takes 150 minutes (p.56). Class size in majority of state high schools are relatively large (30 or over) in Turkey, and this means that, nearly all EFL teachers working in state high schools are fronted with this problem; however, they haven't been trained for the ways to overcome this problem, either.

The challenges of oral language assessment were mentioned in the previous section, in this chapter. Speaking has a momentary and an intricate nature (Ounis, 2017) which tangles the situation for teachers more. Therefore, there is a confusion and an obscurity among teachers which lead most of them to either ignore oral language assessment wholly in their classrooms (Nation & Newton, 2009), or perform only perfunctory activities with the aim of fulfilling the requirements of policy makers. Deriving from the problems about speaking assessment above, the researcher felt the need to scrutinize the current situation for assessment of oral language skills in state high schools from teachers' perspective and reveal their problems and professional training needs for the elimination of the deficiencies in order that they can accomplish an effective speaking assessment.

1.4 Purpose of the Study

The main aim of this thesis is to put the assessment of oral language skills in state high schools in Turkey on the table from teachers' point of view and probe into their knowledge, current practices and challenges confronted in order to ascertain

high school teachers' professional training needs for assessing oral language skills effectively.

In line with the main aim, the goals of this study are to investigate high school teachers' knowledge of constructs, techniques and tasks used in oral language assessment, clarify the influence of high school teachers' contexts on their assessment process, gather data about high school teachers' training background and their professional training needs on oral language assessment. The ultimate aim is to reveal training needs of high school teachers for the assessment of oral language skills in the light of all the data gathered through the questionnaires and interviews.

1.5 Research Questions

In the previous sections, the reasons for choosing this research area and the purposes were clearly explained. In line with the purposes of the study, following research questions were formed:

- 1) What do high school teachers know about oral language assessment?
- 2) What are high school teachers' current practices for oral language assessment?
- 3) What factors in high school teachers' contexts influence their oral language assessment?
- 4) What are high school teachers' professional training needs for assessing their students' oral language performance efficiently?

1.6 Significance of the Study

Teachers have the key role in effective assessment of a language learning process in order to enhance motivation for student learning. (Ur, 1996; Brown, 2000; Fulcher, 2003; Nation & Newton, 2009). Teachers are required to guide students in taking the responsibility of their own learning, tell the students the goals of their work well and negotiate the purpose of learning with them, provide feedback that will foster learning, explain the criteria for the assessment well and assist students so that they can manage to do self-assessment, encourage students for collaboration and guide for giving positive feedback to each other. Teachers, on the other hand, need to minimize the negative effects of tests on motivation for learning (Assessment Reform Group, 2002). As teachers have the key role in assessment of any language skill, as stated above, they also have a leading role in assessment of speaking skills. For that reason, this study is significant as it deals with the teacher factor in speaking assessment. Furthermore, although oral language assessment has been one of the

most highlighted issues in language teaching for the past few years because of the aforementioned reasons, there are very few studies that give voice to high school teachers regarding their current condition, in Turkey, and to the author's best knowledge, none analyzed high school teachers' training needs for assessment of oral language skills.

To sum up, this thesis was aimed to serve as a step to diagnose and define current problem about oral language assessment in high schools in Turkey and shed light to training needs of high school teachers. In other words, it is an attempt to touch upon a sore spot and contribute to the healing of it. Thus, it is significant and a contribution to the literature.

1.7 Definitions

Oral Language Assessment: The process that is operated to evaluate and measure learners' competence and progress in oral language skills.

Formative Assessment: An informal and ongoing assessment type that is usually carried out by teachers in the classroom to obtain information about students' progress in the targeted learning outcomes and to provide feedback for learners.

Summative Assessment: An assessment type that is usually conducted at the end of a period (e.g. a semester) to gain information about the extent to which learning outcomes gained and to grade or certificate learners' success.

Holistic Scoring: The scoring system that the assessor focuses on the assessment of whole work, rather than specific elements.

Analytic Scoring: The scoring system that the assessor divides the whole performance into partitives and assess each separately using a kind of rating scale and obtain a final score out of all.

Validity: The consistency of assessment process and results with the aspects that are meant to be measured in line with the goals of assessment.

Reliability: The consistency of the results over time, in different contexts and with different assessors.

Washback effect: The positive or negative effect of assessment on teaching practices and learning behaviors.

Chapter 2

Literature Review

2.1 Overview

In this chapter, the literature about oral language assessment is reviewed and presented in two main sections. In the first section, relevant concepts were defined and their relation and significance for oral language assessment were reviewed briefly; in the second section, the existing empirical studies which sought for teachers' knowledge, practices, perceptions, and training needs for language assessment in general, and oral language assessment in particular, were carefully reviewed and discussed.

2.2 Fundamental Aspects of Oral Language Assessment

2.2.1 Definitions and significance of oral language assessment. There are various definitions of assessment which provide an insight for the significance of it in classroom contexts. Fulcher and Davidson (2007) refer to it as a part of modern life which is employed in the classrooms to watch the progress in educational systems or quality of schools. Classroom assessment was most clearly and normatively defined by Angelo (1995) as "Assessment is an ongoing process aimed at understanding and improving student learning". In the same vein, Coombe, Folse and Hubley (2007) defined assessment as an umbrella term which covers all kinds of measuring processes in the classroom to evaluate students' performances. In the light of these definitions, assessment goes beyond simply conducting a test to give a grade and comes into prominence with its function to foster classroom learning. Louma (2004) highlighted the significance of oral language assessment claiming that speaking is one of the major abilities for assessment. Önal (2010) asserted that unless oral language assessment is included in overall assessment processes in language teaching, it is not wise to expect students to perceive it as a significant skill and feel the need to develop it. The more students are subjected to oral language assessment tasks and activities, the more proficient they are likely to become in terms of fluency and accuracy in target language. The statements above lay bare that, assessment of oral language performance within overall language assessment processes is significant.

2.2.2 Considerations in assessment of oral language performance. Once

the need for integrating oral language assessment in the overall assessment process has been well accepted by most scholars, researchers, stakeholder and teachers, another issue comes into prominence about it; how oral language assessment can be carried out efficiently in order to foster students' oral language skills. A number of issues have been focused in the literature regarding the effectiveness of language assessment on enhancing learner competence in related skills. Weir (1990) claimed that when teachers assess oral language performance of students, they should be conscious about several issues as such it should be theoretically tenable, consistent with the goals of teaching oral language, motivating for students, promote student interaction, objective and as much authentic as possible. More specifically, Bachman and Palmer (1996) claimed that the most important consideration of language assessment is its usefulness. The authors defined usefulness in terms of six qualities; authenticity, interactiveness, impact, validity, reliability and practicability. Fulcher and Davidson point out the significance of deciding the skills to assess and the way that those skills are assessed, and learning outcomes gained at the end of the assessment process by saying "The importance of *what* we test, *how* we test, and the *impact* that the use of tests has on individuals and societies cannot be overstated" (2007: pp xix). It is inferential with the reviewed literature above that, there are many aspects to consider while planning, designing and processing oral language assessment to gain utmost learning outcomes out of it. These aspects of oral language assessment are discussed in detail in the following sections.

2.2.3 Defining oral language assessment constructs. Now that the agreement on the significance of assessing oral language performance have been reached, another question arises; *what to assess*. This question is related to oral language abilities to be assessed which are also called "constructs" in the literature (Bachman; 1990; Fulcher; 2003 & Louma, 2004; Seong, 2014). As is asserted by many scholars, it is important to define what oral language constructs to assess precisely prior to an assessment process. Fulcher (2003) emphasized the requisite of determining constructs in accordance with the purposes of assessment and design it to serve to those purposes. According to Luoma (2004), constructs are specified and clearly defined in a good test so that they manifest "what it means to be able to speak a language" and shed light to the designation of criteria, rating scales and tasks. Bachman (1990) also claimed that language abilities need to be well-defined whether they are derived from a general language ability theory or a language teaching

syllabus. Bachman drew our attention to a problem which he called “fundamental dilemma”. That is the difficulty in defining language abilities to assess due to the complex nature of language assessment as language being both object and tool of assessment. He, therefore, underlined that language abilities need to be determined in a way that language performance we assess should indicate the features of authentic use of language in real situations (p. 9). Fulcher (2003), similarly asserted that defining constructs is one of the primary problems in speaking assessment.

Even though the specification of language abilities to assess is a problematic issue, there have been a number of oral language assessment constructs agreed in the literature as they were most frequently accounted by the majority of scholars and researchers to inform the assessment criteria, techniques and tasks to be designed accordingly. Harris (1969) defined speaking as a complex skill which includes several abilities such as pronunciation, grammar, vocabulary, fluency, discourse and comprehension ability. Madsen (1983) addressed some subsidiary skills such as listening comprehension or discernment that need to be considered while dealing with the contents of speaking assessment such as pronunciation, fluency, grammar, vocabulary and appropriate expressions. Fulcher (2003) mentioned speaking constructs as accuracy, fluency, pronunciation, intonation, interaction and pragmatic strategies. Burgess and Head (2005) also suggested to assess pronunciation, accuracy and vocabulary reminding the focus must be on communication skills rather than accuracy of grammar and lexis as it is speaking assessment not a grammar or a vocabular test. Brown, Iwashita and McNamara (2005) conducted two empirical studies coordinately with English-for-academic-purposes (EAP) raters first of which investigated their rating orientations. The results from verbal protocols revealed that the raters agreed on four categories when assessing oral language performance: linguistic resources (grammar and vocabulary), phonology (pronunciation, intonation, rhythm and stress), fluency (hesitation and pauses, false starts, repetition, repair and fillers, and speech rate), and content (task fulfillment, the amount, quality and organization of ideas and the framing of the response). Bøhn (2015) conducted a study with 24 EFL teachers in Norway and elicited similar findings as Brown et al. (2005). The teachers’ agreed on the constructs as vocabulary, grammar, phonology, content and communication despite having divergences in the weight they gave to the constructs. The constructs which are commonly approved above are discussed in detail in the following sub-sections.

2.2.3.1 Accuracy and fluency. Accuracy in oral language is defined as correct usage of grammatical structures and lexical units in spoken language. When considered as an oral language assessment construct, it is usually referred to as spoken grammar which is derived from corpus and involves conceptualization of grammar units (Leech, 2000; Luoma, 2004). Yuan and Ellis (2003) also mention the appropriate use of vocabulary and pronunciation with grammar when speaking off accuracy of speech. Fluency, on the other hand, is defined as speaking in a natural flow without redundant hesitations or pauses (Skehan, 1996; Brown et al., 2005). Although these two oral language assessment constructs have long been controversial with regard to native / non-native norms or intelligibility issues, they are widely admitted as essential aspects of oral language competence and usually placed in rating scales (Brown, 2000; Fulcher, 2003). Involving accuracy and fluency in rating scales is also empirically proved to be essential regarding both raters' perspectives and its impact on learners' development in accuracy of article use, tense-marking, third-person-singular verbs, plural nouns, prepositions and in fluency in terms of speech rate, number of unfilled pauses and total pause time in Brown et al. (2005). Iwashita, Brown, McNamara and O'Hagan (2008) revealed the impact of assessing a number of constructs on test-takers' language development, indicating the greatest enhancement in their fluency with vocabulary. Nation and Newton (2009) also remarked involving fluency in assessment stating that it might have an effect on the development of pronunciation.

2.2.3.2 Vocabulary. Vocabulary refers to the words of a language, and word is defined as a unit which consists of one or more sounds that constitutes the benchmark of the meaning of a speech or a text (Pearson, Hiebert & Kamil, 2007). Since vocabulary is a significant part of a speech that carries the meaning to a large extent, it is indispensable to refer to vocabulary as a construct in oral language assessment, as stated by Iwashita et al. (2008). The study also indicated a remarkable enhancement in learners' vocabulary throughout the assessment process. In Brown et al. (2005), following the assessment process, the learners indicated an improvement in their vocabulary in terms of variety, amount and appropriateness. To sum up, in oral language assessment, it is essential to include vocabulary as an assessment construct.

2.2.3.3 Pronunciation. Pronunciation was defined by Ur (1996) as a concept which involves the sounds of a language, stress, intonation and rhythm. The author

emphasized the importance of knowing variety of sounds, stress and intonation patterns as they can have an influence on each other in the natural flow of the speech (p. 47). In the same vein, Derwing and Munro (2005) stated that a competent pronunciation is essential for intelligibility of speech which is necessary for having a normal communication. Therefore, the assessment of pronunciation is also emphasized and suggested (Louma, 2004; Nation & Newton, 2009). Nevertheless, both teaching and assessment of pronunciation is also a hot spot. Louma (2004) describes the assessment of pronunciation as a ‘thorny issue’ as assessors usually tend to base it upon native standards which he states as erroneous. Another researcher who mentioned current pronunciation assessment as erroneous is Rubin (2012). He similarly, stated the expectation of raters for native speaker standards as a problem and suggested the criteria to be identified carefully for a valid pronunciation assessment. Nation and Newton (2009) specifically points to “intelligible pronunciation” emphasizing the efficacy of “meaning-focused input” and “meaning-focused output” for a successful communication. Yoshida (2016) describes assessors’ expectation for “sounding like native speakers” as a problematic issue as it is difficult to say exactly “what a native speaker sounds like” touching upon varieties of English. She states that it must be acceptable to have the accent of your country if your speech is intelligible enough to be understood by both native and non-native speakers of English, and assessors should be conscious about it. To sum up, pronunciation is a major construct in oral language assessment which is strongly suggested to involve in rating scales; however, as is clear with the statements above, majority of scholars and researchers of this topic suggest considering intelligibility of the speech rather than expecting native-like sounding in the assessment of it.

2.2.3.4 Communicative competence. Despite the variety in definitions and its scope in the literature, communicative competence takes a critical place among oral language assessment constructs. Throughout the literature a large number of models have been proposed such as Canale and Swain’s (1980) and Fulcher’s (2003) which focused on eliminating communication breakdowns, Tarone’s (1981) which centered upon negotiation of meaning among interlocutors from an interactional perspective, Levelt’s (1989) which examined the conceptualization of the intended message, and Dörnyei’s (1995) and Nakatani’s (2006) with a boarder emphasis on interaction strategies (Seong, 2014). Bachman and Palmer’s (1996) model divided communicative language competence into two as language competence and strategic

competence. The former involves grammatical and lexical knowledge, and the latter is more related to appropriateness of the language in context. As is seen in the statements above, communicative competence has been an important point in question for long and its place in the assessment process have clearly been emphasized by many scholars. Brown (2000) called the competence in pragmatic goals in interactive conversations as “the benchmark” of language acquisition. Louma (2004) gave emphasis on consideration of interaction in speaking tests. (Chuang, 2009) mentioned the need for assessing students’ attainment in using target language to give their message clearly in various contexts. There are a number of empirical studies which sought for the impact of assessing interactive strategies with positive results (Huang,2013; Barkaoui, Brooks, Swain & Lapkin, 2013). To conclude, reviewed literature suggests the involvement of interactive and pragmatic abilities as components of communicative competence in the assessment of oral language performance.

As is seen in the reviewed literature above, agreed oral language assessment constructs can be elicited as accuracy, fluency, vocabulary, pronunciation and communication strategies. The next question that comes into minds following the determination of oral language constructs is *how to assess* oral language abilities. The scope of the answer is quite broad as it includes a number of aspects such as different assessment techniques used in oral language assessment, the criteria, rating scales and tasks used in assessment processes, issues related to rating the scores such as how to make a valid, a reliable assessment etc. As cited in Bachman (1990), a large number of studies have indicated that assessment methods have an influence on students’ performance (e.g. Bachman & Palmer, 1982; Shohamy, 1984); thus, the question of “how to assess” becomes more of an issue. The related literature is reviewed and presented in the following sub-sections.

2.2.4 Determining criteria and designing rating scales. Once constructs have been designed, they are generally functionalized in the rating scales (Luoma, 2004; Fulcher, 2009). Rating scales are formed according to the determined criteria in relation with the defined constructs. Determining the criteria is essential as they provide the opportunity to assess complex abilities validly and reliably and make the assessment clear to students, parents, and other related people (Herman, Aschbacher, & Winterset, 1992). A number of issues and considerations have been reminded for the designing of rating scales such as prioritizing communicative skills (Weir,1990;

Burgess & Head, 2005), considering learner differences and including various grading categories accordingly (Underhill, 1987), using an appropriate number of levels (4 – 6 were suggested) in order to obtain reliable scores (Louma, 2004) and placing an appropriate number of criteria and describing them well in order to assess the level of students' knowledge or progress precisely (Underhill, 1987; Louma, 2004). It is also recommended to use trial and error method to form the best-fitting rating scale to the context and level of students and (Underhill, 1987, Fulcher, 2003).

2.2.5 Assessment types. As Bailey (2005) claimed, assessment format is important in terms of washback in the classroom which has a big influence on classroom teaching /learning. In the language assessment literature, two prominent types; formative assessment and summative assessment have frequently been discussed (Black, Harrison, Marshall & William, 2003; Hughes, 2003; Fulcher & Davidson, 2007). Additionally, portfolio-assessment, peer-assessment and self-assessment have often been encountered as alternative language assessment (Genesee & Upshur, 1996; McNamara, 1996; Brown & Hudson, 1998). There have been a great amount of debates and many empirical studies on usefulness, pros and cons of these methods and techniques some of which have been reviewed in this review.

2.2.5.1 Formative and summative assessment. In the language assessment literature, formative and summative assessment types have principally been defined together, and they are usually compared in various terms (Bloom, Hastings, Madaus & Baldwin, 1971; William, 2000; Black et al., 2003; Hughes, 2003; Fulcher & Davidson, 2007; Taras, 2008). Formative assessment, on the one hand, has commonly been defined as the informal type of assessment that provide teachers with the information of students' progress on the predetermined learning goals and enable them to revise their future teaching plans accordingly and provide feedback for the students, as well (William & Black, 1996; Hughes, 2003; Black et al. 2003). Summative assessment, on the other, has most frequently been defined as a formal type of assessment which is generally conducted at the end of a semester, a year, or a course to find out to what extent learning outcomes gained by groups or individual with the aim of grading or certificating the progress or evaluating the efficacy of a curriculum (Bloom et al., 1971; William & Black, 1996; Hughes; 2003). Summative assessment implementations have been unfavored by a number of researchers. William and Black (1996), for example, defined summative assessment as “inflexible” and “mechanical” and need to be mitigated. They, still, suggested that

there might be a “common ground” between summative and formative assessment to utilize both as a continuum of one another. In the same vein, Black et al. (2003), prioritized formative assessment in their book mentioning it as “assessment for learning” and emphasizing its function to provide feedback about students’ strengths and weaknesses. Black et al. (2003), additionally, conducted a study with 36 teachers both to investigate the effectiveness of formative assessment practices on student learning and to reveal the consequences of teachers’ efforts to prepare students to an external summative language assessment. The findings indicated positive student-gains as enhanced self-esteem, motivation and learning. It was implied in the study that, students have negative attitudes towards summative assessment; therefore, this type of assessments need to be revised and changed for students to get involved in it actively and feel that they can benefit from it to improve their learning. Tang (2016) which was an experimental study with 115 Chinese students aiming to investigate the influence of formative assessment on students’ anxiety of speaking English and gathered data through an anxiety scale, pre-test, post-test and interviews. The findings revealed that formative assessment increased students’ speaking anxiety in oral language assessment quite contrary to the findings of Black et al. (2003). Some others suggest integration of the two to get the best outcomes both in terms of washback in the classroom and positive impact for future classroom activities, and less anxiety, higher self-esteem and higher scores in summative tests (Hughes, 2003; Fulcher & Davidson, 2003; Wiliam, Lee, Harrison & Black, 2004). Fulcher and Davidson (2007) further mentioned that classroom assessment is usually formative in nature as teachers interact students frequently and interested in development of their skills; thus, they both revise their future plans according to the scores and provide feedback for students. It can be inferred from the reviewed literature that teachers need to be aware of pros and cons of both assessment types and choose the best one according to their context, and if necessary, use them both for better gains.

2.2.5.2 Alternative assessment techniques. Portfolio assessment, peer-assessment and self-assessment, which are also termed as alternative assessments (Wolf, 1989; Ur, 1991; Arter & Spandel, 1992; Herman et al., 1992; Brown & Hudson, 1998), has currently gained more importance (Belgrad, Burke & Fogarty, 2008; Srikaew, Tangdhanakanond, & Kanjanawasee, 2015).

2.2.5.2.1 Portfolio-assessment. Carr and Harris (2001) define portfolios as collections of students’ works and purposeful indicators of their efforts, progress or

gains. Belgrad et al. (2008) mentioned that portfolios serve as a display of students' academic abilities and progress by providing students, teachers and parents with the samples of their work and asserted teachers prefer them as evidence for progress despite being time-consuming and difficult to organize (pp.2). An empirical study which might support the assertion in Belgrad et al. (2008) about teachers' attitudes is Özdemir-Çağatay (2012). The researcher carried out her study with 77 university students and 5 instructors to find out their attitudes towards portfolios and revealed that although some challenges were uttered, both students and teachers showed positive attitudes towards portfolio assessment and suggested it as an alternative assessment method in Turkish context. Srikaew et al. (2015) observed the procedures of a portfolio assessment model, which was used to assess speaking performances of grade 6 students and reported that the assessment comprised of seven steps; planning, preparation, evidence collecting, progress monitoring, improvement of performance, reflection and presenting the works, and activities involved four tasks including oral presentation, interview, storytelling and picture description. Students' performances were rated 1-4 on an analytic rating scale which contains fluency, vocabulary, syntax, pronunciation, cohesion and ideational function as assessment criteria. The researchers claimed that the use of a 4-point rating scale was found reliable and easy. In short, portfolio assessment is an alternative in oral language assessment and recommended as beneficial in several aspects mentioned above.

2.2.5.2.2 Peer-assessment and Self-assessment. There are a number of studies which sought for impact or effectiveness of self- and peer-assessment on language learning outcomes and attitudes of teachers or students towards them with varying results (e.g. Black, 2003; Cheng & Warren, 2005; Le'ger, 2009; Mok, 2011; Grez, Valcke & Roozen , 2012). Both Cheng and Warren (2005) and Grez et al.(2012), for example, found significant differences between teachers' assessment and self- and peer assessments and implied these alternative assessments need improvement in terms of reliability; whereas, they diverged in terms of learner perception. Cheng and Warren (2005) found negative attitudes towards peer assessment as students were uneasy and felt insecure about grading their peers' performance, when Grez et al. (2012) reported positive attitudes from students. In Mok's (2011) case study with secondary school students peer- and self-assessment process was revealed beneficial; however, the students indicated negative attitudes towards those assessments. Black's (2003) and Le'ger's (2009) studies found self-assessment beneficial both in

terms of development in language proficiency (e.g. in fluency and vocabulary) and students' self-esteem and autonomy.

In the studies and books related to alternative assessment methods above, some challenges were mentioned such as reliability / validity issues, student anxiety and negative attitudes, difficulty in implementing; nevertheless, portfolio, peer- or self-assessment have strongly been suggested for a number of benefits with the implication for careful design and implementation. For example, self-assessment was referred to improve learner autonomy and motivation (Blanche, 1988; Brown & Hudson, 1998; Noels, Pelletier, Clement, & Vallerand, 2000), be easy to design and implement (Brown & Hudson, 1998) and develop learner meta-cognitive skills (Black, 2003); portfolio assessment was claimed to provide all stakeholders with the evidence of learning outcomes (Belgrad et al., 2008; Nation & Newton, 2009), foster both process and product learning (Herman et al., 1992; Brown & Hudson, 1998; Louma, 2004), encourage participation (Nation & Newton, 2009) and enable revision and practice (Brown & Hudson, 1998); and peer assessment was admitted to provide richer amount of feedback, enable learning from other students (Topping, 1998) and generate a natural motivation for being praised by peers (Black, 2003). Brown and Hudson (1998), taking all alternative assessment into consideration, claimed that students develop their critical thinking and problem-solving skills, they are provided with plenty of feedback to learn about their strengths and weaknesses and they are confronted with more authentic use of language through alternative assessment methods. It is inferable from the reviewed literature that, choosing appropriate methods or techniques depends to a great extent on teachers' own competence, context and conditions, and it is advisable to integrate them as much as possible. In this sense, it is important to provide teachers with training both for the terminology and the practical ways to use these assessment types (Brown & Hudson, 1998; Topping, 1998; Hughes, 2003; Black, 2003; Taras, 2008).

2.2.6 Oral language assessment tasks. Nunan (1989) defined the term task as “a piece of classroom work which involves learners in comprehending, producing or interacting in the target language while their attention is principally focused on meaning rather than form”. In this definition, language tasks are meant to be interactive, input-output oriented and meaning -focused. Bachman and Palmer (1996) defined tasks as activities that people perform in particular situations to achieve a goal. Willis's (1996) definition is closer to assessment function as he defined it as “a

goal-oriented activity in which learners use language to achieve a real outcome”. A task is commonly described as an activity which people take up to gain an objective (Ellis, 2003; Van de Branden, 2006) and which involves meaningful use of language. As for oral language assessment tasks, in line with the definitions above, they can simply be defined as classroom activities which are utilized to assess students’ performance in oral language skills and require them to use the language in order to find out their progress and to what extent the goals are achieved. One can infer that tasks need to involve authentic use of language recalling “fundamental dilemma” term in (Bachman, 1990) which means that language is both tool and the goal of assessment process. In the same vein, Van den Branden (2006) said “people learn a language in order to use it, but also by using it”. These two statements have the implication of the need to use authentic tasks during classroom practices of both teaching and assessment of language.

2.2.6.1 Oral language assessment task types. In the related literature there are a number of tasks which are commonly associated with oral language assessment such as oral presentation tasks in which students are asked to make a speech a few minutes on a topic they have prepared in advance (Mckay, 2006; Ounis, 2017), discussion tasks by means of which students are required to express their ideas on a certain topic as a part of a question-answer activity or as a learner-learner joint discussions (debates) in which students are paired or grouped to debate for or against a topic (Coombe et al., 2007; Thompson, 2009; Pineda, 2014), learner-learner joint decision-making tasks in which students are involved in a group or pair negotiation activity and asked to think over a situation, discuss alternatives, detect differences or make suggestions (Wang, 2014), oral interviews which are also defined as open-ended question-answer tasks in which students are free to answer as they wish (Madsen, 1983; Underhill, 1987; Weir, 1990; Genesee & Uspur, 1996; Brown & Hudson, 1998; Mckay, 2006, Pineda, 2014; Ounis, 2017), role-playing tasks in which students are provided with the information about the setting, character and the topic and asked to act in compliance with it (Underhill, 1987; Brown & Huddson, 1998; Luoma; 2004; Nation & Newton, 2009; Pineda, 2014), picture-talk tasks which require students to describe a picture or compare two or more pictures (Güllüoğlu, 2004; Pineda, 2014; Ounis, 2017), storytelling tasks in which students are asked to make-up a story out of some pictures, words etc. (Önal; 2010; Ounis, 2017), giving instruction or direction tasks in which students are asked to give instructions,

directions or describe something, like, how to cook a dish or a way to go somewhere (Güllüoğlu, 2004), information-gap tasks which require students to complete an information gap in a dialogue, sentence or a part of a listened conversation orally (Brown & Hudson, 1998; Pineda, 2014; Wang, 2014), summarizing / paraphrasing or re-telling tasks through which students listen or read a dialogue, situation, story etc. and required to summarize, paraphrase or re-tell it (Güllüoğlu, 2004; Pineda, 2014; Tang, 2016).

2.2.6.2 Determining oral language assessment tasks. There are variety of tasks that are used to assess students' oral language skills as mentioned above. The thing is, it is not easy to decide which to use for a teacher who have not been trained on assessment strategies. The effectiveness of task types has long been discussed in terms of effectiveness, validity or practicability. Brown and Hudson (1998), for instance, claimed that debates and role-playing tasks measure students' authentic communication more validly as they involve more interaction, despite being time-consuming to implement. The authors also mentioned that information-gap tasks are advantageous as they are easy to design. A number of empirical studies have examined task types from different point of views. Henning (1983), for example, compared information-gap, imitation and oral interviews in terms of reliability and validity during oral language assessment of 143 adult learners and revealed that imitation tasks got the highest validity in pronunciation scores. Dandonoli and Henning (1990) sought for validity and reliability of oral interviews and got positive results. Foster and Skehan (1996) compared a number of tasks in terms of allocated time and workload and found out that decision-making and re-telling tasks were more demanding and time-consuming. In Sinwongsawat's (2012) study, non-scripted role-plays and face-to-face interviews were compared in terms of effectiveness on students' conversational skills and non-scripted role-plays were found more effective. Moere (2012) found repetition and sentence-building tasks highly reliable in distinguishing the test-takers' proficiency level despite not being favored for not being communicative. Wang (2014) revealed that oral presentations, learner-learner joint discussions and information gaps to encourage students to express their ideas freely, and thus effective in developing their fluency. Tang (2016) suggested debates role-playing and story-retelling tasks emphasizing their communicative function. Ounis (2017) claimed to have noticed that most of the students were quite motivated to speak in the interviews. In some studies, models were meant to be developed with

an intend to provide a multi-dimensional assessment combining some tasks (e.g. Gràcia, Vega, & Galván-Bovaira, 2015). There are some studies which revealed the significance of defining appropriate tasks in oral language assessment process. Huei-Chun (2007) investigated the influence of task type on students' performance in speaking tests in Taiwan. 30 college students were asked to fill in a questionnaire after they took the speaking tests comprised of three types of tasks; question-answer, picture description, and presentation. The results indicated task type has a remarkable influence on students' performances. For example, the students achieved higher scores in fluency in question-answer tasks. Interestingly, they reported they felt more anxious during answering the questions. This being the case, designing appropriate tasks has utmost importance for the assessment of oral language skills. Fulcher (2003) mentioned that it is a significant problem to design tasks in an appropriate way to elicit enough speech to be rated. Pica et al.(1993) asserted that tasks need to be goal-oriented and promote students to work to gain the goals. According to Herman et al. (1992), designed or selected tasks need to be authentic, match with the learning outcomes and have the capacity to demonstrate students' progress in order to carry out a good assessment. Louma (2004) claimed that tasks need to be purposeful, interesting, motivating, interactive and provide positive feedback. Fulcher and Davidson (2007) suggested task definition process need to be cyclical not linear in order to re-evaluate the effectiveness and appropriateness of tasks. Brown and Hudson (1998) further emphasized using multiple-tasks, using clear criteria, triangulation of task-designing process and training teachers. In the same vein, Coombe et al. (2007) stressed that a teacher should compile various assessment types, techniques and activities and make use of a mixture of them in order to get a clearer picture of students' competence and performance in the target language. As is inferred from the above literature, teachers have a lot to do in designing the tasks to assess oral skills in their classrooms and need to have a sufficient training on it, as well.

2.2.7 Analytic and holistic scoring systems. Within oral language performance assessment, two approaches to scoring were discussed by a number of scholars (Hughes, 2003; Fulcher, 2003; Luoma, 2004); analytic and holistic. Holistic assessment involves evaluating overall impression of learners' competence in oral language skills and giving a single grade; hence, it does not comprise specified rating scales including categories for evaluating skills and sub-skills separately (Mertler,

2001; Ounis, 2017). As for analytic assessment, it involves dividing oral language skills in to certain categories such as accuracy, pronunciation, vocabulary, content and evaluating each separately and accounting the final score out of all (Moskal, 2000; Mertler, 2001 & Thornbury, 2005).

2.2.7.1 Pros and cons of analytic and holistic scoring. There are a number of studies which sought for effectiveness of analytic and holistic scoring students' oral language performances most of which compared the two in terms of positive results on learning / teaching processes (Iwashita & Grove, 2003; Chuang, 2009 & Tuan, 2012) and got varying findings. Önem (2015), for example, carried out a study with 24 instructors and 10 students in Turkey using an attitude questionnaire and analytic and holistic speaking tests. She found out that there was no statistically significant difference in the scores of analytic and holistic tests; however, the instructors had more positive attitudes towards holistic assessment. Ounis (2017) reported more positive results from holistic scoring tools in her mixed-method study that she carried out with the purpose of finding out the most suitable way to assess speaking comparing the holistic and the analytic ratings, in Tunisia. Recorded performances of university students were rated via holistic and analytical assessment techniques by five teachers from different universities. The results revealed that analytical and holistic assessment generated quite different scores, and holistic grading was found more useful, reliable and consistent. Knoch (2009), on the other hand, found out that reliability of analytic scores was higher, and raters could provide more detailed feedback as they could distinguish among the categories better with analytic scoring. In the same vein, Huei-Chun (2007) utilized both rating systems in his study about oral language assessment task types and concluded in that holistic measure was not sensitive enough to reveal the results that were detected by analytic measure. As was researched for long, it was revealed that both have strengths and weaknesses (Hughes, 2003; Brown, 2005; Luoma, 2004). Holistic assessment is advantageous as it is quick and feasible, especially with large classes, and practical for the teachers who have limited time for oral language assessment because it does not require to be troubled with keeping all the criteria in mind when listening to each student's performance (Thornbury, 2005). It is also claimed to be more advantageous when an overlap among the criteria in the assessment of various factors (Moskal, 2000). As for the disadvantages, Luoma (2004) asserted that it is not illustrative enough to define strengths and weaknesses in learners' language performance. Analytic

scoring, on the other hand, can provide students with necessary feedback for their strengths and weaknesses in specific skills as they are explicitly placed in the rubrics (McNamara, 1996; Luoma, 2004) and inform them about their progress in time (Moskal, 2000; Tuan, 2012). It is also easier to use an analytic rating scale for inexperienced teachers as standardized criteria is provided with the scale ((Weir, 2005). Additionally, analytic rating scales are advantageous as they help both teachers and learners stay focused to targeted aspects to be enhanced (Arter & McTighe, 2001). Regarding the disadvantages of analytic assessment, it is not always easy to gain intra- and inter-rater reliability (Tuan, 2012) since raters tend to overemphasize grammar accuracy compared to other skills in an analytic scale (McNamara, 1996). In addition, the score of one skill might influence the rater while scoring others in the same rubric (Hughes, 2003) which is called “Halo Effect” (Fulcher, 2009), and it is time-consuming as it requires raters to deal with the interpretation of all the aspects exist in the scale (Thornbury, 2005).

2.2.7.2 Determining which scoring system to use. When it comes to the answer of the question which is better to use, Hughes (2003) claimed that it depends on the objectives, context and assessors; and he recommended to use both interchangeably stating multiple scoring is the most reliable approach. Coombe et al. (2007) also suggested making use of a mixture of all assessment types which they called “multiple measures assessment” to gain most accurate outcomes about students’ progress in language proficiency. It can be inferred that, both approaches have advantages and disadvantages when compared to one another and deciding which is better highly depends on various aspects such as the nature of ability to be assessed, the context, the assessor and the audience. If possible, it is best to make use of both to eliminate the disadvantages to some extent. Many of the studies above have implications for training teachers on these approaches in order to be able to decide which is appropriate for their own context or how they can use them interchangeably

2.2.8 Validity and reliability issues. Bachman and Palmer (1996) described usefulness of an assessment as being authentic, interactive, practical, influential on learning, valid and reliable. Brown (2000) also defined a good oral test as practical, reliable and valid. As seen in the definitions above, validity and reliability are the essentials of an efficient assessment.

2.2.8.1 Validity. Validity of an assessment is defined as the extent to which targeted or claimed aspects have been measured accurately throughout the assessment process (Henning, 1987; Weir, 1990; Hughes, 2003; Brown, 2005; Nation & Newton, 2009). Three types of validity have usually been discussed in the literature; content validity, construct validity and face validity (Brown, 2005; Fulcher & Davidson, 2007; Nation, 2009). Content validity is related to the degree that the content of assessment tool is consistent with the content of targeted skill (Nation & Newton, 2009). In other words, assessment tasks or materials need to be “representative samples” of the real content that is to be measured (Brown, 2005). To achieve that, the skills and content to be assessed must be specified well (Hughes, 2003). O’Sullivan, Weir and Saville (2002) described construct validity as the consistency of the tasks or activities used in the assessment process with the constructs that intended to be measured. Bachman (1990) suggested to define language abilities to be measured precisely in order to provide construct validity. However, it is quite complex for a teacher to be sure about it (Bachman, 1990; Hughes, 2003). Face validity is public opinion about the assessment. It is decided by the extent to which the assessment looks right or fair to other interlocutors such as counterparts, stakeholders or students as the addressees of the assessment (Underhill, 1987; Nation & Newton, 2009). Hughes (2003) suggested to define constructs clearly and write them explicitly, relate scoring directly to the constructs that are meant to be assessed and be sure about the reliability of the assessment reminding that an assessment cannot be valid unless it is reliable.

2.2.8.2. Reliability. If the assessment cannot be valid without being reliable; thus, reliability has a greater importance. Reliability is defined as the consistency of the results over time (Brown, 2005; Coombe et al., 2007). Brindley (2000) claimed that the classroom teachers need to be competent in choosing and designing the appropriate method, know how to administrate, grade and interpret the results accurately, be confident about taking the initiative to decide re-designing her /his teaching and developing the curriculum according to results, be knowledgeable about how to design, administrate and evaluate rating procedures valid, be skillful in discussing the results with students and all other stakeholders and be aware of all ethical considerations in an assessment process. Hughes (2003) suggested training scorers mentioning it might be beneficial especially in subjective assessments like

oral interviews. Fulcher (2003) further asserted that the reliability of oral assessment depends on the role of teachers or raters.

The inferred implication of the reviewed literature above is that teachers need an extensive training in order to cope with all the considerations related to the assessment in order to make a valid and a reliable oral language assessment.

2.2.9 Washback effect. So far, the literature was reviewed to discuss *what* and *how* to assess. The third important issue was the *impact* of assessment on teaching and learning (Fulcher & Davidson, 2007). The positive or negative impact of assessment on classroom teaching /learning is termed washback which is also called backwash (Hughes, 2003). Nation and Newton (2009) emphasized the influence of assessment on classroom learning and stated that very little time is allocated to speaking activities in schools since they do not assess speaking. Black and William (1998) stated that students need to be informed and guided about their strengths and weaknesses in order that learning can be fostered. Hughes (2003) also pointed out the negative impact which occurs when skills are well taught but poorly assessed. He stated that a well-designed assessment might have a positive impact on even poor teaching condition. Teaching and assessment are mentioned to be cyclical rather than linear by many (e.g. Hughes, 2003; Fulcher & Davidson, 2007). A number of empirical studies have been conducted on washback affect in the classroom learning (e.g. Alderson & Wall, 1993; Shohamy, Donitsa-Schmidt & Ferman, 1996; Wall, 1996; Burrows, 2004; Cheng, 2004; Muñoz & Álvarez, 2010) which had various results; therefore, washback effect is still unpredictable (Fulcher & Davidson, 2007). Muñoz and Álvarez (2010), for example, conducted a mix-method experimental study with 14 EFL teachers and 110 university students with the primary aim of searching for washback effect of language assessment on a number of teaching and learning areas. Experimental group teachers received training on goal setting, guiding students in self-assessment, use of rubrics and feedback strategies for creating an effective washback in the classroom. The results indicated that experimental group teachers were more conscious about content validity as they managed tasks more successfully to achieve the goals and better in expanding students' performance providing feedback on communicative effectiveness, grammar, vocabulary, and task completion, and they informed students about the objectives more explicitly. As a result, experimental group students were more target-driven and motivated to participate, they had a better understanding of

assessment criteria, and ultimately, achieved a better success particularly in communicative effectiveness, grammar and pronunciation. This being the case, the authors suggested teachers being provided with constant guidance and support to foster positive washback in classroom. Wall (1996) focused on the factors that hinder positive washback of assessment process on classroom learning. She interestingly found out that, most of the drawbacks such as teachers' erroneous conception of exams and resistance to change, the discrepancy of teachers with test-designers and lack of training were all teacher-related, and she emphasized the significance of teacher training for effective washback in the classroom which might foster learning. Brown and Hudson (1998) also emphasized the value of feedback in their article and concluded in stressing teachers' need for help in deciding among various sources. Likewise, many authors suggested developing examination systems, using authentic tasks, providing a clear explanation of scoring and improving teachers' and students' consciousness for assessment goals, criteria and scoring (Shohamy, 1992; Messick, 1996; Bailey, 1999; Hughes, 2003). It is clear that, teachers should also be aware of washback strategies in order that it is effective on classroom learning.

Up to now, the terms and concepts that have a strong relation with oral language assessment have been reviewed and definitions and significance in oral language assessment processes have been discussed. In almost all aspects reviewed above, the specialists and researchers emphasized the role of teachers for the assessment process to be useful and effective. Muñoz and Álvarez (2010) specifically identified some principles for the assessment of oral language which they called "Oral Assessment System" and made suggestions to classroom teachers for the efficient assessment of oral language. According to the authors, the criteria need to be clear to both teachers and students, variety of authentic tasks should be utilized to address various learning styles in the classroom and the tasks should comply with curriculum objectives and teaching activities, variety grouping systems need to be ensured to foster adequate teacher-student and student-student interaction, different aspects of oral language need to be assessed and grammar assessment should only serve as a component of communicative competence, assessment should be carried out in an ongoing system and alternative assessment techniques need to be adopted such as self- assessment and detailed and clear feedback need to be provided in order to create a positive washback to serve future learning outcomes. The principles elicited by Muñoz and Álvarez (2010) above provides an outline of the reviewed

statements of linguists and researchers related to oral language assessment. As clearly seen, language scholars (e.g. Ur, 1996; Brown & Hudson, 1998; Brown, 2000; Fulcher, 2003; Brown et al., 2005; Fulcher & Davidson, 2007; Nation & Newton, 2009) point to teacher factor in practice of all mentioned aspects of oral language assessment and researchers notably emphasize teachers' key role in almost all mentioned studies above (e.g. Shohamy, 1992; Messick, 1996; Muñoz & Álvarez, 2010; Seong, 2014; Ounis, 2017). This being the case, it is evident that, teachers need to be competent in all aspects of oral language assessment to ensure it to be useful for classroom learning. In the following section, empirical studies which investigated teachers' knowledge, perceptions, practices, challenges and needs of language assessment in general, and oral language assessment in particular have been reviewed.

2.3 Empirical Studies on Teachers Role as Practitioners of Assessment

2.3.1 Teachers' knowledge, practices and training needs for language assessment. In the literature, a number of researchers have thought of giving voice to language teachers as the real practitioners of assessment processes and sought for their knowledge, beliefs, understanding, practices and training needs for an appropriate and effective language assessment in EFL classrooms.

Köksal (2004) focused on EFL teachers' knowledge and practices of assessment and collected and analyzed 56 classroom test samples from different state schools during the academic years 2001-2002 and 2002-2003. Findings revealed that same classroom tests had been used without any revision or editing over and over again. In the tests, targeted audience, abilities, allocated time and scores for correct answers were not specified. In addition, the tests did not have separate sections (grammar, vocabulary, writing, reading...etc.). Reading tasks were too long and they were not contextualized well, unrelated items were included, they contained grammar or spelling mistakes and they were not appropriate to students' level. Most instructions were incorrect or ambiguous and the tests had problems regarding reliability, validity and washback. The author concluded in that EFL teachers need to be trained in testing and assessing language. He suggested some solutions for this problem and also proposed a distance learning course for teachers on the Internet. Shim (2009) took the issues from teachers' viewpoints and examined 86 primary school EFL teachers' perceptions and practices of language assessment by means of questionnaires and interviews. The teachers reported to have known the principles of

language assessment; however, they did not put their knowledge into practice. They claimed to be influenced by large class sizes, heavy workload, the challenges of educational systems, poor investment on language education and negative attitudes of some colleagues, parents and students. In the same vein, Muñoz, Palacio and Escobar (2012) investigated 62 Colombian teachers' perceptions of assessment. The findings indicated that there was a gap between their beliefs and practices. The authors claimed that they needed guidance for reflective teaching, self-evaluation and formative assessment strategies. Likewise, Jannati (2015) examined 18 Iranian ELT teachers' perceptions and practices of assessment and the role of experience on their perceptions and practices by means of interviews. The results indicated that teachers were familiar with the basic terms, however, they could not indicate their perceived knowledge in their practices. She noted that experience was not a major factor that influence their perceptions. In the above studies the discrepancy between teachers' perceptions of their knowledge and their real practices were particularly noted. Sahinkarakas (2012) drew attention to experience factor in her qualitative study. She used metaphor-completion technique to explore language teachers' perception of language assessment and their self-efficacy in practice of assessment. The significant variant was teaching experience as she carried out the study with 53 pre-service teachers and 47 in-service teachers 24 of whom have more than 5 years of experience. In spite of the pressure of high-stake exams in Turkey, she found out that majority of teachers perceive assessment as an instrument to foster learning. She also revealed that, in-service teachers with a relatively higher experience have more tendency to perceive language assessment as a negative factor compared to pre-service teachers or less experienced teachers. In addition, more experienced teachers were found to be less motivated and less self-confident about their assessment practices. The researcher concluded in that experience has an influence on teachers' attitudes towards language assessment. She interpreted the results in her study as pre-service teachers have the motivation to be successful in their teaching career; however, experienced teachers lose that motivation in years. The researcher referred to the need to keep teachers motivated by organizing in-service teacher training to recharge their enthusiasm.

As is clearly seen, there are many implications for teachers to be guided or trained in the studies above. Some others particularly focused on teachers' training needs. Wu's (2014) study, for example, as a replication of Hasselgreen et al. (2004)

and Fulcher (2012), aimed at analyzing training needs of English teachers in Taiwan for language assessment. The researcher utilized an online survey to find out 253 teachers' language assessment training needs. The results indicated that teachers needed training mostly on assessment of listening and speaking skills. Designing tasks or writing test items, validity procedures and interpreting the scores were three top topics that teachers needed training. The author claimed that teachers cannot cope with the challenges in designing tasks or choosing appropriate tests alone, and she suggested to develop textbooks and design more appropriate training programs for language teachers. In the same vein, Lam (2015) investigated five teacher education institutes in terms of adequacy in preparing pre-service teachers for language assessment. The documentations were analyzed, and case studies were conducted in two of the institutes. Teacher education institutes were found inadequate in training pre-service teachers for language assessment in Hong Kong since they fail to build a bridge between theory and practice and there remains a gap between the two. In these last two studies, the inadequacy and inappropriateness of teacher-training programs were emphasized as a reason for teachers' lack of knowledge and competence in language assessment. There are also a number of studies that investigate the efficacy of short-term training programs or workshops for assessment strategies. Mertler (2009), for instance, sought for the effectiveness of a two-week workshop on classroom assessment of 7 in-services teachers. The topics included basic knowledge of classroom assessment focusing on discussion, practice, and performance assessment tasks. The data were collected through pre-test and post-test method and reflective journals. The results revealed that the two-week training was highly effective in teachers' developing their skills in designing assessment constructs and materials, scoring and interpretation of the results, understanding of formative and summative assessment types and various aspects related to classroom assessment. The author claimed that the workshop had the strength of having an intensive nature as it enabled teachers to concentrate better on their own assessment knowledge and practices and they could critically assess what they need to know and do. Throughout the study the teachers admitted that they had had only superficial knowledge on what assessment is as a term and perceived it as conducting a test to give a grade prior to the training. The study had the implication that in-service teacher training sessions might be useful for classroom teachers to develop their assessment competence.

It is evident with the above literature that the most problematic area in teachers' language assessment perceptions and practices is oral language assessment. In the following section teachers' knowledge, competence, perceptions, practices and training needs of oral language assessment is reviewed.

2.3.2 Teachers' knowledge, practices and training needs of oral language assessment. The studies which reported satisfactory results from oral language assessment competence and practices of EFL instructors are very few in number. Önal (2010), for instance, interviewed three speaking and listening instructors in a preparatory university, which got '*European Award for Languages 2009*' for its project with the purpose to improve '*speaking skill*' of students, in order to find out whether they have a satisfactory teaching and testing procedure of speaking skill or not. A 32-item interview was designated to reveal the important aspects of assessment and testing procedures for speaking skill. The study obtained positive results as the institution is successful in teaching and developing students' speaking skills at the end of the study. Likewise, Ounis (2017) conducted a study with the aim of investigating perception and practices of speaking skills assessment of 20 EFL instructors at the tertiary level in Tunisia. The researcher utilized a questionnaire to collect data. In line with the results, the instructors were found to take boosting learner' oral language skills to forefront in their assessment process; thus, they were reported to perform a continuous, well-planned and an insightful assessment process by means of analytic rating scales which include interactive and authentic assessment tasks despite the challenges such as technical disorders and knowledge deficits.

Majority of the studies which focused on teachers' knowledge, competence or practices of oral language assessment in the literature, have found inadequacies, deficiencies and inappropriateness, though. Lee's (2010) study, for example, examined teachers' practices and perceptions of speaking assessment in middle schools in South Korea. 51 teachers filled in questionnaires and 6 teachers were interviewed. The results indicated that a performance-based assessment was carried out in South Korea; whereas, there was a tendency to use traditional formal testing methods in teachers' practices. The study also found out that teachers had pessimistic attitudes towards speaking assessment. The author suggested some practical actions such as cooperative working of teachers with native speaker teachers, reducing the number of students in the classrooms and training teachers. Muñoz, Dary, Crespo, Gaviria, Lopera and Palacio (2003) also sought for 30 teachers' understanding of

speaking assessment via focus group interviews including questions about teachers' understanding of speaking assessment in general, purpose and frequency of assessing speaking and aspects of speaking assessment techniques. The findings indicated that there were deficiencies in teachers' knowledge of speaking assessment and the researchers ended in suggesting an in-service training for teachers in order to develop their knowledge of speaking assessment which is significant for increasing the quality of their assessment practices. Khamkhien (2010) investigated the practice of both teaching and assessment of speaking skills in Thailand. The author examined the students' test results and observed classes. According to the findings, there were critical problems in speaking tests. One of the biggest problems was the inconsistency of teaching methods with the assessment of speaking. Another big problem was about the level of students in both vocabulary and pronunciation. The author mentioned that communication breakdowns and misconceptions were frequently encountered due to inadequate vocabulary knowledge and serious pronunciation errors which resulted in misunderstanding instructions in tests and lack of intelligibility in their speech. Moreover, making mistakes was reported to decrease learners' self-esteem and they become too shy to take turns again after making mistakes. The author pointed out the lack of authentic use of language as the main reason to these problems. He stated that Thai students did not have the chance to speak English outside the classroom; so, teachers' role is significant to create an authentic atmosphere in the classroom for the students. The author emphasized that teachers need an effective training for the solutions of the problems concerning both teaching and assessment of speaking. Another study that has interesting findings about pronunciation in Turkish context was Emekçi (2016). The author compared native and non-native English teachers' assessment of EFL learners' speaking skills in a state university analyzing the scores given to 80 learners by 3 native and 3 non-native English language teachers through an analytic rating scale. The study scrutinized the components of speaking skills such as fluency, pronunciation, accuracy, vocabulary, and communication strategies. The results indicated similar rating behaviors between native and non-native EFL teachers and both group had their own skills and characteristics which can complement each other; however, a significant difference was found in the evaluation of pronunciation. The author explained the lower grades of native-speaker teachers as being due to their expectation of native-like pronunciation.

A number of studies, which examined teachers' knowledge, perceptions and practices more extensively got some contradictory results between teachers' attitudes or perceived knowledge and their practices. Teachers' attitudes or perceived knowledge were found positive; nevertheless, there were inappropriate, inadequate and deficient implementations in practice as in the aforementioned studies on language assessment. Güllüoğlu (2004) sought for the attitudes of 73 instructors and 127 students towards the instruction and testing of speaking skill at Gazi University preparatory school with the aim of finding out whether speaking is given the necessary emphasis or not, by means of questionnaires given both to instructors and students. Additionally, a test was administered in 3 classes to 60 students who were also given the test afterwards. The findings of the study indicated that although both instructors and students have positive attitudes towards speaking tests and believe that speaking instruction and testing need to be included in the curriculum, there is a lack of emphasis on speaking assessment at prep the school. Similarly, Grada (2014) investigated 76 novice teachers' current knowledge and practice of speaking assessment in a Libyan secondary school by means of a questionnaire, semi-structured interviews and classroom observations. The findings of the study indicated that teachers' have various and complicated understandings of their knowledge and practices which are influenced by both their own perception of knowledge of spoken language and their view of the purpose of assessment. They base their assessment more on linguistic content rather than communicative components of speaking and their emphasis is on the content of assessment rather than the assessment process. They usually work individually, most of them prefer summative assessment technique rather than formative and they are highly influenced by contextual factors such as class size, lack of teacher training, prescribed rules about what assessment criteria to use, lack of support by more experienced teachers, lack of standardized oral language assessment, curriculum and limited time. Some respondents admitted that they have to omit some tasks requiring pair and group work due to time pressure. Some researchers, who focused on the details of teachers' way of assessing oral language performance, reported a number of factors that influence teachers' practices giving voice to teachers, including recommendations for the improvement in their conditions and provision with necessary guidance and training facilities. İnceçay and İnceçay (2010) investigated 18 EFL teachers' needs regarding the current language curriculum in general in which they were teaching receptive and productive language

skills at a private university preparatory school. The results from semi-structure interviews revealed that the most problematic issues were related to oral language practices. The first one was about the teachers' perception of speaking skills. Half of the teachers thought that, among the four skills, speaking has the least weight. One most problematic aspect was, again, pronunciation. Students had poor pronunciation because teachers usually speak in Turkish in the classroom; thus, students remained not familiar with the accents of English. The teachers claimed that they had to switch to Turkish frequently due to heavy workload and overcrowded classrooms. The teachers suggested integrating a speaking club activity with native teachers to the curriculum in order to develop speaking skill and pronunciation as a solution. The authors interpreted the results as teachers' needs had not been taken into consideration while designing the curriculum and suggested that teachers' needs should also be considered so that the implementation process can be healthier. Lozovska-Güneş (2010) carried out a mixed method study with the aim of investigating instructors' perception of speaking skills, challenges that they face while assessing oral performance and learners' attitudes towards speaking skills in three Turkish universities with the participation of 25 academics and 137 students. The results related to our concern indicated that the academics mentioned eight criteria; clarity of expression, fluency, pronunciation, vocabulary, intelligibility, body-language, organization and content for oral performance assessment as significant; they preferred six tasks; individual interview, all in combination, role-play, oral presentations, paired interview and class discussion as advantageous and they mentioned the challenges they face while assessing the oral performance from the top to the least in importance. They were subjectivity, criteria for evaluation, students' anxiety, task relevance, equipment for the assessment of speaking skills and time. Teachers had different ideas about whether the criteria had to be given by their institutions in the form of a rating scale or they had better design their own criteria and rating scales; however, they met in the point that it is useful to have a standardized rating scale on which they can make small changes to make it applicable in their local classrooms. What was interesting in the results of this study that, none of the academicians think that they were reliable as raters and majority of them thought they need a special training to assess speaking. Plo, Hornero and Mur-Dueñas (2013) sought for attitudes, beliefs, practices and perceived needs of 63 Aragonese teachers, most of whom taught in state high schools, regarding EFL

teaching in general and developing teaching / assessing oral language skills in particular. A survey was distributed online to the teachers to collect data. The results indicated that oral language skills were not sufficiently placed in both practice and assessment, most of the class time was given to grammar, vocabulary and reading, sufficient time was spent for listening, relatively less time was spared on writing which was covered by individual work outside the classroom. Oral language skills were least practiced skills in the classroom due to several reasons mentioned by the participants such as students' reluctance or insufficient class hours. They suggest solutions such as reducing the number of students in classes, encouraging students to travel abroad, attend exchange programs or intensive language courses and being assisted by native teachers. Teachers usually utilize listening comprehension, oral presentation, picture/photo utilized activities and teacher-student interactive activities such as questions-answer or whole-class work. Role-playing was mentioned by very few teachers and none spoke off interactive group projects or debates as oral language teaching / assessment tasks. Teachers reported their students as competent in grammar accuracy; however, they stated that students were not good at fluency and pronunciation. Authors of this study pointed out a need for teacher training emphasizing the importance of life-long learning which they stated that it was also confirmed by the teachers, themselves. The areas that the teachers indicated / mentioned a need for training were using of technological devices, developing strategies and methods in teaching oral language and how to cope with mixed level classes.

Some studies examined the effectiveness of training programs and workshops designed for the assessment of oral language skills and most got positive findings. Knight's (1992) study sought for the effectiveness of a workshop in helping teachers to specify their criteria according to their contexts. Teachers were shown a video-clip of a student speaking English and asked which aspects they would grade in the speech, first. Then, they were shown a second video-clip and asked to comment on relevance and usefulness of the given criteria and which items they would add or delate to make it more useful. The author reported that teachers avoided commenting in the second cycle and remained silent. The results also indicated that, although some teachers were persistent in their opinion for the need to have standard criteria, most found workshop useful in guiding them to determine their own criteria. In a more recent study which was conducted by Davis (2015), 20 experienced teachers

got training on scoring speaking test. They were asked to score recorded responses from TOEFL iBT speaking test before and after the training sessions. Results indicated that inter-rater reliability considerably increased after training sessions. Koh, E Carol-Ann Burke, Luke, Gong and Tan (2017) endeavored a longitudinal study in which 12 Chinese language teachers from two primary school were provided with a two-year professional training program on designing and using authentic intellectual tasks for classroom activities in Singapore. The teachers were asked to submit each task they designed and used with the samples of students' works including high, moderate and low achievement scores prior to the training program and following each pace of the training program. Within the analysis of the data, which was carried out by both research team and the teachers' counterparts from the other school by means of a particular technique called "authentic intellectual quality" including a set of criteria to guide teachers in rating, remarkable improvements were found both in terms of quality of the tasks designed by the participant teachers and students' works.

In the first section of this chapter, the literature on fundamental components of oral language assessment, which were also taken as main aspects in this study, were reviewed, and definitions, discussions and studies on these components were presented. In the second section, the studies related to perceptions, knowledge, skills, practices of teachers were focused. Additionally, the factors that have influence on their practices and the studies on the areas that needs improving and teachers' training needs were reviewed. In the sense of giving voice to teachers on this subject, there are few studies conducted in Turkish contexts. As for analyzing their training needs for oral language assessment, no studies were encountered in the related literature to researchers best knowledge. This thesis might be a contribution to the literature in terms of shedding light to teachers' training needs from their perspectives for oral language assessment.

Chapter 3

Methodology

3.1 Overview

The primary aim of this thesis was to ascertain the needs of high school EFL teachers for carrying out a successful oral language skills assessment in their classrooms. For doing so, high school EFL teachers' knowledge of speaking assessment constructs, criteria, tasks and techniques, the influence of their contexts on their knowledge of speaking assessment and practices and their professional training needs were investigated by means of questionnaires and semi-structured interviews addressing to following research questions:

- 1) What do high school teachers know about oral language assessment?
- 2) What are high school teachers' current practices for oral language assessment?
- 3) What factors in high school teachers' contexts influence their oral language assessment?
- 4) What are high school teachers' professional training needs for assessing their students' oral language performance efficiently?

This thesis was carried out with 50 EFL teachers, who work at state high schools in Beşiktaş, İstanbul, Turkey, in the Spring Semester of 2017–2018 Academic Year. In this chapter, there exists the information about design, setting, participants, data collection instruments, data collection procedures and data analysis procedures.

3.2 Research Design

In this thesis, mixed-method design, in which both quantitative and qualitative data collection tools and data analysis methods are collated, was used. Mixed-method design has most often been advocated in the literature. Johnson and Onwuegbuzie (2004) argue that mixed research solutions enable many research questions and combinations of questions to be answered ideally and fully. The researcher utilizes the strengths of both qualitative and quantitative research and bases the inquiry on collecting diverse types of data; thus, more insight is gained, and a broader understanding of the research problems is provided than using either of them separately (Creswell, 2009; Creswell, 2014). Reviewing the literature, mixed-method

design was chosen, as it is also the best type of design for this thesis to find out complete answers to the research questions. In order to collect quantitative data, two questionnaires; *EFL Teachers' Knowledge of Speaking Assessment Questionnaire (KOSAQ)* (See Appendix A) and *Training Needs Analysis for Oral Language Assessment Survey (TNAFOLAS)* (see Appendix B) were given to 50 participants. The first questionnaire focused on high school teachers' knowledge of speaking assessment and their perception of self-efficacy in assessing oral language in their classroom. The second questionnaire required high school teachers to think on the areas that they need training for successfully assessing oral language performance of their students.

In-depth qualitative data were collected through semi-structured interviews (*see Appendix C*) conducted to 7 participants, 6 of whom were females and 1 was male. In addition, open-ended sections in both KOSAQ and TNAFOLAS include items that enabled the researcher to gather data about details of high school teachers' knowledge, current practices, influence of their contexts to their knowledge and practice of oral language assessment and their training background on oral language assessment. Detailed information about these two questionnaires and semi-structured interviews was presented in the 'Data Collection Tools' section in this chapter.

3.3. Setting and Participants

The study was carried out in the Spring Semester of 2017-2018 Academic Year, in 15 state high schools, 10 of which are vocational high schools in Beşiktaş, İstanbul, Turkey. Non-probability convenience sampling was adopted in the present study, in which samples are chosen with regard to a number of criteria such as voluntariness, easy accessibility and time availability (Dörnyei, 2007), and high school EFL teachers, who work in Beşiktaş, were chosen as the participants of the present study. By this means, the researcher could visit the high schools and meet the participants face-to-face, give information about the aim and scope of the study and could gather adequate qualitative and quantitative data as was significant for the purposes of this study. The participants of the present study are 50 non-native EFL teachers who work in the state high schools in Beşiktaş, İstanbul. Detailed information about the gender, educational degree, age and teaching experience of the participants have been provided in Table 1 below:

Table 1
Demographic Information of Participants

Demographic Information	Frequency (n)	Percentage (%)	
Gender	Male	8	16
	Female	42	84
Educational Degree	College graduate	40	80
	Master's degree	10	20
Age	25 – 34	9	18
	35 – 44	30	60
	45 or over	11	22
Teaching Experience	2 – 5 years	2	4
	6 – 10 years	7	14
	11 years or more	41	82

As seen in *Table 1*, the participants of the study are 50 non-native EFL teachers, 42 females and 8 males, each is over age of 25 and has varying experience over 2 years. Female teachers outnumbered male teachers which is natural as is convenient with the gender range of EFL teachers in Turkey. Majority of participants are college graduates since teachers are not required to get any further education to teach English in high schools. Majority of them are in the age range of 35 – 44. Teachers are appointed in compulsory districts in their first years of teaching; therefore, teachers with at least 2 – 3 years of teaching experience are usually assigned in Beşiktaş. None of the participants are under the age of 25. The results for teaching experience of high school teachers are also consistent with the age range mentioned above. Vast majority of teachers have more than 11 years of experience and very few of them have the experience of teaching 2-5 years, probably for the similar reasons in the age section.

3. 4. Procedures

3.4.1 Data Collection Instruments. Data collection instruments consist of two questionnaires and semi-structured interviews which sought for data about the high school EFL teachers' current knowledge and practices, the deficiencies in their

knowledge and training, the role of their contexts in their current knowledge and practices and their training needs.

3.4.1.1 Questionnaires. As stated by Dörnyei (2003), questionnaires are quite practical in terms of time, effort and financial resources and a researcher can collect a huge amount of data in a relatively short time, and as also confirmed by Gillham (2000), they are significant for the researchers who are having a full-time job at the same time. Along the line of this thesis, using questionnaires consisting of mostly close-ended items were helpful as they enabled the researcher to collect data from high school teachers, who were too busy to spare much time. They could rate their knowledge of oral language assessment and indicated their professional training needs in a short time; the open-ended items on the other hand, gave the opportunity to probe into the factors affecting high school teachers' knowledge and practices, what they exactly do to assess oral performance in their classrooms and how they do it with more detailed information. As suggested by Oppenheim, (1992) open-ended questions let respondents say what they think spontaneously with greater richness (p.81); therefore, they provide rich, in-depth data for the research.

3.4.1.1.1 Knowledge of speaking assessment questionnaire (KOSAQ). The researcher adapted this questionnaire from Grada (2014) after getting her permission via e-mail correspondence (*see Appendix D*). The aim of using this questionnaire was to reveal high school teachers' knowledge, current practices for oral language assessment and the influence of their contexts on their knowledge and practices.

The adapted version of the questionnaire consisted of 42 items in total and was divided into three sections including both close-ended and open-ended items. The first section consisting of 7 items, was designed for gathering demographic data about the participant high school teachers and providing an overview of their classroom practices for the assessment of oral language skills.

The second section includes three parts, with 23 items in total (13- 35) with 4-point Likert rating scales requiring high school teachers to choose a number from 1 to 4 (1= not good, 2= moderate, 3= good, 4= excellent) to evaluate their knowledge of oral language assessment constructs in *Part A*, knowledge of oral language assessment techniques and tasks in *Part B* and knowledge of assessment criteria used in speaking tests in *Part C*. This section addressed and provided answers to the first research question; *What do high school EFL teachers in Turkey know about oral language assessment?*

Before administrating the questionnaire with the participants of this thesis, the researcher conducted a pilot study with 6 teachers who have similar contexts and conditions in Üsküdar, İstanbul in the fall semester of 2017. The alpha coefficient reliability was measured via SPSS program. The reliability of each part (A, B and C) in Section Two, which involves quantitative data, was processed separately on SPSS. Within each section, the items were internally consistent, which indicated high reliability as shown in *Table 2*:

Table 2

Reliability Statistics for KOSAQ

PART	Number of items	Cronbach's Alpha
A	5	.892
B	13	.899
C	10	.874

The third section consisted of 7 items, 6 of which include titles about factors that might have an influence on high school teachers' knowledge and practices of oral language assessment in their contexts and requires high school teachers to explain how the chosen factors in their contexts affect their knowledge and practices. As for the last item, it required them to explain how they carry out speaking assessment in their classrooms. This section provided detailed information for the second and third research questions; *What are high school teachers' current practices for oral language assessment?* and *What factors in high school teachers' contexts influence their oral language assessment?* The data collected from this section also contributed to the first research question since their explanations shed light on their knowledge of oral language assessment.

3.4.1.1.2 Training needs analysis for oral language assessment survey (TNAFOLAS). This survey was developed by adapting LTA Survey from Wu (2014), again asking her permission kindly via e-mail (*see Appendix E*). She let the researcher adapt the Survey and recommended citing and giving reference carefully to the researchers (Hasselgreen et al., 2004; Huhta et al., 2005 & Fulcher; 2012) to whom she referred while developing LTA Survey. The researcher added some items to the Survey informed by the reviewed literature with the aim of ascertaining the professional training needs of high school teachers for a purposeful, reliable and valid assessment of oral language performance in their classrooms which will have a positive effect on their students' learning, addressing to the third research question;

What are teachers' professional training needs for assessing their students' oral language performance efficiently?

In the first section of the survey, the participants were asked whether they had had any training on oral language assessment, and they were required to rate its adequacy in satisfying their needs in the process of assessment for oral language skills if they had had so. They were also required to write on what aspects of oral language assessment they still have training needs in an open-ended question in this section. This section provided an overview of high school teachers professional training experience and their professional training needs and warmed them up to the scope of the questionnaire.

The second section involved a *5-point Likert Scale* which requires high school EFL teachers to select items from not important to very important with the aim of investigating their thoughts about their professional training needs on prespecified topics. Again, a pilot study for the reliability of the adapted version of the Survey was conducted with the same 6 high school teachers who work in Üsküdar, İstanbul in the Fall Semester of 2017. The results indicated a high reliability as shown in *Table 3* below:

Table 3

Reliability Statistics for TNAFOLAS

Number of items	Cronbach's Alpha
22	.979

This section enabled the researcher to collect detailed data to reveal on what aspects of oral language assessment high school teachers need professional training.

3.4.1.2 Semi-structured interviews. Dörnyei (2007) claims that interviewing is a good way of collecting rich amount of in-depth data in various situations, even by inexperienced researchers, as it is a 'natural and socially acceptable' data collection method by which people might feel more comfortable and researcher might indicate more flexible approaches allowing her/him to scrutinize any issues arising from natural flow of the interview (p.143). In the present study, 7 of the high school teachers (6 females and 1 male), who filled in the questionnaires, volunteered for the semi-structured interviews. They were interviewed by the researcher in their own schools at the time appropriate to them in their mother tongue so that they could feel more comfortable. Each interview took approximately 15 to 20 minutes. For the

semi-structured interviews, 7 questions were prepared, in advance, parallel to the scope of the questionnaires, in accordance with the research questions and complying with the aim of the study. The interviews were efficient in obtaining in-depth data and triangulating the results of the quantitative data gathered by means of questionnaires.

3.4.2 Data Collection Procedures. Having reviewed the literature, two questionnaires were chosen to be utilized while developing the questionnaires to adopt for the purposes of this thesis. After necessary permissions were received from the researchers who developed the questionnaires via e-mail correspondence, the related sections and items were adapted to the present study. In this pace, first, the researcher adapted three sections from Knowledge of Speaking Assessment Questionnaire (Grada, 2014) as these sections completely overlapped the scope of this thesis to develop KOSAQ. Secondly, TNAFOLAS was developed referring to the items related to the scope of the present study in Wu (2014). After the development of the questionnaires, the reliability of the Likert scales was checked through piloting with 6 non-native high school EFL teachers, who worked in a high school in Üsküdar, İstanbul. Once high reliability statistics was obtained, as previously shown on *Table 2* and *Table 3* in this chapter, the researcher received the official permission from National Education Directorate of Beşiktaş District to conduct the study.

The researcher started to gather data in the spring semester of 2017 – 2018 Academic Year. First, high school administrators were contacted to get information about the available time for English teachers to be at their schools. After obtaining the information about the best time to find most of EFL teachers in their schools available to fill in the questionnaires, the researcher started to visit the high schools in person and meet teachers face-to-face. In this way, a brief information about the aim and the scope of the thesis and what they were expected to do in order to contribute to it was given to the high school teachers by the researcher. 50 of them volunteered to contribute to the study and filled in the questionnaires. Visiting the schools and meeting the teachers in person in their own contexts helped the researcher create a warmer atmosphere, have conversations about the assessment of oral language skills in their classrooms, tell the aim of the study better, make explanations for any item that they needed clarification and get an idea about their concept and conditions. The researcher also granted that the participants' identities

will be kept anonymous, the gathered data will only be utilized for the purposes of the current thesis and the information obtained about them will never be used anywhere else for another purpose. Thus, it enabled the researcher to obtain more detailed and more sincere answers. Some of the teachers did not have time to fill in the questionnaires when they were given it first; so, the researcher made the definitions and explanations they needed and gave her contact information in case they needed any additional clarification of the items in the questionnaires and visited their schools again a week later. In the end, 50 questionnaires were filled in and given back to the researcher ready to be analyzed.

The teachers, who had time after filling in the questionnaires, were asked whether they would also like to participate in the interview and 7 of them agreed to be interviewed. The interviews took place in high school teachers' own place, in Turkish which enabled a more comfortable circumstance for the interviewees to answer as they wish. Each interview took approximately 15 – 20 minutes in total and recorded with the permission of interviewees.

It took 3 weeks, in total, to complete visiting the high schools, being introduced to high school EFL teachers by school administrators, delivering and taking back the questionnaires and conducting the interviews. At the end of 3 weeks, all the data gathered from two questionnaires and semi-structured interviews were prepared to be analyzed.

3.4.3 Data Analysis Procedures. In this thesis, both qualitative and quantitative data collection tools were utilized; therefore, the analysis of data is based on both qualitative and quantitative data analysis methods. The data collection tools and data analysis methods were shown in relation to the research questions, in detail on *Table 4* below:

Table 4

Data Collection Tools and Data Analysis Methods

Research Question	Data Collection Tool	Data Analysis Method
1- What do high school teachers know about oral language assessment?	KOSAQ (Items 8 – 35)	SPSS (Frequencies and Percentages)
	Semi-structured interview (Question - 6)	Content analysis (Pattern-coding)

Table 4 (cont.d)

2- What are high school teachers' current practices for oral language assessment?	KOSAQ (Items 5 - 7 & 42)	Content analysis (Pattern-coding)
	Semi-structured interview (Question-3)	Content analysis (Pattern-coding)
3- What factors in high school teachers' contexts influence their oral language assessment?	KOSAQ (Items 36 - 41)	Content analysis (Pattern-coding)
	Semi-structured interview(Question-3)	Content analysis (Pattern-coding)
4- What are high school teachers' professional training needs for assessing their students' oral language performance efficiently?	KOSAQ (Item-41)	SPSS (Frequencies and Percentages) & Content analysis (Pattern-coding)
	TNAFOLAS (Items 1-25)	SPSS (Frequencies and Percentages) & Content analysis (Pattern-coding)
	Semi-structured interview (Question- 4, 5& 7)	Content analysis (Pattern-coding)

The data from the Likert scales were transferred into SPSS, analyzed statistically, and frequency and percentage of ratings for each item were indicated in tables. For the analysis of the qualitative data, recorded interviews were listened by the researcher several times and each were transcribed on a Word document and translated into English. The interviewees were named as T1, T2, T3... to ensure anonymity. The statements of the interviewees were categorized under main topics, considering the research questions and main purpose of the thesis as recommended by Auerbach and Silverstein (2003), such as “the factors influencing high school teachers' knowledge and practices” or “the assessment criteria used by high school teachers” in order to keep focused. After categorizing the answers of each interviewee, pattern-coding was carried out by highlighting most frequently used phrases and words, or the ones with similar meanings.

The data gathered from open-ended questions in both KOSAQ and TNAFOLAS were also transferred on a Word document under related titles, which were comprised of the questionnaire items, and then coded in the same way with semi-structured interviews. In addition, striking quotes from the interviewees, which might serve as illustrative examples or key pieces (Booth, Colomb & Williams, 2003; Lofland, Snow, Anderson & Lofland, 2006; Saldaña, 2009) or might provide

surprising or interesting results (Creswell, 2007), were also underlined as suggested by many researchers (e.g. Layder, 1998; Boyatzis, 1998).

Ultimately, the statistical results were presented on tables, and the scores were interpreted by the researcher in Chapter 4. The findings from semi-structured interviews and open-ended items in the questionnaires were compared and contrasted, as was emphasized by Saldaña (2009), and discussed in relation with the research questions under subtitles which were generated in the light of categories and codes obtained through the analysis of qualitative data.

In this chapter, the setting, the participants, data collection instruments, procedures and data analysis procedures were discussed. The detailed information of the results from the analyzed data are provided in chapter four.



Chapter 4

Findings

4.1 Overview

In this chapter, results from both questionnaires and semi-structured interviews are presented in line with the research questions. The aim of the first research question was to find out what high school EFL teachers know about of oral language assessment, the second research question sought for high school teachers' current practices of oral language assessment, the third research question aimed to investigate influence of high school teachers' contexts to their knowledge and current practices of oral language assessment and the fourth research question was formed to reveal high school teachers' training needs for carrying out an effective oral language assessment.

4.2. Findings

To investigate high school EFL teachers' knowledge and current practices of oral language assessment and to determine their training needs for carrying out an efficient assessment of oral language skills, 50 high school EFL teachers were given *KOSAQ* and *TNAFOLAS* to fill in, and 7 of them were interviewed. Quantitative data from Likert scales were analyzed via SPSS, and in-depth analysis of qualitative data from open-ended questions and semi-structured interviews were carried out through pattern-coding. Both qualitative and quantitative findings were compared and contrasted, and the results were presented in four main sections including knowledge of speaking assessment, current practice of speaking assessment, influence of teachers' contexts on their decision making in oral language assessment process and high school teachers' training needs for assessment of oral language skills. The results from four sections were discussed in accordance with the research questions.

4.2.1 Findings for research question 1: What are high school teachers' knowledge of speaking assessment?

4.2.1.1 Findings from *KOSAQ*. In the light of the first research question, high school EFL teachers' knowledge of speaking assessment was investigated within three parts in *KOSAQ*; knowledge of speaking assessment constructs, knowledge of speaking assessment techniques and tasks and knowledge of assessment criteria used

in speaking assessment. Additionally, the participants' answers to open-ended questions in the same questionnaire and the data from the related questions in semi-structured interviews were also utilized to probe into high school teachers' knowledge, strengths and weaknesses in oral language assessment.

4.2.1.1.1 *Knowledge of speaking assessment constructs.* In the first part (Part A) of Section Two, high school teachers' knowledge of speaking assessment constructs were investigated via participants' ratings for their knowledge of 5 speaking assessment constructs; aspects of intonation, stress and pronunciation, accuracy, fluency, students' interactive skills and ability to take turns in a pair/group work, with a number 1 – 4 on a 4 – point Likert scale, and frequencies and percentage values of each item are shown in *Table 5*.

Table 5

Knowledge of Speaking Assessment Constructs

		<i>1=Not good</i>		<i>2= Moderate</i>		<i>3=Good</i>		<i>4=Excellent</i>	
Items	<i>When I design speaking assessment, my knowledge of considering ...is...</i>	<i>f</i>	<i>%</i>	<i>f</i>	<i>%</i>	<i>f</i>	<i>%</i>	<i>f</i>	<i>%</i>
8	...aspects of intonation, stress, and pronunciation			15	30	23	46	12	24
9	...accuracy	1	2	7	14	26	52	16	32
10	...fluency			5	10	27	54	18	36
11	...students' interactive skills	2	4	7	14	17	34	24	48
12	Students' ability to take turns in a pair/group work	3	6	10	20	20	40	17	34

The results indicate that majority of the participants tended to rate their knowledge of speaking constructs as 'Good' in most of the items to varying degrees, and the range of ratings for 'Excellent' has a relatively high percentage (over 20%), as is apparent in *Table 5* above. On the other hand, 15 participants (30%) rated their knowledge as 'Moderate' for the aspects of intonation, stress and pronunciation (*Item – 8*). Concordantly, only one of the participants (T16) mentioned intonation and none of them mentioned stress in open-ended questions in KOSAQ, when asked about their practices of oral language assessment. Likewise, although majority of the participants mentioned fluency, accuracy and content, very few of them stated that they assess students' pronunciation skill in their answers in KOSAQ. It is evident

that most of the participants of this study do not give a place to pronunciation in their assessment process.

I let students speak about a particular topic for 2-5 minutes. While listening to them, I take notes about fluency, pronunciation, intonation and self-confidence (T16, Feb 2018).

The findings of the interviews were consistent with the results above. Only T6 claimed they had pronunciation as an oral language assessment construct during the interviews as shown in the following quote:

Fluency, accuracy, content, vocabulary and interaction to assess their interactive skills, we also have pronunciation. There is one more, but I cannot remember now (T6, Feb 23, 2018).

Other 6 interviewees did not speak of pronunciation as an ability that they assess. T5 openly said that she omits pronunciation in her assessment process as is seen in the following excerpt:

Accuracy, using vocabulary appropriately... I don't care much about pronunciation, because they pronounce many things incorrectly and when they are corrected too much they are discouraged (T5, Feb 21, 2018).

For students' ability to take turns in a pair/group work (*Item -12*), 26 % of the participants ($n=13$) rated their knowledge as 'Not good' or 'Moderate'. Within the analysis of the qualitative data for a better insight, it was found out that only one participant (T1) out of 50 participants mentioned pair / group work while talking about her assessment process as is indicated in the following excerpt from her interview:

I made a system which I have seen beneficial for years. I have low level students sit with the ones who have higher level of English and I can clearly see the progress with low level students. They gain confidence while working with peers and participate in the activities more, especially when we carry out group or pair work, and they learn from peers (T1, Feb 8, 2018).

In short, the findings from KOSAQ indicated that most participants rated their knowledge as 'Good' for speaking assessment constructs; whereas, a considerable number of the participants rated their knowledge as 'Moderate' for the assessment of aspects of intonation, stress and pronunciation, and 'Not good' or 'Moderate' in assessing students' ability to take turns in a pair/group work. Additionally, the qualitative data collected by means of *Item - 42* in KOSAQ and semi-structured interviews revealed that very few participants assess pronunciation, intonation and stress. It is evident that, high school teachers do not include these aspects in their oral

language assessment. Likewise, vast majority of participants do not seem to assess oral language skills in group / pair work.

4.2.1.1.2 *Knowledge of speaking assessment techniques and tasks.* In the second part (Part B), which was related to speaking assessment techniques and tasks, 53 % of the participants in *Item – 15* (using a portfolio assessment technique), 50 % of the participants in *Item – 14* (using a peer assessment technique), 44 % of the participants in *Item – 16* (using formative and summative assessment techniques) , and 40 % of the participants in *Item – 21* (using learner-learner joint discussion and decision-making tasks) and *Item – 25* (using re-telling story or text from aural stimuli as a speaking task) rated their knowledge as ‘*Not good*’ or ‘*Moderate*’ . In addition, 30 % or more of the participants rated their knowledge as ‘*Not good*’ or ‘*Moderate*’ for using a student self-assessment technique (*Item – 13*), using different speaking assessment tasks (*Item - 17*) and using interview tasks (*Item – 23*) during speaking assessment as shown in *Table 6* below.

Table 6

Knowledge of Speaking Assessment Techniques and Tasks

1=*Not good* 2= *Moderate* 3=*Good* 4=*Excellent*

Items	<i>I rate my knowledge of.....as...</i>	1	2	3	4
		<i>f</i> %	<i>f</i> %	<i>f</i> %	<i>f</i> %
13	...using a student self-assessment technique, when students are being orally assessed	3 6	12 24	25 50	10 20
14	...using a peer assessment technique, when students are being orally assessed	6 12	19 38	14 28	11 22
15	...using a portfolio assessment technique, when students are being orally assessed	6 12,2	20 40,8	15 30,6	8 16,3
16	...using formative and summative assessment in assessing speaking ability	3 6	19 38	18 36	10 20
17	...using different speaking assessment tasks	2 4	13 26	24 48	11 22
18	...time factor when choosing or designing speaking tasks		12 24	25 50	13 26
19	...setting speaking assessment tasks in an appropriate level of difficulty	2 4	7 14	22 44	19 38
20	...using oral presentation tasks in speaking assessment	1 2	9 18	25 50	15 30
21	...using learner-learner joint discussion and decision-making tasks	3 6	17 34	18 36	12 24
22	...using role-play tasks in speaking assessment	6 12	8 16	27 54	9 18

Table 6 (cont.d)

23	...using interview tasks in speaking assessment	6 12,2	12 24,5	23 46,9	8 16,3
24	...using visual stimuli tasks such as pictures to provide a topic of conversation to the students in speaking test	1 2	11 22	21 42	17 34
25	...using re-telling story or text from aural stimuli as a speaking task in speaking assessment	7 14	13 26	21 42	9 18

In Part B, the number of teachers who rated their knowledge as ‘*Good*’ exceeded the number of teachers who used other ratings in general again; however, fewer teachers rated their knowledge as ‘*Excellent*’ compared to Part A (knowledge of speaking constructs), and the teachers who rated their knowledge as ‘*Not good*’ or ‘*Moderate*’ outnumbered the previous part in KOSAQ. According to the statistical findings, majority of the participants perceive their knowledge of using a portfolio assessment technique as not adequately good; half of the participants are not wholly confident about their knowledge of using a peer assessment technique, and a remarkable number of the participants think that their knowledge is not good in carrying out formative and summative assessment, using a student self-assessment technique, using learner-learner joint discussion and decision-making tasks, re-telling story or text from aural stimuli tasks and interview tasks.

4.2.1.1.3 Knowledge of speaking assessment criteria. As is seen in Table 7 below, participants mostly rated their knowledge as ‘*Good*’ or ‘*Excellent*’ in this part (Part C), too, which should indicate that they are confident about their knowledge of speaking assessment criteria. In the last two items (34 and 35), which are related to using ‘analytic’ or ‘holistic’ scores in speaking tests, 19 of the participants (38 %) rated their knowledge as ‘*Not good*’ or ‘*Moderate*’. Ratings of these two items are considerably low when compared to the other items which might mean that a remarkable number of the high school teachers are not familiar with these terms or they do not know how to deal with these scoring systems while assessing oral language performance in their classrooms.

In item 32, 24 % of the teachers rated their knowledge of considering pronunciation as a category of speaking assessment as ‘*Moderate*’, and this is concordant with the rating of ‘*Item – 8*’ in Part A which was, again, related to intonation, stress, and pronunciation. Another item, which 10 participants (20 %) rated as ‘*Not good*’ or ‘*Moderate*’, was related to using learner-learner joint discussion and decision-making tasks.

rated their knowledge as 'Moderate', is 'Item – 33' (considering the use of appropriate expressions as a category of speaking assessment). This might be another aspect of oral language assessment which high school teachers have difficulty in oral language assessment process.

Table 7

Knowledge of Speaking Assessment Criteria

		<i>1=Not good</i>		<i>2= Moderate</i>		<i>3=Good</i>		<i>4=Excellent</i>	
Items	<i>I rate my knowledge of.....as...</i>	<i>f</i>	<i>f</i>	<i>f</i>	<i>f</i>	<i>f</i>	<i>f</i>	<i>f</i>	<i>f</i>
		<i>%</i>	<i>%</i>	<i>%</i>	<i>%</i>	<i>%</i>	<i>%</i>	<i>%</i>	<i>%</i>
26.	...developing assessment criteria used in a speaking test	2	7	32	9	4	14	64	18
27.	...categories used in speaking assessment	2	6	28	14	4	12	56	28
28.	...considering accuracy as a category of speaking assessment		5	29	16		10	58	32
29.	...considering fluency as a category of speaking assessment.		7	27	16		14	54	32
30.	...considering communicative skills as a category of speaking assessment		8	19	23		16	38	46
31.	...considering range of vocabulary use as a category of speaking assessment		4	28	18		8	56	36
32.	...considering pronunciation as a category of speaking assessment		12	21	17		24	42	34
33.	...considering the use of appropriate expressions as a category of speaking assessment		10	28	12		20	56	24
34.	...using an analytic score for each category in a speaking test	2	17	22	9	4	34	44	18
35.	...using a holistic score in a speaking test	1	18	21	10	2	36	42	20

The results from Part C indicate that majority of the participants are confident about their knowledge of oral language assessment criteria in general; nevertheless, a considerable number of participants put down to the fact that they do not think their knowledge is good enough for considering pronunciation as a category of speaking assessment (24%), which is compatible with the related results in Part A, using 'analytic' or 'holistic' scores in speaking tests (38%) and considering the use of appropriate expressions as a category of speaking assessment (20%).

4.2.1.2 Findings of semi-structured interviews for knowledge of oral language assessment.

4.2.1.2.1 Adequacy of high school teachers' knowledge of oral language assessment. All interviewees were asked whether they think they have adequate knowledge to assess oral language assessment and 6 out of 7 interviewees clearly stated that they do not think their knowledge is adequate. T4, for example, answered “No” when she was asked whether she thought her knowledge was adequate for assessing oral language skills in her classrooms or not, and added she was both tired of dealing with the paper work in the assessment process and confused because of the arduous nature of speaking assessment as is shown in the following excerpt:

No, I don't. First, I think I have a kind of professional deformation, because I don't want to deal with all those paper work. Secondly, speaking is not a thing that you can assess clearly writing on a piece of paper [...] Assessment of speaking is ambiguous. You can assess written things in a way and you can show what is incorrect easily, but assessing speaking is too difficult (T4, Feb 19, 2018).

Another interviewee (T3) also answered “No” when he was directed the same question. When the researcher asked about his training needs, T3's answer was interesting. He sincerely said that he even does not know what he needs to know as shown in the following excerpt:

I don't know much about this issue. I haven't searched about it; so, I don't know what exactly I need to know [...] I mean, I have my own ideas, but they do not have an academic basis (T3, Feb 14, 2018).

This might also be the main reason for majority of high school teachers' rating almost all items as *Important* or *Very Important* in TNAFOLAS for their training needs (see Table 14), which contradicts with their high ratings in KOSAQ for their knowledge of oral language assessment. They might have rated their knowledge of concepts written in the questionnaire items in KOSAQ; whereas, they are probably not sure what they really need to know to carry out an efficient oral language assessment in their classrooms. T5 mentioned some consequences of lack of knowledge for oral language assessment in her own practices, clearly. She stated that she cannot stick to the criteria that she has previously determined due to some challenges in her context. In addition, she frankly said that she thinks they cannot assess oral language skills healthily as indicated in the following quote from her interview:

Adequate? No, I believe no one can have adequate knowledge on any subject. We all have shortcomings, but we are trying to cover them up. For example, I determine some criteria, but I cannot stick to them while doing the assessment because of large class sizes and physical inappropriacy of the classrooms and many other factors like that. I don't think we teach or assess speaking healthily (T5, Feb 21, 2018).

In line with the findings above it is evident that, despite high ratings of knowledge of speaking assessment in KOSAQ, high school teachers' have some shortcomings in their knowledge of oral language assessment. They might have had an idea on what the terms they encountered in the questionnaires stand for; however, they do not indicate that they have sufficient knowledge to put them into practice as will also be analyzed within the second research question related to their practices.

4.2.1.2.2 We are required to assess speaking. Assessment of oral language skill is as significant as the assessment of other language skills such as reading, writing and listening as was previously mentioned in this thesis in Chapter 1 and 2. While the researcher was conducting the interviews, she also found out that high school teachers lack knowledge of significance of oral language assessment. Two of the interviewees (T4 and T6), for example, mentioned that they try to make an assessment for oral language skills just because they are required to do it. It is clear in the following statement of T6 that she carries out oral language assessment since she feels obliged:

Well, of course it is better if we can learn the things better. I make an assessment because we have to do it... (T6, Feb 23, 2018).

There were similar statements in KOSAQ. Some participants mentioned that oral language assessment was something that they were obliged to do:

It is compulsory for us to assess students' oral performance. Instead of making an oral exam, it is better to evaluate during the course (T38, Feb 2018).

The statements of the participants above reveal that some high school teachers carry out speaking assessment in their classrooms since they are required to do it by the Ministry of Education. Although they try to make an assessment for oral language skills, they are not aware of its significance or contribution to their classroom instruction processes. Some high school teachers do not even think that it is necessary, and they do it perfunctorily without any objectives in mind as is seen in the following excerpt from T4's interview:

We are required to do this, so we are trying to do this. If you ask me, if we weren't required, we would not do [...] I don't know whether it is necessary or not. To be honest, I don't think it is necessary... (T4, Feb 19, 2018).

As is seen in the statements above, some high school teachers are not aware of significance and benefits of assessing oral language assessment, and some carry it out aimlessly and disgustedly, just because they are required. This indicates that there is also a deficiency in their knowledge of effect of the assessment process in teaching of oral language skills.

The first research question was designed to find out what high school EFL teachers know about the assessment of oral language assessment. In line with this purpose, the qualitative data from both open-ended items and interviews were analyzed and interpreted together with the quantitative results, and ultimate findings were deduced as follows:

Even though majority of the high school teachers rated their knowledge for most of the items as 'Good' in KOSAQ, a remarkable number of them indicate inadequacies in their knowledge of speaking assessment which were revealed as the researcher probed into the items that were rated as 'Not good' or 'Moderate' by many of the participants. These items include the assessment of the aspects of intonation, stress and pronunciation and the use of appropriate expressions, managing pair/group activities as oral language assessment tasks; using variety of assessment types such as portfolio assessment, peer assessment, student self-assessment, formative and summative assessment, utilizing authentic speaking assessment tasks such as learner-learner joint discussion and decision-making tasks, re-telling story or text from aural stimuli and interviews and using 'analytic' or 'holistic' scores in oral language assessment. It is evident that a considerable number of participants are not confident about their knowledge in the above areas of oral language assessment. In addition, in-depth analysis of the interviews revealed that high school teachers think that their knowledge is inadequate for an efficient oral language assessment in general, as is clearly seen in their statements of the interviewees above. Moreover, some high school teachers do it just because they were required and that they do not think it is necessary to assess oral language skills. This indicates that they also lack the knowledge of the place, significance and benefits of assessing oral language skills.

4.2.2 Findings of research question 2: What are current practice of high school teachers for speaking assessment? Participants’ current practices were investigated for the second research question, via related items in KOSAQ (5, 6, 7 and 42) and *Question-3* in semi-structured interviews. The results for the current practices of high school teachers were presented as follows:

4.2.2.1 High school teachers try to assess oral language skills. Vast majority of participants ($n=46$) answered ‘yes’ when they were asked whether they assess oral language skills or not in KOSAQ (*Item – 5*) as shown in *Table 8* below:

Table 8

Participants Who Claimed to Assess Oral Language Skills

	Frequency	Percent	Valid Percent	Cumulative Percent
Do you assess oral language skills?	Yes	46	92	92
	No	4	8	100
	Total	50	100	100

Majority of the participants of this thesis seem to be including the assessment of oral language skills in their teaching process as is evident within the data collected in KOSAQ. In addition, interviewees were asked whether they assessed oral language assessment in their classrooms or not, and 5 out of 7 answered as ‘yes’. T6 said that they make skills exams and they embed speaking in those exams as follows:

Yes, we do. We make skills exams and we include speaking within skills exams (T6, Feb 23, 2018).

It is evident that high school teachers try to assess oral language skills in their classroom; whereas, many of them carry it out due to the regulation of the ministry of education for the past two years as was also mentioned previously in this chapter. T4, for example, stated that she has been assessing oral language skills for the last two years, which coincides the time that the new regulation about it was introduced:

Yes, I do. I have been assessing for the last two years (T4, Feb 19, 2018).

Similarly, T5 mentioned that she is “currently” assessing speaking which indicates that she started to make an assessment for oral language skills due to the regulations of the ministry of education as shown in the following excerpt:

Yes, I am currently assessing speaking. I use different activities for it (T5, Feb 21, 2018).

Although high school teachers have challenges in the assessment process and many of them try to do it in order to fulfil the requirement, they give a place to oral language assessment in their classrooms. Some of high school teachers, on the other hand, completely set it aside for various reasons. Two interviewees and some participants mentioned that they did not assess speaking mentioning a number of reasons for that. T2, who works at a vocational high school, claimed that student level is not appropriate in her current context:

I was at an Anatolian high school before I came here. I had a foreign language class. I was more active there. When I came here, I was disappointed. We cannot teach speaking here, because students cannot form sentences. Students do not have the basic knowledge of English and we have to start all over again. They can understand, but they cannot speak (T2, Feb 9, 2018).

T3 also said that he did not assess speaking skill when asked and, like T2, he said he used to make an assessment in his previous context. He explained that he gives a final grade for oral language skills. It is clear that, he does not perceive what he does as assessment, since probably he does not use rubrics for assessment criteria as seen in his following statements:

I give a kind of final grade. I don't conduct an exam for it. We used to do such exams when I was working at foreign language intensive high school. There were prep classes and they had speaking lessons. We used to assess speaking, then (T3, Feb 14, 2018).

Likewise, there are similar statements of the participants indicating that they do not carry out oral language assessment in KOSAQ. The stated reason is usually the disadvantage of high school teachers' contexts. The following statement is taken from T40's answer for *Item – 42*. T40, who also works in a vocational high school, referred to the disadvantage of his context as a reason for not carrying out speaking assessment just similar to T1 and T3:

I usually don't carry out speaking assessments, because I teach at a vocational high school which has a very poor quality both in terms of students' capacities and school's low standards (T40, Feb 2018).

It is apparent that improving techniques and designing oral language assessment activities and tasks appropriate to student level is an important deficiency in high school teachers' knowledge of oral language assessment. Other reasons

mentioned by the participants were limited time and large classes as is shown in the excerpt from T29 below:

I cannot conduct speaking tests, so I cannot assess their oral language. Because the population of the class is too much, so I have no time to assess speaking skill (T29, Feb 2018).

In brief, findings from both KOSAQ and interviews indicate that majority of high school teachers try to do something to assess oral language skills in their classrooms; however, some of the participants, especially the ones working in vocational schools, cannot make an assessment because of the disadvantages of their contexts such as students' level, class size and inadequate class hours. In addition, some high school teachers, who assess oral language skills in their classrooms, carry it out botchily just because they were required to do.

4.2.2.2 The frequency of high school teachers' oral language assessment.

Item – 6 in KOSAQ was about the frequency of their assessment of oral language assessment. The ones who answered 'No' in *Item – 5* left this item blank; so, 92 % of the participants answered this question. 32,6 % of them answered 'Once', 30,4 % answered 'More than three times' and 28, 3 % of the teachers answered 'Twice' as shown in *Table 9* below. Three times was not popular among teachers.

Table 9

Frequency of The Participants' Oral Language Assessment

		Frequency	Percent	Valid Percent	Cumulative Percent
How many times do you assess oral language skills?	Once	15	30	32,6	32,6
	Twice	13	26	28,3	60,9
	Three Times	4	8	8,7	69,6
	More than three times	14	28	30,4	100
Total		46	92	100	
Missing	System	4	8		
Total		50	100		

Participants answered once, twice and more than three times in a semester almost equally when they were asked about the frequency of their oral language assessment. Very few participants answered three times, probably because they either give a final grade at the end of each semester by means of a summative

assessment or they make an assessment 1 or 2 times each month which corresponds to 2 or 4.

4.2.2.3 How high school teachers assess oral language skills. When high school teachers were asked about the way they assess oral language skills in their classrooms in *Item-42* in KOSAQ, 36 the teachers (72 %) gave a brief information about it. 14 participants (28 %), on the other hand, left the related item unanswered (*see table 10*) although they answered *Item - 5* (Do you assess speaking skills in your classrooms?) as ‘Yes’. This might simply indicate that they glossed over the questionnaire item as they did not want to write; or, it might also be because of a lack in their knowledge of oral language assessment or a deficiency in their practice to apply what they know in practice.

Table 10

Participants Who Answered How They Assess Oral Language Skills

		Frequency	Per cent	Valid Per cent	Cumulative Per cent
How do you assess oral language skills?	answered	36	72	72	72
	not answered	14	28	28	100
	total	50	100	100	

The answers of the participants to *Item 42*, which was an open-ended question asking about the way that high school teachers assess oral language skills and the transcribed answers to *Question 3* in semi-structured interviews were analyzed by pattern-coding and majority of high school teachers were found to assess oral language skills. The data also provided results about the details of current practices of high school teachers such as assessment criteria, assessment techniques and tasks they use in the process of oral language assessment, as discussed in the following categories.

4.2.2.4 Assessment criteria that high school teachers use. When the participants were asked about the assessment criteria they used while carrying out oral language assessment, majority of them (63 %) mentioned they both use ‘outside resources’ and ‘their own criteria’ when 21,7 % answered they develop their own criteria. It seems teachers need to revise the criteria from outside resources to fit in their contexts as shown in *Table 11* below:

Table 11

Sources of Oral Language Assessment Criteria

		Frequency	Percent	Valid Percent	Cumulative Percent
Assessment criteria	I use outside resources	7	14	15,2	15,2
Assessment criteria	I develop my own criteria	10	20	21,7	37
	Both	29	58	63	100
	Total	46	92	100	

4.2.2.4.1. *Intelligibility instead of accuracy.* In semi-structured interviews, *Question 3* sought for whether high school teachers' make oral language assessment or not and what are their current practices of oral language assessment. The ones who answered 'Yes' were directed three more questions the first of which was about the criteria they used in oral language assessment. Three of the interviewees mentioned that they pay much more attention to whether students can give the message that they want or not instead of grammar accuracy. T7, for example, explicitly said that she does not give emphasis to grammar and she tries to improve students' public speech by assessing the intelligibility of their messages as follows:

I don't pay attention to grammar, but I care about whether they can express their ideas, well...I mean, instead of grammar, I assess whether students can give their message to their friends. Indeed, I have been trying to develop their public speech skills rather than measuring the correct grammar. I do not care much about it, lately (T7, Feb 23, 2018).

Similar statements were also found in the answers of the open-ended items in KOSAQ. T30's following statement supported the idea that grammar is not prioritized in practice of oral language assessment, although it was mentioned by most participants in their assessment criteria aspects:

To speak with better ease means that grammar is sometimes left behind, as long as that is understood, is of utmost importance (T30, Feb 2018).

It is evident with the statements of the participants above that, some high school teachers are conscious about the importance of intelligibility of students' speech, rather than its being grammatically accurate in the assessment of oral language skills.

4.2.2.4.2. *Vocabulary is important.* Vocabulary use was one of the highlighted skills as many participants mentioned it in their assessment criteria in KOSAQ.

Likewise, 4 out of 7 interviewees said that vocabulary is important as they were talking about their assessment criteria. T1, for example clearly expressed vocabulary is important to her as assessment criteria, as is seen in the following excerpt:

Grammar is not that important. I do not attach to it much, lately. I pay more attention to students' speaking fluently. [...] Vocabulary might be a little more important... (T1, Feb 8, 2018).

As is clear in the excerpt above T1 emphasizes fluency in oral language assessment in her classrooms. She also mentioned that she did not care much about grammar which supports the idea that in the assessment of oral language skills grammar must be left aside. She added that vocabulary is more important. It indicates that she emphasizes vocabulary use in her assessment process, too. T7 is another interviewee who said vocabulary is important explicitly in her answer to the question about oral language assessment criteria. She also mentioned the importance of repeated use of vocabulary for learning process as in the following excerpt:

...And vocabulary is also important. For example, in one panel, I learnt that a student must encounter a word at least 6 times to learn it... (T7, Feb 23, 2018).

These findings from the interviews indicate that vocabulary is given priority as a component of high school teachers' assessment criteria in oral language assessment in their current practices. Vocabulary as assessment criteria was also proved to be beneficial by empirical studies in the literature (Iwashita et al., 2008; Brown et al., 2005). Some high school teachers indicated their insight for the importance of vocabulary enhancement for improving speaking skills in the current study.

4.2.2.4.3. Fluency, accuracy and content. Majority of the participants stated that they use these three oral language assessment criteria in *Item – 42*. These three criteria were usually mentioned in the same order in most teachers' statements as seen in the quote from T25 below:

In our practice (skills) exams, we have a speaking section. We evaluate their fluency, accuracy, use of English and content when they are speaking on topic we gave (T25, Feb 2018).

Two of the interviewees also claimed that they use these three assessment criteria for oral language assessment when they were asked about their assessment criteria. T6 mentioned them in the first place and added vocabulary, interactive skills and pronunciation as seen in the following quote:

Fluency, accuracy, content, vocabulary and interaction to assess their interactive skills; we also have pronunciation. There is one more, but I cannot remember now (T6, Feb 23, 2018).

Fluency, accuracy and content seem to be the most frequently used assessment criteria as many of the participants mentioned. The reason for most participants mentioning these three criteria together might be due to their presence in outside sources, which high school teachers make use of, in this way.

4.2.2.5 Oral language assessment tasks that high school teachers use.

4.2.2.5.1 *Oral presentation.* Clear majority the participants who answered Item – 42 in KOSAQ ($n=18$) mentioned that they have their students make oral presentations. With participants' statements, it seems clear that they usually determine a number of topics in advance, ask students to choose among those topics and give some time to get prepared. Finally, they have students make a few minutes oral presentation on that topic as described by T37 in the following quote:

We have the students pick up 2 topics out of 5 topics and speak / make a presentation on that topic in about 5 minutes. Then we assess it on the speaking assessment rubric (T37, Feb 2018).

Prepared speech on a topic seems to be the most frequently used task in oral language assessment. Consistently, majority of the interviewees ($n=4$) stated that they use prepared oral presentations for the assessment of oral language assessment. The following excerpt is taken from T7's speech, in which she gave details about their assessment process. It is more or less similar to all participants of this study.

What I do... I give some topics on which students develop their critical thinking. I choose the topics that enable them both they can express their ideas and practice in English. I give some time for preparing, or I give the topics a few days ago and give a deadline which is usually Friday. I let them take small notes to have a look. I give them choices and let them choose whatever they want to do. I also give time. I give 2 minutes each... (T7, Feb 23, 2018).

Prepared oral presentations seem to be the most popular assessment tasks among high school EFL teachers. Probably, it is the only known way of assessing speaking skills by a great number of high school teachers as very few teachers referred to other tasks.

4.2.2.5.2. *Role-playing.* The second most popular technique was role-playing among the participant teachers which is considerably less popular compared to oral presentations, though. 6 participants stated that they use role-playing as assessment

task in Item–42, in KOSAQ. T22 is the one of the participants who spoke of various activities such as interviews, dialogues or role-playing:

I generally carry it out by using role-playing, interviews and dialogues among students (T22, Feb 2018).

T1 was the only interviewee who mentioned role-playing. She said that she gives a place to role-playing activities through the end of the class time as shown in following excerpt:

... Especially, during the class time, through the end of the class, if we have time- something like creating a dialogue related to the topic that they have just learnt, instead of giving it as homework, for example, or a kind of game that they like and role-playing it in front of the class and I see that it is efficient... (T1, Feb 8, 2018).

Role-playing was one of the tasks that high school teachers claimed to use in oral language assessment even though it was not as popular as prepared speech. It was also the only student-student interactive activity that participants mentioned to use as an oral language assessment task.

4.2.2.5.3. *Picture description.* Picture description seems to be one of the task types that is used by a number of high school teachers as mentioned by the participants of this study. T15 was one of the participants who stated in KOSAQ that they used pictures as a speaking assessment activity.

We decide on a subject or we show some pictures to the students and want them talk on that subject or the picture. Sometimes, we want them to do role-playing (T15, Feb 2018).

Similarly, two of the interviewees (T1 and T6) spoke off picture description as an oral language assessment task they used. T1, for example, said that she utilized smartboard to show photographs or videos for picture description task:

...We use pictures. We generally use smartboard. I sometimes show a photograph or a video at the beginning of a class or a topic and I try to make them speak about it... (T1, Feb 8, 2018).

T6 was the other interviewee who claimed that she uses picture description. She said that students were familiar with the pictures as they had carried out similar activity before. It seems they repeat similar activities using the same materials with the same students. This might be because of limited access to materials; nevertheless, it also leads us to the fact that either they are not familiar with the

variety of speaking assessment activities, or they do not / cannot give a place to them in their current practices as seen in the excerpt given below.

We gave them pictures last semester, for example. We assigned 2 minutes for each student. They were familiar with the pictures, because we had done the similar kind of testing activity before. They chose the photographs themselves. I sometimes guide them to choose (T6, Feb 23, 2018).

To sum up, participants mentioned very few task types in their answers to the related questions. Prepared speech on a topic was most frequently used task type. Role-playing and picture description were other two tasks mentioned by the participant teachers. Role-playing was the only activity that includes student-student interaction to a degree, and it was mentioned by only 6 of the participants. High school teachers do not seem to prefer student-student interactive tasks probably because of the challenges in large classes or they do not have adequate knowledge on them.

The second research question was about current practices of high school teachers. The data about current practices of teachers were obtained through *Item - 5, 6, 7 and 42* in KOSAQ. In addition, semi-structured interviewees provided a large amount of data about high school teachers' current practices. The results, which were presented under categories comprised of pattern-coding above, indicate that majority of the participants ($n=46$) answered 'yes' when they were asked whether they assessed oral language assessment in their classrooms or not; however, some of the participants ($n=7$), who answered 'yes' in *Item - 5*, left *Item - 42* blank. This means that, the number of the teachers who really make an assessment might be less. The participants who assess oral language skills stated that they use the criteria taken from outside resources together with their own criteria. Majority of participants mentioned 'fluency', 'accuracy' and 'content' as their assessment criteria; on the other hand, 4 of 7 interviewees and some respondents in KOSAQ stated that grammar accuracy is not that important while making an assessment and they expressed that they care more about intelligibility of students' speech, instead. Vocabulary was also prioritized by high school teachers.

In conclusion, the participants try to assess oral language skills as is required by the ministry of education. They mostly use the criteria from both outside resources and their own criteria. Most of them have vocabulary, fluency, accuracy and content in their rubrics; whereas, they give emphasis to intelligibility more than

accuracy and they value vocabulary in their assessment. As for assessment techniques and tasks, oral presentation is most frequently used task by participants. Role-playing and picture description two are other tasks which are utilized by high school teachers in oral language assessment. High school teachers prefer to use presentation tasks rather than student-student interactive tasks.

4.2.3 Findings of Research Question 3: What are the factors in teachers’ contexts that have influence on speaking assessment? The third research question was about factors in high school teachers’ contexts that influence their oral language assessment. Section Three in KOSAQ involved items (36 – 41) which provide data for the answer of the third research question. Besides, the interviewees were directed a question to investigate whether they have any challenges during the assessment of oral language assessment, and what the challenges they encounter were.

The participants were provided six items (36 – 41) in Section Three, in KOSAQ to guide them about possible factors that might influence their oral language assessment, and they were asked to tick relevant ones to their own contexts and give reasons in the spaces provided below each item. First, the results from the questionnaires were presented in *Table 12*. Then, these results were compared and contrasted with the qualitative data which were gathered from participants’ explanations in KOSAQ and interviewees’ statements. The factors, which influence high school teachers in the assessment process of oral language assessment, were ultimately given under 8 titles comprised of pattern-coding together with interpretation of quantitative data. They are class size, curriculum objectives, having standardized / prescribed criteria, lack of time, student level, student reluctance, subjectivity and lack of teacher training.

Table 12

Influence of Factors in Participants’ Contexts

Items	<i>Which of these following factors do you think have an influence on your speaking assessment?</i>	Answered <i>f</i>	Percentage %	Not answered <i>f</i>	Percentage %
36	Class size	45	90	5	10
37	Curriculum objectives	27	54	23	46
38	Lack of support by more experienced teachers	6	12	44	88

Table 12 (cont.d)

39	Prescribed rules about what assessment criteria to use	18	36	32	64
40	Lack of standardized oral language assessment criteria	18	36	32	64
41	Lack of teacher training for learning how to make oral language assessment	17	34	33	66

Table 12 indicates the results from teachers marking the factors in the related items in KOSAQ. As is seen in the table, vast majority of the participants (90 %) stated ‘class size as a factor influencing their decision making in the oral language assessment process, in their classrooms. 27 participants out of 50 also mentioned curriculum objectives. This means that 54 % of high school teachers are influenced by curriculum objectives. Which was interesting about the participants’ marking in the following items that, the number of teachers who answered *Item – 39* is the same as the ones who answered *Item – 40* ($n=18$). This means that they have different viewpoints about having prescribed criteria or being free in determining their own criteria. These answers were interpreted with the participants’ statements later in this section. $N=17$ teachers, (34 % of total participants) marked *Item – 41* which means that teachers also suffer from lack of teacher training.

4.2.3.1 Class size. As seen in *Table 12*, class size generates the biggest influence on high school teachers’ oral language assessment practices. The statements of the participants in KOSAQ and semi-structured interviews are also consistent with this result. In the following excerpt, T12 mentions the difficulty in creating an atmosphere for an authentic language use in class due to large numbers of students:

Making a real-life speech is very difficult in crowded classes. Unless using language, you can’t learn all skills appropriately (T12, Feb 2018).

44 % of the participants ($n=22$) also stated that they are influenced by the large samples in their classes in *Item – 36* in KOSAQ. The previous section was about current practices of high school teachers and the findings indicated that high school teachers do not tend to use tasks or techniques which include student-student interaction. They prefer oral presentations more which require students to make a

prepared speech in front of the class. This might be because of having large samples of students in their classrooms. Following quote supports the interpretation of this result as having crowded classes is the main cause of avoiding learner-learner interactive tasks:

In large classes it is not always possible to carry out such activities as role-playing, pair work, interviews... (T27, Feb 2018).

It is evident that teachers do not know how to adapt interactive activities or tasks to crowded classes. 5 out of 7 interviewees also stated that having large classes hinders their assessment process. T5 said that she cannot carry out oral language assessment as she has planned because of the difficulties she encounters laying emphasis on large class size, and she frankly mentioned that her knowledge is not adequate to deal with such challenges and carry out oral language assessment healthily.

...I determine some criteria, but I cannot stick to them while doing the assessment because of large class sizes and physical inappropriacy of the classrooms and many other factors like that. I don't think we teach or assess speaking healthily (T5, Feb 21, 2018).

As seen in the statement above, high school teachers have difficulty in making an oral language assessment in large classes. Some high school teachers completely abandon oral language assessment due to having large classes and limited time as shown in T29's excerpt below:

I cannot conduct speaking tests, so I cannot assess their oral language. Because the population of the class is too much, so I have no time to do it (T29, Feb 2018).

The results from the related items in Section Three in KOSAQ and the in-depth analysis of the data gathered from high school teachers' expressions in both *Item – 36* in KOSAQ and semi-structured interviews revealed that teaching to large classes, which is generally 30 – 40 in state high schools in Turkey, generates the biggest challenge for the assessment of oral language assessment.

4.2.3.2 Curriculum objectives. Curriculum objectives gets the second biggest lot as 54 % of the participants expressed that they were influenced by curriculum objectives in their practice of oral language assessment, in *Item – 37*. This indicates that high school teachers are concerned with curriculum objectives while carrying out oral language assessment in their classrooms. Participants' statements in this question provide us with details about their concern. T15, for example, claimed that

he takes curriculum objectives into consideration while designing speaking assessment:

Curriculum objectives are important when I design speaking assessment (T15, Feb 2018).

Another participant, T26, explained that she is trying to include speaking in her classes to fulfil the curriculum objectives.

We, as teachers of English, try to fulfil the curriculum objectives by using speaking in our lessons all the time within the other skills (T26, Feb 2018).

Majority of high school teachers think that curriculum objectives affect their way of making oral language assessment; however, they have different perspectives and some contradictory ideas about the influence of curriculum objectives. T14's statement indicates that she thinks curriculum objectives hinder her oral language assessment process as they constitute an obstacle:

Curriculum is an obstacle (14, Feb 2018).

T29 has parallel ideas about curriculum objectives and thinks that teachers are restricted by curriculum objectives as is seen in the following quote:

Teachers should have more flexible and have more freedom about the curriculum (T29, Feb 2018).

Other participants, contrarily, think that curriculum is a guide that might be helpful in oral language assessment. T19, for example, thinks that curriculum objectives guide her to design activities as is shown in the quote below:

Objectives give us a way to lead the activities accordingly (T19, Feb 2018).

Similarly, T33 states that both her students and she, herself, feel more comfortable when they are provided with the topics. She seems not to prefer to be troubled by searching for appropriate topics and designing materials accordingly:

I use our own curriculum subjects, especially our coursebook topics as the students feel more comfortable (T33, Feb 2018).

T3 answered *Item- 39* in the following way which means that does not have sufficient knowledge on curriculum objectives and he needs to get informed more them:

What is our objective? Our goal, common goal? (T3, KOSAQ, Feb 2018)

To sum up, it is clear with the statistical results of KOSAQ and the expressions of participants above that, although high school teachers have varying viewpoints

about the pros and cons of following curriculum objectives, majority of high school teachers are influenced by the curriculum objectives and most of them think that they have to take curriculum objectives into consideration while making an oral language assessment.

4.2.3.3 Having standardized / prescribed criteria. Prescribed rules about the assessment criteria and lack of standardized criteria have the same rates as influencing factors in teachers' oral language assessment process. The researcher interpreted two items, (*Item – 39* and *Item – 40*), in Section Three in the same pace as both related to having prescribed criteria for oral language assessment. This is another factor that high school teachers have contradictory ideas, as was for curriculum objectives. Two of the interviewees stated that standard criteria must be determined for oral language assessment. T3, for example, expressed the situation very clearly as is seen in the following excerpt from his answer when he was asked about his training needs:

First of all, objectives must be set clearly. I mean, what level in speaking that we are supposed to bring these students up. And, the criteria by which we can make more objective assessment might be included. Determining the criteria according to the objectives... (T3, Feb 14, 2018).

The statements of T3 about the objectives indicates that he believes that standardized criteria need to be determined by which all teachers assess oral language skills along the same line. He also mentioned that the objectives must be common for all high schools for maintaining equal conditions for each in his answer to *Item – 40* in KOSAQ:

What should I look for? And, other teachers must assess with the same criteria to be equal and to reach our goal (T3, KOSAQ, Feb 2018).

T5, similarly, mentioned her concern for objectivity while her training needs as follows:

We all determine our criteria for ourselves. There are a lot of differences. I don't think this is objective. We need to be enlightened about the criteria for oral skills (T5, Feb 21, 2018).

It is apparent that many high school teachers believe that they need to assess with the same criteria in order to provide fairness. There are similar statements in KOSAQ by other participants as one of them is shown below:

There should be a common way of assessing for every teacher (T48, Feb 2018).

Other participants, on the other hand, think that prescribed criteria hinder assessment process for oral language assessment. They mention the drawbacks of assessing via same criteria. T12 stated that students are in varying levels and meeting individual needs is difficult with standardized criteria in her answer to *Item – 39* in KOSAQ:

Students' levels are different, and it is difficult to teach language according to educational system and support individual needs (T12, Feb 2018).

T1 took the issue from another point of view and meant that it might limit teachers' decision making on the criteria appropriate to their own contexts as follows:

Prescribed rules might limit a teacher's decisions on the criteria that s/he uses (T1, Feb 8, 2018).

The data collected from *Item – 39* and *Item – 40* in KOSAQ reveal that teachers have different ideas about having standardized criteria, and the number of participants who are for or against having prescribed criteria is equal, too. The data gathered from interviews provide similar results as interviewees also have various viewpoints about this factor.

4.2.3.4 Lack of time. Having limited time was one of the major challenges in oral language assessment process as was mentioned by 3 out of 7 interviewees. T1 for example said that having limited time with large classes restrains her from making an extensive assessment for oral language skills:

We cannot make an extensive assessment in our classrooms because of large class sizes or limited time... (T1, Feb 8, 2018).

T3 spoke of many challenges that hinder his oral language assessment including inadequate class hours and limited time among them as follows:

...We have a lot of difficulties, like, in catching up with the curriculum, we have difficulties even in making written examinations, we have big class sizes, we have inadequate class hours, we are trying to teach and do a lot of things in a limited time we also overwork as we prepare at home and take students' assignments home to check and, we are paid very little... (T3, Feb 14, 2018).

It is evident that lack of time is perceived as an important problem in oral language assessment processes for many high school teachers. Some participants,

who filled in KOSAQ, mentioned that they do not make an assessment for speaking because of having limited time:

I cannot conduct speaking tests, so I cannot assess their oral language. Because the population of class is too much, so I have no time to do it (T29, Feb 2018).

As confirmed with the participants' statements above, limited time is one of the major factors that influence teachers' oral language assessment process which is mostly mentioned together with the hardship of having large class sizes. It hinders teachers' making an extensive assessment for speaking skills, or they wholly set it aside since they cannot spare time for it.

4.2.3.5 Student level. Student level was not placed in KOSAQ among the factors influencing teachers' oral language assessment processes; however, it was mentioned by both interviewees and other participants under different items in KOSAQ. Two of the interviewees stated that they were influenced by different levels of students while making an oral language assessment. T1, for example, said that there are great differences in students' level in her classes:

...We cannot arrange classes according to students' English level; so, we have different levels in our classes, like, I can say there are students from A1 to C1 in the same class. I need to prepare different things to assess their speaking. What can I do to assess different levels of students? These might be my training needs (T1, Feb 8, 2018).

T1 spoke off this issue while talking about her training needs as shown above. T2 claimed that she had difficulty in even teaching speaking skill because of low levels of students and, she desperately said that they cannot teach speaking in her context:

...We cannot teach speaking here, because students cannot form sentences. Students do not have the basic knowledge of English and we have to start all over again... (T2, Feb 9, 2018).

It is apparent that T2 have difficulty in teaching speaking to students with low level of proficiency, thus naturally she has drawbacks in assessing this skill in her classrooms. In KOSAQ, participants also have statements about student or class level as given in following excerpt from T5 and T46

Personal differences make the activity harder or easier for each student (T5, KOSAQ, Feb 2018).

While making speaking assessment, class level is important (T46, Feb 2018).

It is obvious in the excerpts above that, mixed and /or low levels of English proficiency of students is another issue that high school teachers have difficulty while making an assessment for oral language skills. It is evident that, teachers have challenges in designing activities for low level of students, or they have difficulty in dealing with various levels of students while designing and using their oral language assessment activities.

4.2.3.6 Students' reluctance to speak. Students unwillingness or anxiety to speak was one of the challenges that was found within the data collected from interviews or some items in KOSAQ. This was not one of the factors included ready in KOSAQ, but some participants wrote in their statements under the items that have spaces for teachers' comments. Three interviewees mentioned this factor in the first place when they were asked whether they had any challenges in their contexts while making oral language assessment. T1 believed that the real cause of this reluctance is lack of self-confidence as shown below:

Some students have a negative attitude towards English and some students have self-confidence problem. These might have negative effects. I mean, encouraging the students for speaking and making them believe that they can express themselves takes a long time. In short, we mostly have problems with the students' attitudes, point of view, reluctance which is caused by lack of self-confidence and feeling of inadequacy (T1, Feb 8, 2018).

Two of the interviewees (T4 and T5) clearly mentioned the reason of the reluctance of students to speak as peer pressure as is seen in the following excerpt from T4's interview below:

The most frequent problem is they make fun of each other. I try to stop this, and I try to make them speak even if they make mistakes [...] To be honest, the students are not very willing to speak. The most frequent problem I have observed is the fear of making mistakes (T4, Feb 19, 2018).

T5's following statement is parallel to T4's:

Yes, there are some other challenges. For example, they mock at each other a lot. When a student makes a mistake, others immediately attack on him and laugh at him. They are competing each other because of puberty and this is a big challenge for us (T5, Feb 21, 2018).

As seen in the excerpts above teachers have to deal with the challenge of students' unwillingness to speak due to various reasons such as lack of self-confidence, fear of making mistakes, peer pressure etc. In KOSAQ, there are also statements related to this problem:

Before I start to speak in English, I just need to courage my students. To me, it is the main problem of lack of speaking English (T10, Feb 2018).

Both interviewees statements and other participants' utterances indicate that high school teachers are influenced by unwillingness of students in their oral language assessment processes.

4.2.3.7. Subjectivity. Three of the interviewees (T3, T5 and T6) expressed that they feel uncomfortable with assessing students' performances alone making their own decisions as they think it is not objective. This means that high school teachers mind their assessment to be objective; nevertheless, they have difficulty in ensuring it as is clear with the statement of T6 given below:

Do I have difficulties? Yes, I do. Like, for example, I observe them speaking and evaluate whether they speak fluently or whether they use the words they have learnt, and, how can I say? It looks as if everything is up to my own humble decision; but, I try to give high grades as much as I can (T6, Feb 23, 2018).

T5 also mentioned the same challenge uttering explicitly "not objective". She referred to her training need on this topic saying that she needs to be enlightened on assessment criteria, as seen in the excerpt below:

We all determine our criteria for ourselves. There are a lot of differences. I don't think this is objective. We need to be enlightened about the criteria for oral skills (T5, Feb 21, 2018).

It is evident that, subjectivity is a significant factor that high school teachers concern, and they perceive difficulties in their practices in providing an objective oral language assessment in their classrooms.

4.2.3.8 Lack of teacher training. This factor was placed in KOSAQ and 34 % ($n=17$) of participant teachers stated they are influenced by lack of training on oral language assessment in this item (*Item – 41*) in KOSAQ. In the following statement, for example, a participant states they have not taken training for it:

Teachers haven't taken any training on this subject (T37, Feb 2018).

Besides, high school teachers also clearly mentioned in various parts of the questionnaires that they need to get training for oral language assessment. Some of the statements presented within the quotes from KOSAQ below provide us with an insight for a major challenge; lack of teacher training that influences teachers' knowledge and practices of oral language assessment:

During seminars teachers might be given the methods on this (T32, Feb 2018).

We need to have teacher training seminars to improve our techniques of oral language assessment. I usually search on the Google and learn more about it (T26, Feb 2018).

Another participant (T30) used an idiom to explain her situation which means that she tries to make oral language assessment perfunctorily in order to fulfil the requirements, which was also found and presented in above sections and confirmed with other participants' statements:

If you are not shown, 'fake it, until you make it' situation (T30, Feb 2018).

The following quote from T3's answer to *Item – 41* support that high school teachers have challenges due to lack of professional training, as well, and they need in-service training on oral language assessment:

I don't know or forgot. I need to be reminded or taught (T3, KOSAQ, Feb 2018).

As is seen in the statements of respondents above, there are many teachers who think that they have difficulties because of lack of training on oral language assessment. Their training needs were analyzed in-depth and presented in the following section related to high school teachers training needs.

The third research question was about the factors in high school teachers' contexts that influence their oral language assessment processes. The findings from both quantitative and qualitative data indicated that the biggest challenge they perceive is having large class sizes. Majority of the participants note curriculum objectives as a factor that influence their decision making in speaking assessment, too; however, some perceive it as a positive effect stating that curriculum objectives guide them in the process of oral language assessment while others think that it is an obstacle as it retains their decision making in the assessment process. Another factor, which high school teachers have disagreements about is having standard criteria. The number of the participants who think they need to have standardized criteria in order to make objective assessment is equal to the number of the ones who think it might hinder making assessment considering individual differences and different levels of students. Lack of teacher training is also mentioned by a remarkable number of high school teachers ($n=17$) which comprises the 34 % of the total participants as a factor that influence their oral language assessment. Deriving from the statements of the

interviewees and participants who answered open-ended questions in KOSAQ, other factors influencing teachers' oral language assessment processes are determined as having limited time, working with students with low / various levels of English, student unwillingness to speak and subjectivity of oral language assessment process. Lack of support from more experienced teachers was answered by very few teachers since there are very few inexperienced high school EFL teachers in Beşiktaş and most have relatively long years of experience, themselves.

4.2.4 High school teachers' training needs for oral language assessment.

The aim of the fourth research question was to find out training needs of high school teachers for oral language assessment. The data for training needs of high school teachers were collected through TNAFOLAS, which consists of two sections. In the first section, there are questions about participants' training backgrounds and an open-ended question investigating their beliefs about what aspects of oral language assessment they still need training. The second section, which provided quantitative data about high school teachers' training needs, involves a 5-point Likert scale on which participants were asked to rate each item with a number from 1 – 5 to indicate the importance of the training aspects given in those items to the them. Qualitative data were gathered from the statements of high school teachers within semi-structured interviews and open-ended questions in both KOSAQ and TNAFOLAS. The results were also compared with the results in the previous sections; knowledge of speaking assessment, current practice of speaking assessment and the factors influencing teachers' oral language assessment and an overall interpretation was made at the end.

4.2.4.1 I need / I want training. In the first section of TNAFOLAS participants' training backgrounds and their beliefs on what aspects of oral language assessment they still need training were investigated. Majority of high school teachers (54%) answered *Item -1* in TNAFOLAS “No” which means they have not got any training on oral language assessment. The ones who answered yes were directed another question in item – 2; ‘how much they thought the training they got satisfied their needs in the process of oral language assessment. 69,6 % of the respondents who answered the 1st question as ‘yes’ answered ‘adequate’ in this item, and 13 % of them answered ‘*Limited*’ or ‘*Very limited*’ for the training they have got on oral language assessment as is shown in *Table 13*.

Table 13

Participants' Evaluation of Their Previous Training

		Frequency	Percent	Valid Percent	Cumulative Percent
If you answered 'Yes' for question 1, do you think it satisfied your needs in the process of assessment for oral language skills?	Very limited	1	2,0	4,3	4,3
	Limited	2	4,0	8,7	13,0
	Adequate	16	32,0	69,6	82,6
	Very adequate	4	8,0	17,4	100,0
	Total	23	46,0	100,0	
Total		50	100,0		

This means that vast majority of the respondents believe that they have not got an adequate training for the assessment of oral language assessment. Additionally, when scrutinized, it was found out that the training that some of the teachers stated they got was not directly related to the assessment of oral language skills. During the interviews, the participants, who stated that they got training on oral language assessment, were asked to talk about the details of the training they got, and their answers indicated that the content of the training was far disparate. T2, for example, answered "Yes" when asked whether she got any training on oral language assessment or not while being interviewed; however, she referred to different accents rather than oral language assessment while telling about the training that she got as shown in the following excerpt:

I went to a summer school in England for three months. I got a training for different accents in English, there (T2, Feb 9, 2018).

The training that T2 took seems to be more useful for improving her own oral language skills rather than her assessment skills, especially when her statement about not conducting an assessment for speaking or not teaching speaking in her context is recalled. It seems clear that, although some teachers mentioned that they got training on oral language assessment, the scope of the training they got does not directly pertain to oral language assessment; therewith it is evident that, the number of teachers who have not got any training on oral language assessment is probably more than it was found within the results of the statistical data.

Furthermore, lack of teacher training was mentioned as a challenge by high school teachers in the previous section; therefore, it was coded as one of the factors that influence teachers' oral language assessment. In addition, only two participants answered *Item -3* in TNAFOLAS as they do not need any training. As is consistent with the ratings in TNAFOLAS, majority of participants stated that they need / want to get a training on oral language assessment in their answers to open-ended items in KOSAQ and TNAFOLAS. T29 also mentioned that she is not provided with the opportunity for it, although she needs training for oral language assessment:

I need to have training for oral language assessment, but I have no opportunity (T29, Feb 2018).

Another participant, T26 drew our attention to another point; searching on the Internet for the improvements is not adequate for being competent in oral language assessment and she emphasized the need for training programs:

We need to have teacher training seminars to improve our techniques of oral language assessment. I usually search on the Google and learn more about it (T26, Feb 2018).

T27, likewise, perceives training as a necessity for keeping up-to-date about the improvements in oral language assessment:

We would like to learn more about what kind of improvements there are in the field of oral language assessment (T27, Feb 2018).

It is apparent in the statements of the participants above that high school teachers are conscious about their needs for professional training on oral language assessment; moreover, they are eager to get training to learn more about oral language assessment. The findings from the interviews indicate parallel ideas from the interviewees. All interviewees clearly mentioned that they would be willing to get training on oral language assessment, if one was designed for them as is seen in the following excerpt from T2's speech:

If there was something like training on speaking assessment, I would certainly get it (T2, Feb 9, 2018).

T3 also stated that he needs to refresh his knowledge about oral language assessment and it is clear in his statement that he is willing to get training on it:

I don't know, or I forgot. I need to be reminded or taught (T3, Feb 14, 2018).

To sum up, it is evident with the findings of this study that a great majority of high school teachers try to assess oral language skills in their classrooms either as

being aware of its necessity, or for the requirements of the ministry of education. However, it also seems clear with the results deduced from all data collected by means of KOSAQ, TNAFOLAS and semi-structured interviews that, although majority of high school teachers rated their knowledge of speaking assessment as ‘Good’ or ‘Excellent’ in KOSAQ, they have a lot of deficiencies in their knowledge and there are various shortcomings in their practices of oral language assessment. The results of TNAFOLAS revealed that high school teachers lack training on oral language assessment as a vast majority of them stated they have not got training on it. Participants also sincerely stated that they need training as shown in the excerpts above. All in all, high school teachers lack knowledge on oral language assessment and they need to be provided with in-service teacher training on it.

4.2.4.2 I need training on all aspects of oral language assessment.

Participants’ ratings for the importance of given topics to get training on are shown on Table 14 below. The items which were mostly rated as ‘Important’ and ‘Very important’ were taken into consideration in order to determine the areas that high school teachers need training most.

Table 14

High School Teachers’ Training Needs

1=Not important 2=Slightly important 3=Moderately important

4=Important 5=Very important

Items	Please indicate whether you think you need to get training on this topic.	1 <i>f</i> %	2 <i>f</i> %	3 <i>f</i> %	4 <i>f</i> %	5 <i>f</i> %
4	Theoretical framework for approaches to assessment of oral language performance	5 10	19 38	12 24	11 22	3 6
5	Highlights of oral language skill assessment in a language learning / teaching process	1 2	12 24	9 18	17 34	11 22
6	Benefits of assessing oral language performance	1 2,1	9 18,8	3 6,3	25 52,1	10 20,8
7	Deciding on main purpose of assessing oral language performance	1 2	5 10	11 22	23 46	10 20
8	Specifying objectives in assessing oral language performance		8 16	12 24	22 44	8 16
9	Assessment types for oral language skills	1 2	7 14	11 22	20 40	11 22
10	Determining assessment criteria for oral language performance	1 2	7 14	15 30	11 22	16 32
11	Deciding on what basic oral language skills to assess	1 2	6 12	8 16	20 40	15 30

Table 14 (cont.d)

12	Deciding on what sub-skills to assess	1	7	15	17	10
		2	14	30	34	20
13	Preparing tasks /test items for assessing oral language performance	6	5	23	16	
		12	10	46	32	
14	Deciding on ready-made tasks / test items to be used while assessing oral language performance	7	9	22	12	
		14	18	44	24	
15	Preparing rubrics for oral language assessment	2	5	6	18	19
		4	10	12	36	38
16	Deciding on ready-made rubrics to be used while assessing oral language performance	1	6	9	22	12
		2	12	18	44	24
17	Task /test types that are used in oral language assessment	1	4	7	20	18
		2	8	14	40	36
18	Samples of tasks, tests, rubrics that can be used in the assessment of oral language performance	6	7	22	15	
		12	14	44	30	
19	Evaluating / Rating oral language performance	6	7	21	16	
		12	14	42	32	
20	Interpreting scores	6	13	16	15	
		12	26	32	30	
21	Reliability and Validity issues	7	11	19	13	
		14	22	38	26	
22	Ethical considerations (giving equal chance to students, influence of personal differences, grader's subjectivity...etc)	2	8	8	16	16
		4	16	16	32	32
23	Washback on the classroom	6	19	16	16	
		12,5	20,8	33,3	33,3	
24	Common European Framework of Reference (CEFR)	2	6	14	16	12
		4	12	28	32	24
25	Relating my assessment to CEFR	2	5	15	19	9
		4	10	30	38	18

As is seen in *Table 14*, a vast number of the participants rated the aspects of oral language assessment as ‘*Very important*’ or ‘*Important*’ for their training in all items to varying percentages. Each item got rating of ‘*Very important*’ or ‘*Important*’ from 54% or more of the participants, except for *Item – 4* which is about theoretical framework for approaches to assessment of oral language performance. In TNAFOLAS, 10 items were rated as ‘*Very important*’ and 20 items were rated as ‘*Important*’ by at least 30 % of total participants. This indicates that, participants think that they need to get training on almost all topics given in the scale or they do not have any ideas about what they really need to know. In addition to that, some participants claimed straight that they need training in all aspects of oral language

assessment when they were asked on what aspects of oral language assessment they need training as seen in the following quotes from TNAFOLAS:

Every (T37, Feb 2018).

We, as teachers of English, need training in all aspects of oral language assessment to improve ourselves on this skill (T26, Feb 2018).

These two participants' statements support the findings of needs analysis scale as they openly stated that they feel the need to get an overall training on various aspects of oral language assessment in order to gain competence in making oral language assessment in their contexts.

4.2.4.3 I don't know what I need to know. As mentioned above, teachers were found to have training needs on almost all of the given aspects of oral language assessment. It was also forespoken by the researcher that, some of the teachers might have not known what they really need to know for carrying out a successful oral language assessment. This interpretation was supported by two interviewees (T3 and T4) as they stated that they do not know what they need to learn about. oral language assessment They claimed they cannot tell about their exact training needs for this reason. T3 clearly mentioned that he does not know what he needs to learn when he was asked about his training needs for oral language assessment:

I don't know much about this issue. I haven't searched about it; so, I don't know what exactly I need to know [...] I have my own ideas, but they do not have an academic basis (T3, Feb 14, 2018).

T3's statements indicate that his training needs include a wide range of aspects of oral language assessment such as speaking assessment constructs, criteria, techniques and tasks; however, he cannot tell these areas precisely since he lacks the knowledge of what dimensions or components might oral language assessment involve. Consistently, T4 openly stated that she cannot tell the aspects that she needs training as she lacks the knowledge about it:

I don't know. To tell the truth, I don't know which parts of it we should get training very well. Personally, I think I have to learn many things about how we can include speaking in our classes, but I don't know much about assessment, so I cannot tell the areas that I need training...(T4, Feb 19, 2018).

T4 mentions the need for training about the instruction of speaking skill, however she is not aware of the need to assess it as was also previously found in her statements for the necessity of oral language assessment. She indicates a need for a

bottom-up training on oral language assessment including the significance and benefits of it. T17's statement in his answer to *Item – 3* in TNAFOLAS below summarizes the issue for a considerable number of high school teachers:

It is not up to me to decide (T17, Feb 2018).

As is evident in these statements, some high school teachers might not be aware of deficiencies in their knowledge of oral language assessment and what they need to learn about it. This might also be the main cause of high ratings in KOSAQ, which was inconsistent with the results in TNAFOLAS, as majority of participants claimed that they have not received a sufficient training on oral language assessment and they stated they need training on many aspects of oral language assessment by ticking '*Very important*' or '*Important*' in most items in the scale and / or stating directly they need training.

4.2.4.4 Significance and benefits of assessing oral language skills. *Item – 6*, which is about 'benefits of assessing oral language performance', was one of the items that was mostly rated as '*Very important*' or '*Important*' (73 %). This result is also concurrent with qualitative findings in Section Two, where participants mentioned that they assess oral language skills just for they were required, as is reminded with the following excerpt from T4's interview:

We are required to do this, so we are trying to do this. If you ask me, if we weren't required, we would not do [...] I don't know whether it is necessary or not. To be honest, I don't think it is that necessary (T4, Feb 19, 2018).

Unless high school teachers have the knowledge of significance and use of assessment process in teaching of speaking, it is not wise to expect them to make an effective oral language assessment in their classrooms. As was noticed in their statement given in this section and the previous sections related to this issue, they perform perfunctory activities to pretend to do it. Therefore, it is evident that, high school teachers need to get informed about the benefits of making an assessment for oral language assessment.

4.2.4.5 Defining what skills to assess. *Item – 11*, which was about 'deciding on what basic oral language skills to assess' was rated as '*Very important*' or '*Important*' by 70 % of participants. 15 participants (30 %) rated this item as '*Very important*' and 20 participants (40 %) rated it as '*Important*' which means that teachers also need training on how to decide on what basic oral language skills to assess. In KOSAQ, teachers gave high ratings to their knowledge of 'speaking

assessment constructs'; nevertheless, it is apparent that they need to be trained for how to decide which of those speaking assessment constructs to use in the assessment process of oral language assessment. T3 mentioned it clearly in his answer about his training needs:

...I don't know what exactly I need to know, but it would be better if we could get training on how we should assess, what we should assess, for what we should say "good or not good", how we should guide students... (T3, Feb 14, 2018).

In T3's answer to *Item – 40* of T3 there is also a critical sentence which reveals the situation from high school teachers' point of view:

What should I look for ... (T3, KOSAQ, Feb 2018).

It is quite clear with the findings above that teachers need training on deciding what skills to assess.

4.2.4.6 Specifying objectives. In *Item – 7*, in TNAFOLAS, 33 participants (66 %) stated that deciding on the main purpose of oral language assessment is *Very important* or *Important* for their training needs. Similarly, in *Item – 8*, 30 participants (60 %) responded as '*Very important*' or '*Important*' for their training needs on specifying objectives in assessing oral language performance.

High school teachers were found to have contradictory ideas about being dependent on curriculum objectives with their statements in KOSAQ. Although a considerable number of high school teachers do not want to be tied up with curriculum objectives, it is evident with the results above that they have shortcomings in determining their objectives for oral language assessment and they need training on it. In his following statement T3 clearly explains that he needs for training on objectives of oral language assessment:

First of all, objectives must be set clearly. I mean, what level in speaking that we are supposed to bring these students up. And, the criteria by which we can make more objective assessment might be included. Determining the criteria according to the objectives and applications that we can assess these more practically. If we get training on these, it would be fantastic.... (T3, Feb 14, 2018).

As is clear in the results of *Item – 7* and the statements of T3 that teachers need training on how to decide on the main purpose and objectives for oral language assessment.

4.2.4.7 Techniques and tasks. *Item -13*, which is about 'preparing tasks /test items for assessing oral language performance was rated as '*Important*' by 23

participants, which is nearly half of total number of participants (46 %). This item was also rated as *'Very important'* 16 participants, which is 32 % of total number. Total ratings of *'Very important'* and *'Important'* in this item corresponds to 78 % ($n=39$). This item got the top rating from high school teachers as the most important topic for their training which is consistent with participants' ratings in Section Two in KOSAQ (See Table 6). It was related to knowledge of speaking assessment techniques and tasks. Although participants rated their knowledge as *'Good'* in general in KOSAQ, in this part, more participants rated their knowledge as *'Moderate'* or *'Not good'* compared to the other parts. This might be interpreted as high school teachers need training on preparing tasks /test items for assessing oral language performance most. Within this context, *Item – 14* was rated as *'Very important'* or *'Important'* by 34 participants (68 %) which indicates that high school teachers not only have shortcomings in preparing test / task items for the assessment of, but also, they have trouble in deciding on ready-made tasks or test items to be used while assessing oral language performance and they also need training on this topic. Consistently, *Item – 17*, which was about 'task /test types that are used in oral language assessment, was rated as *'Very important'* by 18 participants which constitutes 36 % of total participants, and it was rated as *'Important'* by 20 participants which is 40 % of the total. This item got the rating of *'Very important'* or *'Important'* by 38 participants in total which is 76 % of total participants. This means that task /test types that are used in oral language assessment was the second most important topic that high school teachers need training. These results reveal that high school teachers also need to get training about task /test types in order to be able to prepare or choose tasks /test items for assessing oral language assessment.

As is, again, consistent with the results from Likert Scale in TNAFOLAS, 13 of the participants claimed that they need training on techniques and tasks that they can use for oral language assessment in *Item 3* in TNAFOLAS, as is indicated in the following quote from T25:

To improve ourselves about types of tasks, techniques (T25, Feb 2018).

The findings of the semi-structured interviews were also consistent. Two interviewees stated that they need training on techniques or tasks. T1 said that she'd better learn about techniques and methods of oral language assessment referring to her training needs:

I think it would be better if I learn more about which methods and techniques that I can use, Whether I do the right things while rating the students and how can I do it correctly... (T1, Feb 8, 2018).

Additionally, *Item – 18*, which is about samples of tasks, tests and rubrics that can be used in oral language assessment, was rated as ‘*Very important*’ by 15 high school teachers, and as ‘*Important*’ by 22 high school teachers. 74 % of the participants ($n=37$) rated it as ‘*Important*’ or ‘*Very Important*’ in total, in accordance with ratings in *Item – 13, 14* and *17*. This means that high school teachers need to be introduced with samples of tasks and techniques for oral language assessment in a training provided for them, as well. T6’s statement supports the need for being introduced or provided with the samples of tasks as she expressed that there might be various samples of tasks in a training program that she might take as seen in the following excerpt:

There might be variety of tasks like in the ILEC speaking tests. I wonder whether we can be as competent as them on assessment... (T6, Feb 23, 2018).

As is clear with all the findings from scales and statements of the participants above, oral language assessment techniques and tasks is a major topic that high school teachers feel the need to improve themselves which reveals their training needs on it.

4.2.4.8 Assessment criteria. Knowledge of speaking assessment criteria was mostly rated as ‘*Good*’ or ‘*excellent*’ in KOSAQ, except for a few items that were rated as ‘*Moderate*’ or ‘*Not good*’ by some of the participants. Concordantly, relatively less participants rated *Item – 10* as ‘*Very important*’ or ‘*Important*’. 11 participants rated it as ‘*Important*’ and 16 participants rated it as ‘*Very important*’ which still corresponds to 54 % of the total number of participants. 15 participants rated this item as ‘*Moderately important*’ while only 2 of them rated ‘*Not important*’.

As rubrics are prepared and used in the light of assessment criteria, items related to preparing or deciding on rubrics (*15 and,16*) might again indicate high school teachers’ training needs on assessment criteria. *Item – 15*, which is about preparing rubrics for oral language assessment, was mostly rated as ‘*Very important*’ by 19 high school teachers (38 %). This item was also rated as ‘*Important*’ by 18 participants (36 %). It is evident that high school teachers think that one of the most important topics they need training on is preparing rubrics for oral language assessment. Similarly, *Item – 16* was rated as ‘*Very important*’ or ‘*Important*’ by 34

participants (68 %) which revealed that high school teachers need to be trained in deciding on ready-made rubrics to be used in oral language assessment process, as well. Besides, the results from *Item – 8* indicates that teachers need to be shown samples of rubrics for oral language assessment.

In the interviews and in the open-ended answers, the need for training for oral language assessment criteria is also evident. T3 and T5, clearly stated that they might need a training on designing or determining the criteria as is seen in the following excerpts:

...And there might be training on the criteria. What kind of criteria can we use? On designing more detailed criteria might be included (T6, Feb 21, 2018).

...The criteria by which we can make a more objective assessment might be included. Determining the criteria according to the objectives and applications that we can assess these more practically (T3, Feb 14, 2018).

In Section Three, in KOSAQ, oral language assessment criteria were the most controversial issue as the participants split in half with their opposing ideas about having standardized oral language assessment criteria. When the statements of participants in *Item -3* in TNAFOLAS was analyzed, it became obvious that, a remarkable number of teachers feel the need to be familiarized with standard samples of criteria for oral language assessment, as is also seen in the following excerpt from T34 and T48:

We need more standardized oral language assessment criteria and more samples (T34, Feb 2018).

I think, if we have to assess speaking during a class time, we should be provided with standardized criteria (T48, Feb 2018).

The excerpts and quotes above provided the researcher with the insight that, teachers are not confident about determining and designing their criteria for oral language assessment even if majority of them rated their knowledge as ‘*Good*’ for in the related item in KOSAQ. From the statements above, it can be inferred that, teachers need training on how to design their assessment criteria, how to determine the criteria to use in their classrooms and how to prepare or choose rubrics accordingly. In addition, they need to be shown the samples of assessment criteria and rubrics including them that can be used in the assessment of oral language skills.

4.2.4.9 Evaluating / Rating. Item – 19 is about ‘evaluating/rating oral language performance’ and this was regarded as ‘*Very important*’ by 16 participants (32 %), and ‘*Important*’ by 21 participants (42%) which constitutes 74 % of total participant number. It is apparent that high school teachers believe that they need training on how to rate students’ performance in oral language assessment process. The need for training on evaluating/rating of students’ oral language performance was also confirmed by the interviewees who mentioned this topic in their statements. T3, for example, stated that he does not know whether he did true or not with his effort to assess oral language:

...I, myself, try to assess it by making observations during the class time. I don’t know much about it. I am not sure whether I do the right thing or not [...] I mean, I have my own ideas, but they do not have an academic basis (T3, Feb 14, 2018).

With this statement, it seems clear that, he wants to be confident about what he does while rating students’ speech for oral language assessment, and he needs his knowledge to have academic basis. It indicates that he is not comfortable with carrying out the assessment according to his own ideas which we can infer that he has hesitations about the ethical concerns of the assessment processes. T1 clearly stated that she might need training on rating students’ performance:

I think it would be better if I learn more about which methods and techniques that I can use, Whether I do the right things while rating the students and how can I do it correctly... (T1, Feb 8, 2018).

Item – 21 and 22, which were again about rating / evaluating students oral language performances as they include topics such as ethical considerations, validity, reliability and subjectivity issues, got high ratings as ‘*Very important*’ or ‘*Important*’ from 32 participants (64 %). The results are consistent with the statement in the interviews as high school teachers expressed their concerns about ethical issues in their assessment process. In her following statement, T6 said ‘everything is up to my own personal decision’ in which she implies her worry for subjectivity in rating, while telling about the difficulties that she has in oral language assessment:

Do I have difficulties? Yes, I do. Like, for example, I observe them speaking and evaluate whether they speak fluently or whether they use the words they have learnt, and, how can I say? It looks as if everything is up to my own humble decision; but, I try to give high grades as much as I can (T6, Feb 23, 2018).

T5 also expressed her concern about the same issue while answering the question about assessment criteria. It seems clear that she is not comfortable with determining everything by herself. She also uttered ‘not objective’ as is seen in the following excerpt:

We all determine our criteria for ourselves. There are a lot of differences. I don't think this is objective. We need to be enlightened about the criteria for oral skills (T5, Feb 21, 2018).

All in all, high school teachers have worries about being left the whole initiative in rating students' performance in oral language skills. Interestingly, all used similar utterances like “my own decision, on my own, ourselves...” indicating that they do not want to be left alone with the conscience of this subjective process. It is apparent that they do not have confidence in making the right decision on this as they feel a lack in their knowledge and training on it. High school teachers need training on how to deal with subjectivity issue in the evaluating and rating process.

4.2.4.10 Washback in the classroom. Item – 23 also got high ratings from 33 participants (66 %). It was about washback on the classroom. High school teachers also reflected on their needs for providing washback in their classroom while assessing oral language assessment by rating this item as ‘*Very important*’ or ‘*Important*’. T7 also stated that she might need training on this topic which is shown in the following excerpt:

...For example, while giving feedback, we only say ‘Nice!’, ‘Good!’, ‘well-done!’ We can learn the alternative ways of giving feedback (T7, Feb 23, 2018).

It seems that high school teachers need to learn more about how to provide effective feedback for their students.

The following three topics are also among the ones that high school teachers need training. They were not included in TNAFOLAS and were found as the qualitative data were analyzed and interpreted.

4.2.4.11 Designing assessment according to student level. In KOSAQ, student level was found as one of the factors that influence high school teachers' oral language assessment practices. Some participants stated that they cannot make oral language assessment because of students' low level of proficiency or capacity. T2 was one of those teachers. She clearly mentioned, that she might need training on carrying out oral language assessment according to students' level while answering

the question that was seeking for whether she thought she might have any training needs on any aspects of oral language assessment as shown in the following excerpt:

I may need training on how we can make the assessment according to students' level... (T2, Feb 9, 2018).

T1 also explained her need for a training in designing assessment techniques / tasks according to different levels of students in her context in detail in her statement for the answer of similar question below:

...We cannot arrange classes according to students' English level; so, we have different levels in our classes, like, I can say there are students from A1 to C1 in the same class. I need to prepare different things to assess their speaking. What can I do to assess different levels of students? These might be my training needs (T1, Feb 8, 2018).

As is apparent within the findings of KOSAQ, TNAFOLAS and semi-structured interviews that, teachers need training on how to cope with the hardship of mixed or low levels of students in target language while designing activities and tasks for assessing oral language skills and in the process of oral language assessment in their classrooms.

4.2.4.12 Motivating students. Student motivation was another factor that was revealed as having an influence on high school teachers' oral language assessment processes in the previous section and it was discussed in detail with excerpts to indicate teachers point of view. In *Item – 3*, in TNAFOLAS some teachers also mentioned it as their need for oral language assessment.

Increasing the motivation of students speaking (T33, Feb 2018).

In addition, during the interviews student motivation was mentioned by interviewees and lack of students' motivation to speak was determined as one of the important factors influencing high school teachers' oral language assessment previously in this chapter. T6 stated clearly that she needs training for oral language assessment in her following statement:

May be for how to motivate students to speak during lessons (T6, TNAFOLAS, Feb 2018).

As is evident with the findings above and recalling T4's and T5's statements about students' anxiety and unwillingness to speak in the previous section, the ways for overcoming students' reluctance and increasing students' motivation is another area that high school teachers need training.

4.2.4.13 Pronunciation and fluency. In KOSAQ, the *Item-8*, which was related to knowledge of aspects of pronunciation, intonation and stress, was rated by many participants as *'Not good'* or *'Moderate'*, and the results were interpreted as high school teachers were less confident about their knowledge in this topic and it was consistent with *Item – 32* which is about considering pronunciation as a category of speaking assessment. Two participants also mentioned it when asked about their training needs in TNAFOLAS as seen in the statement of T23 below:

Pronunciation, fluency and communication (T23, Feb 2018).

T10 stated similar topic as her training need as she claimed that she has problems in pronunciation. What was different about her statement that, she said that she, herself, live pronunciation problems as shown in the following quote:

Sometimes, I live pronunciation problems (T10, Feb 2018).

It can be inferred with the findings of KOSAQ and TNAFOLAS that teachers might be ignoring the assessment of pronunciation skill as they do not feel confident about even their own competence in pronunciation of English. Similarly, T12 said she needed to live in a place that the target language is spoken probably due to the same problem:

I don't need training for assessment of speaking, I need to live in a place that target language is used as first language (T12, Feb 2018).

To sum up, high school teachers' training needs for oral language assessment were investigated via TNAFOLAS with the purpose of providing an answer for the fourth research question. The results from KOSAQ were also utilized in order to interpret the analyzed data better. Qualitative and quantitative data from TNAFOLAS, KOSAQ and semi-structured interviews were compared and contrasted, and overall results were presented harmoniously under the titles that were created via pattern-coding. First of all, participants' training on oral language assessment was investigated in the first section of TNAFOLAS and the results revealed that majority of participants have not got any training on oral language assessment. In addition, some of the participants who got training before, believe that the training they got was not adequate and some others mentioned the training they got on different topics rather than oral language assessment. So, it was interpreted as a vast majority of the participants have not received a sufficient training on oral language assessment. As mentioned at the beginning of this section, all items except for *Item – 4* got ratings of *'Very important'* or *'Important'* from over 54 % of high

school teachers as they were thought important for high school teachers' training needs for the assessment of oral language assessment. Some also stated that they need training in all aspects of oral language assessment. Some others mentioned that they do not know what aspects of it they need training. It is evident that majority of teachers believe that they need an extensive training in which most of the aspects of oral language assessment are included. The most prominent area for participants' training needs was found as 'tasks and techniques' which got relatively high ratings in each item that it was mentioned and as was also found in the statements of participants and other topics followed with a minor difference in ratings.

In line with the findings regarding the fourth research question; the topics that high school teachers need training on were found as significance and benefits of assessing oral language skills, specifying main purpose and objectives for oral language assessment, assessing aspects of intonation, stress and pronunciation, preparing and using group / pair work activities in oral language assessment, preparing and using oral language assessment techniques such as portfolio assessment, peer assessment, student self-assessment, formative and summative assessment, preparing and using different oral language assessment tasks such as retelling story or text from aural stimuli, decision-making tasks, interview tasks and samples oral language assessment techniques and tasks which can be used for oral language assessment, preparing and using assessment criteria and rubrics and samples of assessment criteria / rubrics, using analytic and holistic scores, aspects of intonation, stress and pronunciation as a category of oral language assessment and considering the use of appropriate expressions as a category of oral language assessment, evaluating / rating oral language performance, validity, reliability and subjectivity concerns, providing washback in the classroom, making an oral language assessment according to different student level, increasing the motivation of students for speaking activities and tasks, and how to make oral language assessment in large classes in a limited time.

Chapter 5

Discussion and Conclusion

5.1 Overview

The aim of this thesis was to explore high school EFL teachers' knowledge and practices of oral language assessment, the contextual factors that influence their assessment processes and their needs for an in-service training on assessment of oral language performance. The study was conducted in Beşiktaş, İstanbul, Turkey with the participation of 50 high school teachers from 15 state high schools. Mixed-method design was utilized in this study, in which both quantitative and qualitative data were collected by means of two questionnaires and semi-structured interviews. In this chapter, findings of the study will be discussed in the light of research questions with regard to the related studies in the literature. The limitations of the study will be shared next. Finally, suggestions for further studies will be given place.

5.2 Discussion of Findings for Research Questions

5.2.1 Research question 1: The first research question was designed to find out what high school EFL teachers know about the assessment of oral language skills. The findings indicated that, majority of the high school teachers rated their knowledge for most of the items (*each over %50 except for Item-15*) as “Good” or “Excellent” in KOSAQ. However, deficiencies in their knowledge and inadequacies in their practices in many aspects of oral language assessment were revealed when qualitative data from open-ended items of the same questionnaire and semi-structured interviews were probed. Additionally, in the analysis of TNAFOLAS, majority of teachers were found not to have an adequate training on oral language assessment to satisfy their needs to assess oral language skills effectively. In this survey, teachers could have the opportunity to discern more about the aspects that they have inadequate knowledge; so, they expressed the areas they need training explicitly. These results were consistent with a number of studies in the literature (Güllüoğlu,2004; Shim, 2009; Muñoz et al.; 2012; Jannati, 2015). All of these studies revealed that teachers have a tendency to grade their knowledge as good, mostly because they interpret the term knowledge as their knowledge of basic concepts and rate it, instead of rating their insight of those concepts that should serve their

practices. Therefore, there found a gap between teachers' beliefs of their knowledge and their real knowledge that they can apply to their practices. Both Shim (2009) and Jannati (2015) claimed that teachers had some basic knowledge on the concepts related to language assessment; however, they were unable to put their knowledge into practice. This seems also to be the case in this study, since most participants' implementations found to be very limited in terms of defining tasks and task variety, assessing interactive skills and pronunciation, defining assessment criteria according to various student levels, using variety of assessment types and rating scores. In addition, 6 out of 7 interviewees claimed that they think they have inadequate knowledge on oral language assessment. Some said they do not have an idea what they really need to know (e.g. T3, T4). Some participants mentioned they assess speaking because they were required which means that they were also not aware of the significance of it. T4 stated that she has a kind of professional deformation and she does not want to deal with the assessment procedures. This reminded of experienced teachers' condition in Sahinkarakas (2012). The author found experienced teachers less confident and less motivated compared to prospective teachers and less experienced teachers and she stated that teachers need to be kept motivated. These results contradict with Ounis's (2017) findings, which was one of the rare studies with positive results both in terms of knowledge and practices of EFL teachers in terms of oral language assessment. The instructors were reported to carry out a well-designed, well-versed and continuous oral language assessment using authentic and interactive tasks despite the challenges they had.

Through analysis of the items individually, the aspects that teachers lack knowledge were detected more clearly. Especially the items that were rated as '*Not good*' or '*Moderate*' by relatively more participants indicated the areas that teachers' knowledge needs improving.

The items that include student-student interaction tasks were mostly rated as '*Not good*' or '*Moderate*' by many high school teachers although most (82%) rated their knowledge for assessing students' interactive skills as "Good" in *Item-11*. *Item-12* in KOSAQ was about group and pair work activities for speaking assessment and 13 participants rated their knowledge as '*Not good*' or '*Moderate*' in this item. Similarly, only one teacher (T1) stated that she uses group or pair work in oral language assessment. In addition, more participants rated their knowledge as '*Not good*' or '*Moderate*' in the items related to student-student interactive tasks such as

using learner-learner joint discussion and decision-making tasks (40 %), using interview tasks (36%) and using role-play tasks in (28%). In the same vein, there were very few signs of student-student interactive tasks in the answers of open-ended questions and interviews. Few teachers ($n=6$) mentioned they use role-playing as an assessment task. Others seem to be heavily depending on oral presentations. These results are consistent with Plo et al. (2013), in which very few Aragonese teachers mentioned to use role-playing, and none mentioned interactive group work or discussions. In contrast, in Ounis (2017), instructors were insightful in using interactive assessment tasks. Using interactive tasks and group activities were emphasized by many linguists and researchers in the literature (e.g. Bachman & Palmer, 1996; Brown & Hudson, 1998; Louma, 2004; Muñoz & Álvarez, 2010); therefore, this is an important aspect that needs improvement in the practices of the high school teachers in Turkey, too.

In terms of task variety, 40% of the participants rated their knowledge as '*Not good*' or '*Moderate*' in using re-telling story or text from aural stimuli and 30% of the participants rated their knowledge as '*Not good*' or '*Moderate*' in using different assessment tasks. Majority of participants reported using oral presentations on a prepared topic. Role-playing was mentioned by only a few teachers and picture description or comparison by very few. Similarly, designing tasks or choosing appropriate tests was also one of three top topics that teachers have inadequate knowledge in Wu (2014). Task variety in oral language assessment is emphasized in the literature (Brown & Hudson, 1998; Coombe et al., 2007); it is another aspect that needs improvement in teachers knowledge and practices.

Item-8 and *Item-32* in KOSAQ were both related to pronunciation as a construct or criteria for oral language assessment and they were rated by less participants as "*Good*" or "*Excellent*". In addition, very few teachers mentioned it as an aspect of their assessment in their answers to open-ended questions related to their practices of oral language assessment and during the interviews. T5 explicitly mentioned that she omits the assessment of pronunciation since students have a very poor pronunciation. Likewise, Khamkhien (2010), found pronunciation as the most problematic component in oral language instruction and assessment. The author stated that having poor pronunciation caused communication breakdowns in students' interaction and a decrease in their self-esteem. It is evident that this is one of the areas that high school teachers lack knowledge. What is more, teachers'

statements in some items in both KOSAQ and TNAFOLAS revealed that they believe pronunciation need to be assessed based upon native-speaker standards. Some high school teachers stated that foreign teachers might be helpful in solving the problems related to diverse accents. Similarly, İnceçay and İnceçay (2010) found out that teachers did not use target language as classroom languages mentioning crowded classes and their heavy schedule as an excuse. The teachers in their study also thought that working coordinatively with native-speaker teachers would be a solution to this problem. These findings support Louma (2004), Nation and Newton (2009) and Yoshida (2016) as all stated that the biggest problem about the assessment of pronunciation is the expectation of native-like mastery in it.

In the items related to using various assessment types, participants were the least confident in this study. 53% of the participants for using portfolio assessment, 44% for using summative assessment, 50% for using peer assessment and 30% of the participants for using self-assessment rated their knowledge as “*Not good*” or “*Moderate*”. None of the interviewees spoke of these assessment types while talking about their assessment processes and none of the participants motioned these assessment types in their answers to the items about their practices. In addition, in TNAFOLAS, Item-9 was about their training needs on different assessment types and this item was rated by 62 % of the participants as “*Very Important*” or “*Important*”. It seems clear that, high school teachers admit that they do not have adequate knowledge on these assessment types. It is also evident with their ratings as “*Very Important*” or “*Important*” in TNAFOLAS that, they are aware of the significance of using variety of assessment types as stated by many linguists in the literature (Brown & Hudson, 1998; Louma. 2004; Belgrad et al., 2008; Nation & Newton, 2009). These results comply with Muñoz et al. (2012) in terms of teachers’ needs for guidance, which also suggested teachers training on various assessment types at the end; whereas, differ in terms of awareness of the lack in their knowledge. To sum up, teachers lack knowledge and need to gain more insight in using various assessment types and strategies.

In KOSAQ, high school teachers mostly rated their knowledge as good in knowledge of assessment criteria. In two items (*Item-34* and *35*) related to ‘analytic’ or ‘holistic’ scores in the related part, more participants (38%) rated their knowledge as “*Not good*” or “*Moderate*” indicated that they are less confident about using these assessment types appropriately. In Emekçi (2016), the results were more

positive in terms competence of using analytic scales for both native and non-native EFL teachers as they were consistent in using analytic scales except for one item. In this study both native and non-native teachers were interpreted as having their own skills in using analytic scoring system. In the literature making use of the strengths of both rating systems is suggested for obtaining the highest reliability (Hughes, 2003; Brown, 2005; Luoma, 2004). The strengths and weaknesses are highly related to the teachers as the assessors in the classroom contexts (McNamara, 1996; Hughes, 2003; Weir, 2005; Fulcher, 2009; Tuan, 2012).

In Item-33 in KOSAQ which was related to high school teachers' knowledge of considering the use of appropriate expressions as a category of oral language assessment was rated as "*Moderate*" by 10 participants. This means that 10 out of 50 teachers do not think they have adequate knowledge in the assessment of these skills. Concordantly very few of the teachers mentioned that they assess these skills in their classrooms. In fact, the communicative strategies seem to be ignored in all participants' assessment processes. Grada (2014), similarly, found that teachers were more focused on linguistic components rather than communicative components in their oral assessment practices. Bachman (2000) defined pragmatic aspects as "the benchmark" of language acquisition to emphasize its significance. It is also highlighted by many other scholars (Bachman and Palmer, 1996; Fulcher's, 2003; Louma, 2004). The skills related to appropriate expressions should be given place in the assessment of oral languages and teachers need to know how to assess them appropriately.

5.2.2 Research question 2. By the second research question, teachers' current practices of oral language assessment were investigated. Vast majority of teachers mentioned they assessed oral language skills in their classrooms; however, not all gave details about their assessment processes in KOSAQ. This is probably because of the lack in teachers' practical knowledge of oral language assessment as discussed within the first research question. Majority of the participants (56%) mentioned that they assess oral language skills once or twice a semester which implies summative assessment rather than formative. It was the same in Grada (2014) in which teachers carry out the summative assessment individually. It is clear in the participants' statements that, even if they determine the criteria with other EFL teachers, the participants of this study also assess oral language individually. Only T6 used "we" while talking about their assessment out of 7 interviewees.

Fluency, accuracy, vocabulary and content were the components of their criteria which also included in many studies in similar scope (Brown et al., 2005; Lozovska-Güneş, 2010; Bøhn, 2015; Emekçi, 2016); however, pronunciation was mentioned in the present study by very few teachers. The problems about the involvement of pronunciation in the assessment process was mentioned within the discussion about teachers' knowledge of pronunciation as an assessment construct. It is evident that assessment of pronunciation is ignored by most of the high school teachers as a problematic skill to assess. T5 mentioned that she cannot assess pronunciation because students' pronunciation competence is very poor. The statement of T5 is consistent with the Aragonese teachers' in Plo et al. (2013). They also mentioned they mentioned their students' poor pronunciation, as well. As was also revealed in this study, pronunciation has often been found to be the most troublesome aspect of oral language assessment in the related literature (e.g. Khamkhien, 2010; İnceçay & İnceçay, 2010; Plo et al., 2003; Emekçi, 2016), and it was also admitted by many researchers as a problematic issue (Louma, 2004; Yoshida, 2016).

Vocabulary is prioritized as mentioned to be important by 4 out of 7 interviewees as such in a number of studies in the literature (Iwashita et al., 2008; Brown et al., 2005). In this study, most high school teachers seem to have internalized the importance of intelligibility of the given message rather than grammatical correctness in oral language assessment since 3 out of 7 interviewees spoke of and some participants mentioned in KOSAQ. Similarly, in Lozovska - Güneş (2010) intelligibility was mentioned within the criteria that teachers' commonly mentioned instead of grammar accuracy. It is clear that, teachers are getting more conscious about the assessment of oral language as a communication device and leaving the traditional assessment approach which strictly focused on grammatical errors.

As for assessment techniques and tasks, most of high school teachers in this study seem to be heavily dependent on prepared oral presentations rather than using variety of activities that involve student-student interaction as mentioned in the discussion of the first research question. In this study, the mentioned tasks were oral presentation, role-playing and picture-description, but the latter two were quite rare. Similar findings were reported by Plo et al. (2013) as teachers spoke of a few tasks such as listening comprehension, oral presentation and picture-utilized tasks and their

practices were focused on oral presentations. In the literature, various tasks were discussed and empirically researched (Henning, 1983; Dandonoli & Henning, 1990; Foster & Skehan, 1996; Sinwongsuwat, 2012; Moere, 2012; Wang 2014) and eventually using various tasks was suggested (Brown & Hudson,1998; Coombe et al., 2007) to gain better outcomes in terms of reliability and learning outcomes. The high school teachers in this study also need to be informed about designing or choosing various interactive and authentic tasks according to their objectives, context and student levels.

5.2.3 Research question 3. The third research question was about the factors in high school teachers' contexts that influence their oral language assessment processes. The findings from both quantitative and qualitative data indicated that the biggest challenge they perceive is having large class sizes. Teaching to crowded classes has always been the biggest challenge for most teachers in all contexts in assessment of language skills as also mentioned in Chapter 2 in the current study (Shim, 2009; Lee, 2010; Plo et al., 2013; Grada, 2014). İnceçay and İnceçay (2010), a study conducted a study in Turkish context, revealed that teachers give up even using target language in the classroom because of this problem. Educational authorities should stop overlooking this problem and take necessary steps to solve it before anything else.

In the present study, majority of the participants (54%) note curriculum objectives as a factor that influence their decision making in speaking assessment; however, some perceive it as a positive effect stating that curriculum objectives guide them in the process of oral language assessment while others think that it is an obstacle as it retains their decision making in the assessment process. In Grada (2014), 32.9% of the participants were mentioned to be affected from curriculum objectives in a Libyan context. It is clear that, in Turkey, teachers feel more pressure from curriculum objectives. Another factor, which high school teachers have different perspectives, is having standard criteria. The number of the participants ($n=19$) who think they need to have standardized criteria in order to make objective assessment is equal to the number of the ones who think it might hinder making assessment considering individual differences and different levels of students (*both* 38%). The percentage value of the teachers who were uncomfortable with prescribed rules (60.5%) were quite higher than the ones who preferred a common way of assessment (36.8%) in Grada (2014). It is also higher than the ones in the present

study. It is inferable that Turkish teachers feel more comfortable when they are free of making too many critical decisions. Three of the interviewees mentioned their concern on subjectivity of the assessment as they decide everything about their assessment processes by themselves. They claimed that there need to be a common way to assess oral language skills. These results were consistent with Lozovska-Güneş (2010) as the author found similar diverse ideas about having the criteria given by their institutions or determining their own criteria; whereas, in Lozovska-Güneş (2010) instructors have agreed on the idea of having a standardized rating scale on which they can make small changes while applying in their own classroom. In Knight's (1992) study, while majority of teachers were positive about determining their own criteria some teachers insisted that they need to have standard criteria following the workshop on determining assessment criteria. It is evident that some teachers need to be secure in using standardized criteria and think that their assessment is more objective in that way. Koh et al. (2017) got positive results in their study that examining a longer training program on determining criteria. It can be inferred that extensive training might be beneficial for teachers in determining criteria.

Deriving from the statements of the interviewees and participants who answered open-ended questions in KOSAQ, other factors influencing teachers oral language assessment processes are determined as having limited time, working with students with low/various levels of English, student unwillingness to speak and subjectivity of oral language assessment process. Lack of time seems to be one of the major problems in oral language assessment in most contexts (Lozovska-Güneş, 2010; Plo et al.,2013; Grada, 2014). Grada (2014) similarly found this factor within the qualitative data through the interviews. Some participants in her study claimed that time pressure cause them to abandon the use of some tasks requiring group and pair work. In this study, even worse some participants (e.g. T2, T29) admitted that they wholly abandon speaking assessment due to limited time which supported Nation and Newton (2009) who stated that teachers do not assess speaking due limited time allocated to speaking activities in schools. Thornbury (2005) suggest using holistic assessment as it is less time-consuming. In short, teachers have limited class-time, and this is a challenge in their assessment process; therefore, they need to improve their knowledge to use practical ways of making assessment in limited time. Another factor that revealed to have influence teachers' oral language assessment

was low/various levels of students. T1, for example, claimed that there were diverse levels of students in her classrooms, so it was difficult to design or find activities that will fit each. T2, on the other hand, claimed it was not possible to assess speaking in her classes as her students' level was quite low. In practices of oral language assessment in Plo et al. (2013), Aragonese teachers were also found to fail to deal with mixed-level classes, and the authors noted it as a major issue that teachers needed training. In a Turkish context, Köksal (2004) conducted a study with state school teachers and the teachers were similarly reported to fail preparing tasks appropriate to students' level. Language scholars stress designing or choosing both tasks and assessment criteria according to student level (Underhill, 1987; Fulcher, 2003); hence, it is an important shortcoming for high school teachers, and it needs improving. Three of the interviewees and some other participants mentioned that students' unwillingness to speak was a major challenge in the oral language assessment process. T1 and T10 for example mentioned lack of intrinsic motivation of students when T4 and T5 pointed to the anxiety derived from peer pressure. T5 also spoke of the age factor in high schools stressing the effects of puberty in peer interaction. Khamkhien (2010), similarly, mentioned the decrease in the self-esteem of Thai learners when they make mistakes, and they were mentioned to restrain themselves from taking turn again. When the context is a high school it is more difficult as mentioned by T-5. Thus, this is another important issue that high school teachers need to learn how to deal with.

Lack of teacher training is also mentioned by a remarkable number of high school teachers as a factor influencing their assessment process. It will be discussed within research question 4 in the next section.

5.2.4 Research question 4. The fourth research question was aimed to investigate high school teachers' training needs for the assessment of oral language skills. TNAFOLAS was utilized to find out the training needs of high school teachers on oral language assessment. The findings from KOSAQ and semi-structured interviews were also utilized. The findings indicated that majority of participants have not got any training, and some of the participants who mentioned that they got training do not think that it was adequate. Lack of teacher training is also mentioned by a remarkable number of high school teachers ($n=17$) which comprises the 34 % of the total participants as a factor that influence their oral language assessment. It indicates these participants are conscious about their shortcomings on oral language

assessment knowledge and practices, and they are aware of the need for a professional in-service training. When the participants were asked for the importance of the aspects for their training needs, almost all items were rated as *'Very important'* or *'Important'* from over 50 % of high school teachers. The ratings in TNAFOLAS contradicts with the ratings in KOSAQ. As we discussed within research question 1, participants might have rated their superficial knowledge while rating in KOSAQ. In the third section of the questionnaire they were asked about their practices and in semi-structured interviews interviewees were asked whether they think their knowledge was enough to manage the assessment process effectively. In addition, they were asked whether they think their training satisfied their needs in their oral language assessment processes. In this way, they were more oriented to think about their practical knowledge that they need to apply in their classroom assessments. In TNAFOLAS, some participants explicitly mentioned that they need a top-down training in all aspects. Some interviewees said that they do not know what they need to learn to make their assessment more effective. In Mertler (2009), teachers admitted that they had had superficial knowledge of language assessment as a concept and they needed to know more about it in order to be competent in the process language assessment only after getting a two-week workshop. In Grada (2014), lack of training was mentioned as a factor influencing oral language assessment processes by more teachers (75%). It is comprehensible when the experience factor is considered, as novice teachers in Grada's (2014) study were more open to admit having training needs compared to the participants of this study who have relatively long years of experience. The present study further analyzed teachers' needs in detail by means of TNAFOLAS and semi-structured interviews and found out what aspects that teachers thought they need training. Many researchers also prioritized providing teachers with training on the terminology and practical ways of using various assessment types (Brown & Hudson, 1998; Topping, 1998; Hughes, 2003; Black, 2003; Taras, 2008).

The major aspects that teachers need training on were found as defining objectives, assessing intonation, stress and pronunciation, preparing and using group / pair work tasks, using portfolio assessment, peer assessment, self-assessment and using formative and summative assessment, preparing and using different tasks such as re-telling story or text from aural stimuli, decision-making tasks and interview tasks, preparing and using assessment criteria and rating scales, using analytic and

holistic scores, assessing communicative strategies, providing washback in the classroom, assessing different levels in the same classroom, increasing motivation of students, evaluating / rating oral language performance how to cope with validity, reliability concerns and challenges related to their contexts. In line with their answers in KOSAQ and interviews, it is also clear that they need to be motivated and informed about the usefulness and benefits of assessing oral language skills.

In the literature, there are not many studies that have examined teachers' knowledge, practices and training needs on oral language assessment in the same pace; nevertheless, there are some studies that have found similar aspects for the needs of EFL teachers professional training on oral language assessment. Köksal (2004), for example, having found many deficiencies in the assessment practices of EFL teachers, suggested an overall training in language assessment. Köksal (2004) received these results by analyzing the tests that were prepared by teachers; in the present study on the other hand, the overall training is voiced by teachers' themselves which seems to be a better gain in terms of consciousness throughout time. Muñoz and Álvarez (2010) highlighted training teachers for defining goals, guiding self-assessment in the classroom, using rating scales and providing washback in the classroom after the trial of a training program in an experimental study. The present study found similar aspects by giving voice to high school teachers. Sahinkarakas (2012) revealed a lack of motivation and self-esteem with experienced teachers and suggested they need to be kept motivated by in-service teacher training. In the current study, a number of teachers were found to be unaware of the benefits of assessment processes, and some claimed that they carry it out due to the requirement of the ministry. İnceçay and İnceçay (2010), likewise, found half of the participants of their study thought that speaking is the least important skill among four macro skills; so, they do not give a place to the instruction and assessment of it in their classrooms. Thus, this is another issue that high school teachers lack motivation; therefore, in the present study, teachers need to be trained on the practical uses and benefits of assessment in their classroom teaching so that they can be more motivated to assess speaking skills. In Lozovska-Güneş (2010), academicians have similar concerns to the present study about the challenges and difficulties they have in the assessment process such as subjectivity, defining assessment criteria, students' anxiety, limited time and reliability of their ratings and they voiced they need a training on speaking assessment in general. It is clear that the

present study also found major aspects of teacher training and the common concerns of the instructors in the assessment of oral language skills regardless the context. Plo et al. (2013), with quite similar findings with the present study, specifically emphasized the need for teachers' training on handling mixed method classes which was also voiced by many high school teachers in this study. Wu's (2014) study was one of the rare studies which undertook an analysis of teachers' training needs for language assessment, although not specifically for oral language assessment, and found out that teachers need training especially on speaking skills assessment in terms of designing and preparing tasks and tests, validity issues and interpreting the scores. Likewise, these three aspects were also found to be significant for high school teachers need training in the present study. One of the similarities of the current study with Wu's (2014) is that teachers are unable to cope with contextual factors' influence and need training on it. Önal (2010) was one of rare studies whose results contradicted the current study and many others in the literature in terms of the knowledge of instructors and their successful practices. In it explainable when the context that the author conducted the study. It was an award-winning private university on language teaching; hence, the study might not represent the divergent EFL contexts in Turkey.

In addition to the studies discussed with similar and diverse findings above, there are suggestions of some scholars and other researchers for teachers' training on oral language assessment that match to the findings of this study. Weir (1990), for example, emphasized that teachers need to know several issues such as how to make a defensible, objective and authentic assessment, consistent with goals, promoting student interaction and motivation which were all detected as their own professional training needs by the high school teachers in this study. Yoshida (2016) stressed the need for training teachers to assess the intelligibility of pronunciation which was also voiced by the participants of this study. Likewise, Herman et al. (1992) on determining criteria consistent with the constructs, Underhill (1987) for designing assessment according to diverse levels of students, Wall (1996) for providing positive washback and Fulcher (2003) for carrying out a reliable assessment.

5.3 Limitations

The findings of the current study brought high school teachers' knowledge, practices, contextual challenges and training needs of oral language assessment to

light to a great extent by giving voice to 50 high school teachers who work in state high schools in Beşiktaş, İstanbul. The study provided a picture of current situation of oral language assessment in high school context; however, it still comprised some limitations related to generalizability of the results. First of all, the study was conducted in the state schools in Beşiktaş, as it was more convenient to reach to the researcher. This was, on the one hand, advantageous since the researcher had the opportunity to visit each high school and meet almost all participants in person; thus, each participant could be well-informed about the aim and the scope of the study and encouraged to answer the questions sincerely seeing that the researcher was in the same boat as them; on the other hand, it has the limitation of giving voice of the divergent regions in İstanbul, especially more unprivileged ones in suburbs. Conducting the study in Beşiktaş high school constituted a second limitation for the study. The high school teachers in Beşiktaş have relatively long years of experience, since less experienced teachers are usually appointed in compulsory districts. In this study most of the participants were at the age range of 35- 44 and had over 11 years of experience. 7 participants had 6-10 years of experience, which is relatively high again, and only two participants had 2-5 years of experience. Thus, the study limits the voice of younger and less experienced teachers, either. The findings might have been different, if there were more younger and less experienced high school teachers from divergent regions of Istanbul. Additionally, high school teachers were asked whether they got any training on oral language assessment; whereas, their college education on this topic were not investigated as the questionnaires and semi-structured interviews did not include additional questions for the details of their training background. The data about their pre-service training in the college on this topic, would provide better insight for their present knowledge and practices of oral language assessment.

5.4 Recommendations for Further Studies

Future studies might be conducted in divergent districts of Turkey and might compare the knowledge, practices and training needs of high school teachers who work in more disadvantageous contexts. Further research might also focus on less experienced teachers' point of view in terms of oral language assessment knowledge, practices and training needs and provide a comparison of findings. The experience factor might be investigated as a variable for teachers' perceptions, practices and perceived training needs in further research. In addition, EFL teachers' pre-service

education on oral language should also be investigated in order to shed light to the adequacy of their college training on oral language assessment.

5.5 Conclusion

The purpose of the present study was to investigate Turkish high school EFL teachers' knowledge, practices and training needs on oral language assessment. 50 high school teachers filled in two questionnaires (KOSAQ and TNAFOLAS) including both close-ended and open-ended questions related to their perceived knowledge, practices, contextual factors that influence their oral language assessment knowledge and practices and their needs. 7 participants were interviewed through semi-structured interviews in order to collect further qualitative data to probe into their current knowledge, practices, influencing factors and their training needs. Within the analysis of the data collected by means of both questionnaires and semi-structured interviews, the following findings were attained:

Majority of the high school teachers rated their knowledge as good or excellent in the items related the aspects of oral language assessment such as speaking assessment constructs, techniques, tasks and criteria in KOSAQ; whereas, when they were asked about their practices of oral language assessment, they indicated very limited use of their knowledge in practice. Most high school teachers' assessment practices seemed to depend on only 2-5-minute oral presentations that they ask students to prepare once or twice a semester. In addition, through the analysis of some items related to oral language assessment constructs, techniques, tasks and practices it was revealed that, many high school teachers lack knowledge of various assessment techniques such as formative assessment and summative assessment, portfolio-, self- and peer-assessment and failed to assess pronunciation, intonation and stress as speaking constructs, students' interactive skills through student-student interactive tasks such as interviews, role-playing, learner-learner discussions and decision making tasks, using group and pair activates as speaking assessment tasks, utilizing story-retelling activities from an aural stimuli or a written text, using analytic and holistic scores in the assessment of oral language skills. The analysis of the data from interviews and the answers of the open-ended items also revealed that, teachers think that they have inadequate training and practical knowledge to apply successfully in their practices. Some high school teachers also lacked knowledge of the necessity and benefits of the assessment on teaching and learning activities since they claimed to perform perfunctory activities because they were required to assess

speaking skills. In their practices of oral language assessment high school teachers have shortcomings in specifying the objectives of their assessment processes, determining constructs, criteria, tasks or test items for the assessment of oral language performance, carrying out a valid and a reliable assessment of oral language skills, designing tasks or criteria according to diverse levels of students and providing positive washback for the classroom teaching and learning of oral language skills.

As for the practices of high school teachers, majority of the participants claimed to assess speaking skills in their classrooms, although some mentioned to do it just because they were required, and some others explained they cannot make an assessment due to contextual factors. Majority of high school teachers mentioned to have fluency, accuracy, vocabulary and content as assessment criteria and oral presentations as speaking assessment tasks. Very few participants mentioned pronunciation in their assessment criteria and different tasks such as role-playing and picture description or comparison. High school teachers seem to be dependent on oral language performance tasks that they use to assess students' fluency, accuracy, vocabulary and content knowledge in their performances once or twice a semester. High school teachers think that vocabulary is an important construct to assess and they are conscious about grammar accuracy should not be prior to the intelligibility of the speech. The participants of this study mentioned the factors that affect their assessment processes as class size, lack of teacher training, having inadequate class hours, diverse and poor language proficiency levels of the students, reluctance of students to speak. Curriculum objectives were also mentioned by majority of the participants; however, they have different ideas about them. Some think that they are helpful as a guide to follow, others think that they restrict their practices. Similarly, contradictory ideas were also revealed about having prescribed criteria, since some of the participants think that they need to be free in determining their own criteria according to their students' level, others think there must be common standardized criteria to assess regarding objectiveness.

The findings eventually indicated that high school teachers have shortcomings on many aspects of oral language assessment and they need training on various topics such as how to specify the objectives of their assessment practices in line with curriculum objectives and how to determine their criteria accordingly, how to assess intonation, stress, pronunciation, students' interactive skills and communicative

strategies through various interactive and authentic tasks such as interviews, student-student discussion and decision making tasks and pair or group activities, how to apply various assessment techniques such as portfolio-assessment, self-assessment and peer-assessment to their oral language assessment, how to use analytic and holistic scoring systems, how to deal with issues related to rating students' performance and interpreting the scores such as validity, reliability and objectivity, how to make oral language assessment in mixed level classes and/or crowded classes, how to carry out oral language assessment in limited time, how to motivate students and guide them to over their speaking anxiety and how to provide positive washback in the classroom.



REFERENCES

- Angelo, T. A. (1995). Reassessing and defining assessment. *American Association for Higher Education*, 48 (3), 7-9.
- Alderson, J. C. & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14, 115- 129.
- Arter, J., & McTighe, J. (2001). *Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance*. New York: Corwin Press.
- Arter, J., & Spandel, V. (1992). Using portfolios of student work in instruction and assessment. *Educational Measurement: Issues and Practice*, 11(1), 36-44.
- Auerbach, C. E., & Silverstein, L. B. (2003). *Qualitative data: An introduction to coding and analysis*. New York: New York University Press.
- Bachman, L. F & Palmer, A. S. (1982). The construct validation of some components of communicative proficiency. *TESOL Quarterly*, 16(4), 449 – 465.
- Bachman, L.F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. New York, NY: Oxford University Press.
- Bailey, K. M. (1999). *Washback in language testing. TOEFL monograph series*. Princeton, New Jersey: Educational Testing Service
- Bailey, K. M. (2005). *Practical English language teaching: Speaking*. New York, NY: McGraw-Hill.
- Barkaoui, K., Brooks, L., Swain, M., & Lapkin, S. (2013). Test-takers' strategic behaviors in independent and integrated speaking tasks. *Applied Linguistics*, 34(3), 304-324.
- Belgrad, S., Burke K. & Fogarty, R. (2008). *The portfolio connection*. Thousand Oaks, CA: Corwin Press.

- Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74.
- Black, P. (2003). The nature and value of formative assessment for learning. *Improving Schools*, 6(3), 7-22.
- Black, P., Harrison, C., Lee, C., Marshall, B., & William, D. (2003). *Assessment for learning: Putting it into practice*. England: Open University Press.
- Blanche, P. (1988). Self-Assessment of foreign language skills: Implications for teachers and researchers. *RELC Journal*, 19(1), 75-93
- Bloom, B. S., Hastings, J. T. & Madaus, G. F. & Baldwin, T. S. (1971). *Handbook on the formative and summative evaluation of student learning*. New York, NY: McGraw-Hill.
- Booth, W. C., Colomb, G. G., & Williams, J. M. (2003). *The craft of research* (2nd ed.). Chicago: University of Chicago Press.
- Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. Thousand Oaks, CA: Sage.
- Brindley, G. (2002). *Assessment and the language teacher: Trends and transitions. The language teacher online*. Retrieved from <http://langue.hyper.chubu.ac.jp/tlt9/sep/brindley.html>
- Brown, J.D. (1995). *The elements of language curriculum: A systematic approach to program development*. Boston, Massachusetts, USA: Heinle & Heinle Publishers.
- Brown, J. D. & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly*, 32(4), 653-675.
- Brown, H. D. (2000). *Teaching by principles: An interactive approach to language pedagogy*. (2nd ed.). New York: Pearson Education.
- Brown, A., Iwashita, N., & McNamara, T. (2005). *An examination of rater orientations and test-taker performance on English-for-academic-purposes speaking tasks (TOEFL MS-29)*. Princeton, NJ: Educational Testing Service.
- Brown, J. D. (2005). *Testing in language programs*. Upper Saddle River, New Jersey: Prentice Hall Regents.
- Burgess, S. & Head, K. (2005). *How to Teach for Exams*. UK: Pearson Education Ltd.

- Burrows, C. (2004). Washback in classroom-based assessment: A study of the washback effect in the Australian adult migrant English program. In L. Cheng, & Y. Watanabe, (Eds.), *Washback in language testing* (pp.113–128). Mahwah, NJ: Lawrence Erlbaum.
- Bøhn, H. (2015). Assessing spoken EFL without a common rating scale: Norwegian EFL teachers' conceptions of construct. *SAGE Open*, 1(2), 1–12.
- Canale, M. & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics* 1(1) 1-47.
- Carr, J. F. & Harris, D.E. (2001). *Succeeding with standards: Linking curriculum, assessment and action planning*. Alexandria, VA: Association of Supervision and Curriculum Development.
- Cheng, L. (2004). The washback effect of a public examination change on teachers' perceptions toward their classroom teaching. In L. Cheng, & Y. Watanabe & A. Curtis (Eds.), *Washback in language testing: Research context and methods*. (pp.147–170). Mahwah, NJ: Lawrence Erlbaum.
- Cheng, W. & Warren, M. (2005). Peer assessment of language proficiency. *Language Testing*, 22 (1) 93–121.
- Cohen, A. (1980). *Testing language ability in the classroom*. Rowley, MA: Newbury House.
- Coombe, C., Folse, N. & Hubley, N. (2007). *A practical guide to assessing English language learners*. USA: The University of Michigan Press.
- Creswell, J. W. (2009). *Research design: qualitative, quantitative, and mixed method approaches*. (3rd ed.). Thousand Oaks, CA: SAGE Publications.
- Creswell, J. W. (2014). *Research design: qualitative, quantitative, and mixed method approaches*. (4th ed.). Thousand Oaks, CA: SAGE Publications.
- Chuang, Y. Y. (2009). Foreign language speaking assessment: Taiwanese college English teachers' scoring performance in the holistic and analytic rating methods. *The Asian EFL Journal Quarterly*, 11 (1), 152-175.
- Dandonoli, P., & Henning, G. (1990). An investigation of the construct validity of the ACTFL proficiency guidelines and oral interview procedure. *Foreign Language Annals*, 23(1), 11-22.
- Davis, L. (2015). The influence of training and experience on rater performance in scoring spoken language. *Language Testing*, 33(1) 117–135.

- Dörnyei, Z. (1995). On the teachability of communication strategies. *TESOL Quarterly*, 29(1), 55-85.
- Derwing, T. & Munro, M. (2005). *Second language accent and pronunciation teaching: A research-based approach*. *TESOL Quarterly* 39 (3), 379–398.
- Dörnyei, Z. (2003). *Questionnaires in second language research: Construction, Administration and processing*. London: Lawrence Erlbaum Associates Inc. Publishers.
- Dörnyei, Z. (2007). *Research Methods in Applied Linguistics: Quantitative, Qualitative, And Mixed Methodologies*. Oxford: Oxford University Press.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford, UK: Oxford University Press.
- Emekçi, E. (2016). Comparison of native and non-native English language teachers' evaluation of EFL learners' speaking skills: Conflicting or identical rating behaviour? *English Language Teaching*, 9 (5), 98- 105.
- Foster, P. & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18(3), 299-323.
- Fulcher, G. (2003). *Testing second language speaking*. Harlow: Longman/Pearson.
- Fulcher, G., Davidson, F. (2007). *Language Testing and Assessment: An Advanced Resource Book*. UK: Routledge.
- Fulcher, G. (2009). *Rating scales and the Halo Effect*. Retrieved from <http://languagetesting.info/gf/glennfulcher.php>
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113-132.
- Genç, B. (2007). *An Analysis of Communication Strategies Employed by Turkish Speakers of English*. (Unpublished Doctoral Dissertation). Çukurova University, Adana.
- Genesee, F. & Upshur, J. A. (1996). *Classroom-based evaluation in second language education*. UK: Cambridge University Press.
- Gràcia, M., Vega, F. & Galván-Bovaira, M. J. (2015). Developing and testing EVALOE: A tool for assessing spoken language teaching and learning in the classroom. *Child Language Teaching and Therapy*, 31(3) 287 –304.
- Gillham, B. (2000). *Developing a questionnaire*. London: Continuum.

- Grada, T.K.A. (2014). *An investigation into Libyan EFL novice secondary school teachers' current knowledge and practice of speaking assessment: A socio-cultural perspective*. (Unpublished doctoral dissertation), University of Exeter, Exeter, U.K.
- Grez, D. L., Valcke & M. Roozen, I. (2012). How effective are self-and peer assessment of oral presentation skills compared with teachers' assessments? *Active Learning in Higher Education* 13(2) 129 –142.
- Güllüoğlu, Ö. (2004). *Attitudes towards testing speaking at Gazi University Preparatory School of English and suggested speaking tests*. (Unpublished Master's Thesis), Gazi University, Ankara, Turkey.
- Harris, David P. (1969). *Testing English as a second language*. New York: McGraw Hill.
- Hasselgreen, A., Carlsen, C., & Helness, H. (2004). European survey of language testing and assessment needs (Report: Part one—General findings). Retrieved from <http://www.ealta.eu.org/documents/resources/survey-report-pt1.pdf>
- Henning, G. (1983). Oral proficiency testing: Comparative validities of interview, imitation, and completion methods. *Language Learning*, 33(3), 315-332.
- Henning, G. (1987). *A guide to language testing: development - evaluation research*. Rowley, Massachusetts: Newbury House.
- Herman, J. L., Aschbacher, P. R., & Winters, L. (1992). *A practical guide to alternative assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Hosseinia, A.S. & Azarnooshb, M. (2014). Iranian EFL instructor's oral assessment practices: purposes, methods, procedures. *Procedia - Social and Behavioral Sciences*, 98, 653 – 658.
- Huang, L. (2013). *Cognitive processes involved in performing the IELTS speaking test: respondents' strategic behaviours in simulated testing and non-testing contexts*. Retrieved from <https://www.ielts.org/teaching-and-research/research-reports/online-series-2013-1>
- Huei-Chun, T. (2007). *A study of task type for L2 speaking assessment*. Retrieved from <https://eric.ed.gov/?id=ED496075>
- Hughes, A. (2003). *Testing for language teachers*. (2nd ed.) Cambridge: Cambridge University Press.

- Iwashita, N. & Grove, E. (2003). Comparison of analytic and holistic scales in the context of a specific purpose speaking test, *Prospect*, 18 (3), 25 – 35.
- Iwashita, N., Brown, A., McNamara, T. & O'Hagan, S. (2008). Assessed levels of second language speaking proficiency: How distinct? *Applied Linguistics*, 29 (1), 24–49.
- İnceçay, G. & İnceçay, V. (2010). A case study on needs assessment of English language teachers. *Procedia Social and Behavioral Sciences*, 2, 317–321.
- Jannati, S. (2015). ELT Teachers' language assessment literacy: Perceptions and practices. *International Journal of Research in Teacher Education*. 6(2), 26-37.
- Johnson, R. B. & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33 (7), 14-26.
- Khamkhien, A. (2010). Teaching English speaking and English-speaking Tests in the Thai context: A reflection from Thai perspective, *English Language Teaching*, 3 (1), 184-190.
- Knight, B. (1992). Assessing speaking skills: A workshop for teacher development. *ELT Journal*, 46(3), 294- 302.
- Knoch, U. (2009). Diagnostic assessment of writing: A comparison of two rating scales. *Language Testing*, 26(2), 275-304.
- Koh, K., Carol, L. E. - Burke A. L., Gong, W. & Tan, C. (2017). Developing the assessment literacy of teachers in Chinese language classrooms: A focus on assessment task design. *Language Teaching Research*, 1 –25.
- Köksal, D. (2004). Assessing teachers' testing skills in ELT and enhancing their professional development through distance learning on the net. *Turkish Online Journal of Distance Education TOJDE*, 5 (1). Retrieved from <https://files.eric.ed.gov/fulltext/ED494459.pdf>
- Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*, 32(2) 169 –197.
- Layder, D. (1998). *Sociological practice: Linking theory and research*. London: Sage.
- Lozovska-Güneş, A. (2010). Differences and challenges involved in the assessment of speaking skill: the case of three universities in Ankara. (Unpublished Master's Thesis). Middle East Technical University, Ankara, Turkey.

- Lee, S. (2010). *Current practice of classroom speaking assessment in secondary schools in South Korea*. (Unpublished Master's Thesis). The University of Queensland, Australia.
- Leech, G. (2000). Grammars of spoken English: New outcomes of corpus oriented research. *Language learning* 50(4), 675-724.
- Le'ger, D.S. (2009). Self-assessment of speaking skills and participation in a foreign language class. *Foreign Language Annals*, 42(1), 158-179.
- Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge: The MIT Press.
- Lofland, J., Snow, D., Anderson, L., & Lofland, L. H. (2006). *Analyzing social settings: A guide to qualitative observation and analysis* (4th ed.). Belmont, CA: Thomson Wadsworth.
- Louma, S. (2004). *Assessing speaking*. Cambridge: Cambridge University Press.
- Madsen, H.S. (1983). *Techniques in testing*. Oxford: Oxford University Press.
- McKay, P. (2006) *Assessing young language learners*. Cambridge: Cambridge-University Press.
- McNamara, T. (1996). *Measuring second language performance*. London: Longman.
- Mertler, C. A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research and Evaluation*, 7(25), 1-10.
- Mertler, C. A. (2009). Teachers' assessment knowledge and their perceptions of the impact of classroom assessment professional development. *Improving Schools*, 12(2), 101-113.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing* 13(3), 241-256.
- Mok, J. (2011). A case study of students' perceptions of peer assessment in Hong Kong. *ELT Journal*, 65(3), 230–239.
- Moere, V. A. (2012). A psycholinguistic approach to oral language assessment. *Language Testing* 29(3) 325 –344.
- Morrow, K. (1991). Evaluating communicative tests'. In Anivan, S.(Eds.) *Current Developments in Language Testing*. (pp.56-58). Singapore, Regional Language Centre.
- Moskal, B.M. (2000). Scoring rubrics: What, when and how? *Practical Assessment, Research & Evaluation*, 7 (3). 1 – 5.

- Muñoz, A., Dary, L., Crespo, F., Gaviria, S., Lopera, L. & Palacio, M. (2003). Assessing spoken language in EFL: Beliefs and practices, *Medellin: Universidad EAFIT*, 129, 63-73.
- Muñoz, A. P. & Álvarez, M. E. (2010). Washback of an oral assessment system in the EFL classroom. *Language Testing* 27 (1) 33–49.
- Muñoz, A. P., Palacio, M., & Escobar, L. (2012). Teachers' beliefs about assessment in an EFL context in Colombia. *Profile: Issues in Teachers' Professional Development*, 14(1), 143 -158.
- Nakatani, Y. (2006). Developing an oral communication strategy inventory. *Modern Language Journal*, 90(2), 151–168.
- Nation, I. S. P. & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. New York: Routledge/Taylor and Francis.
- Noels, K. A., Pelletier, L. G., Clement, R., & Vallerand, R. J. (2000). Why are you learning a second language? Motivational orientations and self-determination theory. *Language Learning*, 50(1), 57–85.
- Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge: Cambridge University Press.
- Oppenheim, A. N. (1992). *Questionnaire design, interviewing and attitude measurement*. Continuum: New York.
- Ounis, A. (2017). The assessment of speaking skills at the tertiary level. *International Journal of English Linguistics*, 7 (4), 95 – 112.
- Ounis, M. (2017). A comparison between holistic and analytic assessment of speaking. *Journal of Language Teaching and Research*, 8 (4), 679-690.
- O’Sullivan, B., Weir, C.J. & Saville, N. (2002). Using observation checklists to validate speaking-test tasks. *Language Testing*, 19 (1) 33–56.
- Önal, A. (2010). *Testing and assessment of speaking skills in preparatory*. (Unpublished Master’s Thesis), Selçuk University, Konya, Turkey.
- Önem, E. E. (2015). *Instructors’ attitudes towards assessing speaking holistically and analytically*. (Unpublished Master’s Thesis). İhsan Doğramacı Bilkent University, Ankara, Turkey.
- Özdemir-Çağatay, S. (2012). *Speaking portfolios as an alternative way of assessment in an EFL context*. (Unpublished Master’s Thesis). Bilkent University, Ankara, Turkey.

- Pearson, P. D., Hiebert, E. H., Kamil, M. L. (2007). Vocabulary assessment: What we know and what we need to learn. *Reading Research Quarterly*, 42(2), 282–296.
- Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and using communication tasks For second language research and instruction. In G. Crookes & S. M. Gass (Eds.), *Tasks and second language learning* (pp. 9-34). Clevedon, UK: Multilingual Matters.
- Pineda, D. (2014). The feasibility of assessing teenagers' oral English language performance with a rubric. *PROFILE Issues in Teachers' Professional Development*, 16 (1), 181-198.
- Plo, R., Hornero, A. & Mur-Dueñas, P (2013). Implementing the teaching/learning of oral skills in secondary education in Aragón: Gauging teachers' attitudes, beliefs and expectations. *IJES*, 14 (1), 55-77.
- Prodromou, L. (1995). The backwash effect: From testing to teaching. *ELT Journal*, 49(1). 13-25.
- Rubin, D. (2012). The power of prejudice in accent perception: Reverse linguistic stereotyping and its impact on listener judgements and decisions. In J. Lewis & K. LeVelle (Eds.), *Proceedings of the 3rd Pronunciation in Second Language Learning and Teaching conference* (pp.11-17). Ames, IA: Iowa State University.
- Saldaña, J. (2009). *The coding manual for qualitative researchers*. London: SAGE Publications.
- Seong, Y. (2014). Strategic competence and L2 speaking assessment. *TESOL & Applied Linguistics*, 14 (1), 13-24.
- Sinwongsawat, K. (2012). Rethinking Assessment of Thai EFL learners' speaking skills. *Language Testing in Asia*, 2(4), 75-85.
- Shim, K. N. (2009). *An investigation into teachers' perceptions of classroom-based assessment of English as a foreign language in Korean primary education*. (Unpublished doctoral dissertation). University of Exeter, Exeter.
- Shohamy, A. (1984). Does the testing method make a difference? *Language Testing*, 1(2), 147 – 170.
- Shohamy, E. (1992). Beyond proficiency testing: A diagnostic feedback testing model for assessing foreign language learning. *Modern Language Journal*, 76(4), 513–521.

- Shohamy, E., Donitsa-Schmidt, S., & Ferman, I. (1996). Test impact revisited: Washback effect over time. *Language Testing*, 13(3), 298–317.
- Skehan, P. & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, 49(1), 93-120.
- Srikaew, D., Tangdhanakanond, K. & Kanjanawasee, S. (2015). Development of an English-speaking skill assessment model for grade 6 students by using portfolio. *Procedia - Social and Behavioral Sciences*, 191, 764 – 768.
- Sahinkarakas, Ş. (2012). The role of teaching experience on teachers' perceptions of language assessment. *Procedia - Social and Behavioral Sciences*, 47, 1787 – 1792.
- Tang, L. (2016). Formative assessment in oral English classroom and alleviation of speaking apprehension. *Theory and practice in language studies*, 6(4), 751-756.
- Taras, M. (2008). Summative and formative assessment: Perceptions and realities. *Active Learning in Higher Education*, 9(2), 172-192.
- Tarone, E. (1981). Some thoughts on the notion of communication strategy. *TESOL Quarterly*, 15(3), 285- 295.
- Thompson, G. (2009). Reevaluating the test specifications of an oral proficiency test. *The Journal of Kanda University of International Studies*, 21, 233-260.
- Thornbury, S. (2005). *How to teach speaking*. England: Pearson Education Limited Longman.
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research* 68(3), 249-276.
- Tuan, L. T. (2012). Teaching and assessing speaking performance through analytic scoring approach, *Theory and Practice in Language Studies*, 2 (4), 673 – 679.
- Underhill, N. (1987). *Testing spoken language. A handbook of oral testing techniques*. Cambridge: Cambridge University Press.
- Ur, P. (1996). *A course in language teaching: Practice and theory*. UK: Cambridge University Press.
- Van den Branden, K. (2006). Introduction: Task-based language-teaching in a nutshell. In K. Van den Branden (Eds.), *Task-based language education: From theory to practice* (pp.1–16). Cambridge, UK: Cambridge University Press.

- Wall, D. (1996). Introducing new tests into traditional systems: Insights from general education and from innovation theory. *Language Testing*, 13(3), 334–354.
- Wang, Z. (2014). Developing accuracy and fluency in spoken English of Chinese EFL learners. *English Language Teaching*, 7(2), 110-118.
- Weir, C. J. (1990). *Communicative Language Testing*. UK: Prentice Hall.
- William, D. & Black, P. (1996). Meanings and consequences: A basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal*, 22(5), 537-548.
- William, D. (2000). *Integrating summative and formative functions of assessment*. Retrieved from <http://journals.sagepub.com/doi/pdf/10.1177/1469787408091655>
- William, D., Lee, C., Harrison, C. & Black, P. (2004). Teachers developing assessment for learning: impact on student achievement. *Assessment in Education: Principles, Policy and Practice*, 11 (1), 49–65.
- Willis, J. (1996). *A framework for task-based learning*. Harlow, UK: Pearson Education.
- Wolf, D. P. (1989). Portfolio assessment: Sampling student work. In C. L., Artur (Revised Eds.,1). *Developing Minds: A Resource Book for Teaching Thinking*. (pp.351). Association for Supervision and Curriculum Development Alexandria, Virginia University.
- Wu, J.R.W. (2014). Investigating Taiwanese teachers' language testing and assessment needs. *English Teaching & Learning*, 38 (1), 1-27.
- Yoshida, M. T. (2016). *Beyond repeat after me: Teaching pronunciation to English learners*. USA: TESOL Press.

APPENDICES

A. Knowledge of Speaking Assessment Questionnaire

Section One: Background Information and Practices of Speaking Assessment

1. What is your gender?

Male Female

2. What is your age?

Under 24 25- 34 35 – 44 45 or over

3. What is your graduation degree?

College graduate Master’s Degree Doctoral Degree

4. How long have you been teaching?

1 year or less 2-5 years 6-10 years 11 or more

5. Do you assess speaking skills in your classrooms?

Yes No

6. How many times do you carry out speaking assessment during a semester?

Once Twice Three times More than three times

7. Do you use language assessment criteria provided by outside sources and/or develop your own criteria when you assess students’ speaking performances?

a. I use language assessment criteria provided by outside resources

b. I develop my own language assessment criteria

c. Both

Section Two: Teachers’ knowledge of speaking assessment

This section is divided into three parts. Please describe the degree of your awareness by placing a tick under a category which best reflects your knowledge of oral language assessment. (1= not good 2= moderate 3= good 4= excellent)

Part A: Knowledge of speaking assessment constructs

No	<i>When I design speaking assessment, my knowledge of considering...is...</i>	1	2	3	4
8.	...aspects of intonation, stress, and pronunciation				
9.	...accuracy				
10.	...fluency				
11.	...students’ interactive skills				
12.	...students’ ability to take turns in a pair/group work				

Part B: Knowledge of speaking assessment techniques and tasks

No	<i>I rate my knowledge of.....as...</i>	1	2	3	4
13.	...using <i>a student self-assessment technique</i> , when students are being orally assessed				
14.	...using <i>a peer assessment technique</i> , when students are being orally assessed				
15.	...using <i>a portfolio assessment technique</i> , when students are being orally assessed				
16.	...using <i>formative and summative assessment types</i> in assessing speaking ability				
17.	...using different speaking assessment tasks				
18.	...time factor when choosing or designing speaking tasks				
19.	...setting speaking assessment tasks in an appropriate level of difficulty				
20.	...using oral presentation tasks in speaking assessment				
21.	...using learner-learner joint discussion and decision-making tasks				
22.	...using role-play tasks in speaking assessment				
23.	...using interview tasks in speaking assessment				
24.	...using visual stimuli tasks such as pictures to provide a topic of conversation to the students in speaking test				
25.	...using re-telling story or text from aural stimuli as a speaking task in speaking assessment				

Part C: knowledge of assessment criteria used in speaking tests

No	<i>I rate my knowledge of.....as...</i>	1	2	3	4
26.	...developing assessment criteria used in a speaking test				
27.	...categories used in speaking assessment				
28.	...considering <i>accuracy</i> as a category of speaking assessment				
29.	...considering <i>fluency</i> as a category of speaking assessment.				
30.	...considering <i>communicative skills</i> as a category of speaking assessment				
31.	...considering <i>range of vocabulary use</i> as a category of speaking assessment				
32.	...considering <i>pronunciation</i> as a category of speaking assessment				
33.	...considering the use of <i>appropriate expressions</i> as a category of speaking assessment				
34.	...using an <i>analytic score</i> for each category in a speaking test				
35.	...using <i>a holistic score</i> in a speaking test				

Section Three: Influence of context on teachers' knowledge and Practices of oral language assessment.

Which of these following factors do you think has an influence on your decision making when you design speaking assessment? Please tick all the boxes that are relevant to you and explain why in the space provided below each item.

36. Class size

.....
.....

37. Curriculum objectives.

.....
.....

38. Lack of support by more experienced teachers.

.....
.....

39. Prescribed rules by school policy / educational system about what assessment criteria to use.

.....
.....

40. Lack of standardized oral language assessment criteria.

.....
.....

41. Lack of teacher training for learning how to make oral language assessment.

.....
.....

42. How do you carry out speaking assessment? Please explain your answer in the space provided.

.....
.....
.....
.....

B. Training Needs Analysis for Oral Language Assessment Survey

Section One: Training for Oral Language Assessment

1. Have you got any training on oral language skills assessment?

Yes

No

2. If you answered “Yes” for question 1, do you think it satisfied your needs in the process of assessment for oral language skills?

1

2

3

4

Very Limited

Limited

Adequate

Very adequate

3. On what aspects of oral language assessment do you still have training needs?

.....

Section -2) Teachers’ Perceived Training Needs

Please indicate whether you think you need to get training on this topic in a training program that you would attend on the assessment of oral skills.

(1= Not Important 2= Slightly Important 3= Moderately Important 4= Important 5= Very Important)

	1	2	3	4	5
1. Theoretical framework for approaches to assessment of oral language performance					
2. Highlights of oral language skill assessment in a language learning / teaching process					
3. Benefits of assessing oral language performance					
4. Deciding on main purpose of assessing oral language performance					
5. Specifying objectives in assessing oral language performance					
6. Assessment types for oral language skills					
7. Determining assessment criteria for oral language performance					
8. Deciding on what basic oral language skills to assess					
9. Deciding on what sub-skills to assess					
10. Preparing tasks /test items for assessing oral language performance					
11. Deciding on ready-made tasks / test items to be used while assessing oral language performance					

12. Preparing rubrics for oral language assessment					
13. Deciding on ready-made rubrics to be used while assessing oral language performance					
14. Task /test types that are used in oral language assessment					
15. Samples of tasks, tests, rubrics that can be used in the assessment of oral language performance					
16. Evaluating / Rating oral language performance					
17. Interpreting scores					
18. Reliability and Validity issues					
19. Ethical considerations (giving equal chance to students, influence of personal differences, grader's subjectivity...etc)					
20. Washback on the classroom					
21. Common European Framework of Reference (CEFR)					
22. Relating my assessment to CEFR					

C. Semi-Structured Interview

1. How long have you been teaching English?
2. How long have you been teaching English in a high school?
3. Do you assess speaking skills of your students?

**If the answer is "Yes" for the 2nd question;*

What criteria do you use for assessing speaking skills?

What kind of tasks do you use in your assessment?

Do you have any challenges in your assessment process?

** If yes,*

What kind of challenges do you have?

**If the answer is "No" for the 2nd question;*

What are the reasons for not assessing speaking skills?

4. Have you got any training for the assessment of oral language skills?

**If the answer is "Yes"*

Did it satisfy your needs for assessing speaking skills in classrooms you teach? Why / Why not?

5. Do you think you have adequate knowledge to assess speaking skills of your students?

**If the answer is "Yes"*

How and why? Can you explain in detail?

6. What do you think your training needs are for assessing speaking skills?

D. E-Mail Correspondence for the permission to use KOSAQ

10.05.2018 Gmail - A Request for "EFL Teachers' Knowledge of Speaking Assessment Questionnaire"

 Gmail Ayten Kaplan <ayten0329@gmail.com>

A Request for "EFL Teachers' Knowledge of Speaking Assessment Questionnaire"

3 messages

Ayten Kaplan ayten0329@gmail.com Tue, Dec 19, 2017 at 6:19 PM
To: tazizk_haled@hotmail.com

Dear Ms Taaziz Khaled,

I could get into contact with you through the kind assist of Susan Riley. I am writing my Master's Thesis on "Knowledge and Professional Training Needs of EFL Teachers for Speaking Assessment". I prepared a needs analysis scale; however, I still need a questionnaire to measure their current knowledge. While reviewing the literature, I came across your thesis and I read it. It was really helpful for me to draw my way better. I would really appreciate if you could grant permission to adapt the questionnaire, that you designed in your thesis "An Investigation into Libyan EFL Novice Secondary School Teachers' Current Knowledge and Practice of Speaking Assessment: A Socio-cultural Perspective", to my study.

Best Regards.
Ayten Kaplan

taziz khaled tazizk_haled@hotmail.com Wed, Dec 20, 2017 at 10:47 AM
To: Ayten Kaplan ayten0329@gmail.com

Dear Ayten,

Dr. Riley contacted me about that.
You can use the questionnaire and I hope you find it helpful. Your area of research seems quite interesting. I wish you all the best with your dissertation. Should you have further queries, please do not hesitate to contact me.

Kind regards
Taaziz

Sent from my iPhone

E. E-Mail Correspondence for the permission to use TNAFOLAS

10.05.2018

Gmail - A Request for "The LTA Survey"



Ayten Kaplan <ayten0329@gmail.com>

A Request for "The LTA Survey"

3 messages

Ayten Kaplan ayten0329@gmail.com

Tue, Dec 19, 2017 at 6:39 PM

To: jw@lrtc.ntu.edu.tw

Dear Ms Wu,

I am writing my Master's Thesis on "Knowledge and Professional Training Needs of EFL Teachers for Speaking Assessment". I am preparing a needs analysis scale to measure teachers' needs for speaking assessment. While reviewing the literature, I came across your study and I read it. It was really helpful for me to draw my way better. I would really appreciate if you could grant permission to adapt the related parts of the questionnaire that you designed in your study "Investigating Taiwanese Teachers' Language Testing and Assessment Needs" to my study.

Thank you for your interest and any help you can do in advance!

Best Regards.

Ayten Kaplan

Jessica Wu jw@lrtc.ntu.edu.tw

Wed, Dec 20, 2017 at 8:34 AM

To: Ayten Kaplan ayten0329@gmail.com

Dear Ayten Kaplan,

Thanks for your query. You are welcome to adapt my questionnaire for my study "Investigating Taiwanese Teachers' Language Testing and Assessment Needs". But, if you read the paper carefully, the questionnaire was developed based on previous literature. Please make it clear and cite all related literature properly in your thesis.

Best regards,

Dr Jessica Wu

R&D Program Director

Language Training & Testing Center

F. Curriculum Vitae

Personal Information

Surname, Name: Kaplan, Ayten

Nationality: Turkish (T.C.)

Date and Place of Birth: 10 October 1979, Kastamonu

Marital Status: Single

Phone: +90 212 261 63 03

Fax: +90 212 260 67 05

email: ayten0329@gmail.com

Education

Degree	Institution	Year of Graduation
BS	Dokuz Eylül University	2002

Work Experience

Year	Institution	Enrollment
2002-...	MEB	English Teacher

Foreign Languages

English