SELF-ASSESSMENT OF STUDENTS' SPEAKING SKILLS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY
HELIN BAŞAK

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

FEBRUARY 2019

Approval of the Graduate School of Social Sciences

_____
Prof. Dr. Tülin Gençöz
Director

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

_____
Prof. Dr. Çiğdem Sağın Şimşek
Head of Department

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

_____
Prof. Dr. Çiğdem Sağın Şimşek
Supervisor

**Examining Committee Members**

Assoc. Prof. Dr. Nurdan Gürbüz (METU, FLE)          _____

Prof. Dr. Çiğdem Sağın Şimşek  (METU, FLE)          _____

Assoc. Prof. Dr. Suzan Kavanoz    (Yıldız Teknik Uni., FLE)          _____

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 : Helin BAŞAK

Signature           :

# ABSTRACT


## SELF-ASSESSMENT OF STUDENTS' SPEAKING SKILLS


BAŞAK, Helin

M.A., English Language Teaching

Supervisor: Prof. Dr. Çiğdem Sağın Şimşek


February 2019, 131 pages

This study has the main aim to examine whether there is a significant difference between students' self-assessment of their speaking skills and teachers' assessment of students' speaking skills. In addition, the study also aims to determine whether students and teachers assess different components of speaking skill such as grammar, vocabulary, pronunciation, and fluency significantly different from each other or not. Finally, the study aims to investigate whether students' self-assessment show differences in terms of their proficiency level in English, gender, and exposure to English outside the school. The data were collected via a self-assessment scale given to the participants and English language teachers' assessment of the participants based on a speaking assessment scale. The results show that the teachers and the students assess the speaking skills significantly differently from each other. The mean score of the teachers' assessment is 79.2 out of 100 while the mean score of the students' self-assessment is 61.7. These scores conclude that the teachers assess the students' speaking skills higher than the students themselves. In a nutshell, students evaluate their vocabulary, grammar, fluency, pronunciation and communication skills lower than their teachers. The results also show how insecure

the students feel about their speaking skills as a whole. In addition, the results show that the students' self-assessment of their speaking skills and teacher assessment of speaking skills change depending on the students' level of proficiency, gender and exposure to English outside the classroom.

**Keywords**: Language assessment, self-assessment, learner autonomy

# ÖZ

## ÖĞRENCİLERİN KONUŞMA BECERİLERİNİ ÖZ DEĞERLENDİRMESİ

BAŞAK, Helin

Yüksek Lisans, İngiliz Dili Eğitimi Bölümü

Tez Yöneticisi: Prof. Dr. Çiğdem Sağın Şimşek

Şubat 2019, 131 sayfa

Bu çalışma, öğrencilerin konuşma becerilerinin öz değerlendirmesi ve öğretmenlerin öğrencilerin konuşma becerilerini değerlendirmesi arasında anlamlı bir fark olup olmadığını incelemeyi amaçlamaktadır. Buna ek olarak, bu çalışma aynı zamanda, öğrencilerin ve öğretmenlerin dilbilgisi, kelime bilgisi, telaffuz ve akıcılık gibi farklı konuşma becerilerini birbirinden önemli ölçüde farklı değerlendirip değerlendirmediklerini belirlemeyi amaçlamaktadır. Son olarak, bu çalışma öğrencilerin öz değerlendirmelerinin İngilizce yeterlilik düzeyleri, cinsiyet ve okul dışında İngilizce'ye maruz kalma açısından farklılık gösterip göstermediğini araştırmayı amaçlamaktadır. Veriler katılımcılara verilen bir öz değerlendirme ölçeği ve İngilizce öğretmenlerinin, katılımcıları bir konuşma değerlendirme ölçeğine göre değerlendirmesi yoluyla toplanmıştır. Sonuçlar, öğretmenlerin ve öğrencilerin konuşma becerilerini birbirinden önemli ölçüde farklı olarak değerlendirdiğini göstermektedir. Öğretmenlerin değerlendirmelerinin puan ortalaması 100 üzerinden 79,2 iken, öğrencilerin öz değerlendirme puan ortalaması 61,7'dir. Bu puanlar, öğretmenlerin öğrencilerin konuşma becerilerini öğrencilerden daha yüksek değerlendirdiği sonucuna varmıştır. Özetle, öğrenciler kelime,

dilbilgisi, akıcılık, telaffuz ve iletişim becerilerini öğretmenlerinden daha düşük olarak değerlendirirler. Sonuçlar ayrıca öğrencilerin bir bütün olarak konuşma becerileri hakkında ne kadar özgüvensiz hissettiklerini gösterir. Buna ek olarak, sonuçlar öğrencilerin konuşma becerilerini öz değerlendirmelerinin ve öğretmenlerin konuşma becerilerini değerlendirmelerinin, öğrencilerin yeterlilik, cinsiyet ve sınıf dışındaki İngilizce'ye maruz kalma düzeylerine bağlı olarak değiştiğini göstermektedir.

**Anahtar sözcükler**: Dil değerlendirmesi, Öz değerlendirme, Öğrenci Özerkliği

*To my father who never left me, and*
*to my mother who did more than my right hand.*

# ACKNOWLEDGEMENTS

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# LIST OF ABBREVIATIONS

| | |
|---|---|
| **CEFR** | Common European Framework of Reference |
| **EFL** | English as a Foreign Language |
| **ESL** | English as a Second Language |
| **IELTS** | International English Language Testing Standards |
| **TOEFL** | Test of English as a Foreign Language |

# CHAPTER 1

# INTRODUCTION

In this chapter of the study, the purpose, background and the significance of the study, research setting, research questions and assumptions are presented.

## 1.1. Purpose and Background to the Study

The main aim of this study is to investigate the contribution of self-assessment to students' learning and assessment process. Self-assessment is an alternative assessment type which requires students to assess their language learning process; how successful they are during the learn process and whether they achieve their goals or not. According to Nunan (1988) and Oscarson (1989), self-assessment, as an assessment tool, increases students' understanding of their own language skills by enabling them to realize their strengths and weaknesses with a reflective approach. Besides, self-assessment enables learners with plenty of chances to reflect on and assess the methods and techniques utilized during their learning process.

Numerous studies conducted on self-assessment highlight that including self-assessment in student evaluation process increases learner autonomy. In the recent history of language learning, learner autonomy, which can briefly be defined as reflecting and being in charge of learners' own learning process, has been a significant concern (Holec, 1981; Dickinson, 1987; Little, 1991; Dam, 1995; Brown, 2007; Balçıkanlı, 2010). The notion of learner autonomy was first put forward by Henri Holec (1981) who described learner autonomy as a necessary skill that requires students to take responsibility for their own learning through natural or educational means. According to Holec, autonomous students are able to implement

their language competency and skills outside the classroom. Therefore, learner autonomy goes beyond the classroom atmosphere and permanently develops the learner's linguistic awareness (Najeeb, 2013).

It is alleged by Holec (1981), Allwright (1990) and Little (1991) that autonomous students are able to reflect on their own language learning process with their knowledge about their learning and are enthusiastic to collaborate with their peers. These students determine the goal of the learning process, take charge of their learning, share knowledge, plan and manage learning techniques and review them. In order to achieve all of these, students are required to be qualified and authorized so as to apply their autonomy. An important characteristic of autonomous students is their ability to make reflections on their own language learning process, and to make realistic assessments of their success during this process. Najeeb (2013) suggests that for a fruitful autonomy process, continuous assessment of the learning process done by teachers, peers and through self-assessment is essential.

Self-assessment does not only contribute to the development of learner autonomy, but it also assists students to get involved in the learning process and in turn increase their motivation (Nunan, 1988; Oscarsson, 1989). Since students have to make judgements about their own learning, this may lead to positive attitudes, and therefore, higher motivation towards learning (Nurov, 2000). It was also stated by Gardner (1996) that self-assessment could serve numerous purposes such as building confidence and increasing motivation. Besides, according to Tudor (1996), increased student motivation and the awareness that self-assessment contributes to language education might help students to evaluate their language skills more accurately (Tudor, 1996).

Studies also report that self-assessment lessens the negative emotional influences of the traditional assessment methods. For instance, negative emotional outcomes such as anxiety, stress and fear of being assessed by teachers or other evaluators become irrelevant when students assess their own performance (Nurov, 2000).

In addition, the practicality and cost-effectiveness of self-assessment are two more reasons for language institutions to implement self-assessment in their programmes. Self-assessments are regarded as more practical than the traditional assessment methods because they are easier to design, construct, and apply and the process takes less time. (Brown and Hudson, 1998; LeBlanc and Painchaud, 1985).

Furthermore, a great number of studies shows that self-assessment can be a reliable and valid alternative assessment tool and that there is frequently a statistically good correlation between teachers' and students' assessments. Also, the outcomes of research on self-assessment indicate that self-assessment can serve as a valid and reliable measurement tool (Bachman & Palmer, 1989; Buck, 1992; MacIntyre, Noels, & Clement, (1997); Ross (1998).

Self-assessment of language skills has also been one of the major concerns in the field of language assessment research. There are numerous studies conducted on different language skills with participants from nearly all ages and with almost all proficiency levels in various learning environments. For instance, there are studies correlating self-assessments of receptive (reading and listening skills) and productive (writing and speaking skills) skills (LeBlanc and Painchaud, 1985), and correlating reading, writing, listening, speaking skills, grammar and vocabulary development separately with one another (Clark, 1981). However, most studies on self-assessment seem to investigate self-assessment of reading, listening and writing skills but neglect speaking skills. In one of the few studies focusing on self-assessment of speaking skills, Harris (1997) states that self-assessment of oral abilities is harder to design, organise and apply.

Thus, studies on self-assessment do not, usually, focus solely on the self-assessment speaking skills but are generally converged all four skills (Krausert, 1991; Ross,1998; Deville and Deville, 2003). Furthermore, these few studies conducted on self-assessment of speaking skills present contradicting results particularly about the validity and reliability of the self-assessment of speaking abilities. Therefore, it is crucial to conduct a study based only on the self-assessment of speaking skills to contribute to the continuous discussion in self-assessment research.

## 1.2. The Significance of the Study

Review of the relevant literature displays that research on self-assessment shows inconsistent results. From one point of view, a considerable amount of research indicates that self-assessment can be acceptable in terms of its reliability and validity, and there is frequently a good correlation between teachers' and students' assessments (Bachman & Palmer, 1989; Buck, 1992; MacIntyre et.al, 1997; Ross, 1998). However, other studies conducted on self-assessment point out that self-assessment may not be able to display a valid image of students' skills (Blanche, 1990; Heilenman, 1990; Nurov; 2000). In respect to this, it has not been absolutely confirmed whether self-assessment can be accepted as a valid and reliable tool to assess students' language abilities or not.

As far as skills-based self-assessment research is concerned, the literature shows that there are more studies conducted on receptive skills than productive skills. In other words, there is plenty of research concentrated on listening and reading skills and less on speaking and writing.

There are very few studies conducted on self-assessment of language skills in the Turkish context. These studies are generally focused on the reliability of self-assessment tools and students' beliefs about self-assessment. There is only one study on students' self-assessment of their speaking skills in Turkish context completed by Burcu Karakaya in 2017. This study is conducted with only intermediate level (B1) students, and its aim is to investigate the effect of self-assessment application on students' speaking skills improvement. That's why conducting research that focuses on the concerns mentioned above is both significant and essential. The present study aims to determine whether there is a distinction between the self-assessment of the English speaking skills of Turkish learners and teachers' assessments of the learners' English speaking skills. Another purpose of the study is to investigate the effect of students' proficiency level in English, students' gender and their exposure to English via being engaged with extracurricular activities on students' self-assessment of their speaking skills.

To conclude, the present study aims to shed light on whether students and teachers assess speaking skills in English differently from each other or not and how students assess their speaking skills in English in general and in particular for different components of speaking skills. Thus, the present study aims to contribute to the past research by providing insights about self-assessment of speaking skills.

## 1.3. Research Setting

The present study is conducted at TOBB University of Technology and Economic (TOBB ETU). TOBB ETU was established in 2003 by the TOBB Foundation. It has six faculties, 21 departments and approximately 6000 students. Every year, about 1000 students are registered to study their programmes. The university has a three-term educational year, which is different from the other universities in Turkey. Students are supposed to have three internships for every $3^{rd}$ semester, starting from their sophomore year at the university.

All departments, except the Department of Turkish Language and Literature and the Department of Visual Communication and Design, are instructed in 30% English and 70% Turkish. That's why students registered to their departments with English instruction have to attend English Preparatory Programme. In other words, the students who are registered to Turkish Language and Literature and Visual Communication and Design departments can directly start their education in their departments. Other students have to start with the English Preparatory Programme.

The English Preparatory Programme is applied by the Department of Foreign Languages. There are five different levels in the programme: AF (beginner), A (elementary), B (Pre-Intermediate), C (intermediate) and CR (Intermediate students who repeat the programme for the second time.). These levels are determined according to the exams that students who are registered to the programme have to take at the begginig of the education year. The preliminary exam is the first one that is conducted at the beginning of the programme and it assesses listening comprehension (60 questions), structure (40 questions of grammar & vocabulary) and reading comprehension (25 questions). The proficiency exam is the second

exam and is applied after the preliminary exam. The students who have a score of 65 out of 100 from the Preliminary exam are able to take the Proficiency exam. The Proficiency exam is TOEFL Institutional Testing Programme (ITP) which tests listening comprehension (50 questions), structure (40 questions of grammar) and reading comprehension (50 questions). Students who score 500 out of 677 from TOEFL ITP are exempt from the preparatory programme and are able to continue to their departmental courses. Students who have a score lower than 500 are obliged to start to the preparatory programme.

## 1.4. Research Questions

The present study aims to provide answers to the following research questions:

- **Is there a significant difference between students' self-assessment scores and teachers' assessment scores regarding students' speaking skills in English?**

- **Is there a significant difference between students' self-assessment scores and teachers' assessment scores regarding components of speaking skills?**

- **Do students' self-assessment of speaking skills scores vary as students' proficiency level increases?**

- **Is there a significant difference between female and male students' self-assessment of their speaking skills?**

- **Do students' self-assessment scores of speaking skills vary in relation to students' daily use of English in extracurricular activities?**

## 1.5. Assumptions

This study on the self-assessment of speaking skills investigates whether there are any similarities or differences between students' self-assessments of their speaking skills and teachers' assessment of the participants' speaking skills in English. The study has the following assumptions.

To begin with, significant differences between the teachers' and students' assessment of the students' speaking skills are expected. Relevant literature reveals that teacher assessments are not in accordance with students' assessments simply because of the self-assessment's variation in validity (Nurov, 2000). In line with the research findings presented in the literature, it is expected that teachers' assessment scores will be higher than the students' self-assessments.

Another important difference is expected in terms of proficiency level. In the study, there are students with different proficiency levels of English which might lead to differences in the way they assessed themselves. The students with higher proficiency levels will probably assess their speaking better than the students with lower levels because the higher proficiency of English will enable the students to have a better perspective of what it means to have good speaking skills.

In addition, the self-assessment scores of students' speaking skills might show significant differences within the groups due to gender differences. In accordance with the previous research, it is expected that male students will have significantly higher mean scores of self-assessments than female students. Relevant literature indicated that male students tend to be more autonomous and confident towards learning. In a study by Wright and Houck (1995), it was stated that the male students may have a tendency to evaluate their verbal and numerical abilities higher than the female students (Wright and Houck, 1995).

The students' use of English in extracurricular activities might create differences in the students' self-assessment scores because use of English in after-school activities may enhance their English speaking skills and accordingly, the more they are exposed to the English language, the more self-confident they might become.

# CHAPTER 2

# LITERATURE REVIEW

## 2.1. Introduction

The purpose of this study is to determine whether there is a significant difference between students' self- assessment of their own English speaking skills and teachers' assessment of students' speaking skills. In this chapter, the description, aims and techniques of both language assessment and self-assessment in particular will be examined. In addition, self-assessment of language skills, ways of assessing speaking skills and previous studies focusing on self-assessment will be presented in order to provide in-depth information about self-assessment and its related concepts.

## 2.2. Assessment and Testing

Assessment is the process of evaluating students' learning process and attained performance (Schmitt and Schmitt, 2014). Although assessment and testing are inclined to be misunderstood as synonymous terms, they refer to different concepts. Testing is a procedure that is prepared for administrative purposes and happens periodically throughout a curriculum. On the other hand, assessment is an ongoing and a much broader process (Brown, 2004). For instance, teachers consciously or subconsciously assess their students whenever students answer a question, propose a new comment or try to make a sentence with a newly-acquired grammar point in class. Although these assessments can be deliberate or casual, a successful teacher never stops to assess their students. Assessment is an inevitable component of the

learning process as it allows teachers to monitor their students' progress of learning. However, assessment is a superset of tests, and tests are not necessarily the mere devices for teacher assessments. Despite the fact that they can be beneficial media of assessment, they are one of the devices and techniques that teachers utilize so as to assess their students' learning process and attained performance (Brown, 2004).

Assessment types can be grouped as formative assessment and summative assessment. It will be beneficial to explain these assessment types before moving onto self-assessment.

### 2.2.1. Summative Assessment

To start with, summative assessment is a type of assessment which is used after the learning process in order to understand what students have learnt (Brown, 2004). It provides information about whether students achieved learning targets or not. Summative assessment is generally applied to determine whether students are qualified to be given certificates, complete school or receive a grade etc. It is generally used to make decisions about further education (OECD Centre for Educational Research and Innovation, 2008: 1). For instance, a teacher may change his/her teaching techniques after the learning process is completed based on the results of summative assessment.

One of the most widely used summative assessment types is standardized tests. There are serious debates about the advantages and disadvantages of standardized tests. In spite of the fact that authorities and many teachers do not approve of summative tests to be used to gain information about students' success, standardized tests are widely used. Moreover, many teachers generally think that the data gathered via summative tests do not provide authentic information (Kordes et. al., 2014).

Despite the disadvantages, there are some reasons why institutions and teachers prefer summative tests. Qu and Zhang (2013) summarize some advantages and disadvantages of summative assessment as follows. The first advantage of

summative assessment is that teachers have the chance to examine the exam results to monitor future lesson plans since summative assessment enables fairly precise quantitative data for analyzing teaching techniques. For instance, in this year's final examination, teachers found out that students obtained the lowest average grades from vocabulary. That's why, they decided to consolidate students' vocabulary knowledge in the upcoming semester. Secondly, it can also assist learners to figure out their weaknesses and modify their study habits accordingly. On the other hand, summative assessment has disadvantages, too. To begin with, learners may prefer to guess the correct answer in a test when they run out of time or do not know the correct answer. Therefore, it can be stated that learners may have the help of their luck and the results of the test may not show the real knowledge or abilities of the students, which can also be misleading for teachers. (Qu and Zhang, 2013).

Some types of summative assessment are examinations (major, high-stakes exams), final examination (a truly summative assessment), student evaluation of the course (teaching effectiveness), and instructor self-evaluation ("Formative and Summative Assessment", n.d.: 3). As seen, summative assessment provides information about the final judgments of students' achievement and the effectiveness of the learning process.

## 2.2.2. Formative Assessment

The other assessment type is formative assessment. Formative assessment is the act of assessing students during the process of learning language components and skills with the aim of assisting them to maintain their language development (Brown, 2004). According to Brown (2004), approximately every type of assessment is (or should be) formative simply because of practical reasons since the primary focus of formative assessment is on the ongoing growth of the student's language. Thus, when a student is given a comment or a suggestion or correction on an error, feedback is given with the intent of improving students' language skills. A teacher may change his/her teaching techniques during the learning process thanks to the results of formative assessment. Taking that fact into account, it may be claimed that formative assessment contributes to the learning process more than summative

assessment (Garrison and Ehringhaus, 2010; Kordes, Sicherl, & Holcar, 2014). Kordes et.al. (2014), in their study, favoured formative assessment over summative assessment. According to the results of the study, the students have positive attitudes toward assessment when formative assessment procedures such as peer-assessment and descriptive assessments are used. When quality feedback is provided, improvement of independent learning and learning achievements are supported. This situation helps in developing meta-knowledge, in promoting self-assessment, self-regulation and further learning. According to the authors, learning is improved when students experience and discover information themselves with the help of research, group work, aiming and taking their interests into account. Learners are encouraged by real-life learning and by understanding the sense behind the learning process (Kordes et.al., 2014).

Some types of formative assessment, also types of alternative assessment, are term papers (drafts submitted throughout the semester would be a formative assessment), projects (project phases submitted at various completion points could be formatively assessed), portfolios (could also be assessed during its development as a formative assessment) and performances ("Formative and Summative Assessment", n.d.:3).

## 2.3. Self-Assessment

Self-assessment can be considered as a part of formative assessment. In the self-assessment process, students make judgments about the degree of success of their work. During that process, goals and evaluation criteria must be taken into account. Teachers should make the assessment criteria clear and understandable for the students so that students are able to understand the strengths and weaknesses of their work and revise it accordingly. Quality of work is determined by students (Spiller, 2012). It is emphasized that the principles of formative assessment are put into operation with the intent to control the learning progress and to help give corrective feedback to enhance learning (Gronlund and Cameron, 2004).

Students use self-assessment to assess their own performance and to determine their language skills and competencies (Brown, Andrade, & Chen, 2015). According to

Blachford (1997), self-assessment is a process including comparisons of a learner's own success with other learners. It is also defined as an appraisal of students on their own learning performance by Montgomery (2001). Self-evaluation and self-monitoring are two other terms that are usually related to self-assessment. The former requires students to judge the quality of their own performance, which depends upon evidence and explicit criteria with the intent of future language growth (Rolheiser and Ross, 2000). The latter is a term, put forward by Dickinson (1987), referring to the processes that involve record keeping applied with learning diaries, record cards and sheets (Chalkia, 2012).

Self- assessment is also regarded as an alternative assessment method which many scholars suggested to be applied in the field of language assessment. Self-assessment has been implemented as an indispensable part of assessment and an important element of autonomous language learning programs by many scholars around the world because a significant part of language programs has become more student-centred (Nurov, 2000).

The implementation of self-assessment in the classroom is thought to provide plentiful advantages for not only teachers but also students. Since learning a language is an ongoing process, students ought to be able to progressively evaluate their language performance and skills in everyday situations such as interacting with native speakers, comprehending a newspaper, seeing a TV program or producing a text in the target language (Dickinson, 1996). That's why, Cameron (2004) stated that learners who know how to assess their own performance transform from being "other-regulated" to "self-regulated and/or autonomous".

They also gain the ability to control their own progress, assess their competence, manage their learning and determine how to make use of the tools and facilities provided within and outside the classroom (Chalkia, 2012). Hence, self-assessment makes students active participants in their own learning processes (Joyce, Weil, and Calhoun, 2009).

Furthermore, Gardner (2000) claimed that confidence and motivation is enhanced by a prospering self-assessment application. In other words, self-assessment has a motivational effect. It is also highlighted that student motivation enhanced by self-assessment affects classroom activities and leads students to achieve communicative goals (Oscarsson, 1989), which in turn improves their intrinsic motivation.

Another advantage of self-assessment is that teachers themselves can benefit from it for practical purposes. Since students are involved in the assessment process, the assessment burden is shared between students and their teachers. This decreases teachers' workload and assists them to deal with the other elements of teaching such as creating resources and materials (Blue, 1988). Self-assessment is also suggested to be practical for teachers since self-assessment tools are easier to prepare and administer compared to other traditional assessment types (Nurov, 2000).

In addition to the reasons and aims to utilize self-assessment, it has also some benefits for learners. Spiller (2012) states that the first of these benefits is improved employability. Self-assessment helps students get better information on their performances. The other benefits include higher motivation for learning and advancement, equality of opportunity, active involvement, raised confidence and self-respect, increased value for learners' comments, and views and support of the learning programme (Spiller, 2012).

Additionally, McMillen and Hearn (2008) express that self-assessment is a vital skill that raises learners' motivation and success because it helps a learner self-monitor, self-evaluate and determine correct methods to improve their learning. Several authors claim that when students have a particular aim which can assist them with their developed knowledge and then determine criteria, self-assess and reflect on their performance and produce techniques for learning, they will demonstrate developed competencies with higher motivation. It can be inferred that developing students' self-efficacy and self-confidence in learning process is the main aim and this could be accomplished by self-assessment (McMillen and Hearn, 2008).

Self-assessment has benefits not only for students but also for other shareholders. O'Malley and Pierce (1996) state that

> The self-assessment provides a means for stimulating ongoing dialogue and collaboration between program directors and instructional staff around the interrelated goals of effective classroom practices, professional development, and positive learner outcomes. Not every teacher in every programme will need to become proficient in each competency. (O'Malley and Pierce, 1996).

They also add that under some circumstances, not all the competencies may be relevant. For instance, if teachers are in an environment in which there is no computer or any other technological devices, it would be insensible to expect competencies on technology.

Teachers can also utilize self-assessment to evaluate their own performance, as well. To improve one's professional development, it can also be assigned as a confidential instrument. This instrument might motivate teachers to reflect on their own progress and to determine their strengths and weaknesses. It might, also, be beneficial to create priorities for the professional development of teachers (O`Malley and Pierce, 1996).

There are a number of studies conducted on self-assessment in EFL/ESL contexts. To start with, Huang (2016) aimed to examine the effect of self-assessment and self-feedback of university students' performance captured by audio files. Participants were asked to listen to, analyse and transcribe their own performances. According to the author, "results indicated that learners' self-feedback was far reaching and multifaceted. Through self-feedback, learners' identified discrepancies, answered feed up, feedback and feed forward questions, and inspected performance at task, process, self-regulation and self-levels. Much of the feedback involved reflections on past learning history, other areas of learning, deviation of performance from preparation and learner personality traits. The self-feedback went largely beyond most teachers' feedback capacity and bore great potential for learning and instruction. In particular, contrary to theoretical presumptions, self-level feedback seemed quite enlightening. Whether the observed quality self-feedback could

actually help learners improve their performance, however, was not clear" (Huang, 2016).

Another study conducted by McDonald and Boud (2003) analysed the question of whether self-assessment practices can develop the quality of the performances of high school students. The researchers randomly selected teachers from a sample of high schools and trained them how to improve learners' self-assessment abilities. Ten high schools were chosen from different success rates according to national examination results and the students received training on self-assessment by their regular classroom teachers. The experimental group was comprised of 256 students who received the training, but the control group did not receive any training. The results revealed that there was a significant difference between the two groups in favour of the group who was trained on self-assessment. In other words, the study claims that the self-assessment training may influence the quality of students' work (McDonald and Boud, 2003).

Another study on self-assessment analyses a different aspect of self-assessment, namely; attitudes of students and teachers towards self-assessment. There has been less research in this aspect when compared to learners' involvement in language testing and evaluation. The paper aimed to compare students' and teachers' attitude towards self-assessment. The participants consisted of 20 learners of English and 2 English language teachers in the UK. A structured questionnaire and a follow-up interview were used as data collection tools, and the data was analysed through SPSS. The results showed that not only the students but also the teachers found self-assessment application useful, and they had positive attitudes towards self-assessment (Ibberson, 2012).

Last but not least, LeBlanc and Painchaud (1985) examined whether self-assessments could be used as placement tests. The participants consisted of 200 students studying in Ottowa University with either English or French as a second language. The data collection tools were a self-assessment questionnaire and a proficiency test taken by the students. The results showed that the overall scores of the self-assessment survey and the scores of the proficiency tests correlated. When

the subset scores were analysed, it was found that the listening, reading and cloze test parts of the proficiency test and the questionnaire also highly correlated. As a result, it was stated that self-assessment may be used as a placement test when proper monitoring can be established (LeBlanc and Painchaud, 1985).

### 2.3.1. Reliability and Validity of Self-Assessment

Reliability and validity are two important criteria for any assessment method to be accepted and applied in a real language evaluation environment. Reliability refers to the consistency of the results in repeated measurements. In other words, reliability is the extent to which a measurement produces the same results with the same test-takers and under the same conditions. With respect to self-assessment, the question is whether self-assessment is a reliable measurement tool or not. Namely, to what extent that student can be objective with assessing their language abilities and how reliable students' assessment can be are also other questions that should be asked at this point. Therefore, it can be claimed that self-assessment could be used as academic objectives of an EFL classroom as long as students can provide reliability in their self-assessment (Karakaya, 2017). As reliability in EFL context is illustrated by Huerta-Macias (1995), if a student is asked to evaluate his/her understanding of a movie or a part of his/her speaking performance based on the same criteria now and after a while, the two assessments will not be very different from each other.

As for validity, it is related to the accuracy of the measurement tools. Simply, validity refers to whether the assessment tool measures what it is required to measure or not. When it comes to the validity of self-assessment tools, the question is whether self-assessment is able to assess what it should assess or not. According to Gardner (1996) and Heilenman (1990), since self-assessment has the potential to model the learning environment and authentic assignments that students deal with in everyday situations, it has the capacity to be really reflective and valid under teachers' control (Gardner, 1996; Heilenman, 1990). However, some researchers have concerns about the validity of self-assessment in the EFL/ESL classroom. For example, it is alleged that students may have difficulties in self-assessing themselves

16

accurately because they may not be aware of their own certain behaviours no matter how well trained they are for self-assessment.

Also, subjectivity and peer pressure are other factors that can affect validity of self-assessment. Students may want to assess themselves as high as they can in order to look better in the classroom (Karakaya, 2017). Moreover, if students are not provided with accurate criteria and a rubric for self-assessment, it may have a negative impact on accuracy of self-assessment (Cohen, 1994). Lastly, Rolfe (1990) claimed that students, from time to time, are likely to take self-evaluation more seriously than they are supposed to and they may give lower scores to themselves than their teachers would give (Rolfe, 1990).

Reviewing the relevant literature, it is noticed that there are some studies conducted on the validity and reliability of self-assessment and their results revealed that students may not be consistently accurate with self-assessment of their own language abilities.

To exemplify such as case, in a study conducted by Moritz (1997), the cognitive processes and the social impact behind self-assessment of students' foreign language skills were investigated concentrating on the process instead of the outcomes of the self-evaluation. The participants were 28 learners of the French language with different proficiency levels. These students were asked to complete a self-assessment questionnaire about the four language components (reading, writing, listening and speaking). Apart from the questionnaire, the participants were also asked to elicit verbal reports through a think-aloud protocol and an interview. The data was analysed and the results revealed that the learners showed different assessment for their own language skills. Some of the participants thought that their speaking skills were better than others although their achievement scores were lower in speaking. Also, there were some students who thought that they were better at writing than speaking while the scores were vice versa. The author claimed that these discrepancies between students' assessments and achievements scores may be due to the limited function of questions in the survey which may be not measuring

specific language skills. The author stated that students' understanding of their language skill is complicated and influenced by many elements. (Moritz, 1997).

In addition, in Ross's study (1988), the validity of self-assessment in terms of language skills was examined. The participants were 236 "just-instructed" ESL learners and they were asked to fill in a self-assessment questionnaire. The other data collection tool was teachers' assessments of these 236 students. When the data was analysed, the results showed that there were valid results for self-assessments compared to the teacher assessments and the listening parts of these two assessments were correlated, as well. It can also be stated that learners were more accurate in self-assessment in terms of functional skills. If the skills are abstract, the self-assessment success rate decreases (Ross, 1988).

Another study was conducted in Turkey by Nurov (2000). The relationship between students' self-assessment and teacher assessment was investigated and students' achievement results were also compared with these two variables in the study. The data collection tools are two questionnaires prepared for students to assess their own language abilities and for teachers to assess their students' English proficiency. The achievement test scores were also used as another measure for language skills assessment. The study showed that the students' self-assessments differ from the teacher assessments and the achievement test results. When language components are taken into account, there was no correlation between these three measurements in terms of vocabulary, for example. Also, gender did not create a significant effect in students' self-assessments even though students' proficiency level did. Students with higher success on the achievement test were apt to underevaluate their language skills when compared to their peers with lower success (Nurov, 2000).

In their study, Babaii, Taghaddomi & Pashmforoosh (2015) stated that the learning success or failure may be affected by the mismatches between teachers and students and self-assessment may be used as a tool to decrease these mismatches. That's why the authors investigated whether the students' and teachers' assessment of students' speaking skills would match after the students were provided with a practice session of a self-assessment application. There were 29 learners of English and 6 English

language teachers in the study as participants. First, the students listened to and assessed their audio-record speaking performances. Later, they were provided with the assessment criteria and they were asked to assess their performance again and write reflection papers on self-assessment application. In the meantime, teachers also assessed the students' speaking performance with the same criteria. According to the results of the study, there was a significant difference between students' first and second assessment of their own speaking skills. The study showed that the discrepancies between the teacher and student assessment were decreased after the students were provided with the assessment criteria. The students' reflection papers also revealed that students were positive towards the application of self-assessment (Babaii et al. 2015).

Moreover, Gardner and Lambert (1972) examined the correlation between students' self-assessment and French tests including reading, vocabulary and grammar components. The participants were 561 American students learning French in high schools in the states of Lousiana, Maine and Connecticut. The results showed that the correlation was not statistically significant confirming that self-assessment was not a reliable and valid tool for assessment (Gardner & Lambert,1972).

In his research, Blue (1988) studied whether there is a meaningful relationship between students' self-assessments and teachers' assessments. The participants were 117 students from different backgrounds and different majors, studying EAP (English for Academic Purposes) courses who were asked to fill out a self-assessment survey in the first and last lesson of a language course. The teachers' assessments were analysed as external criteria. At the end of the study, it was found out that the association between teachers' and students' assessment is rather weak (Blue,1988).

In conclusion, it can be pointed out that various restraining and interfering variables influence the self-assessment process which is already a complex construct. The majority of the studies indicate that self -assessment is a reliable and valid alternative assessment tool. However, its reliability and validity may be relatively thwarted under certain circumstances and different classroom settings. There are a

variety of causes leading to this problem such as the students' training. If students are trained on how to apply self-assessment, they will be able to make logical judgements about their language learning. Another reason is students' individual differences such as cultural background, level of education, and length of exposure to the target language. Also, methodological inadequacies might result in reliability and validity issues in research. Finally, researches do not apply reliability and validity test s on their data collection tools and might also use different statistical tests for correlation and comparing data, which might also lead to conflicting results in research (Nurov, 2000).

### 2.3.2. Models of Self-Assessment

According to Dickinson (1996), Rea-Dickins & Germaine (1996) and O'Malley & Pierce (1996), there are different methods for self-assessment skills such as "learner-prepared self- assessment, self-assessment checklists, informal self-assessment devices, and learner record keeping procedures." In addition, stronger and weaker models of self-assessment are asserted by Taras (2010) such as "self-assessment with integrated tutor feedback, learner contract design, self-marking and sound standard." Finally, other models are alleged by Saito (2003) such as "performance-oriented self-assessment and development-oriented self-assessment." In this section, these methods of self-assessment will be explained.

### 2.3.2.1. Learner-Prepared Self-Assessment

Tests which are prepared by students constitute the learner-prepared self-assessment method. These tests include comprehensive skills questions (questions for reading and listening), self-assessment forms for writing skills and self-monitoring questions. The questions are prepared according to the pattern laid out in published books and can be responded by learners themselves or by their peers after some time. In addition, learners themselves may evaluate their written works such as essays, journals or diaries (Dickinson, 1996). Moreover, in self-monitoring procedures, students are supposed to compare their performance with other models

and alter theirs accordingly so that they can measure their own performance closer to the model provided by their teachers.

### 2.3.2.2. Self-Assessment Checklists

Self-assessment checklists are a type of tests which is comprised of a list of assignments and questions that students are supposed to answer. To illustrate, a survey requiring learners to evaluate their performance of their language skills with a specific rating scale may be one such self-assessment checklists.

### 2.3.2.3. Informal Self-Assessment Devices

Informal self-assessment tools are developed and applied by learners to evaluate their performances in everyday life situations and include elements such as speaking with a native speaker they may encounter in real life, reading a newspaper, watching a TV show and listening to a radio broadcast in the target language (Gardner, 1996; Oscarsson, 1989).

### 2.3.2.4. Learner Record-Keeping Self-Assessment

In the learner record keeping process, some kind of self-assessment and implementation of it have been applied in the students' recent past and a self-appraisal scheme is included in the recording of the material. Record keeping is categorized into three devices by McNamara and Deane (1995): writing letters, keeping a dairy on language learning and keeping a portfolio. Also, learners' aim cards are suggested as a self-assessment recording tool by (Smolen, Newman, Wathen, & Lee, 1995).

### 2.3.2.5. Strong and Weak Models of Self-Assessment

According to Taras (2010), there are some stronger and weaker models of self-assessment. Stronger models of self-assessment are defined by Taras (2010) as follows:

a) Self-assessment with integrated tutor feedback: This model requires learners to integrate tutor and peer feedback before students self-assess.
b) Learning contract design: The learning contract design was developed in the context of self-directed learning (and is self-assessing by the students' own criteria).

These models have been regarded as strong models by Taras (2010) because the rationale behind them is that when students assess themselves, they are free from the emotional stress of getting a certain score and are able to make use of tutor and peer feedback. Furthermore, students are permitted to assess their own performances and learning despite strict curriculum limitations, which makes them more autonomous learners. On the other hand, weaker models are defined as follows:

a) Self-marking: In self-marking, students use a model answer(s) with criteria (and possibly mark sheets) to compare to their own work.
b) Sound standard: The tutor provides an objective descriptor of 'sound standard' that is of medium level work.

The standard model of self-assessment, on the other hand, requires learners to use criteria to judge, provide feedback and grade their work prior to submission to tutors. Students were asked to declare their expected mark when submitting work, and to accompany it with a note of the strengths and weaknesses which they had perceived and which had contributed to that judgment. Tutors marked in the usual way, but commented in endorsement of, or in addition to, the students' comments (Taras, 2010).

These models have been asserted as weaker models by Taras (2010) since in these models students implement the identical process as their teachers' assessment process. These models can also be regarded as behaviourist models because there is a sample question and correct answer frame that is used to compare and contrast students' answers in a systematic way (Taras, 2010).

When these strong and weak models are taken into account, it can be stated that the self-assessment process requires students to have some important skills. For instance, students must have the ability and experience of self-monitoring in order to make it a beneficial self-assessment process. Also, the self-assessment process may not provide accurate results about weaknesses, strengths, progress and learning techniques when student are not able to monitor their own learning process. During self-monitoring, students are required to concentrate on the process and to pay attention what they are doing for an effective self-monitoring. While doing this, evaluation standards which are not determined by students but determined by teachers and authorities, must be prepared thoroughly and revised if needed (McMillen and Hearn, 2008).

First, self-assessment can be seen as difficult for a student. He/she may make many mistakes when assessing their performance. Consequently, teachers have to monitor the process and provide corrective feedback about the mistakes made by students during the self-assessment process, which may be considered a real burden for the teacher. However, when time passes, students will get used to evaluating their development. They will take the responsibility for their assessment process so teachers can spend less time on assessment.

### 2.3.2.6. Performance-Oriented and Development-Oriented Self-Assessments

Apart from Taras's models, there are other self-assessment models as well. To start with, one is "Performance-oriented self-assessment" (Saito, 2003). Outcomes related to placement, selection, achievement, certification, diagnosis tests etc. are measured via performance-oriented assessment. For example, students are asked about their abilities and competencies during performance-oriented self-assessment

at a placement test (Saito, 2003). Performance-oriented self-assessment has some benefits in terms of safety, cost and time issues. These advantages attract test administrators to carry out self-assessment in their institutions. However, it may be influenced by elements such as the wording of the items, language skills to be evaluated, students' proficiency levels and cultural backgrounds (Saito, 2003). That's why it may require teachers or administrators to prepare and pilot the process before the actual self-assessment.

It is also beneficial to explain what the development-oriented self-assessment is in order to understand self-assessment types more clearly. In development-oriented self-assessment, the process of learning, generally in a classroom atmosphere, is measured in cooperation with self-managed activities. Dörnyei (2001) stated that development-oriented self-assessment can be used as a review for the learners to notice the differences and patterns of improvement in a long time period. Many experimental researches show that learners have shown greater competence, autonomy, and motivation and less failure when they are exposed to development-oriented self-assessment (Saito, 2003). This type of self-assessment both enhances learner autonomy and assists teachers to evaluate learners' development throughout the learning process. It may also be of assistance as an auxiliary means to conventional evaluation instruments. The table below summarizes the abovementioned self-assessment models.

Table 2.1.
Models of Self-Assessment

| Model | Detail | Author |
|---|---|---|
| **Learner-Prepared Self-Assessment** | Prepared by students and including comprehensive skills questions, self-assessment forms for writing skills and self-monitoring questions. | Dickinson (1996), Rea-Dickins & Germaine (1996) and O'Malley & Pierce (1996) |
| **Self-Assessment Checklists** | A type of tests which are comprised of a list of assignments and questions that students are supposed to answer. | Dickinson (1996), Rea-Dickins & Germaine (1996) and O'Malley & Pierce (1996) |
| **Informal Self-Assessment Device** | Developed and applied by learners to evaluate their performances in everyday life situations. | Dickinson (1996), Rea-Dickins & Germaine (1996) and O'Malley & Pierce (1996) |
| **Learner Record-Keeping Self-Assessment** | Some kind of self-assessment and implementation of it have been applied in the students' recent past and a self-appraisal scheme is included in the recording of the material. | Dickinson (1996), Rea-Dickins & Germaine (1996) and O'Malley & Pierce (1996) |

Table 2.1. (cont'd)

| Model | Detail | Author |
|---|---|---|
| **Self-Assessment with Integrated Tutor Feedback** | Requiring learners to integrate tutor and peer feedback before students self-assess. | Taras (2010) |
| **Learner Contract Design** | Developed in the context of self-directed learning (and is self-assessing to students' own criteria). | Taras (2010) |
| **Self-Marking** | Students use a model answer(s) with criteria (and possibly mark sheets) to compare to their own work. | Taras (2010) |
| **Sound Standard** | The tutor provides an objective descriptor of 'sound standard' that is of medium level work. | Taras (2010) |
| **Performance-Oriented Self-Assessment** | Outcomes related with placement, selection, achievement, certification, diagnosis tests etc. are measured via performance-oriented assessment | Saito (2003) |
| **Development-Oriented Self-Assessment** | The process of learning, generally in a classroom atmosphere, is measured in cooperation with self-managed activities. | Saito (2003) |

In order to have an efficient self-assessment process, students must have an idea about targets and they must be good at self-monitoring and self-judgment. That process is illustrated at Figure 2.1.



Figure 2.1. Self-Assessment Process for Students (Mcmillen and Hearn, 2008: 41)

For self-assessment to be successful, teachers should play a significant role by guiding, monitoring and facilitating the process for the students. Joyce et.al. (2009) held a project to study teachers' roles and ideas about self-assessment. They wanted to understand how self-assessment is transformed into the classroom environment and what teachers think about self-assessment. Self-assessment strategies and teacher beliefs were the main areas of the study. A small case study and focus group interviews were used as data collection methods. The results of the study revealed

that the beliefs of teachers determine whether they will use self-assessment or not in the classroom environment. The study also showed that the teachers are the ones who are responsible for establishing a classroom culture suitable for self-assessment. Joyce (2009) also stated that teachers must teach students how to assess their works and performance. It is not easy to apply self-assessment at first. However, thanks to scaffolding, better results can be obtained. Teachers must be a leader in establishing self-assessment culture. It was also stated that students must think they are able to assess themselves, and they need teachers' support. Lastly, students need time to become good at self-assessment (Joyce et. al. 2009). All in all, teachers seem to have a significant role in the self-assessment process.

### 2.3.3. Aims of Self-Assessment

Literature review presents that scholars express different aims for self-assessment. To start with, according to Sutalo (2011), the major aim of assessment is to determine the current situation of learners' performances and learning process, to compare the current state of self-assessment with good examples of practice, and to determine potential improvement and systematic quality developments (Sutalo, 2011).

Joyce et.al. (2009) expresses that self-assessment aims to increase intrinsic motivation of students regarding taking responsibility of their own learning. Thanks to self-assessment, students are able to understand how to apply steps related to developing their own language competencies since they are able to access better than any other assessor to their database on their own language competencies and to assess them. Kordes et al. (2014), having a different perspective, state that self-assessment is used simply because it greatly contributes to personal development and has a significant impact on learning. According to the authors, by means of self-assessment, students can develop their learning skills, evaluation skills, and critical thinking abilities. Moreover, they can also improve their self-confidence when they understand assessment procedures. As assessment may be seen as a mystery for students, self-assessment may demonstrate to these students how assessment is held and what aspects must be taken into account in order to understand what is expected.

Students will be able to understand what is expected and what they must learn (Kordes et.al., 2014).

According to Spiller (2012), first of all, students have a natural tendency to examine their own progress, therefore; using self-assessment strategies learner responsibility is fostered. Secondly, self-assessment provides students with an understanding of what they need to learn after the learning process. Using self-assessment strategies, students will see their weaknesses and will continue learning and compensating for their weaknesses. The author also states that self-assessment motivates further learning, encourages reflection on learners' own learning, and provides an opportunity to control their own learning process. Furthermore, many students may focus on just results. However, self-assessment may encourage students to focus on process, accommodate diversity of learners` readiness, experience and backgrounds, and may be a turning point from teacher-centred learning to student centred learning. Therefore, students may feel independent and are encouraged to "have" own learning. Last but not least, Spiller (2012) believes that self-assessment develops student's skills for lifelong learning. When students understand how to assess their own work and performance, they will be able to understand what they must study next.

According to Dickinson (1996) and Oscarson (1989), self-assessment may most properly be used with achievement and diagnostic objectives. Tests that are prepared according to the objectives of a language programme and assess whether students accomplish these objectives or not are achievement tests (Hughes, 1989). Progress achievement tests that are implemented during a course and final achievement tests administered at the end of a course are two distinct types of achievement tests. To supply students with feedback about their language performance is the main aim of achievement tests (Dickinson, 1996). Because of achievements tests, learners are able to revise and evaluate their own performances periodically and write reflections on their language learning process (Harris, 1997). Since one of the main aims of the self-assessment process is to make students able to assess their language learning and develop and revise their skills accordingly, carrying out self-assessment with the purpose of achievement tests will accomplish this aim. Another field that self-

assessment is most suitably applied to is diagnostic testing (Dickinson, 1996; Harris, 1997). Since diagnostic test are used for the purpose of identifying problems in the language learning process, using self-assessment for diagnostic purposes can assist students to notice the language components that they need to work on. Harris (1997) states that there may be students from different backgrounds in EFL and ESL settings, and therefore teachers and administrators can learn differences among these students with the help of self-assessment and can adapt and modify their instructions and orientations (Harris, 1997).

### 2.3.4. Self- Assessment of Language Skills

In this section, self-assessment is examined in relation to the development of four language skills.

### 2.3.4.1. Reading

There are various reading assessment techniques used to determine the amount of learning and the competency level that students reach in reading. Some of the frequently used reading assessment techniques are reading comprehension, cloze tasks, language comprehension tasks, and decoding tests. Reading comprehension tests include a text passage to be read by learners and detailed questions to be answered about the text. Cloze task is another technique in which words or some parts of the words are removed from the passage and learners are expected to fill in the blanks with appropriate words. In Language comprehension tests information is presented orally to the learners and learners are expected to answer questions related to the provided information. Finally, decoding is an assessment type in which learners try to guess the words which are related to a text or hints given via pictures.

With regard to self-assessment of reading skills, studies suggest that students take note of what they are reading and why they are reading, chart their progress on various skills, explain the targets of the reading, make reading choices, list reading strategies which they want to use, and also assess their reading portfolios,

considering some criteria. In addition, it is crucial to create an environment for students to reflect on their own language learning process. For instance, prior to doing a reading task, it may be beneficial to ask students to predict how easy or challenging the text is going to be and what factors make a text difficult (such as kind, length, subject or style of a text). After doing the task, they can think about their performance and reflect on their strengths and weaknesses (Harris, 1997).

Students may develop greater self-awareness about reading skills when their self-assessment procedures are discussed and reviewed. Self-assessment of reading skills encourages students to take part in the assessment process of reading and they develop opinions about the strengths and weaknesses of their reading abilities (Cambridge University Press, n.d.:363).

In addition, self-assessment of reading skills teaches students to observe and evaluate their own reading skills. In a reading lesson, when teachers assess students, it may be difficult for students to understand how to develop their own reading skills. Students are generally graded for their reading performance and provided with feedback. When students take part in the assessment process, they may have a chance to see their faults. Moreover, they may also have a chance to solve their own problems (Johansson, 2013).

Regarding the connection between self-assessment and reading habits, Moheidat (2011) presents a study aiming to investigate the impact of students` self-assessment on their reading performance in English. There were two groups in the study; an experimental group and a control group. There were 39 students in each group who were studying at Moosa bin Nusseir School for General Education in Muscat, Oman. A reading test adapted from previous General Certificate Exams in the Sultanate of Oman was developed. A technique of students` self-assessment through rating scale sheets and one-minute papers were used as a data collection tool. According to the results of the study, the self-assessment of students` had a positive effect on reading performance in English. The study recommended that it would be beneficial to train students about how to self-assess their reading habits and processes which would develop the reading skills of students. The study also suggested that teachers must

be aware of the positive effects of self-assessment for the development of reading skills (Moheidat, 2011).

Another study was held by Shahrakipour in 2014 in order to analyse whether there was a significant difference between the students who assessed their own reading and listening skills and the ones who did not. The participants were 120 Iranian students whose proficiency levels were beginner and intermediate. They were divided into an experiment and a control group. Both received a pre-test, then, the experiment group received a training on the self-assessment process, and they assessed their performances of reading and listening skills. Next, a post-test was given to both groups. When the pre-test and post-test results were compared, it was reported that there was a significant difference between the scores of the beginner participants on both reading and listening tests. It was also stated that the influence of self-assessment on reading skills was much more than on listening skills (Shahrakipour, 2014).

In another study, Johansson (2013) aimed to explore the credibility of Swedish third-grade students' self-assessments of their reading achievement by relating those assessments to two different criteria—teachers' judgments and students' reading test scores. Students' gender and socioeconomic status (SES) were introduced to determine to what extent, if any, these variables were associated with the accuracy of the self-assessments, once students' attitudes toward reading had been controlled for. In the study, students ($N = 5,271$) and teachers ($N = 351$) were the participants. The main method of analysis was two-level structural equation modelling (SEM) with latent variables. Johansson (2013) concluded that although the participants were young, the self-assessment of reading skills of third-grade students could be regarded as highly reliable. In Sweden, after spending 3 years together, teachers and students found it easier to have shared comprehension of literacy knowledge and skills (Johansson, 2013).

In brief, the above mentioned studies showed that self-assessment is beneficial for developing reading skills. Also, the correlation between teachers' and students' assessments shows that self-assessment may be a reliable way of assessment.

## 2.3.4.2. Listening

Listening is generally described as attentive hearing or active hearing, which is more than just hearing in a way that it is a receptive skill. According to Thomlinson (1984), listening is crucial for effective communication and used to interpret meaning of a conversation and thus, requires knowledge of semantics, syntax and surely phonology. In addition, listening for comprehension requires critical thinking, analytic thinking, prior knowledge and concentration. There are some important skills for language learners to acquire in order to listen well. Firstly, they should know how to make use of different strategies to create meaning out of context while listening. Secondly, being able to answer what is stated is another skill. Finally, they should have the ability to convey meaning in different ways (Shen et. al., 2007). Apart from these skills, nonverbal listening skills such as tone of voice, gestures, mimics, and body language comprise an important part of listening skills when taken into account as a whole.

The nature of listening skills make test developers assume that language learners practice recognizable listening skills, which can be defined as lower and higher skills, so as to understand various recordings and. The lower skills refer to skills that are used for comprehending "utterances in literal level" and the higher skills refer to the skills that are used to make inferences and use critical thinking. Listening tests are based on these skills and that's why they include parts about understanding main ideas, listening for specific information and inferring the speakers' meaning (Brindley, 1989).

Therefore, there are different assessment methods for listening skills. One of the most commonly used techniques is listening comprehension tests in which students are supposed to listen to a recording and answer the main idea or detail information questions related to the recording. Another technique is the performance test. Performance tests require students to listen to a recording and perform an action or select a picture based on what they hear. Apart from these, to assess listening skills, teachers can observe students with checklists, write down notes on learners' questions, answers, discussions or inferences, make use of directed listening

activities or use listening comprehension quizzes in class, arrange conferences led by students, and make use of listening response journals to observe students listening skills (Decker, 2004).

As an alternative to formal listening assessment techniques, self-assessment can also be an effective tool for assessing listening skills. Students can evaluate their own listening skills and data gathered from their own evaluation may provide students with ideas regarding development of their listening skills or may provide data about students' proficiency level in a placement test. Students may make plans and learn where to start and how to proceed in learning. (Baleghizadeh and Rahimi, 2011). Moreover, according to Harris (1997), students can self-assess their listening skills by making some predictions before they perform the task. They can try to predict if the recording will be manageable or demanding. Then, they can think of their own performance and try to comprehend how successful they were and why the task was easy or difficult (Harris,1997). Teachers can also extend this reflection by making students assessing how they listen and the strategies they use to assist them while listening, such as prediction, note-taking, summarizing and selective attention (O'Malley and Chamot, 1990).

Literature review demonstrates that there are several studies focusing on self-assessment and listening. To start with the earliest one, Oskarsson (1978) conducted a study on the self-assessment of listening and reading skills of adult students of EFL in Sweden, as well as teacher assessments of these skills. Participants were adult learners from 120 Swedish language schools and organisations. The questionnaire prepared by the author was sent to these schools. The results showed that self-assessments of students were correlated positively with teacher assessments. Also, the listening scores were also consistent with the written test scores of the students.

On account of the studies mentioned above, it can be stated that the self-assessment of listening skills can enhance students' listening abilities and make them aware of their own strengths and weaknesses.

## 2.3.4.3. Grammar

According to the Oxford Dictionary (2018), grammar is "the whole system and structure of a language or of languages in general, usually taken as consisting of syntax and morphology (including inflections) and sometimes also phonology and semantics" ( p.128). Hence, grammar is an integral part of language learning and in order to use language, a learner must be able to use grammar accurately. Since grammar is the "structural glue of a language" that puts four skills together, it has been of great importance in teaching and assessing language since the middle ages (Purpura, 2004).

Aforementioned, grammar is one of the most widely assessed skills in language learning and both formative and summative assessment types are used for assessing grammar. Testing grammatical knowledge is generally done via some techniques such as unscrambling, fill-in-the-blanks, sentence completion, error correction, elicited imitation, and cloze tests.

According to Liepa (2011), "in mastering accuracy the involvement of the language learner in correcting mistakes, in setting grammar tasks is substantial. Formal accuracy is increased through language use, though, taking into account students' needs, students' awareness of grammar structures is raised by means of pinpointing them when reading, listening to clarifications, fulfilling the appropriate tasks and consulting the teacher" (p.93). In the summative assessment type, students have an idea about how they have performed. However, understanding their mistakes and correcting them might not be easy. Students correct mistakes better when they find out the reasons of their mistakes and correct them individually. Due to their increased language awareness, searching and correcting their own grammar mistakes would produce better results than listening to them from a teacher (Kreuther, 2015).

Even though there are not many studies concerning self-assessment of grammar skills, in Nurov's aforementioned study (2000), self-assessment of grammar was one of the variables analysed. The study showed that the lowest correlation between

teachers' and students' evaluation turned out to be the grammar section and this correlation is not a significant one. Also, the correlation between students' assessment and the test regarding grammar section was also one of the lowest correlations. These results might state that grammar is a more difficult and complex component of language for students to learn and teachers to assess their other components of language.

## 2.3.4.4. Writing

"Writing is far from being a simple matter of transcribing language into written symbols: it is a thinking process" (White and Arndt, 1991). As highlighted by the definition, writing is a productive skill which permits learners to put their emotions and thoughts on paper, to use their insights and beliefs to create persuading reasons and to deliver a message via a well-structured text. It is becoming more and more important to be able to write effectively in second language education. Therefore, second language learners must not only be able to write correctly in terms of their use of grammar and vocabulary but they also must be aware of how to organize texts.

The importance of writing skills has led to a growing interest into the techniques used for assessing them. According to Cooper (1984), direct and indirect assessment techniques have emerged; the first is basically writing an essay, generally on a given topic and the second is to answer multiple-choice test questions. That is, direct assessment requires learners to produce in the target language whereas indirect assessment is regarded as a recognition measure (Cooper, 1984). These measurements give students an idea about the development of their writing skills and they are generally supported with feedback. If effective feedback is provided, students will understand and correct their mistakes, which will help them to self-assess their written texts. Since writing is a skill in which students actively take part in, feedback they receive becomes a significant tool for their development. It is not only the feedback provided by their teachers of pair but also feedback that comes from their own evaluation process that may provide good results. When students see their own progress, they try to increase their performance by revising their texts

(Salimi and Ahmadpour, 2015). According to de Leeuw (2014), the process of revising their texts may be interpreted as self-development. Students take responsibility of their own writing by being "involved" in the process. When students are involved in their own writing process, they realize their weaknesses and try to improve them. As a result of increasing their control on their writing skills, students become more autonomous learners (Assis, 2012), which in turn facilitate their metacognitive development (Taghizadeh, 2015).

During self-assessment process, student must be aware of how to assess the texts they write. In other words, students must be trained about the expected qualities of the text type before they start self-assessment of their writing skills. After (Lin-Siegler, Shaenfield, and Elder, 2015). In order to create effective self-assessment sessions, students must be aware of the components tested in writing assessments such as writing simple texts, detailed texts, presentations, reports and essays about students' field of interests, making an outline of a text whose target and format will be taken into account, and checking the text for cohesion and clarity, reviewing punctuation, related terminology and the use of linking words (Escribano and McMahon, 2010). Hence, as it was mentioned before, teachers have an important role in the preparation and implementation of self-assessment. According to Harris (1997), after students are shown the criteria, self-assessment may become a crucial part of the writing process. The criteria can be used by students as a checklist for their first draft and an ultimate self-assessment for their writing. They can compare this last self-assessment with their peers' or teachers' assessment. Also, they can make a list of common mistakes and use it as another checklist for self-editing (Harris, 1997).

There are a few studies that have been conducted on the self-assessment of writing skills. In the study held by Lin-Siegler et.al. (2015), the authors express that accurate self-assessment is beneficial for the academic achievement of students. It is not easy for students to assess themselves. Students must overcome difficulties faced during self-assessment. In order to overcome these difficulties, teachers must support students. In the study, the authors compared two stories one of which included poorly written stories and the other that included well-written stories, Fifty-three

6th-grade students in two history classrooms were randomly assigned to assess these two different types of stories and write a new story on their own. At the end of the study, the group given the poorly written stories performed better than the group given well-written stories as they had the chance to see the mistakes in poorly written stories and learned how to write the text and what the assessment criteria were (Lin-Siegler et.al., 2015).

In another study, Bing (2016) intended to explore the effects of self-assessment of students related to their writing skills in order to get more insight about self-assessment in writing instruction. Writing assignments, questionnaires, an analytic scoring rubric and interviews were used to collect data The results showed that the judgments made by the students about the quality of their own writing skills were in consistence with the assessments of their teachers. In addition, the students enhanced their writing skills considerably in terms of content, organization and mechanics of writing after they were exposed to the self-assessment procedure even though their development of vocabulary skills and language skills remained slightly less developed (Bing, 2016). Bing (2016) expressed that the results also indicated that the students' attitudes towards self-assessment of their writing skills were positively affected (Bing, 2016)

Another study on the self-assessment of writing was conducted by Fahimi and Rahimi (2014) in order to explore the effect on students' self-assessment of writing skill development. There were 41 female participants in the study who were not informed about self-assessment because the assessment had been done by the teacher before the study. At first, the participants were asked to write a text and assess it without any instructions of assessment. In the following weeks, the participants were informed about the assessment procedure and the teachers assessed their papers, as well. The data were collected with these writing assessments and another self-assessment survey was applied before and after the assessment procedure. The results showed that the writing skills of the students developed in time and positive attitudes towards self-assessment, both by the teachers and students were observed (Fahimi & Rahimi, 2014).

To conclude, it can be claimed that self-assessment has great importance for all skills. According to Heisigova (2015), although it is difficult and complicated to utilize effective self-assessment in the classroom and the outcomes require a long time to be observed, teachers should make the attempt at facilitating self-assessment. If learners can impartially asses their skills and performance, they may then target attainable goals which can serve for a long time. Finally, teachers can regard self-assessment as a future investment simply because student-teacher collaboration may facilitate learning when students can monitor their own learning process (Heisigova, 2015: 72).

## 2.4. Testing Speaking

The purpose of the present study is to find out whether or not there is a significant association between students' self-assessment and teachers' assessments of students' English speaking skills. This section of the study is devoted to the definition of speaking skills and ways of testing them including self-assessment of speaking skills.

As a skill, speaking is a productive skill which enables speakers to converse, to express their ideas, beliefs and emotions in verbal language and to convey information. Burns and Joyce (1997) define speaking as "an interactive process of constructing meaning that involves producing, receiving and processing information" which suggests that speaking is a communicative process in which the interlocutors take the roles of speaker and listener interchangeably and which also includes verbal and nonverbal elements. These properties of speaking make it an inevitable part of language learning, hence, it is important to know how to assess speaking skills.

### 2.4.1. Ways of Testing Speaking

There are different ways of testing speaking such as interviews, information gap techniques, picture tests, role-plays, cloze tests, semi-direct tests, band descriptors

and technology. Figure 2 summarizes those testing ways (Source: Helin Başak (2018):
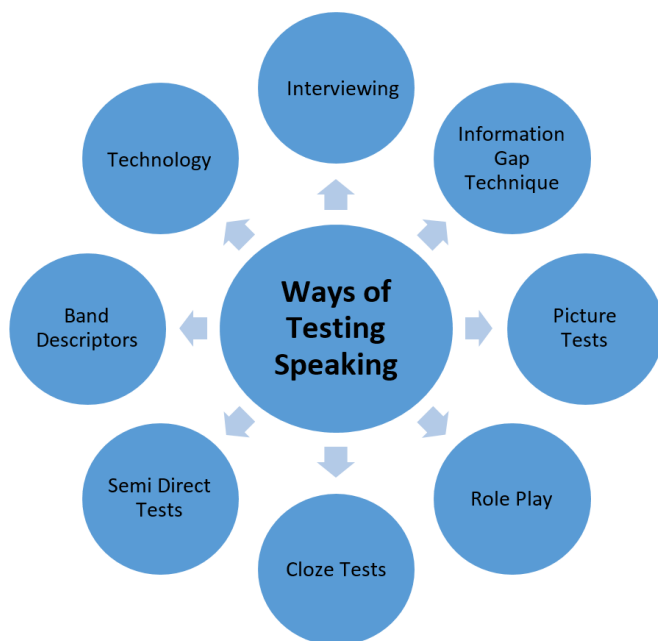


Figure 2.2. Popular Ways of Testing Speaking

One of the most widely used ways of testing speaking is interviewing. Interviewing is a communicative process in which students are asked a prepared series of questions by teachers and are expected to answer these questions in the target language. During interviews, students have an opportunity to express their ideas and answer questions and teachers assess whether students can use language properly during an interview. (Goh and Aryadous, 2010). Interviews allow teachers/assessors to evaluate students' speaking performance on specific topics or in situations better than other assessment types. Also, information that interviews provide can be used as a source to revise or enhance curriculum and instruction.

During interviews, the speaking performance of a student can be recorded in order to be evaluated later. Teacher can also evaluate students' performance immediately during the interview. According to Hall and Hope (2016),

> An interview can be scripted, where all or most of what the interviewer says and asks the test taker to do is prepared and written down, or it can be unscripted. Scripted and semi-scripted interviews are more consistent and reliable, and they allow the examiner to focus on evaluating the speech sample

more than on the interview process. They can, however, involve a lot of work to prepare, and they may result in a less realistic conversation. (Hall and Hope, 2016).

In other words, when the interview is unscripted, it will be easier for teachers to apply as it will require less time. Moreover, it will be more authentic thanks to unexpected direction or questions provided by the interviewer. (Hall and Hope, 2016).

In addition, interviews may be long or short. Especially during long interviews, it would be beneficial to use an evaluation chart or rubric. The chart must include the points regarding the aims of the speaking course. Thanks to these charts, teachers' assessment process has become more effective and informative.

Another traditional way of testing speaking is using the information gap technique. In the information gap technique, there is a speaker and listener or listeners. It is important that listeners need information about a particular topic (listeners do not have any idea about what the speaker will speak about). There are four mainly used information gap techniques. They are picture tasks, re-telling a picture story, discussion presentation and guided role plays (Kombercova, 2006).

The picture test is another method for testing speaking. In fact, the picture test is an easy way of testing speaking skills since students are presented with something visible to talk about (Kanga, 2012). For example, students can choose only a part of a picture as a speaking topic or generate new ideas to talk about while looking at the whole picture (Kanga, 2012).

During a role play session, speakers plan, assess and execute speech acts with a native speaker. Speakers may prepare self-reports and observations about the process. Moreover, teachers can also make plans about how to assess and how to manage the process. When students are aware of what to talk about, they have an opportunity to prepare for the speaking tasks. They may also develop strategies for the speaking session. Despite the fact that students can prepare speaking tasks, different topics may be spoken about during a role play (Seong, 2014).

In addition to traditional speaking tests, there are some newly utilized tests for assessing speaking. One of those tests is band descriptors as applied in the IELTS (The International English Language Testing Systems) exam. In the IELTS exam, there is no pass or fail score, but the exam score shows your performance through these band descriptors, which are the subskills corresponding to explanatory presentation of English competence of a test taker. Band descriptors have four equally-weighted but different criteria; fluency and coherence, lexical resource, grammatical range and accuracy and pronunciation (British Council, 2012). These band descriptors have been created in order to assist the test takers to comprehend their performance level in English. Using these points, speakers are graded under nine band scores. These cores are shown in the Table 2.2.

In addition to the IELTS, the TOEFL test (Test of English as a Foreign Language), which is another widely used standard test around the world, also has band scores. In the TOEFL exam, test takers are supposed to fulfil six different tasks. These tasks include both independent and integrated tasks. Independent tasks are about responding to a prompt orally. Integrated tests are orally responding to questions including reading passages and listening recordings. Test takers are evaluated according to four different component of speaking skills: general description, delivery, language use and topic development. The TOEFL exam speaking band descriptors are summarized in Table 2.3.

The other important descriptors for speaking skill are decided by Common European Framework of Reference (CEFR) which is an international standard for defining language skills. The framework was constructed in order to enable language learners with a clear, systematic and broad base for designing language courses syllabuses, preparing materials for language teaching and for evaluating language proficiency levels of learners. It has a 6-level scale starting from A1 (beginner) to C2 (proficient) for language abilities. For speaking skill, it has 5 different criteria shown in the Table 2.4.

Table 2.2.
Band Descriptors for IELTS Speaking Exam

| | |
|---|---|
| *9 Expert Users* | Have fully operational command of the language: appropriate, accurate and fluent with complete understanding. |
| *8 Very Good Users* | Have fully operational command of the language with only occasional unsystematic factual errors and inappropriate use. |
| *7 Good Users* | Get operational command of the language, though with occasional inaccuracies, inappropriate use and misunderstandings in some situations. |
| *6 Competent Users* | Enclose generally effective command of the language despite some inaccuracies, inappropriate use and misunderstandings |
| *5 Modest Users* | Acquire partial command of the language, coping with overall meaning in most situations, though is likely to make many mistakes. |
| *4 Limited Users* | Are considered as the basic competence which is limited to familiar situations; also they have frequent problems in understanding and expression. |
| *3 Extremely Limited Users* | Convey and understand only general meaning in very familiar situations. Frequent breakdowns in communication occur. |
| *2 Intermittent Users* | Here, no real communication is possible except for the most basic information using isolated words or short formulae in familiar situations and to meet |
| *1 Non User* | Essentially, students have not the ability to use the language beyond possibly a few isolated words (British Council, 2012). |

Table 2.3.
TOEFL Exam Speaking Band Descriptors

| *SCORE* | *General Description* | *Delivery* | *Language Use* | *Topic Development* |
|---|---|---|---|---|
| 4 | The answer meets the requirements of the task with minor mistakes. It is fairly comprehensible and presents coherence. | The speech is generally paced well and clear. It may contain minor pauses, minor pronunciation or intonation mistakes that do not affect general comprehension. | The answer shows impressive use of grammar and vocabulary. It presents a high level of using both basic and complex structures with minor mistakes that do not affect general meaning. | The answer is supported and adequate for the question. It is mainly well-developed and rational. Connections between notions are clear. |
| 3 | The answer fulfils the task bıt it is not well developed. Overall, it is understandable and rational despite the pauses in the explanation. An answer at this level is categorized by at least of the following: | The answer is mainly clear with minor mistakes in intonation, pronunciation and pacing the speech. Listener's effort is required at times. | The answer shows relevant ideas and good use of grammar and vocabulary. It may present some incorrect vocabulary or grammar. Range of grammar structures may affect fluency in general, but not the delivery of the message. | The answer is mainly rational and maintained with relevant ideas. General development may be limited to basic ideas. Connections between ideas are not clear from time to time. |
| 2 | The answer addresses the question with limited development. Despite mistakes with delivery and coherence, it is understandable. | The answer is fundamentally understandable but listener's effort is required due to vague articulation and intonation, unorganized pace and unclear meaning. | The answer shows basic use of grammar and vocabulary, which hinders delivery of the ideas. Most part of the speech is fluent but also unclear. | The answer is related to the question, however, ideas are not elaborated adequately. Repetitions and vague expressions are sometimes noticed. Relations between ideas are unclear. |

Table 2.3. (cont'd)

| SCORE | General Description | Delivery | Language Use | Topic Development |
|---|---|---|---|---|
| 1 | The answer is minimally related to the task and the content is limited and not clear. | Continuous pronunciation, intonation and articulation mistakes are noticed. Frequent lapses and hesitations exist in the fragmented delivery. | Delivery of the ideas and connections are prevented by limited range of grammar and vocabulary. Formulaic expressions are noticed in low-level answers. | The answer is addressed with limited related content. It mainly lacks complex ideas and is sustained with repeated prompts. |
| 0 | Test-taker doesn't make any attempt to answer the question or gives irrelevant answers. | | | |

Table 2.4.
CEFR Band Descriptors for Speaking Skills

| | RANGE | ACCURACY | FLUENCY | INTERACTION | COHERENCE |
|---|---|---|---|---|---|
| C2 | Presents a great flexibility in communicating, uses complex language structures and idiomatic expressions. | Uses constantly correct grammatical structures. | Conveys his/her ideas with a natural flow speed without any hesitation or pauses. | Interacts easily with utilizing body language clues without any difficulty. Turn-taking, reference making and allusions are natural. | Discourse is created coherently and cohesively and proper use of transitions and accurate organisation of speech exist. |
| C1 | Owns a good knowledge of language range and can express ideas in a wide range of topics such as academic, professional or general. | A broad range of grammar structures are maintained and nearly no mistake is made. If made, they are corrected. | Fluent and spontaneous expressions are used but difficult concept may slow down the conversation. | A proper phrase for a specific discourse can be selected from a good range of discourse functions. | Speaks clearly, smoothly and in a well-structured way and transitions and cohesive phrases are used. |
| B2 | Owns an adequate range of language knowledge, most general topics are available to talk and some complex sentences are used. | High degree of grammatical accuracy is shown and mistakes leading to misunderstandings are not made. Can correct the majority of the errors. | There are obvious pauses in the speech and hesitations may occur despite the even tempo of the speech. | Starts a conversation, takes turn and ends a conversation when he needs to but not always in a proper way. | Uses a restricted range of transitions and organisational patterns. |
| B1 | Owns enough language knowledge and vocabulary to survive, expresses ideas on topics such as family, hobbies, work or travel. | Frequently used language phrases are utilised in expected situations. | Pauses for correct grammar and vocabulary occurs but keeps the speech comprehensible. | Basic face to face talks are maintained on familiar topics but needs to repeat him/herself for confirmation. | Shorter sentences are connected with basic use of transitions. |

Table 2.4. (cont'd)

|  | RANGE | ACCURACY | FLUENCY | INTERACTION | COHERENCE |
|---|---|---|---|---|---|
| A2 | Memorised sentences and words are used to communicate for daily life situations. | Simple grammatical structures are accurately used but there are some errors, too. | Can make him/herself understandable in short sentences with frequent pauses and hesitations. | Responds simple questions and shows that s/he can follow the speech adequately. | Connects sentences only with "and", "but" and "because". |
| A1 | Owns very simple words to use in speeches on personal information. | Only a restricted control of a few simple grammar structures and memorised sentences. | Controls very short and memorised sentences with a lot of pauses and hesitations. | Has questions and answers about personal information but the speech mostly depends on repetition. | Connects words and phrases with only "and" or "then". |

In conclusion, it is also possible to assess speaking using various technological tools. Since technology is rapidly developing and affecting all aspects of life, including education, technological tools are getting more and more popular in education and different types of learning environments. Tools from videos to interactive online courses are being used by institutions and teachers. Consequently, technology is also being used as a tool for assessment not only for speaking skills but also for other skills. Recording and listening to one's own and others speaking performance, responding to on-line questions and many other ways can be tools for testing speaking. According to Sullivian (2012), "as technology improves, we are beginning to see a new generation of technology-driven tests (e.g. the Pearson Test of English) which are attempting to use technology in new ways. While the actual content of these new tests remains quite traditional, the fact that performance is assessed automatically by machine is certainly both innovative and controversial" (Sullivian, 2012: 268).

### 2.4.2. Components Tested in Speaking

As much as how to test, it is as important to understand what to test in speaking skill. As summarized in Table 2.5, specific competencies underlying speaking skills such as grammatical accuracy, vocabulary knowledge, organization of opinions, fluency and pronunciation abilities are tested.

Table 2.5.
Important Points to Be Tested About Speaking (Vigoya, 2000: 98)

| | |
|---|---|
| **ACCURACY** Ability to master a language as a system. | **FLUENCY** Ability to use language for communication. |
| **PRONUNCIATION** Ability of students to produce sounds and follow stress and intonation. Those abilities must be in an acceptable manner. | **MECHANICAL SKILLS** Ability of students to use punctuation, sentence length, pauses, rhythm and speed. |
| **VOCABULARY** Ability of students to use necessary words and to understand the vocabulary in questions. | **LANGUAGE USE** Ability of students to talk. Showing a mastery of semantic and grammatical sources of language. |
| **GRAMMAR** Ability of students to produce suitable syntactical and morphological patterns in a given speech. | **JUDGEMENT SKILLS** Ability of students to talk about a variety of things. Ability of students to select, order, organize and develop thoughts. |
| **COMPREHENSION** Ability of students to negotiate, transmit and share information in a fluent and accurate manner. | |

To begin with, while testing speaking, accuracy of grammar must be tested simply because grammar is vital for expressing thoughts correctly. Another important thing to be assessed during speaking is vocabulary knowledge. Vocabulary is a necessity for making meaningful sentences, answering questions and expressing ideas. Speakers cannot express their ideas without using the vocabulary related to the spoken topic (Mertova, 2013).

Skills such as organization of opinions and developing a good outline are not only expected to be shown during writing but also in speaking. During speaking, it is important to express knowledge and ideas in an organized manner supported by correct transitional words (Mertova, 2013: 7-8).

It is natural to expect fluency during speaking. Agung (2012) describes fluency as the ability to converse fluently and accurately in the target language. Many language learners' main objective is to speak the target language fluently. A reasonably fast pace of speaking and a small number of pauses such as "ums" and "errs" are the indicators of fluency. Another indicator of fluency is dealing with hesitations. In fact, a great number of hesitations make it boring for listeners to listen to the speaker (Mertova, 2013). These indicators show that speakers do not need a lot of time to think about required vocabulary or grammar structure to construct sentences to deliver the message when she/he speaks in the target language (Agung, 2012). Hence, this makes fluency an indispensable component of speaking assessment.

Pronunciation, which is another speaking assessment component, is the act and manner of saying words by correctly producing the sounds of speech with correct stress and intonation. Teaching and testing pronunciation in language classes is important for making sure that native speakers or other second language learners understand non-native speakers (Pramujunarta, 2007).

### 2.4.3. Self – Assessment of Speaking Skills

To self-assess one`s own performance during speaking activities is a difficult process. In order to make a well-established self-assessment process, there are four steps to be followed by teachers. First of all, teachers should establish a set of criteria so that both teachers and students would be informed about the inclusion of the self-assessment process. Secondly, it is better to ask students to reflect on their achievements and the inadequacies they perceive rather than asking them to grade themselves which may make them nervous. Moreover, asking students to grade themselves may also create objectivity and reliability-related problems. Thirdly, self-assessment should take place immediately after the completion of a task, although audio and video recording may prove helpful for students to remember their performance. Last but not least, it is important to complement self-assessment results with the teacher's information based on observations and tests feedback in order to create a more reliable assessment tool.

Moreover, self-assessment of speaking is an alternative assessment type and is also a metacognitive strategy. According to Forbes and Fisher (2015),

> Cognitive strategies are related to the processing of information, social-affective strategies are concerned with interaction with others, and metacognitive strategies involve thinking about the learning process, planning for learning, monitoring of comprehension or production while it is taking place, and self-evaluation after the learning activity has been completed. (Forbes and Fisher, 2015).

Next, an objective measurement tool should be created by the teachers and students should be informed about the important components to be assessed (Trofimovich, 2016). One suggestion to develop such a tool is provided by Fulcher (1993).

Formulating statements starting with "I can" the desired skills of speaking can be assessed by students. Table 2.6 presents an example of a tool including "I can" statements:

Table 2.6.
"I can" Statements for Self-Assessment of Speaking Skills

| | |
|---|---|
| **I can…** | ask someone how I can go to a hotel, restaurant or any other place |
| | buy items like ticket, clothes etc. |
| | order a meal |
| | arrange a hotel room |
| | make social introductions |
| | ask and answer simple questions |
| | describe my job |
| | make conversations about some topics |
| | interview with employee and make reports |
| | talk about future plans and hopes |
| | describe a place or person |
| | talk about a trip |
| | write messages or leave voice messages |
| | communicate with native speakers |
| | speak about a professional subject |
| | defend personal opinions |
| | solve problems about various topics |
| | talk about topics affecting my life or civilization |
| | talk as well as my mother tongue |
| | prepare or give a lecture |
| | persuade people |
| | express other people`s point of views |
| | direct a conversation, I naturally integrate appropriate cultural and historical references in my speech. |
| **Other statement samples** | Native speakers generally understand me |
| | I feel that I am good at speaking, |
| | There is not much grammar topic I avoid using, |
| | I have a good vocabulary and I can use many vocabulary items |
| | I am able to adjust my speech to suit my audience |
| | I never make grammar mistakes, |
| | I am rarely unable to finish a sentence |

During the self-assessment process of speaking, teachers must listen to and observe the process carefully and provide specific feedback to students when necessary. Giving feedback should include some important points. First of all, students must be informed about what mistakes they made. Secondly, they must be encouraged for speaking, assessing speaking and correcting their mistakes after receiving feedback. Moreover, teachers must also appreciate students when they correct their mistakes, develop themselves, use a new vocabulary item or an amusing role-play, produce an accurate sentence, make an interesting contribution to a discussion and become

eager to speak and assess speaking. Teachers may correct their students' mistakes immediately or at the end of the speaking process (Rocio, 2012).

In conclusion, thanks to the self-assessment of speaking skill, students will be able to understand how well they speak. Assessing their performance, they become aware of the strengths and weaknesses of their performance. For instance, if students can record speaking process, they can find mistakes and understand speaking strengths after listening/watching the performance. (Yang and Tseng, 2015). Hence, self-assessment of speaking skills will help students to become more autonomous in terms of developing their speaking skills since self-assessment itself has been a core point making students more autonomous and contributing to learner-centred language learning.

## 2.5. Previous Studies on Self-Assessment of Speaking Skills

Research on other areas of education was investigated by scholars substantially earlier than self-assessment in language learning. The initial studies were conducted in the 1970s. In these studies, students' beliefs about self-assessment and its relationship with their language learning were investigated. The studies in which self-assessment was used as a variable were conducted by Gardner and Lambert (1972), Vigil, Baca and Oiler (1977). These scholars are regarded as pioneers in this field of research.

In the literature of language testing research, there are numerous studies concerning self-assessment and producing various results. While self-assessment is regarded as a practical alternative to formal foreign language evaluation for placement or achievement tests, the variation in self-assessment validity proposes possible hardships in definite interpretation (Ross, 1998). If the findings of all the experimental studies concerning self-assessment were harmonious, it would be exemplary. Nevertheless, they revealed various outcomes which have left researchers, test administrators and teachers with a query of whether they should implement self-assessment in language assessment or not (Saito, 2003).

Although numerous scholars recognised the contributions of self-assessment procedures to the learning and assessment process, studies conducted on the topic do not provide rich information for the use self-assessment in the development of language skills. As for the main topic of this study, the self-assessment of speaking skills, research studies are scarce. In this section of the present study, studies on the self-assessment of speaking will be examined.

Hooshangi, Mahmoudi, and Yousofi (2014) state that self-assessment has been one of the most popular assessment topics to be discussed because of its importance for autonomous learners and student-centred education environments. In their study, authors aimed to examine effect of self-assessment on Javanroodian foreign language learners' oral performance ability. Video typing, assessment, feedback, practice and training sessions were utilised during the process. There were 20 English learner participants in the study. They were divided into two groups as an experimental group and a control group, taking oral performance pre-tests into account. A self-assessment checklist including interaction with the audience, grammar rules, vocabulary use, pronunciation, content and presentation was utilized in the study. At the end of the study, it was seen that participating in the self-assessment process increased the oral performance ability of learners. Authors suggest that the results may be used by policy makers, teachers, developers, material designers and students (Hooshangi, et.al., 2014).

Moreover, Leger (2009) conducted a 12-week study investigating the ways that language learners could improve the perception of their speaking skills. 32 learners of French with advanced proficiency level participated in the study. They were asked to assess their speaking skills and participation in communicative practices and goal setting in French. The results indicated that the participants' self-perception improved especially in terms of fluency, vocabulary and self-confidence while speaking French. Furthermore, setting goals was also suggested as beneficial for learners to take the responsibility for their own learning process. In other words, benefits of self-assessment, not only cognitively but also affectively, was highlighted in the study (Leger, 2009).

Chalkia (2012) concentrated on self-assessment as an alternative way to assess speaking skills of foreign language learners. 18 Greek learners of English, aged from 11 to 12, participated in the study. The students were asked to fill in a questionnaire including questions related to self-assessment and were also observed during the process of completing the questionnaire. The results revealed that the participants showed positive attitudes towards self-assessment. This activity helped students to be more motivated to be more active in speaking and to realize their strong and weak sides in their speaking abilities in spite of the difficulties they faced during the self-assessment application (Chalkia, 2012).

According to the results of another study, held by Shahrakipour (2014), the results indicated that the students with an intermediate level of language proficiency improved their listening and speaking abilities with the help of the self-assessment process to which they were exposed. Moreover, the statistical analysis showed that the speaking scores of self-assessment were higher than those of the listening skills scores which showed that the impact of self-assessment was higher in speaking than in listening and reading skills (Shahrakipour, 2014).

In one of the few studies about the self-assessment of speaking skills, Singh (2015) stated that verbal skills are ignored when it comes to assessment and having good grades from reading and writing skills does not necessarily mean having good performances in speaking skills. For this reason, he drew attention to the self-assessment of speaking skills and tried to find out which guidelines would be beneficial for students to assess their own speaking skills. 15 learners of English participated in the study and were asked to assess their speaking skills. They were given two different guidelines, namely; the American Council on Teaching of Foreign Languages (ACTFL) and the Foreign Language Oral Skill Evaluation Matrix (FLOSEM). Teachers also evaluated students' speaking skills on these two same criteria. The results showed that there were differences between students' self-assessments and teachers' assessments. The students overestimated their speaking skills when compared to teachers' assessments (Singh, 2015).

Smith (2015) investigated the accuracy of self-assessment of speaking proficiency. The study was held in Japan with undergraduate students. Before the study, the students were informed about the importance of moderation, modesty and group harmony. They were also about self-assessment protocol. At the end of the assessment, self-assessment results and assessment results of the teachers were compared to reach inter-rater reliability. According to the results of the study, there were 11 cases and 10 of those cases were within the boundaries of inter-rater reliability. Moreover, the results also showed that self-assessment is a reliable assessment type as 8 out of 10 completed self-assessment results were within the same boundaries (Smith, 2015: 41).

In their study, Babaii et.al. (2015) stated that learning success or failure may be affected by the mismatches between teachers and students and that self-assessment may be used as a tool to decrease these mismatches. Due to this, the authors investigated whether the students' and the teachers' assessment of students' speaking skills would match after the students were provided with a practice session of a self-assessment application. There were 29 learners of English and 6 English language the teachers in study as participants. Firstly, the students listened to and assessed their audio-record speaking performances. Later, they were provided with the assessment criteria and were asked to assess their performances again and write reflection papers on the self-assessment application. In the meantime, the teachers also assessed the students' speaking performances with the same criteria. According to the results of the study, there was a significant difference between the students' first and second assessment of their own speaking skills. It was also shown that the discrepancies between the teacher and student assessment were decreased after the students were provided with the assessment criteria. The students' reflection papers also revealed that students were positive towards the self-assessment application.

Last but not least, in the study conducted by Karakaya (2017) in the Turkish EFL context the aim was to find out the function and effect of self-assessment of English speaking skills. The participants were 46 Turkish students with intermediate proficiency level. In the study, the data was collected via students' and teachers' assessments of speaking skills, students' interview, think-aloud protocols and a

motivation questionnaire for L2 speaking. After the data was analysed, it revealed that the implementation of self-assessment helps students to gain awareness on their own learning process and motivate them for speaking English. Also, the outcomes indicate that students regard self-assessment process as a stimulating and engaging method to develop their speaking skills and to get involved in their own learning process. As a result, the author interprets that self-assessment of speaking skills can be implemented as an influential learning strategy to improve students' speaking skills and motivation (Karakaya, 2017).

In the light of studies above, it can be stated that self-assessment contributed to the ongoing debate of its own use in the EFL/ESL context. Generally, the results show that self-assessment of speaking skills enhances students' oral abilities and increases their motivation and self-confidence for speaking in the target language. Also, it can be concluded that students have positive attitudes towards the self-assessment of speaking skills. When teacher and student assessments are taken into account, some studies indicate that teachers and students assess speaking skills differently from each other (Singh, 2015; Babaii et.al, 2015) while another study shows that there are not significant differences between teacher and student assessments of speaking skills (Smith, 2015). In conclusion, the present study aims to contribute to the ongoing debate on use of self-assessment as an alternative tool to evaluate speaking abilities in EFL/ESL contexts.

# CHAPTER 3

# METHODOLOGY

## 3.1. Introduction

This chapter presents a brief overview of the study and the setting, the research questions, participants, data collection tools, piloting and data collection procedures of the present study.

## 3.2. Aims of the Study

This study aims to examine whether there is a significant difference between students' self-assessment of their speaking skills and teachers' assessment of students' speaking skills. In addition, the study also aims to determine whether students and teachers assess different components of speaking skills such as grammar, vocabulary, pronunciation, and fluency or not. Finally, the study aims to investigate whether students' self-assessment differs in terms of their proficiency level in English, gender, and exposure to English in extracurricular activities.

## 3.3. Setting

The present study was conducted in TOBB ETU Department of Foreign Languages. The department aims to improve students' capability to understand written texts, to think of their skills critically and to make them aware of their educational performances. The department expects students to speak English fluently, to

understand and interact in English, both in educational and conversational situations when they finish the preparatory programme.

There are five different levels in the programme: AF (beginner), A (elementary), B (Pre-Intermediate), C (intermediate) and CR (Intermediate students who repeat the programme for the second time.). These levels are determined according to the exams that students who are registered in the programme have to take. The preliminary exam is the first one that is conducted at the beginning of the programme and it assesses listening comprehension (60 questions), structure (40 questions of grammar and vocabulary) and reading comprehension (25 questions). The proficiency exam is the second one that is applied after preliminary exam. The students that have a score of 65 out of 100 from the Preliminary exam are able to take the Proficiency exam. The Proficiency exam is TOEFL Institutional Testing Programme (ITP) which tests listening comprehension (50 questions), structure (40 questions of grammar) and reading comprehension (50 questions). Students who score 500 out of 677 from the TOEFL ITP are exempted from the preparatory programme and are able to continue to their departmental courses. Students who have a score lower than 500 are obliged to start to the preparatory programme. In the first semester, the student levels are classified as shown in the table below:

Table 3.1.
Proficiency Level of Students at TOBB ETU in the first term

| AF Beginner | A Elementary | B Pre-Intermediate | C Intermediate | CR Intermediate |
|---|---|---|---|---|
| Students who don't take the Preliminary exam or get a score between 0-49. | Students who take a score between 50-64 from the Preliminary exam or who get a score between 0-399 from TOEFL ITP. | Students who get a score between 400-449 from TOEFL ITP. | Students who get a score between 400-449 from TOEFL ITP. | Students who are in their second year in the preparatory programme and get a score between 400-449 from TOEFL ITP |

There are nearly 900 students registered to in the programme per year. As for their contact with English, Beginner (AF) level students have 30 contact hours in a week, but the other levels (A, B, C, and CR levels) have 25 contact hours in a week. These

lessons are divided into 3 different subcategories: Main Course (15 hours for AF students, 10 for others), (Reading and Writing 10 hours for all levels) and Listening and Speaking (5 hours for all levels).

Since the research was conducted in the second term of the education year, it is important to explain the levels in the programme. In the second semester, students' levels are determined according to the average of the grades that they take from the quizzes and midterm exams during the first semester. Therefore, there are 5 levels in the programme in the second semester: AR level (students who repeat the elementary level in the second semester, R stands for repeat), B level (intermediate students), BR (Students who repeat the intermediate level in the second semester, R stands for repeat), C (Upper-Intermediate level), CR (students who repeat the upper-intermediate level in the second term, R stands for repeat). The table below explains the levels in detail;

Table 3.2.
Proficiency Level of Students at TOBB ETU in the second term

| AR Elementary | B Intermediate | BR Intermediate (Repeat) | C Intermediate | CR Intermediate (Repeat) |
|---|---|---|---|---|
| Students who were in A or AF levels in the first term and unable to collect 70 points from the exams during the first term. | Students who were in A or AF levels in the first term and able to collect 70 points from the exams during the first term. | Students who were in B level in the first term and unable to collect 65 points from exams during the first term. | Students who were in B level in the first term and able to collect 65 points from exams during the first term. | Students who were in C level in the first term and unable to collect 65 points from exams during the first term. |

When it comes to developing speaking skills, students have the opportunity to improve their speaking skills inside and outside classroom. In class, they have a spare speaking lesson - although it is not limited to only one lesson- and they have student clubs such as English Drama Club, English Friends Club, Documentary Club, TOBB Masters Club and Grammar Boost Club outside the classroom.

As for the assessment of speaking skills, in the speaking exams, students are expected to introduce themselves and to be able to answer general questions related to their personal interests, leisure time activities, opinions about their school,

54

teachers, friends etc. Also, they are expected to describe a picture that they choose during the exam; that's why, they should be able to use proper adjectives while describing the picture. In the final part of the exam, students draw a question from the box and answer it. In order to answer these questions and be successful in the exam, students should be able to think critically during the exam to produce rational and meaningful responses. To achieve this, they should be able to use appropriate grammar and vocabulary in addition to being fluent and having correct pronunciation. The week before the exam, students are shown sample questions and pictures to get familiarized with the exam and get prepared for it.

## 3.4. Research Question

The present study aims to provide answers to the following research questions:

- Is there a significant difference between students' self-assessment scores and teachers' assessment scores regarding students' speaking skills in English?
- Is there a significant difference between students' self-assessment scores and teachers' assessment scores regarding components of speaking skills?
- Do students' self-assessment of speaking skills scores vary as students' proficiency level increases?
- Is there a significant difference between female and male students' self-assessment of their speaking skills?
- Do students' self-assessment scores of speaking skills vary in relation to students' daily use of English in extracurricular activities?

## 3.5. Participants

The participants of this study are 549 Turkish learners of English with the age range of 18-25 studying at TOBB Economy and Technology University in Ankara, Turkey. All of the participants are native speakers of Turkish with different levels of proficiency in English. They are all graduates of state or private secondary schools where they had English for 4-8 hours a week. During the data collection

period, the participants were at the TOBB ETU preparatory school. The language proficiency level of the participants was determined by the preliminary exam given by Department of Foreign Languages in TOBB ETU at the beginning of the 2016-2017 academic year. This preliminary exam was given in order to place new students into different classes based on their proficiency level in English. The exam included items testing grammar, listening skills and reading skills and the students who scored between 0-49 points out of 100 were determined as beginner (Group AR in the study), 50-65 points out of 100 were determined as elementary students in TOBB ETU preparatory classes. Intermediate (Group B-BR) and Upper-Intermediate (Group C-CR) levels were determined by the success rate of the students during the preparatory programme. The distribution of the participants based on their proficiency levels is presented in Table 3.3.

Table 3.3.
Distribution of the Participants based on Proficiency Levels

| Proficiency Levels | Number |
|---|---|
| AR | 53 |
| B | 348 |
| BR | 12 |
| C-CR | 132 |

## 3.6. Data Collection Tools and Procedure

Data collection tools that are used in this study include a self-assessment rubric, the participants' self-assessment scores of their speaking performance and their teachers' assessment scores of the students' speaking performance. The self-assessment rubric consisted of questions about the participants' demographic information, their use of English in after-school activities and their performance in speaking skills. The teachers' assessment scores of the students' speaking performance were obtained through their assessment given to the students in the speaking exam throughout the 2016-2017 spring semester.

### 3.6.1. The Self-Assessment Rubric

The rubric which was developed to collect data about how the students self-assess their speaking performance consisted of two sections. The first section included 19 questions about the participants' demographic information and their daily use of English outside the classroom. In the second section, there were 25 questions asking the participants to assess their speaking performance in English in terms of their use of grammar, vocabulary knowledge, fluency in English, pronunciation accuracy and their communication skills. In order to ensure content consistency in the student and the teacher participants' evaluation of the speaking performances, the rubric presented to the students was prepared according to the rubric which was used by the teachers in the speaking exam in which the participants were assessed (see Appendix 2 for the rubric). The items in the rubric were prepared in the form of "I can…" statements to be more comprehensible (Interagency Language Roundtable, n.d.). The rubric was piloted and the actual data were collected two weeks after the piloting process. The participants were given approximately 20 minutes to assess their speaking skills with the rubric as it was understood that this duration would be enough based on the piloting experiences.

### 3.6.2. Piloting The Self-Assessment Rubric

Before the actual data collection was applied, the rubric was read and evaluated by three different native speakers of English, who are also English language instructors, for the items' grammaticality, appropriateness and naturalness of the questions. Next, the test was piloted on 19 intermediate students from the Preparatory programme from the institution where the actual data collection took place. The researcher administered the piloting process in order to quarantee that the students were instructed properly and accurately.

The participants of the piloting process were given the rubric and were instructed about how to fill it in. Since there were some participants who were not familiar with such a rubric providing a 5-scale option, the familiarisation process was

necessary. They were told about the Likert-Scale survey type and the meaning of values in the scale. The students were also informed about the focus of the study and they were asked about the clarity of the items.

The piloting process revealed that some items were perceived as complicated for the participants to comprehend because of wording or the vocabulary used in the items. Even though the participants were intermediate students of English, some of the items in the rubric included vocabulary that they did not know. For instance, the item related to the vocabulary knowledge included the word "hesitation" and some of the students did not know the meaning of the word. That's why, considering that the elementary level students would not be able to respond to the rubric items as they could not comprehend the items in English, the rubric was translated into Turkish by the researcher in order to eliminate any misunderstandings and enhance the reliability of the data (see Appendix 1 for the rubric). The Turkish version of the rubric was read and evaluated by the researcher and a Turkish native English instructor so as to detect any typo or ambiguous expressions. After proofreading, the Turkish version of the rubric was piloted again with an intermediate and an elementary class in order to find out if there was any word or sentence which could lead to misunderstandings that could affect the reliability of the data. The students were also instructed about how to fill in the rubric and informed about the study for the purpose of clarifying any potential question in students' minds. After the second piloting process, the students were asked for feedback on the rubric and necessary changes were made in the wording of some items before data were collected.

### 3.6.3. Teachers' Speaking Exam Grades

The other data collection tool was the speaking exam which TOBB ETU Preparatory students take once in every academic semester. The exam was held as an interview in which the students were invited individually to be assessed by two teachers. The interview procedure was as follows. First, the student introduced him/herself and gave information about him/herself and the teachers randomly asked some questions about their daily activities as a warm up to reduce the students' anxiety. Secondly, the students were presented with a picture and were asked to describe it. While

describing the picture, the students could be asked some leading questions to proceed with the conversation. After the picture description, the students were asked to pick a random question out of several questions and asked to answer it. Each student was given one minute to think about and to prepare an answer to the question. During this third part, the teachers did not explain the question or ask other questions to proceed with the conversation. After a student answered the question, s/he left the room and the teachers assessed his/her performance according to the rubric that they were given before the exam.

### 3.6.4. Data Collection Procedure

After the piloting process, the actual data collection with the rubric was implemented. The first step was to inform teachers about the study and the implementation of the rubric in their classes with an email sent by a school administration staff. The teachers were asked to inform students about the rubric and the Likert-scale question forms. Also, they were told that they should distribute the rubric to the students, give instructions about how to assess their speaking with the rubric and allocate 20 minutes from their lessons to the self-assessment process.

The second step was to copy the rubrics and prepare the packs that the teachers could obtain and bring to their classes. Although the teachers were informed through email, the researcher put some critical notes such as giving clear instructions, helping students in case of a misunderstanding and timing into the packs in which there were the rubrics for each class. After the implementation in classes, the researcher collected the packs from the teachers.

### 3.7. Data Analysis

After the data were completely collected, all data were entered into Statistical Package for the Social Sciences (SPSS Version 22) to analyse. While entering the data, the scores from the self-assessment rubric and students' speaking exam grades were converted. To clarify, the maximum score that could be obtained from the self-

assessment rubric ls 125 because there are 25 questions in the rubric and the maximum score from one questions is 5. On the other hand, the maximum score that could be obtained from the speaking exam is 100. Hence, the scores that are out of 125 and that the students gave themselves through the self-assessment rubric were converted into the scores out of 100 in order to make the data analysis more accurate.

In order to analyse the relationship between students' self-assessment and teachers' assessment of students' speaking skills, some statistical tests in SPSS programme were run. To start with, Paired Sample T-test is used when determining whether the mean difference between two paired variables is statistically different or not. The present study's main research question is to find out the relationship between the students' and teachers' assessments of speaking skills, therefore; Paired Sample T-test was used to analyse and answer this question.

Independent Sample T-Test is another test that was used to determine if the mean scores between two independent groups are statistically different or not. One of the research questions of this study is whether gender creates a difference in students' self-assessment of their speaking skills. Since gender is an independent variable, Independent Sample T-test was applied to answer this question.

In addition to Independent Sample T-test, One-Way ANOVA (Analysis of Variance) was also applied. ANOVA tests are applied when more than two groups' mean scores are needed to be compared to find out the statistical differences within the groups. The present study's research questions are trying to determine significant differences between groups such as groups with different English proficiency levels and groups exposed to English outside the school for different length of time. Thus, ANOVA tests were used to find out mean differences within these groups. Along with the ANOVA, Post-Hoc tests were also used to determine which of the means are not equal within the groups. Since homogeneity is also important to run Post-Hoc tests, Levene tests were run in order to understand if the data showed homogeneity or not. According to results of the Levene test, Tukey HSD Post-Hoc Test was applied in order to understand the group mean differences.

# CHAPTER 4

# RESULTS

## 4.1. Introduction

The purpose of the present study is to find out whether there is a significant difference between the Turkish learners' self-assessment of English speaking skills and the teacher assessments of the learners' English speaking skills. The data were collected via a self-assessment scale given to the participants. Also English language teachers' assessment of the participants was based on a speaking assessment scale. The data were analysed using the Statistical Package for the Social Sciences (SPSS) programme. In this part of the study, the results of the data analyses are presented. Each section explains the relationship between the data obtained by the two data collection tools and the data analysis related to the research questions. The presentation of the data and the analyses are organised according to the order of the research questions.

## 4.2. Relationship Between the Students' Self-Assessment Scores and the Teachers' Assessment Scores

The first research question of the study was as follows.

**4.2.1. Is There a Significant Difference Between Students' Self-Assessment Scores and Teachers' Assessment Scores Regarding Students' Speaking Skills in English?**

With the first research question we aimed to examine whether there is a significant difference between the assessments of speaking skills evaluated by the students themselves and their teachers. To answer this question, the data gathered from the students' self-assessment scale and their speaking exam scores given by the teachers were analysed with Paired Sample T-Test in SPSS.

Table 4.1.
Differences between Teacher and Student Assessments of Students' Speaking Skills

|  | N | x̄ | Min. | Max. | Df | Std.D. | t | P |
|---|---|---|---|---|---|---|---|---|
| **Teacher Assessment** | 545 | 79.269 | 3 | 100 | 544 | 12.597 | 146.89 | 0.00* |
| **Student Self-Assessment** | 548 | 61.7752 | 20 | 100 | 547 | 15.255 | 94.790 | |

*: Correlation is significant at the 0.05 level

As can be seen in Table 4.1, while the mean score for the students' self-assessment of their speaking skills was 61.7752 in general, the mean score for the teachers' assessment of students' speaking skills was 79.269. According to the results of Paired Sample T-Test there was a significant difference between these two groups' assessment of the students' speaking skills (P= 0.00). The results indicated that the students underevaluated their speaking skills as the students' self-assessment mean score was lower than the teachers' assessment mean score, which might also mean that the students were not aware of their own speaking performances and tended to believe that they were not competent in speaking skills. In other words, the way the participants perceived their speaking skills and the way the teachers perceived and assessed the students' speaking skills were not in line with each other.

### 4.2.2. Self-Assessment and Teacher Assessment of Components of Speaking Skills

The second research question of the study was as follows:

<u>Is there a significant difference between students' self-assessment scores and teachers' assessment scores regarding components of speaking skills?</u>

This research question aimed to investigate whether the students' and the teachers' assessment of different components/subskills of speaking skills differ from one another. In order to find an answer to this question, the data were analysed with Paired Sample T-Tests and the results are presented in the tables below. Each language component assessed separately in the student self-assessment scale and in the teachers' assessment scale are analysed and presented separately.

#### 4.2.2.1. Assessment of Grammar

In this section, the self-assessment and teacher assessment of students' use of English grammar during the speaking exam are examined. Questions 6 to 10 in the rubric are considered to grammar component of speaking skills. Teachers' rubric for speaking exam involves a section for grammar, too. They are shown in the table below:

Table 4.2
Questions (6-10) related to Grammar and Teachers' Rubric for Grammar

| Students' Self-Assessment Rubric | Teachers' Rubric for Speaking Exam |
|---|---|
| 6. I can use grammar correctly in general. | Almost all structures are accurate, even complex ones. (5) |
| 7. I can use grammar correctly in a classroom discussion. | Most structures are accurate but complex ones cause some difficulty. (4) |
| 8. I can use grammar correctly in a speaking exam. | Simple structures are mostly accurate but little attempt at complex ones is made. (3) |
| 9. I can use complex grammatical structures. | Simple structures are often inaccurate, more complex ones not even attempted. (2) |
| 10. I can use transition words when I speak. | Simple structures almost always distort and there is little awareness of any grammatical rules. (1) |

Table 4.3 below shows the mean scores for the assessment the use of grammar regarding the students' proficiency levels.

Table 4.3.
Results of Paired Sample T-test and Descriptive Statistics for the Assessments of Students' use of Grammar

| Proficiency Level | Student Assessment | | Teacher Assessment | | | 95% CI for Mean Difference |
| | M | SD | M | SD | N | P |
| --- | --- | --- | --- | --- | --- | --- |
| AR | 2.6852 | .8675 | 3.3148 | .7604 | 54 | *0.00 |
| B | 2.8265 | .8253 | 3.6391 | .6303 | 340 | *0.00 |
| BR | 2.6545 | .4009 | 3.5909 | .4908 | 11 | *0.00 |
| C-CR | 2.9712 | .7697 | 3.7235 | .6218 | 132 | *0.00 |
| Total | 2.8443 | .8127 | 3.6263 | .6478 | 537 | *0.00 |

*: Correlation is significant at the 0.05 level

Table 4.3 shows the differences between the mean scores of the speaking skill assessment made by the teachers and the students in terms of students' the use of grammar. It can be stated that as the level of the students increased, the mean scores for the use of grammar increased as well. To illustrate, the AR group had the lowest proficiency level, so the mean score of their use of grammar was the lowest. These results revealed that the students with lower proficiency levels tended to underestimate the quality of their own grammar use. It could be stated that the evaluation process improved as the familiarity with and exposure to the target language increased. Also, the teachers assessed the use of grammar similarly to the way students self-assessed their use of grammar. The students with a higher level of English received higher scores for their use of grammar, which is a natural result caused by their language knowledge. Likewise, the students with lower proficiency levels received lower scores for their use of grammar, as expected.

When the mean scores in total were compared, the results revealed that the teachers assessed the students' use of grammar with a higher mean score (M=3.6263) than the students self-assessed their skills (M=2.8443). Paired Sample T-Test showed that there was a significant difference between the two groups' assessments of the use of grammar during their speaking performance (P=0.00). When analysed according to level, there were significant differences between the teachers' assessment and the students' assessment at all levels (P=0.00). As a result, it could

64

be concluded that the self-assessment process improved as the familiarity with and/or exposure to the target language increased.

**4.2.2.2. Assessment of Vocabulary**

In this section, the assessment of vocabulary knowledge in speaking performances by the students and the teachers are examined. In the rubric, the vocabulary component of speaking skills is explored by the researcher in questions 11 to 16. Also, teachers' rubric for the speaking exam includes a part devoted to vocabulary. They are shown in the table below.

Table 4.4.
Questions (11-16) related to Vocabulary and Teachers' Rubric for Vocabulary

| Students' Self-Assessment Rubric | Teachers' Rubric for Speaking Exam |
|---|---|
| 11. I can guess the meaning of unknown words from the context when I am given prompts. | Appropriate level and variety of vocabulary to deal with topic, avoiding repetition. Able to paraphrase effectively if necessary. (5) |
| 12. I can use words and phrases when I speak English. | Level and variety of vocabulary satisfactory, but experiences some difficulty in word choice and usage, occasional repetition. Some ability to paraphrase when necessary. (4) |
| 13. I can remember the proper words when I speak English. | Vocabulary used is basic but sufficient to express ideas, some repetition. Little evidence of extended range. Limited ability to paraphrase and some difficulty experienced. (3) |
| 14. I can use different words when I speak. | Vocabulary use tends to be inaccurate, with inappropriate or irrelevant use, making communication awkward, though shows some attempts. (2) |
| 15. I can deal with my hesitations in classroom activities when I cannot remember a word. | Almost all of the vocabulary is inaccurate, inappropriate or irrelevant, making communication difficult or with no attempt at all. (1) |
| 16. I can use appropriate vocabulary. | |

Table 5 shows the mean scores for the student self-assessment and teacher assessment of students' vocabulary knowledge regarding the students' proficiency levels.

Table 4.5.
Results of Paired Sample T-test and Descriptive Statistics for Assessments of
Students' Vocabulary knowledge

| Proficiency Level | Student Assessment | | Teacher Assessment | | | 95% CI for Mean Difference |
|---|---|---|---|---|---|---|
| | M | SD | M | SD | N | P |
| **AR** | 3.0 | .8782 | 3.6296 | .7908 | 54 | *0.00 |
| **B** | 3.2349 | .8047 | 3.9660 | .6679 | 338 | *0.00 |
| **BR** | 2.9273 | .5951 | 3.8182 | .6431 | 11 | *0.04 |
| **C-CR** | 3.3682 | .7783 | 4.0076 | .6269 | 132 | *0.00 |
| **Total** | 3.2378 | .8081 | 3.93 | .6777 | 535 | *0.00 |

*: Correlation is significant at the 0.05 level

In Table 4.5, the mean scores of the teacher and the student assessments were analysed in terms of vocabulary knowledge and use reflected during the speaking performance of the students. As can be inferred from the table, there was a significant difference between the students' and the teachers' assessment of vocabulary knowledge and use in speaking assessment in total (P=0.00). It was revealed that the students (M=3.2378) undervalued their vocabulary knowledge when compared to their teachers' assessments (M=3.93).

In addition, Table 4.5 showed the mean scores of the students' assessments according to their proficiency levels. Apart from BR level students, it can be seen that the students' proficiency level increased as the mean scores for vocabulary assessment increased. For instance, the students in AR classes had the lowest mean score (M= 3.0) while the students in C-CR classes had the highest mean score (M=3.3682). It might be claimed that the proficiency level can be an important factor for the self-assessment of the use of vocabulary during their speaking performance. Furthermore, the teachers assessed the vocabulary knowledge of the students' speaking performances as the students did. The teachers assessed the students with higher proficiency levels with higher scores (M= 4.0076) for C-CR level) and the students with lower proficiency levels with lower scores (M=3.6296), which might be a natural result when the students' levels of proficiency were taken into account.

Moreover, these mean scores were analysed with Paired Sample T-test and it was revealed that the mean scores of the student assessments were significantly different

from the mean scores of the teacher assessment in all levels of proficiency (p=0.00). These results demonstrated that the higher the students' levels were, the more successful they were at assessing their knowledge and use of vocabulary.

### 4.2.2.3. Assessment of Fluency

In this section, results regarding student self-assessment and teacher assessment of fluency during speaking performances are presented. There is only one question in the rubric about fluency, question 4; "I can speak English fluently". (see Appendix). Teachers' rubric for the speaking exam has a fluency section explaining how to grade students' fluency skills. Table 4.6 shows these criteria below:

Table 4.6.
Teachers' Rubric for Fluency

| Fluency |
| --- |
| Almost no hesitation other than native-like search for ideas. The speech is very natural. (5) |
| Most of the speech is without hesitation, which does not disturb the natural flow. (4) |
| Some hesitation with some short pauses, but does not disturb the flow for the listener. (3) |
| Speech is halting with some long pauses. Frequent hesitation, needs prompting but shows attempt. (2) |
| Speech is disconnected and difficult to follow. (1) |

Table 4.7 shows the mean scores for the assessment of speaking fluency in relation to the students' proficiency levels.

Table 4.7.
Results of Paired Sample T-test and Descriptive Statistics for Assessments of Students' Fluency

| Proficiency Level | Student Assessment | | Teacher Assessment | | | 95% CI for Mean Difference |
| --- | --- | --- | --- | --- | --- | --- |
| | M | SD | M | SD | N | P |
| AR | 2.4630 | .9257 | 3.6944 | .8378 | 54 | *0.00 |
| B | 2.7778 | 1.0262 | 4.0307 | .7284 | 342 | *0.00 |
| BR | 3.00 | .7746 | 4.0455 | .7230 | 11 | *0.04 |
| C-CR | 3.1212 | 1.0486 | 4.1970 | .6928 | 132 | *0.00 |
| Total | 2.8349 | 1.0331 | 4.0380 | .7417 | 539 | *0.00 |

*: Correlation is significant at the 0.05 level

Table 4.7 shows how the students assessed their fluency in speaking English and how their teachers evaluated their fluency in the speaking exam. It can be understood from the mean scores that the teachers assessed the students with a higher score (M= 4.0307) than the students self-assessed themselves (M= 2.8349). Paired Sample T-test results showed that there was a significant difference between the students' and the teachers' assessment of fluency in speaking skills in total (p=0.00).

It was also clear that the students who have different proficiency levels assessed their fluency with different mean scores as shown in Table 4.7. The students with lower levels underestimated their fluency in English when compared to the students with higher proficiency levels. Paired Sample T-test results showed that there were significant differences between the students' and the teachers' assessments of students' fluency in relation to their proficiency levels (p=0.00).

These results showed that the students believed that they were not fluent in speaking English -as they had lower proficiency levels- and vice versa for the students with higher level of proficiency. In other words, both the students with low proficiency and the students with high proficiency regarded themselves as not fluent in English, which is not correlated with teachers' assessments.

### 4.2.2.4. Assessment of Pronunciation

In this section, the students' and the teachers' assessment of students' pronunciation during their speaking performances are presented. There is only one question the questionnaire referring to pronunciation, question 5; "I can pronounce the words accurately when I speak English" (see Appendix A). Criteria related to pronunciation skills are also included in teachers' rubric for speaking exam. They are in table 4.8 below:

Table 4.8.
Teachers' Rubric for Pronunciation

| Pronunciation |
|---|
| Almost all individual sounds are well articulated, with effective and natural use of stress and intonation. Easy to understand and follow. |
| Most individual sounds well-articulated, occasional difficulties do not disrupt comprehension. |
| Some individual sounds are not clearly articulated but not in a way that disrupt comprehension. |
| Many individual sounds poorly articulated and pronunciation puts strain on listeners and causes misunderstanding. |
| Impossible to understand at all. |

Table 4.9 presents the mean scores for the pronunciation component of the student self-assessment and the teachers' assessment with regards to the students' proficiency levels.

Table 4.9.
Results of Paired Sample T-test and Descriptive Statistics for Assessments of Students' Pronunciation

| Proficiency Level | Student Assessment | | Teacher Assessment | | | 95% CI for Mean Difference |
|---|---|---|---|---|---|---|
| | M | SD | M | SD | N | P |
| AR | 2.8148 | .9727 | 3.9444 | .6491 | 54 | *0.00 |
| B | 3.2953 | 1.0263 | 4.1886 | .5528 | 342 | *0.00 |
| BR | 3.1818 | .8738 | 4.1818 | .6030 | 11 | 0.24 |
| C-CR | 3.5530 | 1.0137 | 4.3220 | .5283 | 132 | *0.00 |
| Total | 3.3080 | 1.0319 | 4.1967 | .5656 | 539 | *0.00 |

*: Correlation is significant at the 0.05 level

Table 4.9 demonstrates that the mean score of the teachers' assessment of the students' pronunciation (M= 4.1967) was higher than the students' self-assessment of their pronunciation abilities in total (M=3.3080). When these means were compared with Paired Sample T-Test, it was found out that these two mean scores were significantly different from each other, which showed that the teachers and the students assessed the pronunciation abilities of the students' differently. Thus, it could be inferred that the students underestimated their pronunciation performance which might be due to the fact that they were not proficient enough to be conscious of their pronunciation abilities.

Table 4.9 also shows the mean scores of the pronunciation assessment for different proficiency levels. The mean scores showed that there were differences among the students' proficiency levels and their self-assessment of their pronunciation.

Namely, the students with a lower level evaluated their pronunciation lower than the students with higher level of proficiency. The teacher and the students' mean scores were analysed through Paired Sample T-test and the results revealed that the teachers' and the students' assessment differ significantly from each other at all levels, except in the BR level (p=0.24).

### 4.2.2.5. Assessment of Communication Skills

In this section, results concerning the self-assessment and teacher assessment of students' communication skills during speaking performances are presented. For the assessment of the communication skills, the researcher of the present study used questions 20 to 25 in the rubric. While the questions were prepared, the rubric that the teachers used in the speaking exam was taken into consideration and in the content part of the rubric there were explanations on how well students communicated, answered questions, and supported their ideas. The students assessed their communication skills through these questions. The questions are shown in the table below:

Table 4.10.
Questions (20-25) related to Communication Skills and Teachers' Rubric for Communication Skills

| Students' Self-Assessment Rubric | Teachers' Rubric for Communication Skills |
| --- | --- |
| 20. I can communicate effectively in speaking exams. | Communicates effectively, responds well to the topic. Explains well, gives effective examples. (5) |
| 21. I can respond to questions in speaking exams. | Most of the speech responds to topic. However, some supporting details are weak. (4) |
| 22. I can support my opinions in speaking exams. | Responds in a basic way to the topic but lacks relevant supporting details. (3) |
| 23. I can communicate effectively in classroom activities. | States the topic by giving his idea without any supporting details. (2) |
| 24. I can respond to questions in classroom activities. | Unable to respond to the topic or makes almost no attempt. (1) |
| 25. I can support my opinions in classroom activities. | |

Table 4.11 presents the mean scores of the self-assessment and teacher assessment of students' communication skills with regards to the students' proficiency levels.

Table 4.11.
Results of Paired Sample t-test and Descriptive Statistics for Assessments of
Students' Communication Skills

| Proficiency Level | Student Assessment | | Teacher Assessment | | | 95% CI for Mean Difference |
|---|---|---|---|---|---|---|
| | M | SD | M | SD | N | P |
| **AR** | 3.1852 | .8264 | 3.7778 | .8559 | 54 | *0.00 |
| **B** | 3.1998 | .8843 | 4.1418 | .8057 | 342 | *0.00 |
| **BR** | 3.1364 | 1.0269 | 3.5455 | 0.9342 | 11 | 0.126 |
| **C-CR** | 3.5429 | .8961 | 4.1705 | .8284 | 132 | *0.00 |
| **Total** | 3.2811 | .8947 | 4.1002 | .8282 | 539 | *0.00 |

*: Correlation is significant at the 0.05 level

Table 4.11 shows the students' and the teachers' assessment of communication in students' speaking skills. When these scores were analysed, it was clear that the students assessed their communication skills lower than teachers did, which means that the teachers' mean score (M=4.1002) was higher than that of the students in total (M=3.2811). When these two mean scores were compared via Paired Sample T-Test, it was found that the students and the teachers differed in assessing the communication skills of students significantly (p=0.00).

Furthermore, in Table 4.11, it can be also seen that the mean scores for the communication skills assessment made by the teachers and the students were presented according to the students' level of proficiency. There are not considerable differences between the mean scores of students with different levels of proficiency. However, when they were compared with the teachers' assessment, it could be stated that the mean scores were significantly different from each other in all proficiency levels, except for the BR level (p= 0.126). Apart from this, it could also be stated that the students with higher levels such as the C-CR level assessed their communication skills higher (M= 3.5429) than the students' with lower proficiency levels did. Again, it became clear that the proficiency level played an important role in the assessment process of the students' communication skills.

**4.2.2.6. Assessment of Overall Speaking Skills**

In this section, results concerning the self-assessment and teacher assessment of students' overall skills during speaking performances are presented. By overall, questions 1 to 3 in the rubric are referred to (see Appendix). As stated before, the rubric used by the teacher for the assessment included a section for an overall assessment of the students' speaking skills. The first three questions of the rubric asked how the students assessed themselves in general regarding all other components included in the rubric. They are shown in Table 4.12 below:

Table 4.12.
Questions (1-3) related to Overall Skills and Teachers' Rubric for Overall Skills

| Students' Self-Assessment Rubric | Teachers' Rubric for Overall Skills |
|---|---|
| 1. I can speak in a discussion. | Superior (5) |
| 2. I can speak during a role play. | Good (4) |
| 3. I can speak in a speaking exam. | Average/Satisfactory/Acceptable (at the level) (3) |
| | Weak (2) |
| | Very Poor (1) |

Table 4.13 presents the mean scores of self-assessment and teacher assessment of students' overall skills with regards to the students' proficiency levels.

Table 4.13.
Results of Paired Sample T-test and Descriptive Statistics for Assessments of Students' Overall Speaking Skills

| Proficiency Level | Student Assessment | | Teacher Assessment | | | 95% CI for Mean Difference |
|---|---|---|---|---|---|---|
| | M | SD | M | SD | N | P |
| **AR** | 2.5864 | .8494 | 3.6481 | .7244 | 54 | *0.00 |
| **B** | 2.9737 | .9152 | 3.9898 | .6559 | 342 | *0.00 |
| **BR** | 3.1515 | 1.0259 | 3.7727 | .6466 | 11 | *0.033 |
| **C-CR** | 3.4091 | .9085 | 4.0712 | .6520 | 132 | *0.00 |
| **Total** | 3.0451 | .9378 | 3.9711 | .6507 | 539 | *0.00 |

*: Correlation is significant at the 0.05 level

Table 13 shows how the students and the teacher assessed the students' speaking skills overall. The mean scores of these first three questions and the teachers' mean scores for overall evaluation were compared with the help of Paired Sample T-test. It was observed that the teachers assessed the students' overall speaking skills

(M=3.9711) higher than the students did (M=3.0451) and the difference between the two mean scores proved to be statistically significant according to the results of the T-Test.

When it comes to analysing the overall speaking assessment with regards to the proficiency levels, Table 4.13 shows the mean scores of the teacher and the student assessments for overall speaking assessment. It can be understood from Table 18 that there were differences between the mean scores regarding the students' proficiency levels. When Paired Sample T-test was applied, the results showed that the differences between the mean scores of the teachers' and the students' assessments were significantly different from each other at a significance of p=0.00 at all levels, except for the BR level whose mean is P= 0.033. This result could indicate that the higher the students' proficiency level was, the higher they assessed themselves and the more successful they were in assessing themselves.

In summary, there were significant differences between the students' self-assessment and the teachers' assessments of different language components of speaking skills. Although these differences were statistically equal (p= .00), the numerical difference was the greatest between the scores of fluency. This might indicate that the students were most unconfident when it came to the self-assessment of their fluency in speaking English. Also, they did not regard themselves fluent enough even though they had higher levels of proficiency.

## 4.3. Relationship Between Self-Assessment, Teacher Assessment and Language Proficiency Levels

The third research question was formulated as:

Do students' self-assessment of speaking skills scores vary as students' proficiency level increase?

The purpose behind this research question was to determine whether the students with different proficiency levels regard their speaking performances significantly different from one another or not. In order to answer this research question, data were collected from four groups of participants whose English language proficiency

levels were determined as AR (Pre-Intermediate), B (Intermediate), BR (Intermediate and Repeat students), C (Upper-Intermediate) and CR (Upper-Intermediate and Repeat students). The data were analysed and tested with Paired Sample T-test and Analysis of Variance (ANOVA) in SPSS. Table 4.14 presents the results of these tests.

Table 4.14.
Results of Paired Sample T-test and Descriptive Statistics for Assessment Scores of Teachers and Students Based on Their Levels

| Proficiency Level | Student Assessment | | Teacher Assessment | | |
| | M | SD | M | SD | N |
| --- | --- | --- | --- | --- | --- |
| **AR** | 57.931 | 14.656 | 73.840 | 13.422 | 53 |
| **B** | 60.717 | 15.296 | 79.582 | 12.033 | 348 |
| **BR** | 58.666 | 76.708 | 76.708 | 10.575 | 12 |
| **C-CR** | 66.287 | 14.703 | 80.856 | 13.388 | 132 |

As can be seen, Table 4.14 is divided into two sections showing both the students' self-assessment and the teachers' assessment mean scores in total for four groups of participants with different proficiency levels. The group with the highest self-assessment mean score was the C-CR group, with the highest proficiency level (M=66.287), and the group with the lowest self- assessment mean score was the AR Group, with the lowest proficiency level (M=57.931). Similarly, the teachers' assessments showed that the lowest proficiency level group had the lowest mean score (M=73.84) and the highest proficiency level group had the highest mean score (M=80.856). These results revealed that the proficiency level of the students could be an important factor for the self-assessment skills. Since the students with higher proficiency levels might be assumed to be more aware of their language skills, they may have been more successful at assessing their own speaking skills. The teachers' assessments also showed that proficiency is an important indicator of developing speaking skills.

Furthermore, when the mean scores were analysed, the results indicated that there were differences among the mean scores of each group. Table 4.15 presented the results of ANOVA test conducted so as to test whether these differences were statistically significant within the groups.

Table 4.15.
Differences between Students Self-Assessments in relation to Students'
Proficiency Levels

| | Source of Variance | Sum of Squares | N | Mean Square | F | P |
|---|---|---|---|---|---|---|
| **Students Self-Assessment** | Between groups | 4036.85 | 3 | 1345.618 | 5.938 | 0.001* |
| | Within groups | 123274.32 | 544 | 226.607 | | |
| | Total | 127311.182 | 547 | | | |

*: Correlation is significant at the 0.05 level

As can be seen in Table 4.15, the mean score of the students' self- assessment showed a significant difference in terms of the proficiency level variable. To determine within which groups the differences existed, Post-hoc Test was conducted. The differences among groups are shown in the Table 4.16 below:

Table 4.16.
Differences between Students Self-Assessments in relation to Students'
Proficiency Levels (Post-hoc Test Analysis)

| | Mean Difference AR | P | Mean Difference B | P |
|---|---|---|---|---|
| **C-CR** | 8.355 | *0.04 | 5.570 | *0.02 |

*: Correlation is significant at the 0.05 level

As it is shown in Table 4.16, the self- assessment mean scores of the students in the C-CR group were significantly higher than those of the students in the AR group and the B group. In addition, it was also seen that there was no significant difference among the C-CR group and the BR group when compared to other groups.

These result demonstrated that the students with higher proficiency levels tended to assess their speaking skills more accurately than the students with lower proficiency levels. Based on these results, it could be concluded that the students become more successful at assessing their speaking skills when they have higher levels of English proficiency skills.

When it comes to the teachers' assessments of the speaking skills of the participants, Table 4.14 shows that the teachers evaluated the AR group with the lowest score and the C-CR group with the highest score, which is similar to the results of the students' self-assessment mean scores. The mean scores' analysis showed that there

were slight differences among the mean scores of the teacher assessments. ANOVA test was conducted to examine whether these differences were statistically significant and the results are shown in Table 4.17.

Table 4.17.
Differences between Teacher Assessments in relation to Students' Proficiency Levels

|  | Source of Variance | Sum of Squares | Df | Mean Square | F | P |
|---|---|---|---|---|---|---|
| **Teacher Assessment** | Between groups | 2007.573 | 3 | 669.191 | 4.293 | 0.005* |
|  | Within groups | 84328.797 | 541 | 155.876 |  |  |
|  | Total | 86336.370 | 544 |  |  |  |

*: Correlation is significant at the 0.05 level

As can be seen in Table 4.17, the teachers' evaluation mean scores change significantly in terms of proficiency level variable. To determine within which groups the differences existed, Post-hoc Test was conducted. The differences among are shown in Table 4.18.

Table 4.18.
Differences between Teacher Assessments in relation to Students' Proficiency Levels (Post-hoc Test Analysis)

|  | Mean Difference | P |
|---|---|---|
|  | AR |  |
| **B** | 5.7423 | *0.02 |
| **C-CR** | 7.0164 | *0.03 |

*: Correlation is significant at the 0.05 level

In Table 4.18, it is shown that the mean scores given in the teacher assessment of speaking skills of the students in the C-CR group and the B group were significantly higher than those of the students in the AR group. In addition, there was no significant difference among other groups' comparisons.

The difference between the C-CR group and the AR group was similar to the students' self- assessment mean scores. Whereas there was no statistically significant difference between the B group and the AR group in terms of their self-assessment mean scores, there was a significant difference between the teachers' assessment of these groups' mean scores. These results showed that the teachers'

assessments of English speaking skills were in accordance with the students' self-assessments with regards to their proficiency levels.

## 4.4. Relationship Between Self-Assessment and Gender

The next research question is whether gender has a significant effect on students' self-assessments and teachers' assessments of speaking skills. These assessments are analysed separately. The first research question related to gender is:

### 4.4.1. Is There a Significant Difference Between Female and Male Students' Self-Assessment of Their Speaking Skills?

The question asked whether the students' gender creates a difference in their self-assessment of speaking skills. To answer this question, the data gathered from the questionnaire were tested with Independent Sample T-Test and the results are given in Table 4.19 below:

Table 4.19.
Differences between Students Self-Assessments in relation to Students' Gender

|  | Group | N | $\bar{x}$ | Df | Std.D. | T | P |
|---|---|---|---|---|---|---|---|
| **Student Self-Assessment** | Female | 282 | 58.7773 | 541.084 | 16.0223 | -4.854 | 0.00* |
|  | Male | 266 | 64.9534 |  | 13.7295 |  |  |

*: Correlation is significant at the 0.05 level

Table 4.19 shows that the mean score of the female students' self-assessment of their speaking skills was 58.77 whereas the male students evaluated their speaking skills with a higher mean score (64.95). According to the results of Independent Sample T-Test conducted in order to analyse whether this difference was significant, the male students assessed their speaking skills statistically significantly higher that the female students. In other words, the male students perceived their English speaking skills to be better than the female students', which was also supported by the studies in relevant literature.

Moreover, whether gender has an effect on teachers' assessment of speaking skills is also analysed. The second research question related to gender is:

### 4.4.2. Do Teachers' Assessment Scores Regarding Students' Speaking Performance Show a Significant Difference in Relation to Students' Gender?

The question asks whether the students' gender creates a difference in teachers' assessment of speaking skills. To answer this question, the data gathered from the questionnaire were tested with Independent Sample T-Test and the results are given in Table 4.20 below:

Table 4.20.
Differences between Teacher Assessments in relation to Students' Gender

|  | Group | N | $\bar{x}$ | Df | Std.D. | T | p |
|---|---|---|---|---|---|---|---|
| **Teacher** | Female | 279 | 78.418 | 543 | 13.2612 | -1.618 | 0.10 |
| **Assessment** | Male | 266 | 80.162 |  | 11.8215 |  |  |

*: Correlation is significant at the 0.05 level

According to the table, the average score given by teachers to female and male students is very close to each other. The independent samples showed no significant difference between the grades given to the male and female students as a result of the t-test. Accordingly, it can be said that the teachers evaluate students fairly in terms of gender variable. This result is in line with the studies in the relevant literature.

### 4.5. Extracurricular Activities' Effect on Self-Assessment

Another aim of the present study was to investigate whether the students assessed themselves better or worse according to their use of the English language outside the classroom. The students were asked about their use of English outside the school with questions 14-17. They are shown in the table below:

Table 4.21.

Questions (14-17) related to Extracurricular Activities in the Rubric

| | | | |
|---|---|---|---|
| 14. Do you watch TV series in English? | | **YES** ☐ | **NO** ☐ |
| If yes, how many hours in a day? | | | |
| 1-1 | 2-4 | 4-6 | 6 and more |
| 15. Do you listen to radio or music in English? | | **YES** ☐ | **NO** ☐ |
| If yes, how many hours in a day? | | | |
| 1-1 | 2-4 | 4-6 | 6 and more |
| 16. Do you play video games in English? | | **YES** ☐ | **NO** ☐ |
| If yes, how many hours in a day? | | | |
| 1-1 | 2-4 | 4-6 | 6 and more |
| 17. Do you use websites in English? | | **YES** ☐ | **NO** ☐ |
| If yes, how many hours in a day? | | | |
| 0-1 | 2-4 | 4-6 | 6 and more |

As it can be seen in Table 4.21 above, watching TV series, listening to music/radio, playing online/video games, and/or using websites in English are the context that students may use English outside the classroom. Each of these language use contexts is examined separately.

## 4.5.1. Relationship Between Watching TV Series in English and Self-Assessment of Speaking Skills

When Table 4.21 is examined, it can be stated that most of the students who watched TV series in English spend 2 to 4 hours for watching TV. Only six students in the group stated that they watched TV series for more than 6 hours in a day. The mean scores the students gave to themselves change depending on how many hours they watch TV series. As it can be inferred from Table 4.22, the more they watched TV series in English, the higher scores the students gave to themselves (M=75.33). This might be due to the fact that the students were exposed to and were able to learn daily and colloquial language used in the TV series they watched. They may have tried to use them in their own speaking and in turn they may have counted themselves as successful in speaking.

Table 4.22.
Students Self-Assessments and time spent on Watching TV Series

|  | N | x̄ | Std.D. | Min. | Max. |
|---|---|---|---|---|---|
| **Never** | 84 | 58.18 | 16.774 | 23.2 | 95.2 |
| **Up to 2 hours** | 201 | 60.65 | 14.934 | 20 | 100 |
| **2-4 hours** | 237 | 62.88 | 14.662 | 20.8 | 100 |
| **4-6 hours** | 19 | 70.27 | 10.852 | 49.6 | 86.4 |
| **Over 6 hours** | 6 | 75.33 | 19.644 | 54.4 | 100 |
| **Total** | 547 | 61.73 | 15.242 | 20 | 100 |

ANOVA test was applied to find out whether this difference among these groups was significant or not. There were differences among groups with reference to the amount of the time that they watched English TV series. Based on Levene Test result (p=.06) conducted in order to determine which groups showed differences, Tukey HSD test was applied:

Table 4.23.
Students Self-Assessments and time spent on Watching TV Series

|  | **Mean Difference** | **P** |
|---|---|---|
|  | **Never** |  |
| 4-6 hours | 12.092 | *0.001 |

*: Correlation is significant at the 0.05 level

The Students who watched English TV series for 4-6 hours gave higher scores to themselves when compared to those who never watched English TV series. However, there was no significant difference among the other groups.

## 4.5.2. Relationship Between Listening to Music/Radio in English and Self-Assessment of Speaking Skills

In this section, listening to music/ radio in English outside the classroom was analysed. Table 4.24 shows that the mean scores that the students gave to themselves increased as the time the students' spend on listening to the music/radio in English increased. In other words, while the mean score for the students who listened to the music/radio in English more than 6 hours in a day is 71.40 that of students who never listened to music/radio is 56.61.

Table 4.24.
Students Self-Assessments and time spent on Listening to Music/Radio

|  | N | x̄ | Std.D. | Min. | Max. |
|---|---|---|---|---|---|
| **Never** | 44 | 56.61 | 13.995 | 32.8 | 92.8 |
| **Up to 2 hours** | 239 | 59.86 | 15.344 | 20 | 96.8 |
| **2-4 hours** | 184 | 63.23 | 14.405 | 20.8 | 100 |
| **4-6 hours** | 56 | 65.45 | 15.505 | 27.2 | 92 |
| **Over 6 hours** | 24 | 71.40 | 15.732 | 41.6 | 100 |
| **Total** | 547 | 61.81 | 15.237 | 20 | 100 |

In order to find out whether this difference among the groups was significant or not, ANOVA test was applied and the results showed that there were significant differences among the groups in relation to the amount of the time they listened to English songs. Based on Levene Test result (p=.76) conducted in order to determine which groups showed differences, Tukey HSD test was applied (Table 4.25).

Table 4.25.
 Students Self-Assessments and time spent on Listening to Music/Radio

|  | Mean Difference | P | Mean Difference | P |
|---|---|---|---|---|
| Hours | Never |  | Up to 2 hours |  |
| **4-6 hours** | 8.83896 | *0.00 | - | - |
| **Over 6 hours** | 14.78182 | *0.00 | 11.53389 | *0.00 |

*: Correlation is significant at the 0.05 level

The students who listened to English songs for 4-6 hours gave higher scores to themselves when compared to those who never listened to English songs. Also, the students who listened to English songs more than 6 hours in a day gave higher scores to themselves in comparison to both those who never listened to English songs and to those who listened to English songs for at most 2 hours. However, there was not any significant difference among the other groups. Depending on these results, it became clear that listening to English songs for longer hours could be related to higher scores that students gave themselves.

## 4.5.3. Relationship Between Playing Online/Video Games in English and Self-Assessment of Speaking Skills

When the role of playing online/video games in English on the students' self-assessment of their speaking skills was analysed, it was revealed that the students

assessed their speaking skills better if they played online/video games for longer hours. In other words, as the amount of the time spent playing video game increased, the self-assessment scores rise. As it can be seen in Table 4.26, there were differences among the students in terms of their self- assessment scores based on the time they spent on playing video games in English.

Table 4.26.
Students Self-Assessments and time spent on Playing Online/Video Games

|  | N | x̄ | Std.D. | Min. | Max. |
|---|---|---|---|---|---|
| Never | 316 | 58.164 | 15.640 | 20 | 100 |
| Up to 2 hours | 105 | 64.403 | 13.410 | 24 | 100 |
| 2-4 hours | 88 | 68.218 | 11.126 | 31.2 | 97.6 |
| 4-6 hours | 16 | 63.150 | 20.415 | 26.4 | 89.6 |
| Over 6 hours | 20 | 74.640 | 10.330 | 54.4 | 100 |
| Total | 545 | 61.740 | 15.261 | 20 | 100 |

ANOVA test was applied in order to reveal if these differences were significant or not and to determine in which groups these significant differences existed, Post-Hoc test was applied. Based on Levene Test result (p=.00), it was identified that the variances did not show homogeneity, therefore; Dunnet's T3 test was conducted and the results are shown in Table 4.27.

Table 4.27.
Students Self-Assessments and time spent on Playing Online/Video Games

|  | Mean Difference | P | Mean Difference | P |
|---|---|---|---|---|
| Hours | Never |  | Up to 2 hours |  |
| 2-4hours | 10.05362 | *0.00 | - | - |
| Over 6 hours | 16.47544 | *0.00 | 10.23619 | *0.00 |

*: Correlation is significant at the 0.05 level

The students who played video games in English for 2-4 hours gave higher scores in their self-assessment when compared to those who never played video games. Moreover, the students who played video games more than 6 hours in a day gave higher scores to themselves in comparison to those who never played video games in English and to those who played video games for at most 2 hours. However, there was not any significant difference among the other groups. As the results show, playing video games in English for longer hours could be related to a better self-assessment process and higher scores.

### 4.5.4. Relationship Between Visiting Websites in English and Self-Assessment of Speaking Skills

As it can be seen in Table 4.28, the scores which the students gave to themselves increased as the time they spend on visiting websites in English in a day increased. The highest mean score was given by the students who read websites for more than 6 hours (M=82.20) whereas the lowest mean score was given by the students who never used websites in English (M=54.12).

Table 4.28.
Students Self-Assessments and time spent on Using Websites in English

|  | N | x̄ | Std.D. | Min. | Max. |
|---|---|---|---|---|---|
| **Never** | 135 | 54.121 | 16.512 | 20 | 92.8 |
| **Up to 2 hours** | 302 | 63.711 | 113.592 | 21.6 | 100 |
| **2-4 hours** | 95 | 65.044 | 14.671 | 26.4 | 100 |
| **4-6 hours** | 11 | 66.109 | 15.406 | 36.8 | 81.6 |
| **Over 6 hours** | 4 | 82.200 | 13.582 | 71.2 | 100 |
| **Total** | 547 | 61.759 | 15.265 | 20 | 100 |

The ANOVA test was applied to analyse group differences and to determine in which groups these significant differences existed, Post-Hoc test was applied. Based on Levene Test result (p=.03), Dunnet's T3 test was conducted and the results are shown in Table 4.29.

Table 4.29.
Students Self-Assessments and time spent on Using Websites

| **Mean Difference** |  | **P** |
|---|---|---|
|  | **Never** |  |
| **Up to 2 hours** | 9.58978 | *0.00 |
| **2-4 hours** | 10.92273 | *0.00 |

*: Correlation is significant at the 0.05 level

As it can be seen in Table 4.29, those who used English websites up to 2 or 4 hours a day had significantly higher self-assessment scores than those who never used English websites. In brief, it can be deduced that using websites in the target language might have a positive effect on the self-assessment of the students' speaking performance.

In conclusion, the results generally show that students and teachers assess speaking skills of English significantly differently from each other. The data is analysed both for general speaking skills and components of speaking skills. It is understood that students assess their speaking skills significantly and relatively lower than teachers do. The same results are obtained for assessment of grammar, vocabulary, fluency, pronunciation, communication and overall speaking skills.

It is also revealed that the effect of proficiency on self-assessment is significant. The relationship between self-assessment and student proficiency is strong because the students with higher level of proficiency assess their speaking skills better than the ones with lower proficiency. It can be concluded that students with higher proficiency are more accurate and successful in self-assessment of speaking skills.

Moreover, there is also a statistically significant difference in self-assessment regarding students' gender. Male students assess their speaking skills with higher scores n contrast to the female students.

Finally, the results also reveal that there is a significant relationship between the use of English outside the classroom and self-assessment. The more students are exposed to English, the better they are at assessing their speaking skills. It can be inferred that students are able to self-assess better when they are more exposed to the target language, in turn, are proficient in English.

# CHAPTER 5

# DISCUSSION AND CONCLUSION

## 5.1. Introduction

In this chapter of the study, the summary and discussion of the results, conclusion, limitation and further implications for future research are given.

## 5.2. Overview of the Study

The main aim of the present study is to examine the relationship between students' self-assessment and teacher assessment. In particular, the aim is to determine whether students and teachers assess students' speaking skills of English statistically differently from each other or not. Since an analytical approach is adopted to assess the students speaking skills, the scores that students and teachers have given for different components of speaking performance such as the use of grammar and vocabulary, fluency and pronunciation are also analysed to determine whether the students and teachers assess these components significantly differently from each other or not. Last but not least, the study investigates whether there is a relationship between extracurricular activities and self-assessment of speaking skills. To answer these questions, a self-assessment rubric including items related to the assessment of speaking skills is prepared and applied to 549 students who study in TOBB ETU English Preparatory Programme. The rubric includes items about the participants' demographic information and their daily use of English outside the classroom in the first section. In the second section, there are 25 items asking the participants to assess their speaking performance in English in terms of their use of grammar,

vocabulary knowledge, fluency in English, pronunciation accuracy and their communication skills. The items in the rubric are prepared according to the rubric which is used by the teachers in the speaking exam in which the participants are assessed (see Appendix 2 for the rubric) in order to ensure content consistency in the student and the teacher participants' evaluation of the speaking performances. The data gathered through the rubric are compared to the students' speaking exam scores that are given by their teachers. The results are summarized in the following paragraphs.

### 5.2.1. The Relationship Between the Students' Self-Assessment and the Teachers' Assessment

The first research question aims to analyse the relationship between the students' self- assessment and the teachers' assessment of the students' speaking skills in English language. The results show that the teachers and the students assess the speaking skills significantly differently from each other. The mean score of the teachers' assessment is 79.2 out of 100 while the mean score of the students' self-assessment is 61.7. These scores allow us to conclude that the teachers assess the students' speaking skills better than the students themselves. This may be because of the fact speaking is one of the hardest language skills to assess since it is a productive skill (Harris, 1997) because speaking assessment includes a mixture of abilities that may not be correlated with each other. For instance, there may be some students with good pronunciation but not with sensible sentences or students with imperfect grammar and pronunciation but they still can deliver the message to the listener (Kiato and Kiato, 1996). Also, it is important to know that speaking includes components which may not be assessed objectively. Although it is plausibly not possible to refrain from subjectivity, developing a system of assessing is essential to apply assessment procedures as objectively as possible (Kiato and Kiato, 1996).

Another reason for the difference between the students' self-assessment and the teachers' assessment may be related to the students' lack of confidence in their speaking abilities. In Turkey, English is taught as a foreign language, as a result; students do not have adequate opportunities to produce English for communicative

purposes outside the classroom despite the fact that they are exposed to the target language intensively in preparatory classes. Lack of opportunity to use English naturally limits their development of speaking skills. Therefore, it can be stated that students may have difficulty in performing orally in the target language, consequently they may not feel confident when they speak and consequently, they underevaluate their speaking abilities.

They may also not be aware of their speaking competence because they don't perform in target language, especially, when they have the chance to talk in their mother tongue, they prefer not speaking English.

Related literature shows us similar results for the difference between students' and teachers' assessments. For instance, in a study conducted by Babaii et.al. (2015), the correlation between the students' and the teachers' assessment of speaking skills is investigated. Their results show that there was a significant difference between these two assessments (Babaii et. al, 2015). As it is interpreted by the authors, the difference between the assessments indicates that the students are not able to evaluate their own abilities before they are informed about the assessment criteria. It is also suggested that improved comprehension of scoring criteria may cause higher accuracy for self-assessment of speaking (Babaii et. al, 2015). Another study by Blue (1988) is conducted on the relationship between the teacher and the student assessment of their language abilities such as communication skills. The results show that the association between the teacher and student assessment is not strong and the difference is significant. The author interpreted these results as a conclusion of students' being unsure of the actual aim of self-assessment application and, surely, students' inexperience. It may be necessary for students to practice more before they are trained to assess their language abilities (Blue, 1988).

The results of the present study are also in line with a study conducted by Nurov (2000) in the Turkish context on the correlation between the teacher assessment and the students' self-assessment in the students' language skills such as grammar, vocabulary, reading, listening and writing. The results show that the students assess themselves differently from their teachers in all the skills above. According to Nurov

(2000), the results indicate that the students may not be knowledgeable enough about their truly existing language skills stated above. He also added that the results may be because of the fact that students encountered such an uncommon way of assessment for the first time. Thus, the students may have had difficulties to correctly evaluate themselves owing to the unfamiliarity of the assessment tool and inadequate training with it (Nurov, 2000).

Singh (2015)'s study on self-assessment of speaking skills investigated whether there is a significant difference between the students' and the teachers' assessment of speaking skills. The results also showed a significant difference, however; in this study, the students overrated their speaking skills, which is a different result from the present study. Singh alleged that the results may be due to the fact that students may have given higher scores to themselves because of peer pressure. Also, it would become easier for students to assess themselves if they were given some guidelines or criteria beforehand. This would also help students to notice the elements that they should enforce (Singh, 2015).

To sum up, these results may express that the students do not assess their speaking skills as their teachers do. The present study confirms the results of the previous research and it can be stated that self-assessment of speaking skills may be applied with proper training and teacher surveillance. Moreover, students are not accustomed to using self-assessment tools in language education. In Turkish education system, the grading system is based on teacher assessment. According to the Rules and Regulations of Foreign Language Examinations by Ministry of Education (published in official gazette in 2016), a minimum of two written exams are required for each course regardless of the number of hours per week. The number of exams and dates are determined at the beginning of each semester and determined by the school head after the approval of the school principal. The necessary measures for the examinations are taken by the school directorate. Also, the exams of foreign language courses are conducted in written and applied form to measure listening, speaking, reading and writing skills (MoE, 2016). Therefore, it can be stated that students' assessment may have been influenced by lack of self-assessment experience and training.

### 5.2.2. Self-Assessment of the Components of Speaking Skills

The third concern of the present study is assessment of speaking skill components. Components of speaking skills studied in the study are grammar, vocabulary, fluency, pronunciation and communication skills. The results show that students tend to underevaluate all of the components of speaking skills mentioned above. When the mean scores of students' self-assessment and teachers' assessments are compared, the analysis indicates that the differences between these mean scores are statistically significant (p= 0.00) and the mean score of the teachers' assessment is relatively higher than the students' assessment. Moreover, the scores of these components change according to the students' level of proficiency. The mean scores increase as the proficiency levels increase. To clarify, students' self-assessment of grammar is rated lowest (M=2.68) by the group who has the lowest level of proficiency (AR) and/or students with the highest proficiency (C-CR) rate their grammar higher than the other groups (M=2.97). This result is also supported by Nurov's (2000) study. In his study, it was revealed that self-assessment of grammar is not correlated with the teacher assessment and the students' exam scores.

Another study by Gardner and Lambert investigates self-assessment of vocabulary and grammar components. The results show that the relationship between the student assessment and external assessment tools is relatively weak. For vocabulary, another study by Zareva (2005) can be shown as an example. In the study, self-assessment of lexical knowledge is analysed and the results show that the students with higher language competence are more successful at assessing their vocabulary knowledge, which is in line with the present study's results. To conclude, it can be said that students are not aware of their vocabulary and grammar knowledge of English. Although upper-intermediate students are more successful at assessing them, there is still a significant difference.

Fluency is the most differently assessed component of speaking skill because the highest difference between the mean scores of teachers' and students' assessment is in fluency component. This result shows that the students are not confident enough about their fluency in speaking. Moreover, regardless of their proficiency levels, the

students believe that they are not fluent in the English language when their mean score is compared to the teachers' assessment. This may be because of the belief that speaking is an intimidating and complicated skill to acquire in any target language (Harris, 1997; Kaito, 1996). Students may never be satisfied with their fluency until they perceive themselves competent enough in English. This is also stated in Delgado (1999) in which self-assessment of Spanish-English bilingual students is investigated. The study reported that the students are not accurate in assessing their overall language fluency in English although they are highly proficient students (Delgado, 1999). The author interpreted these results as a possible conclusion of the fact that students generally have more opportunities to receive frequent feedback and practice on reading and writing than speaking skills. Similarly, teachers may consider correcting students' minor speaking errors unnecessarily punishing; hence, students do not receive corrective or direct feedback on their speaking abilities (Delgado, 1999). Moreover, Leger (2009) stated that the students report fluency as the most difficult skill to manage and their self-confidence in their fluency was quite weak.

Pronunciation is another component of the speaking skill that is analysed in the present study. The results show that the students' self-assessment of pronunciation is lower than that of their teachers. The difference between these two assessments of pronunciation is significantly different. To be more specific, the students evaluate their pronunciation skills lower than their actual ability perceived by their teachers. This may be because of the fact that the students feel insecure about their pronunciation when they hear native speakers or their teachers in the classroom. They have the opportunity to compare their pronunciation and their teachers' or natives'. This result is also confirmed by Raasch (1979) which alleged that pronunciation is the component that the students underevaluate consistently and that the correlations are low with more objective assessments such as teachers' assessments (Raasch, 1979). It is also interpreted in the study (Raasch, 1979) that the accuracy of students' self-assessment frequently changes subject to the language abilities included their self-assessment. Since students find assessing skills demanding to evaluate (Harris, 1997), the discrepancy between teachers and students can be interpreted as a result of assessing pronunciation.

Also, Dlaska and Krekeler (2008) expressed in their study that even advanced students think that pronunciation is challenging to assess and their assessment of pronunciation is not accurate in comparison to their teachers' evaluation (Dlaska & Krekeler,2008).

In a nutshell, students evaluate their vocabulary, grammar, fluency, pronunciation and communication skills lower than their teachers. All in all, the results show how insecure the students feel about their speaking skills as a whole.

### 5.2.3. The Relationship Between the Self-Assessment of Speaking Skills and Proficiency Levels of the Students

The second research question is about the relationship between self-assessment of speaking skills and proficiency levels of students. The results show that the students' self-assessment of their speaking skills changes depending on the students' level of proficiency in English. To illustrate, the group with the highest proficiency level has a mean score of 66.287 while the mean score of the group with the lowest level is 57.931. The difference between these scores is statistically significant and meaningful. Teacher assessment of speaking skills changes depending on students' levels, as well. The mean score for the highest level students is 80.856 and for the lowest is 73.84. Teacher assessment also shows significant differences in terms of students' proficiency levels.

When the total mean score of the teacher assessment is taken into account (M= 79.269), it can be stated that the students with higher level of proficiency have a closer mean score (M=66.287) than the students with lower level of proficiency. As a result, it can be said that the students with higher proficiency level are better at assessing their speaking skills than the students with lower proficiency level. This may be related to their developing language awareness. The higher the proficiency level is, the more aware the learners become. Since they are more aware of what they have learned and how much they can produce, it is easier for students to assess their speaking abilities. It is also suggested by Blue (1994) that language awareness is associated with proficiency and, in turn, with accuracy of self-assessment. Thus,

students advance in evaluating their language skills as they get more proficient in the target language (Blue, 1994).

The difference between the groups shows that proficiency is an important factor in self-assessment. Therefore, self-assessment can be better applied in high-level classrooms on the condition that students are trained beforehand and teachers guide them and implement self-assessment in a controlled way to make it more reliable and valid. By this means, students could be more motivated, become autonomous and therefore, be more involved in the learning process.

Although there is little research on the relationship between self-assessment and proficiency level, the results of the existing studies share similarities with a number of research. For instance, Heilanman (1990) studied self-assessment of language skills and compared the students' responses in reference to their proficiency levels and it was found out that the advanced students are more successful at rating themselves than the lower-level students. In addition, in Blue's study (1994), the self-assessment of students who attend to English for Academic Purposes classes is investigated. It was revealed that self-assessment is attributed to proficiency and the students improve in assessing their language knowledge as they learn more and have higher levels of proficiency (Blue, 1994). Furthermore, another study by Benson (1991) examines self-assessment of English skills of Japanese students whose level of proficiency is low. The study shows that these students with low proficiency levels rate their speaking ability as the lowest because of little exposure to English language.

In brief, proficiency level is an important determiner in self-assessment and, to meet the utmost accuracy, we suggest that self-assessment should be applied to students with higher levels of English.

## 5.2.4. The Relationship Between Self-Assessment and Gender

Another question of the present study is related to how gender influences the students' and the teachers' assessment of students' speaking skills of English. The

results showed that the male students assess their speaking skills as better than the female students. In other words, the male students think that they have better English speaking skills (M=64.9) than the female students do (M=58.7). This result shows that the male students trust in and feel more confident about their speaking skills than the female students. This may be related to the idea that male students may be more autonomous and confident towards language learning. In one study, which is not in ESL/EFL field, it was revealed that the male students might be more inclined to assess their verbal and numerical abilities higher than the female students (Wright and Houck,1995). This result is in line with the present study' result related to gender variable.

However, in Nurov's study (2000), it is reported that there is no significant difference between the male and female students in terms of assessing their English language skills such as grammar, vocabulary, reading, listening and writing (Nurov, 2000). This result is not in agreement with the present study, which states that there are significant differences between male and female students in terms of self-assessment of speaking skills. The difference between Nurov's study (2000) and the present study may be due to the participants' different backgrounds. In this study the students only assess their speaking skills but in the other study students are supposed to assess their other language skills, except speaking. Because the potential differences between genders in self-assessment of speaking skill have not been investigated, the question remains unanswered.

When it comes to the relationship between the teachers' assessment and the students' gender, there is no significant difference between how the teachers' assess male and female students' speaking exam scores. Similar results are reported in relevant literature. For instance, in Bijani and Khabiri's study (2017), it is revealed that there is no relationship between gender and the students' scores of speaking skills (Bijani & Khabiri, 2017). Motallebzadeh's study (2011) also reported no significant difference between female and male students' speaking abilities (Motallebzadeh, 2011). The authors infer from the results that teachers evaluate students' speaking abilities objectively and fairly.

93

### 5.2.5. The Relationship Between Self-Assessment and Extracurricular Activities

The present study is also concerned about the effect of extracurricular activities on students' self-assessment of their speaking skills. These extracurricular activities include watching English TV series, listening to English songs/radio podcasts, playing video/online games in English, using websites in English. The results show that all of the abovementioned activities have an effect on the students' self-assessment. Namely, the students who do these activities for longer hours than the others assess their speaking skills as better and more accurately than the others who spend less time for these activities. The mean score for each activity increases as the hours spent on these activities increase. For instance, the students watching TV series for 4 or 6 hours assess their speaking skills significantly different from the other students who watch TV series for less than 4 hours in a day. In brief, it can be stated that the students who watch English TV series/films outside the school time assess their English speaking skills better than the others who do not. This outcome might result from the students' perception about their English speaking skills simply because they receive more input of the target language, which in turn increases their awareness of their developing competence and their ability to differentiate their accurate and inaccurate use of English while speaking.

Similar results are also found for listening to English songs/radio podcasts, playing video games, and using English websites. As the time spent on listening to English songs increases, playing video games, and using English websites, the mean scores of self-assessment increases, as well. This may be due to the fact that the students understand lyrics of the songs as they listen to them more. They get familiar with the phonotactics of the language which might contribute to the development of their pronunciation. Therefore, they can better perceive their language use and assess their performance. For playing video/online games, it can be stated that the students learn different vocabulary and communicate in English when they play online or video games, therefore; they consider their English speaking skills to be in an upper level. Finally, it can be deduced that using websites in the target language might have a positive effect on the self-assessment of the students' speaking performance,

which might result from the students' exposure to target language more and students' awareness of speaking elements in target language.

In literature, there is no study that examined the relationship between students' involvement in extracurricular activities and their self-assessment of speaking skills. Consequently, the results of the present study cannot be compared to any other studies. However, it can be stated that the length of exposure as a potential contributor of higher proficiency level, affects how students perceive their language competencies and, in turn, self-assessment their performance.

## 5.3. Conclusion

The outcomes of the study, in general, show that there are significant differences between the students' self-assessments and the teachers' assessment of the students' speaking skill and its components such as grammar, vocabulary, fluency, pronunciation and communication skills. Furthermore, the study also revealed that proficiency, gender and spending longer time on extracurricular activities have a significant impact on assessing speaking skills.

The previous studies conducted on the accuracy of the self-assessment showed inconsistent outcomes. Some of the studies show strong relationships between the students' and teachers' assessments (Bachman & Palmer, 1989; Buck, 1992; MacIntyre et.al, 1997; Ross, 1998) whereas others reveal that there are significant differences between these evaluations (Blue, 1988; Blue, 1994; Blanche, 1990; Heilenman, 1990; Nurov; 2000; Babaii et.al, 2015; Singh, 2015). There are a variety of factors leading to this situation, one of which is students' readiness to assess their own performance. In other words, if students are trained on how to apply self-assessment, they are able to make logical judgements about their language learning. Students' lack of knowledge of components such as grammar and vocabulary may lead to inconsistent assessments between students and teachers. In order to have a better self-assessment process, students should be informed about these components. Also, the rubrics that are used in the exams can be shown to students

and the criteria of assessment can also be made familiar to the students so that students can understand how to assess different components.

Another reason for the inconsistent results in the literature review might be related to students' individual differences such as their gender, age, cultural background, level of education and/or their length of exposure to the target language. In this study, among the individual differences, the relationship between the students' self-assessment and their gender and their involvement in extracurricular activities are examined. As indicated above, our results show that gender is significantly influential on students' assessment of speaking skills, namely, the male students tend to assess themselves as better speakers of English compared to the female students. Additionally, the time spent on extracurricular activities has a significant effect on students' assessments. As the time spent on these activities increases, the students assess their speaking skills as better and more accurately due to the increased length of exposure to the target language. That's why, the new motto in language learning and assessment process should be "exposure makes perfect". It can be also alleged that extracurricular activities should be promoted and students should be encouraged to attend and do these activities in order to make them more autonomous, aware and motivated, which may lead them to self-assess themselves better. Schools and language institutions should create different extracurricular activities for this and it may be a part of their language education and assessment. These results reveal that there is a need for further research to consider the role of individual differences on students' self-assessment.

The present study also shows that proficiency is an important factor that affects the assessment process. It is demonstrated that the students with higher proficiency levels are better able to assess their speaking skills than the students with lower proficiency levels. It can be stated that student improve their self-assessment as they become more proficient in the target language. Therefore, if self-assessment is to be implemented in language classrooms, it should be applied with highly proficient, even with advanced students, in order to accomplish a successful assessment process and receive reliable results.

On account of the results stated above, it can be stated that with the help of self-assessment process, students could become more motivated, autonomous and therefore, be more involved about their language learning process. Since autonomous students have the ability of reflecting on their language abilities and collaborating with their peers, they can make realistic assessments of their success during their language learning process. It is essential for both students and teachers to assess the learning process unceasingly for a fruitful autonomy process (Najeeb, 2013).

Moreover, it may be argued that self-assessment assists students to get involved in the learning process and, in turn, increase their motivation. (Nunan, 1988; Oscarsson, 1989) as much as it contributes to the autonomy. As stated by Nurov (2000), making judgements about their own learning may lead to positive attitudes which, in turn, higher motivation and confidence building towards the learning process (Gardner, 1999; Nurov, 2000). In conclusion, increased student motivation and an awareness that self-assessment contributes to language education might help students deal with evaluating their language skills more accurately.

## 5.4. Limitations and Implications for Further Research

As in any study, there are some limitations of the present study. To start with, this study is conducted only in one university. Although the number of participants is relatively high, there is only one context where the students receive the same language education and instruction. Other schools can also be included in further research so that the effect of studying in different language education contexts such as in state schools and private schools, or at secondary or tertiary levels, or in institutions where the medium of instruction is in native language or in target language can also be analysed.

The present study used only two data collection tools, a questionnaire used by the students to self-assess their speaking performance and the students' speaking exam results provided by their teachers. Different data collection tools such as interviews that could be conducted with the students or language learning diaries kept by the

students could have provided a broader picture and a better understanding of the students' self-assessment of their speaking skills. If the rubric and the exam results had been combined with interviews, the research design would have been a blended one. Blended studies give better insights to understand the problem and, in this case, it would give better understandings of why students assess their speaking skills lower than their teachers or why students do not have confidence in themselves when it comes to speaking English.

Another limitation of the study is about the data collection process. Data from the self-assessment rubric were collected during the semester, before or after the speaking exam. This could not be controlled because speaking exams were not administered at one time. Students from different levels took the speaking exam in different weeks of the semester. For instance, intermediate level students took the exam in the 6th week of the semester, but the upper-intermediate level students took the exam in the 8th week. However, the self-assessment rubric had to be applied at one time because of the institution's intense curriculum and exam schedule. The teachers who applied the rubric in their classes were given the rubrics and asked to administer in a specific time which was decided by the school administration. Therefore, when self-assessment rubric was administered, students may have had the speaking exam before or after the self-assessment rubric completion process. This situation could have been controlled by implementing the rubric in different levels in different times before they took the speking exam if the curriculum and exam schedule had provided the ideal conditions.

Further research might also focus on factors that are related to students' extracurricular activities and individual differences such as age, gender, social and cultural background, their attitudes towards learning English or their motivation to learn English. The relationship between such factors and the students' self-assessment would help us better understand the reasons behind how they assess their speaking skills.

# REFERENCES

Allwright, D. (1990). Autonomy in language pedagogy in CRILE. *Working Paper 6. Centre for Research in Education:* University of Lancaster.

Alonso, R. S. (2012). *The Importance of Teaching Listening and Speaking Skills*. Unpublished Masters' Thesis. Murcia University, Department of Language and Literature Didactics, Murcia.

Assis, B. M. (2012). *Self-Assessment of Writing Skills: A Reliable And Valid Tool In An Efl Classroom.* Master's Thesis. New University of Lisbon, Faculty of Social and Human Sciences, Lisbon.

Babaii, E. Taghaddomi, S. & Pashmforoosh, R. (2015). Speaking Self-Assessment: Mismatches Between Learners' and Teachers' Criteria. *Language Testing* 1–27.

Bachman, L. F. & Palmer, A.S. (1989). The construct validation of self-ratings of communicative language ability. *Language Testing, 6,* 14-30.

Balçıkanlı, C. (2010). Learner Autonomy in Language Learning: Student Teachers' Beliefs. *Australian Journal of Teacher Education, 35*(1), 90-103.

Baleghizadeh, S., & Rahimi, A. (2011). The relationship among listening performance, metacognitive strategy use and motivation from a self-determination theory perspective. *Theory and Practice in Language Studies, 1,* 61-67.

Benson, M. (1991). Attitudes and Motivation Towards English: A Survey of Japanese Freshmen. *Relc Journal*, 22, 34-48.

Bijani,H. & Khabiri, M. (2017). The Impact of Raters' and Test Takers' Gender on Oral Proficiency Assessment: A Case of Multifaceted Rasch Analysis. *Journal of Teaching Language Skills*, *36*(3), 1-32.

Bing, X. (2016). A Study of the Effects of Student Self-assessment on the EFL Writing of Chinese College Students. *Journalism and Mass Communication, 6*(2), 91-107.

Blanche, P. (1990). Using standardised achievement and oral proficiency tests for self-assessment purposes. The DLIFLC study. *Language Testing, 7,* 202 – 229.

Blatchford, P. (1997) Students' Self-Assessment of Academic Attainment: Accuracy and Stability From 7 to 16 Years and Influence of Domain and Social Comparison Group. *Educational Psychology*, *17*(3), 345-359.

Blue, G. (1988). Self-assessment of listening comprehension. *International Review of Applied Linguistics*, *16*, 149–156.

Blue, G. (1994). Self-assessment of foreign language skills: Does it work? *CLE Working Paper*, *3*, 18–35.

Blue, G. M. (1988). Self-assessment: The limits of learner independence. In A.

Brindley, G. (1989). Assessing achievement in the learner-centred curriculum. *National Centre for English Language Teaching and Research.* Retrieved from https://files.eric.ed.gov/fulltext/ED386045.pdf.

Brookes, A. & Grundy, P. (Eds.). *Individualization and autonomy in language learning* 100 – 118. London: Modern English Publications in Association with The British Council.

Brown, D. (2004). *Language assessment, principles and classroom practices.* New York: Pearson Education. CEFR. Retrieved from http://www.coe.int/t/dg4/linguistic/source/Framework_EN.pdf.

Brown, G., Andrade, H., Chen, F. (2015). Accuracy in student self-assessment: Directions and cautions for research. *Assessment in Education Principles Policy and Practice* DOI:10.1080/0969594X.2014.996523.

Brown, H. D. (1987). *Principles of language learning and teaching* (2nd edition). Englewood Cliffs, NJ: Prentice Hall Regents.

Brown, H. D. (2007). *Principles of language learning and teaching* (5th edition.). New York: Longman.

Brown, J. & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly, 32,* 653 – 675.

Buck, G. (1992). Listening comprehension: Construct validity and trait characteristics. *Language Learning, 43,* 313 -357.

Burns, A. & Joyce, H. (1997). *Focus on Speaking*. Sydney: National Centre for English Language Teaching and Research.

Chalkia, E. (2012). Self-assessment: an alternative method of assessing speaking skills. *Research Papers in Language Teaching and Learning,* 3(1), 225-239.

Clark, J. L. D. (1981). Language. In T. S. Barrows (Ed.), *A survey of global understanding: Final report* (pp. 87–100). New Rochelle, NY: Change Magazine Press.

Cohen, A. (1994). *Assessing language abilities in the classroom.* Boston: Heinle & Heinle Publishers.

Cooper, P.L. (1984). The Assessment of Writing Ability: A Review Of Research. Graduate Record Examinations. https://doi.org/10.1002/j.2330-8516.1984.tb00052.x.

Dam, L. (1990). Developing awareness of learning in an autonomous language learning context. In R. Duda & P. Riley (Eds.), *Learning styles* 189-197. Presses Universitaires de Nancy.

de Leeuw, J. (2016). Rubrics and Exemplars in Writing Assessment. In Leadership of Assessment, Inclusion, and Learning (pp. 89-110). *Springer International Publishing*, DOI: 10.1007/978-3-319-23347-5_4.

Decker, M.A. (2004). Incorporating guided self-study listening into the language curriculum. *The Language Teacher*, 28(6), 5-9.

Delgado, P., Guerrero G, Goggin, J. & Ellis, B. (1999). Self-assessment of linguistic skills by bilingual Hispanics. *Hispanic Journal of Behavioral Sciences, 21,* 34-46.

Deville, M. & C. Deville. 2003. Computer Adaptive Testing in Second Language Contexts. *Annual Review of Applied Linguistics,* 273–299.

Dickinson, L. (1987). *Self-instruction in language learning*. Cambridge: Cambridge University Press.

Dickinson, L. (1996). *Self-instruction in language learnings* Cambridge: Cambridge University Press.

Dlaska, A. & Krekeler, C. (2008). Self-assessment of pronunciation. *System*, 36(4), 506-516.

Dörnyei, Z. (2001). *Motivational strategies in the language classroom*. Cambridge: Cambridge University Press.

Escribano, D. & Mcmahon, J. (2010). Self-assessment Based on Language Learning Outcomes: A Study with First Year Engineering Students. *Revista Alicantina de Estudios Ingleses* 23, 133-148.

Fahimi, Z. & Rahimi, A. (2015). On The Impact of Self-Assessment Practice On Writing Skill. *Procedia - Social and Behavioral Sciences*, *192*, 730- 736.

Formative and Summative Assessment (n.d.) Retrieved from https://www.niu.edu/facdev/.

Fulcher, G. (1993). The Construction and Validation of Rating Scales for Oral Tests in English as a Foreign Language (Doctoral Thesis) Retrieved from http://languagetesting.info/articles/store/FulcherPhD.pdf .

Gardner, D. (1996). Self-assessment for self-access learners. *TESOL Journal, 5, 3,* 18-23.

Gardner, D. (2000). Self-assessment for autonomous language learners. *Links and Letters*, 7, 49-60.

Gardner, R. & Lambert, W. (1972). *Attitudes and motivation in second language learning.* Rowley, MA: Newbury House.

Garrison, C. & Ehringhaus, M. (2007). *Formative and summative assessments in the classroom.* Westerville, OH: Association for Middle Level Education.

Goh, C. & Aryadoust, V. (2010). Investigating The Construct Validity of MELAB Listening Test Through The Rasch Analysis And Correlated Uniqueness

Modelling. *Spain Fellowship Working Papers in Second Of Foreign Language Assessment,* 8. 31-68.

Gronlund, E. & Cameron, I. J. (2004). *Assessment of student achievement.* Toronto: Pearson.

Hall, C.S. & Hope, A.K. (2016). Tips for Testing Speaking. *TESOL Connections.* DOI:10.1018/0969594X.2014.996523.

Harris, M. (1997). Self-assessment of language learning in formal settings. *ELT Journal, 51,* 12 – 20.

Heilenmann, K.L. (1990). Self-assessment of second language ability: The role of response effects. *Language Testing, 7,* 174-201.

Heisigová, T. (2015). *Self-assessment in English lessons at grammar schools.* Unpublished Master's Thesis. Masaryk University, Department of English and American Studies, Brno.

Holec. H (1981). *Autonomy in Foreign Language Learning.* Oxford: Pergamon.

Hooshangi, R., Mahmoudi, O. & Yousofi, N. (2014). Oral Performance and Self-Assessment Process. *Iranian Rehabilitation Journal*, 12. Retrieved from http://irj.uswr.ac.ir/article-1-350-en.html.

Huang, C. (2016). Achievement goals and self-efficacy: A meta-analysis. *Educational Research Review*, *19*, 119–137.

Huerta-Macias, A. (1995). Alternative assessment; Responses to commonly asked questions. *TESOL Journal, 5, 1,* 8-12

Hughes, D. (1989). *Testing for language teachings* Cambridge: Cambridge University Press.

Ibberson, H. (2012). An Investigation into Learners' and Teachers' Attitudes Towards Self-Assessment According to CEFR Scales (pp 13-24). *Language at The University of Essex Proceedings.*

Johansson, S. (2013). *On The Validity of Reading Assessments Relationships Between Teacher Judgements, External Tests and Pupil Self-Assessments.*

Master's Thesis. University of Gothenburg. Department of Education and Special Education. Gothenburg.

Joyce, B., Weil, M. & Calhoun, E. (2009). Models of teaching (8th ed.). Boston, MA: Pearson Education, Inc.

Kanga, K.N. (2012). *Individual and Paired Oral Proficiency Testing: A Study Of Learners' Preference*. Unpublished Master's Thesis. Michigan State University. TESOL. East Lansing, Michigan.

Karakaya, B. (2017). *Self-Assessment in Second Language Speaking Ability: An Evaluation of Students' Accuracy and Perception of Growth.* Master's Thesis. Bahçeşehir University, English Language Teaching, İstanbul.

Kiato, S. K. & Kiato, K. (1996). Testing communicative competence. *The Internet TESL Journal*, *11*(5).

Kombercová, A. (2006). *Testing Speaking*. Unpublished Master's Thesis. University of Pardubice, Department of English and American Studies, Pardubice.

Kordes, U, Sicherl K. B. & Holcar B. A. (2014) A Model of Formative Assessment in Music Education. *Athens Journal of Education,* 1(4), 295-307.

Krausert, S.R. (1991). *Determining the Usefulness of Self-Assessment of Foreign Language Skills: Post-Secondary ESL Students' Placement Contribution.* Ph.D. Dissertation. University of Southern California, Los Angeles, California.

Leger, D. S. (2009). Self-assessment of speaking skills and participation in a foreign language class. *Foreign Language Annuals*, *42*(1), 158–178.

LeBlanc, R. & Painchaud, G. (1985). Self-assessment as a second language placement instrument. *TESOL Quarterly, 19,* 673 – 687.

Liepa, D. (2011) Self -Assessment Criteria of Integrated Foreign Language Acquisition. *Problems of Education in The 21st Century*, 29.

Lin-Siegler, X., Shaenfield, D. & Elder, A. (2015). Contrasting case instruction can improve self-assessment of writing. *Educational Technology Research & Development, 63*(4), 517-537.

Little, D. (1991) *Learner autonomy. 1: Definitions, issues and problems.* Dublin: Authentik.

McDonald, B. & David, B. (2003) The Impact of Self-assessment on Achievement: The Effects of Self-Assessment Training on Performance in External Examinations. *Assessment in Education: Principles, Policy & Practice,* 10(2), 209-202.

MacIntyre, P., Noels, K. & Clement, R. (1997). Biases in self-ratings of second language proficiency: The role of language anxiety. *Language Learning* Retrieved from https://doi.org/10.1111/0023-8333.81997008.

McMillan, J.H. & Hearn, J. (2008). Student Self-Assessment: The Key to Stronger Student Motivation and Higher Achievement. *Educational Horizons*, 87, 40-49.

McNamara, M. & Deane, D. (1995). Self-assessment activities: Toward autonomy in language learning. *TESOL Journal, 5*(1), 17-22.

Moheidat, A.S. & Baniabdelrahman, A.A. (2011). The Impact of Omani Twelfth Grade Students' Self-Assessment on Their Performance in Reading in English. *Asian EFL Journal*, *13*(1), 48-84.

Montgomery, K. (2001). *Authentic assessment: a guide for elementary teachers*. New York: Longman.

Moritz, C. (1997). *Student Self-Assessment of Language Proficiency: Perceptions of Self and Others.* Retrieved from https://files.eric.ed.gov/fulltext/ED399771.pdf.

Motallebzadeh, K. & Nematizadeh, S. (2011). Does Gender Play a Role in the Assessment of Oral Proficiency? *English Language Teaching*, 4. DOI:10.5539/elt.v4n4p165.

Nair, S., Patil, A. & Mertova, P. (2009). Re-engineering graduate skills - A case study. *European Journal of Engineering Education*, 34, 131-139.

Najeeb, S. S. R. (2013). Learner Autonomy in Language Learning. Procedia - *Social and Behavioral Sciences*, 70, 1238–1242.

Nurov, A. (2000). *Self-Assessment of Foreign Language Achievement: The Relationship Between Students' Self-Assessment, Teachers' Estimates and Achievement Test*. Master's Thesis. Bilkent University, Teaching English as a Foreign Language, Ankara.

Nunan, D. (1988). *The Learner-centred curriculum.* Cambridge: Cambridge University Press.

OECD Centre for Educational Research and Innovation (2008) Retrieved from Http://Www.oecd.org/Site/educeri21st/40601796.Pdf.

Oiler, J., Baca, L. & Vigil, A. (1977). Attitudes and attained proficiency in ESL: A sociolinguistic study of Mexican Americans in the south-west. *TESOL Quarterly, 11,* 173-183.

O'Malley, J. M. & Chamot A. U. (1990). *Learning Strategies in Second Language Learning.* Cambridge: Cambridge University Press.

O'Malley, J. M. & Pierce, L. V. (1996). *Authentic assessment for English language students: Practical approaches for teachers.* New York: Addison-Wesley Publishing Company.

Oscarson, M. (1989). Self-assessment of language proficiency; Rationale and applications. *Language Testing, 6,* 1-14.

O'Sullivan, B. (2012). Assessment Issues in Languages for Specific Purposes. *The Modern Language Journal.* Retrieved from https://doi.org/10.1111/j.1540-4781.2012.01298.

Purpura, J.E. (2004). *Assessing Grammar.* Cambridge: Cambridge University Press.

Putri, G. (2012). *Personalizing The Contents of Speaking Class: A Pursuit of Humanistic Approach in English Education Department.* Master's Thesis. State University of Malang, English Language Teaching, Malang.

Qu, W. & Zhang, C. (2013). The Analysis of Summative Assessment and Formative Assessment and Their Roles in College English Assessment System. *Journal of Language Teaching and Research*, 4. DOI: 10.4304/jltr.4.2.335-339.

Raasch, A. (1980). Self-evaluation in adult education. (L'auto-Evaluation Dans L'enseignement des Adultes). *Recherches et Echanges* 5, 85–99.

Rea-Dickens, P. & Germaine, K. (1996). *Evaluation.* Oxford: Oxford University Press.

Rolheiser, C. & Ross, J.A. (2000). Student Self-Evaluation: What Do We Know?. *Orbit, 30*(4), 33-36.

Rolfe, T. (1990). Self-and-peer-assessment in the ESL curriculum. In G. Brindley (Ed.), *The second language curriculum in action* (pp. 163-186). Sydney: National Centre for English Language Teaching and Research: Macquarie University.

Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experiential factors. *Language Testing, 15,* 1 – 20.

Saito, Y. (2003). The use of self-assessment in second language assessment. Working paper in TESOL and Applied Linguistics (Online). Retrieved from http://www.tc.columbia.edu/academic/tesol/WJFiles/pdf/Saito_Forum.pdf.

Salimi, A. & Ahmadpour, M. (2014). The Effect of Direct Vs. Indirect Written Corrective Feedback on L2 Learners Written Accuracy in EFL Context. *International Journal of English Language and Literature Studies,* 4, 10-19.

Schmitt, N. & Schmitt, D. (2014). A Reassessment of Frequency and Vocabulary Size in L2 Vocabulary Teaching. *Language Teaching*, 47(4), 484–503.

Seong, Y. (2014). Strategic Competence and L2 Speaking Assessment. *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics,* 14(1), 13-24.

Shahrakipour, H. (2014). On The Impact of Self-Assessment On EFL Learners' Receptive Skills Performance. *ASEAN Journal of Teaching and Learning in Higher Education, 6.* 1-13.

Shen, L., Guizhou. G., Wichura, W. & Kiattichai, S. (2007). *The Use of Websites for Practicing Listening Skills of Undergraduate Students: A Case Study at Suranaree University of Technology*, Thailand. Retrieved from https://files.eric.ed.gov/fulltext/ED500929.pdf.

Singh, S. (2015). Self-Assessment of Oral Proficiency among ESL Learners. *ELT Voice - India, 5*(1), 1-7.

Smith, D. (2015). Enhancing Management Students' Professional Presentations Skills Through Self and Peer Assessment: Calibrating Judgment Using The 3D Presentation Framework. *QUT Business School and ANZAM.* Retrieved from https://www.anzam.org/wp-content/uploads/2017/02/Grant-Smith-Cathcart-2015.pdf.

Smolen, L., Newman, C., Wathen, T. & Lee, D. (1995). Developing student self-assessment strategies. *TESOL Journal, 5*(1), 22 – 28.

Spiller, D. (2012). *Assessment Matters: Self-assessment and Peer Assessment.* Teaching Development Unit, University of Waikato, New Zealand.

Taghizadeh, M. (2015). Diagnosing the Iranian L2 Writing Ability Using Self-Assessment and Level Specific Approaches. *Journal of Teaching Language Skills, 34*(1), 145-173.

Taras, M. (2010). Student Self-Assessment: Processes and Consequences. *Teaching in Higher Education, 15*(2), 199-209.

Thomlison, T. (1984). *Relational Listening: Theoretical and Practical Considerations*. The AnrivAl Meeting of the International Listening Association.

Trofimovich, P., Talia, I., Sara, K., Kazuya, S. & Crowther, D. (2016). Flawed Self-Assessment: Investigating Self- and Other-Perception of Second Language Speech. *Bilingualism: Language and Cognition,* 19, 122-140.

Tudor, I. (1996). *Learner-centeredness as language education.* Cambridge: Cambridge University Press.

Vigoya, F. (2000). Testing Accuracy and Fluency in Speaking Through Communicative Activities. *HOW Journal, 5*(1), 95-104.

White, R. & Arndt, V. (1991) Process Writing. Essex: Addison Wesley Longman Ltd.

Wright, C. & Houck, J. (1995). Gender Differences among Self-Assessments, Teacher Ratings, Grades, And Aptitude Test Scores for A Sample of Students Attending Rural Secondary Schools. *Educational and Psychological Measurement, 55,* 743–753.

Yang, C. & Tseng, S. (2015). A Blended Learning Environment for Individualized English Listening and Speaking Integrating Critical Thinking. *Computers & Education*, 63, 285–305.

Zareva, A. (2005). Models of Lexical Knowledge Assessment of Second Language Learners of English at Higher Levels of Language Proficiency. *System*, 33, 547-562.

## APPENDIX A: THE SELF-ASSESSMENT RUBRIC

**İngilizce Konuşma Becerisini Değerlendirme Anketi**

Değerli katılımcı,

Bu ankette İngilizce öğrenen öğrencilerin kendi İngilizce konuşma becerilerini nasıl değerlendirdikleri ölçülmeye çalışılmaktadır. Bu anketten elde edilen bilgilerin tamamı sadece araştırma amacı ile kullanılacak olup, başka hiç kimseyle paylaşılmayacaktır. İsminizin istenmesinin sebebi, bu anketten elde edilen bilgilerin konuşma sınavı notları ile karşılaştırılacak olmasıdır.

Ankete katılım tamamen gönüllülük esasına dayanır ve katılımcı istediği zaman ankete katılımdan vazgeçebilir. Araştırma sonuçlarını öğrenmek için hbasak@etu.edu.tr adresine mail atabilirsiniz.

   Bu ankete gönüllü olarak katılıyorum ve bilgilerin tamamen gizli tutulacağının farkındayım. ☐

İmza

   **Helin Başak**

**Birinci Bölüm**

1. Adınız/Soyadınız:
2. Sınıfınız:
3. Bölümünüz:
4. Yaşınız (Lütfen Seçiniz): 17-20 ☐    21-25 ☐    30 üstü ☐

5. Cinsiyetiniz: Kadın ☐     Erkek ☐

6. Ana diliniz:

7. Bilinen diğer diller (İngilizce hariç):

8. Bu benim TOBB ETÜ hazırlık programındaki

9. **birinci yılım** ☐     **ikinci yılım** ☐

10. Daha önce İngilizce eğitim aldınız mı? (lütfen seçiniz): **Evet** ☐ **Hayır** ☐

    10.1. Eğer evet ise, toplamda ne kadar süre bu eğitimi aldınız: _____ ay / yıl

11. Daha önce İngilizce konuşulan bir ülkede yaşadınız mı? **Evet** ☐ **Hayır** ☐

    11.1. Eğer evet ise, toplamda ne kadar süre yaşadınız: _____ ay/ yıl

12. Etrafınızda İngilizce konuşan arkadaşlarınız var mı? **Evet** ☐     **Hayır** ☐

13. İngilizceyi bir gün içerisinde yaklaşık kaç saat kullanıyorsunuz? (okul saatleri dışında)

    1-1          2-4          4-6          6 ve daha fazla

14. İngilizce TV dizisi izliyor musunuz? **Evet** ☐     **Hayır** ☐

    14.1. Eğer evet ise, günde kaç saat?

    0-1          2-4          4-6          6 ve daha fazla

15. İngilizce müzik dinliyor musunuz? **Evet** ☐     **Hayır** ☐

    15.1. Eğer evet ise, günde kaç saat?

    0-1          2-4          4-6          6 ve daha fazla

16. İngilizce video oyunu oynuyor musunuz? **Evet** ☐     **Hayır** ☐

    16.1. Eğer evet ise, günde kaç saat?

    0-1          2-4          4-6          6 ve daha fazla

17. Okulunuzdaki İngilizce konuşma kulüplerine gidiyor musunuz?

    **Evet** ☐     **Hayır** ☐

    17.1. Eğer evet ise, günde kaç saat?

    0-1          2-4          4-6          6 ve daha fazla

18. İngilizce gazete veya dergi okuyor musunuz? **Evet** ☐     **Hayır** ☐

    18.1. Eğer evet ise, günde kaç saat?

    1-1          2-4          4-6          6 ve daha fazla

19. İngilizce internet siteleri kullanıyor musunuz? **Evet** ☐     **Hayır** ☐

    19.1. Eğer evet ise, günde kaç saat?

    0-1          2-4          4-6          6 ve daha fazla

**20.** Daha önce kendi İngilizce becerilerinizi değerlendirmenize izin verildi mi?

      **Evet**   ☐    **Hayır**   ☐

## İkinci Bölüm

Lütfen aşağıdaki soruları, İngilizce konuşma becerilerinizi göz önünde bulundurarak *Çok Kötü* ve *Çok İyi* arasındaki seçenekleri işaretleyiniz.

    **1.** Sınıf içi tartışmalarda İngilizce konuşabilirim.

Çok Kötü    1     2     3     4     5   Çok İyi

    **2.** Rol oynamalarda İngilizce konuşabilirim.

Çok Kötü    1     2     3     4     5   Çok İyi

    **3.** Konuşma sınavlarında İngilizce konuşabilirim.

Çok Kötü    1     2     3     4     5   Çok İyi

    **4.** İngilizce konuştuğum zaman akıcı şekilde konuşabilirim.

Çok Kötü    1     2     3     4     5   Çok İyi


    **5.** İngilizce konuştuğum zaman kelimeleri doğru telaffuz edebilirim.

Çok Kötü    1     2     3     4     5   Çok İyi

    **6.** İngilizce konuştuğum zaman genelde dilbilgisini doğru kullanırım.

Çok Kötü    1     2     3     4     5   Çok İyi

    **7.** Sınıf içi tartışmalarda İngilizce dilbilgisini doğru kullanırım.

Çok Kötü    1     2     3     4     5   Çok İyi

    **8.** Konuşma sınavlarında İngilizce dilbilgisini doğru kullanırım.

Çok Kötü    1     2     3     4     5   Çok İyi

    **9.** İngilizce konuştuğum zaman karmaşık dilbilgisi yapılarını kullanabilirim.

Çok Kötü    1     2     3     4     5   Çok İyi

    **10.** İngilizce konuştuğum zaman doğru bağlaçları kullanabilirim.

Çok Kötü    1     2     3     4     5   Çok İyi

    **11.** İpucu verildiği zaman, bilmediğim sözcüklerin anlamlarını tahmin edebilirim.

Çok Kötü    1     2     3     4     5   Çok İyi

    **12.** İngilizce konuştuğum zaman sözcük ve söz öbekleri kullanabilirim.

Çok Kötü    1     2     3     4     5   Çok İyi

    **13.** İngilizce konuşurken uygun sözcükleri hatırlayabilirim.

Çok Kötü    1    2    3    4    5    Çok İyi

**14.** İngilizce konuştuğum zaman çeşitli sözcükler kullanabilirim.

Çok Kötü    1    2    3    4    5    Çok İyi

**15.** Sınıf içi aktivitelerde bir sözcüğü hatırlayamadığım zaman bununla başa çıkabilirim.

Çok Kötü    1    2    3    4    5    Çok İyi

**16.** İngilizce konuşurken doğru ve uygun sözcükler kullanabilirim.

Çok Kötü    1    2    3    4    5    Çok İyi

**17.** Konuşurken fikirlerimi genelde düzgün organize edebilirim.

Çok Kötü    1    2    3    4    5    Çok İyi

**18.** Anlık konuşmalarımı organize edebilirim.

Çok Kötü    1    2    3    4    5    Çok İyi

**19.** Konuşma sınavlarında konuşma konum için iyi bir taslak geliştirebilirim.

Çok Kötü    1    2    3    4    5    Çok İyi

**20.** Konuşma sınavlarında etkili bir şekilde iletişim kurabilirim.

Çok Kötü    1    2    3    4    5    Çok İyi

**21.** Konuşma sınavlarında sorulara cevap verebilirim.

Çok Kötü    1    2    3    4    5    Çok İyi

**22.** Konuşma sınavlarında fikirlerimi destekleyebilirim.

Çok Kötü    1    2    3    4    5    Çok İyi

**23.** Sınıf içi aktivitelerde fikirlerimi destekleyebilirim.

Çok Kötü    1    2    3    4    5    Çok İyi

**24.** Sınıf içi aktivitelerde etkili bir şekilde iletişim kurabilirim.

Çok Kötü    1    2    3    4    5    Çok İyi

**25.** Sınıf içi aktivitelerde sorulara cevap verebilirim.

Çok Kötü    1    2    3    4    5    Çok İyi

# APPENDIX B: TEACHERS' RUBRIC FOR SPEAKING EXAM

**Grammatical Accuracy and Range**

**5**   Almost all structures are accurate, even complex ones.

**4**   Most structures are accurate but complex ones cause some difficulty

**3**   Simple structures are mostly accurate but little attempt at complex ones is made.

**2**   Simple structures are often inaccurate, more complex ones not even attempted.

**1**   Simple structures almost always distort and there is little awareness of any grammatical rules.

**Vocabulary**

**5**   Appropriate level and variety of vocabulary to deal with topic, avoiding repetition. Able to paraphrase effectively if necessary.

**4**   Level and variety of vocabulary satisfactory, but experiences some difficulty in word choice and usage, occasional repetition. Some ability to paraphrase when necessary.

**3**   Vocabulary used is basic but sufficient to express ideas, some repetition. Little evidence of extended range. Limited ability to paraphrase and some difficulty experienced.

**2**   Vocabulary use tends to be inaccurate, with inappropriate or irrelevant use, making communication awkward, though shows some attempts.

**1**   Almost all of the vocabulary is inaccurate, inappropriate or irrelevant, making communication difficult or with no attempt at all.

**Content**

**5**   Communicates effectively, responds well to the topic. Explains well, gives effective examples.

**4**   Most of the speech responds to topic. However, some supporting details are weak.

**3**   Responds in a basic way to the topic but lacks relevant supporting details.

**2**   States the topic by giving his idea without any supporting details.

**1**   Unable to respond to the topic or makes almost no attempt.

**Fluency**

**5**   Almost no hesitation other than native-like search for ideas. The speech is very natural.

**4**   Most of the speech is without hesitation, which does not disturb the natural flow.

**3**   Some hesitation with some short pauses, but does not disturb the flow for the listener.

| **2** | Speech is halting with some long pauses. Frequent hesitation, needs prompting but shows attempt. |
|---|---|
| **1** | Speech is disconnected and difficult to follow. |

**Pronunciation**

| **5** | Almost all individual sounds are well articulated, with effective and natural use of stress and intonation. Easy to understand and follow. |
|---|---|
| **4** | Most individual sounds well-articulated, occasional difficulties do not disrupt comprehension. |
| **3** | Some individual sounds are not clearly articulated but not in a way that disrupt comprehension. |
| **2** | Many individual sounds poorly articulated and pronunciation puts strain on listeners and causes misunderstanding. |
| **1** | Impossible to understand at all. |

**Overall**

| **5** | Superior |
|---|---|
| **4** | Good |
| **3** | Average/Satisfactory/Acceptable (at the level) |
| **2** | Weak |
| **1** | Very Poor |

# APPENDIX C: ETİK İZİN FORMU

ORTA DOĞU TEKNİK ÜNİVERSİTESİ
MIDDLE EAST TECHNICAL UNIVERSITY

Sayı: 28620816 / 2465

05 ARALIK 2016

Konu:     Değerlendirme Sonucu

Gönderilen: Doç.Dr. Çiğdem Sağın ŞİMŞEK,

        Eğitim Fakültesi

Gönderen: ODTÜ İnsan Araştırmaları Etik Kurulu (İAEK)

İlgi:     İnsan Araştırmaları Etik Kurulu Başvurusu

Sayın Doç.Dr. Çiğdem Sağın ŞİMŞEK;

Danışmanlığını yaptığınız yüksek lisans öğrencisi Helin BAŞAK'ın "İngilizce Öğrenen Bireylerin Konuşma Becerilerini Öz-Değerlendirmeleri" başlıklı araştırması İnsan Araştırmaları Kurulu tarafından uygun görülerek gerekli onay **2016-EGT-160** protokol numarası ile **12.12.2016-30.03.2017** tarihleri arasında geçerli olmak üzere verilmiştir.

    Bilgilerinize saygılarımla sunarım.

Prof. Dr. Canan SÜMER

İnsan Araştırmaları Etik Kurulu Başkanı

Prof. Dr. Mehmet UTKU

İAEK Üyesi

Prof. Dr. Ayhan Gürbüz DEMİR

İAEK Üyesi

Yrd .Doç .Dr. Pınar KAYGAN

İAEK Üyesi

Prof. Dr. Ayhan SOL

İAEK Üyesi

Doç.Dr. Yaşar KONDAKÇI

İAEK Üyesi

Yrd. Doç. Dr. Emre SELÇUK

İAEK Üyesi

# BU BÖLÜM, İLGİLİ BÖLÜMLERİ TEMSİL EDEN İNSAN ARAŞTIRMALARI ETİK ALT KURULU TARAFINDAN DOLDURULACAKTIR.

Protokol No: 2016-EGT-160

## İAEK DEĞERLENDİRME SONUCU

Sayın Hakem,

Aşağıda yer alan üç seçenekten birini işaretleyerek değerlendirmenizi tamamlayınız. Lütfen "Revizyon Gereklidir" ve "Ret" değerlendirmeleri için gerekli açıklamaları yapınız.

Değerlendirme Tarihi: 02-12-2016

Ad Soyad:

---

☑ Herhangi bir değişikliğe gerek yoktur. Veri toplama/uygulama başlatılabilir.

☐ Revizyon gereklidir

   ☐ Gönüllü Katılım Formu yoktur.

   ☐ Gönüllü Katılım Formu eksiktir.

      Gerekçenizi ayrıntılı olarak açıklayınız:

   ☐ Katılım Sonrası Bilgilendirme Formu yoktur.

   ☐ Katılım Sonrası Bilgilendirme Formu eksiktir.

      Gerekçenizi ayrıntılı olarak açıklayınız:

   ☐ Rahatsızlık kaynağı olabilecek sorular/maddeler ya da prosedürler içerilmektedir.

      Gerekçenizi ayrıntılı olarak açıklayınız:

   ☐ Diğer.

      Gerekçenizi ayrıntılı olarak açıklayınız:

---

☐ Ret

   Ret gerekçenizi ayrıntılı olarak açıklayınız:

7

## GİRİŞ

Bu çalışmanın temel amacı, öz değerlendirmenin öğrencilerin öğrenme ve değerlendirme sürecine katkısını araştırmaktır. Öz değerlendirme, öğrencilerin dil öğrenme sürecini değerlendirmelerini; öğrenim sürecinde ne kadar başarılı oldukları nı ve hedeflerine ulaşıp ulaşamadıklarını gösteren alternatif bir değerlendirme türüdür. Öz değerlendirme, biçimlendirici değerlendirmenin bir parçası olarak düşünülebilir. Öğrenciler kendi performanslarını değerlendirmek ve dil beceri ve yeteneklerini belirlemek için öz değerlendirme kullanırlar (Brown, Andrade ve Chen, 2015). Blachford'a (1997) göre, öz değerlendirme, bir öğrencinin kendi başarısının diğer öğrencilerle karşılaştırılmasını içeren bir süreçtir. Montgomery (2001) tarafından ise, öğrencilerin kendi öğrenme performansları hakkında bir değerlendirme olarak da tanımlanmıştır.

Öz değerlendirme ayrıca, dil değerlendirmesinde birçok bilim adamı tarafından önerilen alternatif bir değerlendirme yöntemi olarak kabul edilmektedir. Öz değerlendirme, dünya çapında birçok akademisyen tarafından özerk dil öğrenme programlarının önemli bir unsuru olarak değerlendirmenin vazgeçilmez bir parçası haline gelmiştir, çünkü dil programlarının önemli bir kısmı öğrenci merkezli hale gelmiştir (Nurov, 2000).

Ek olarak, McMillen ve Hearn (2008), öz değerlendirmenin öğrencilerin motivasyonunu ve başarısını arttıran hayati bir beceri olduğunu, çünkü öğrenenin kendi kendini izlemesine yardımcı olduğunu, kendi değerlendirmesini ve öğrenmelerini iyileştirmek için doğru yöntemleri belirlediğini ifade etmektedir. Bazı yazarlar, öğrencilerin gelişmiş bilgileri ile onlara yardımcı olabilecek ve daha sonra kriterleri belirleyen, kendi değerlendirmelerini yapabilecekleri ve performanslarını yansıtabilecek ve öğrenme teknikleri üretebilecek belirli bir amaca sahip olduklarında, daha yüksek motivasyonla gelişmiş yetkinlikler sergileyeceklerini iddia etmektedirler. Öğrencilerin öğrenme sürecinde öz yeterlik ve özgüven

geliştirmenin temel amaç olduğu ve bunun öz değerlendirme ile sağlanabileceği sonucuna varılabilir (McMillen ve Hearn, 2008).

Öğrencilerin kendi öğrenmeleri hakkında karar vermeleri gerektiğinden, bu durum olumlu tutumlara ve dolayısıyla öğrenmeye yönelik yüksek motivasyona yol açabilir (Nurov, 2000). Gardner (1996) tarafından da, öz değerlendirmenin güven oluşturma ve motivasyonu artırma gibi birçok amaca hizmet edebileceği belirtilmiştir. Ayrıca, Tudor'a (1996) göre, öğrenci motivasyonunun artması ve öz değerlendirmenin dil eğitimine katkıda bulunduğunun bilinmesi öğrencilerin dil becerilerini daha doğru bir şekilde değerlendirmelerine yardımcı olabilir (Tudor, 1996). Nunan (1988) ve Oscarson (1989)'a göre, bir değerlendirme aracı olarak öz değerlendirme, öğrencilerin kendi güçlü ve zayıf yönlerini yansıtıcı bir yaklaşımla gerçekleştirmelerini sağlayarak kendi dil becerilerini anlamalarını artırır. Ayrıca, öz değerlendirme, öğrenenlerin öğrenme sürecinde kullandıkları yöntemleri ve teknikleri yansıtmaları ve değerlendirebilmelerini mümkün kılar.

Çalışmalar, öz değerlendirmenin geleneksel değerlendirme yöntemlerinin olumsuz duygusal etkilerini azalttığını da bildirmektedir. Örneğin, öğrenciler kendi performanslarını değerlendirdiklerinde, kaygı, stres ve öğretmen veya diğer değerlendiriciler tarafından değerlendirilme korkusu gibi olumsuz duygusal sonuçlar önemsizleşmektedir (Nurov, 2000).

Ek olarak, öz değerlendirmenin pratikliği ve maliyet etkinliği, dil kurumlarının kendi programlarını ve kendi değerlendirmelerini yapmaları için iki sebeptir. Öz değerlendirme, geleneksel değerlendirme yöntemlerinden daha pratik olarak kabul edilir çünkü tasarımı, oluşturması ve uygulaması daha kolaydır ve süreç daha az zaman alır. (Brown ve Hudson, 1998; LeBlanc ve Painchaud, 1985).

Öz değerlendirme sürecinde, öğrenciler çalışmalarının başarı dereceleri hakkında karar verirler. Bu süreçte amaç ve değerlendirme kriterleri dikkate alınmalıdır. Öğretmenler, değerlendirme kriterlerini öğrenciler için açık ve anlaşılabilir kılmalı, böylece öğrenciler becerilerinin güçlü ve zayıf yanlarını anlayabilmeli ve buna göre gözden geçirebilmelidir.

Değerlendirmenin ana amacı, öğrencilerin performanslarının ve öğrenme sürecinin mevcut durumunu belirlemek, mevcut öz değerlendirme durumunun iyi uygulama örnekleri ile karşılaştırılması ve potansiyel iyileştirme ve sistematik kalite gelişimlerinin belirlenmesidir (Sutalo, 2011). Joyce ve diğ. (2009), öz değerlendirmenin öğrencilerin kendi öğrenmelerinin sorumluluğunu alma konusundaki içsel motivasyonlarını artırmayı amaçladığını ifade etmektedir. Öz değerlendirme sayesinde, öğrenciler kendi dil yeterliliklerine, ilgili diğer değerlendiricilerden daha iyi bir şekilde erişebildikleri ve bunları değerlendirebildikleri için kendi dil yeterliliklerini geliştirme ile ilgili adımları nasıl uygulayacaklarını anlayabilmektedirler. Kordes ve diğ. (2014), farklı bir perspektife sahip olarak, öz değerlendirmenin kişisel gelişime büyük ölçüde katkı sağladığı ve öğrenme üzerinde önemli bir etkisi olduğu için kullanıldığını belirtmektedir. Değerlendirme, öğrenciler için bir gizem olarak görülebileceği için, öz değerlendirme, bu öğrencilere, kendilerinden nelerin beklendiğini anlamak için değerlendirmenin nasıl yapıldığını ve hangi yönlerin dikkate alınması gerektiğini gösterebilir. Bu yolla öğrenciler, nelerin ölçüldüğünü ve neleri öğrenmeleri gerektiğini anlayabileceklerdir (Kordes ve diğ., 2014).

Öz değerlendirmenin başarılı olması için, öğretmenler, öğrenciler için süreci yönlendirme, izleme ve kolaylaştırmada önemli bir rol oynamalıdır. Joyce ve diğ. (2009), öğretmenlerin öz değerlendirme hakkındaki rollerini ve fikirlerini incelemek için bir proje düzenlemiştir. Öz değerlendirmenin sınıf ortamına nasıl dönüştüğünü ve öğretmenlerin öz değerlendirme hakkında ne düşündüklerini anlamak istemişlerdir. Öz değerlendirme stratejileri ve öğretmen inançları çalışmanın ana alanlarıdır. Veri toplama yöntemleri olarak küçük bir vaka çalışması ve odak grup görüşmesi kullanılmıştır. Araştırmanın sonuçları, öğretmenlerin inançlarının sınıf ortamında öz değerlendirmeyi kullanıp kullanmayacaklarını belirlediğini ortaya koymuştur. Araştırma ayrıca, öğretmenlerin öz değerlendirme için uygun bir sınıf kültürü oluşturmaktan sorumlu olduğunu göstermiştir. Joyce (2009), öğretmenlerin öğrencilere eserlerini ve performanslarını nasıl değerlendireceklerini öğretmeleri gerektiğini belirtmiştir. Öz değerlendirmeyi ilk başta uygulamak kolay değildir. Ancak, destek sayesinde daha iyi sonuçlar elde edilebilir. Öğretmenler öz değerlendirme kültürünün oluşturulmasında lider olmalıdır. Öğrencilerin

kendilerini değerlendirebildiklerini ve öğretmenlerin desteğine ihtiyaçları olduğunu düşünmeleri gerektiği de belirtilmiştir. Son olarak, öğrencilerin öz değerlendirmede iyi olmaları için zamana ihtiyaçları vardır (Joyce ve diğ., 2009). Sonuç olarak, öğretmenlerin öz değerlendirme sürecinde önemli bir rolü var gibi görünmektedir.

Güvenilirlik ve geçerlilik, gerçek bir dil değerlendirme ortamında kabul edilip uygulanacak herhangi bir değerlendirme yönteminin iki önemli kriteridir. Güvenilirlik, tekrarlanan ölçümlerde sonuçların tutarlılığını ifade eder. Başka bir deyişle, güvenilirlik, bir ölçümün, aynı test katılımcıları ile ve aynı koşullar altında aynı sonuçları vermesidir. Öz değerlendirme ile ilgili soru ise, öz değerlendirmenin güvenilir bir ölçme aracı olup olmadığıdır. Yani, öğrencinin dil becerilerini değerlendirmek için ne ölçüde objektif olabileceği ve öğrencilerin değerlendirmesinin ne kadar güvenilir olabileceği, bu noktada sorulması gereken başlıca sorulardır. Bu nedenle, öz değerlendirmenin öğrencilerin kendi değerlendirmelerinde güvenilirlik sağlayabilmesi koşuluyla bir yabancı dil olarak İngilizce sınıfının akademik hedefleri olarak kullanılabileceği iddia edilebilir (Karakaya, 2017). Yabancı dil olarak İngilizce bağlamında güvenilirlik şu örnekle tanımlanabilir; bir öğrenciden kendi konuşma performansının bir kısmını şimdi ve bir süre sonra aynı kriterlere dayalı olarak değerlendirmesi istenirse, iki değerlendirme birbirinden çok farklı olmayacaktır (Huerta-Macias, 1995).

Geçerlilik, ölçme araçlarının doğruluğu ile ilgilidir. Basitçe, geçerlilik, değerlendirme aracının ölçmesi gerekeni ölçüp ölçmediğini ifade eder. Gardner (1996) ve Heilenman (1990) 'a göre, öz değerlendirme, öğrencilerin günlük ortamlarda ele aldığı öğrenme ortamını ve otantik ödevleri modelleme potansiyeline sahip olduğundan, öğretmenlerin kontrolü altında gerçekten yansıtıcı ve geçerli olma kapasitesine sahiptir (Gardner, 1996; Heilenman, 1990). Bununla birlikte, bazı araştırmacılar yabancı veya ikinci dil olarak İngilizce sınıfında öz değerlendirmenin geçerliliği konusunda endişeleri vardır. Örneğin, öğrencilerin kendilerini doğru olarak değerlendirmede zorluk yaşadıkları iddia edilmektedir, çünkü öz değerlendirme için ne kadar iyi eğitilmiş olsalar da, öğrenciler kendi davranışlarının farkında olmayabilirler.

İlgili alanyazın incelendiğinde, öz değerlendirmenin geçerliliği ve güvenilirliği üzerine yapılan bazı çalışmaların olduğu ve sonuçlarının öğrencilerin kendi dil yeteneklerinin öz değerlendirmesiyle doğru bir şekilde tutarlı olmayabileceğini ortaya koyduğu görülmüştür.

Bir açıdan bakıldığında, önemli miktarda araştırma öz değerlendirmenin güvenilirliği ve geçerliliği açısından kabul edilebilir olduğunu ve öğretmenlerin ve öğrencilerin değerlendirmeleri arasında iyi bir korelasyon olduğunu göstermektedir (Bachman ve Palmer, 1989; Buck, 1992; MacIntyre ve diğ., 1997; Ross, 1998). Öte yandan, öz değerlendirme üzerine yapılan diğer çalışmalar, öz değerlendirmenin öğrencilerin becerilerinin geçerli bir yansıması olamayabileceğine işaret etmektedir (Blanche, 1990; Heilenman, 1990; Nurov; 2000). Bununla ilgili olarak, öz değerlendirmenin öğrencilerin dil yeteneklerini değerlendirmek için geçerli ve güvenilir bir araç olarak kabul edilip edilmeyeceği kesinlikle doğrulanmamıştır.

Ancak, çalışmaların çoğunluğu öz değerlendirmenin güvenilir ve geçerli bir alternatif değerlendirme aracı olduğunu göstermektedir. Bununla birlikte, belirli koşullar ve farklı sınıf ortamlarında güvenilirliği ve geçerliliği nispeten engellenebilir. Öğrencilerin eğitimi gibi bu soruna yol açan çeşitli nedenler vardır. Öğrenciler kendilerini değerlendirmeyi nasıl uygulayacakları konusunda eğitildilerse, dil öğrenimleri hakkında mantıklı kararlar verebileceklerdir. Bir başka sebep ise öğrencilerin kültürel geçmiş, eğitim düzeyi ve hedef dile maruz kalma gibi bireysel farklılıklarıdır. Ayrıca, metodolojik yetersizlikler araştırmalarda güvenilirlik ve geçerlilik sorunları ile sonuçlanabilir. Son olarak, araştırmalar, veri toplama araçları üzerinde güvenilirlik ve geçerlilik testleri uygulamamakta ve ayrıca, araştırmalarda çelişen sonuçlara yol açabilecek olan, korelasyon ve verilerin karşılaştırılması için farklı istatistiksel testler de kullanabilmektedir (Nurov, 2000).

Dil becerilerinin öz değerlendirmesi, dil değerlendirme araştırması alanındaki en önemli ilgi alanlarından biri olmuştur. Her yaştan katılımcı ile farklı dil becerilerinde ve çeşitli öğrenme ortamlarında hemen hemen tüm yeterlilik seviyelerinde gerçekleştirilen çok sayıda çalışma bulunmaktadır. Mesela, algısal (okuma ve dinleme becerileri) ve üretimsel (yazma ve konuşma becerileri) becerilerin (LeBlanc ve Painchaud, 1985) öz değerlendirmelerini ve okuma, yazma,

dinleme, konuşma becerileri, dilbilgisi ve kelime bilgisi gelişimini hem birbirleriyle hem de ayrı ayrı ilişkilendiren çalışmalar vardır (Clark, 1981). Bununla birlikte, öz değerlendirme üzerine yapılan pek çok çalışma okuma, dinleme ve yazma becerilerinin öz değerlendirmesini araştırırken, konuşma becerilerini ihmal eder. Konuşma becerilerinin öz değerlendirmesine odaklanan az sayıda çalışmadan birinde, Harris (1997), sözlü yeteneklerin öz değerlendirmesinin tasarlanması, organize edilmesi ve uygulanmasının zorlaştığını belirtmektedir. Bu nedenle, öz değerlendirme üzerine yapılan çalışmalar, genellikle, yalnızca öz değerlendirme konuşma becerilerine odaklanmamakta, ancak genellikle dört beceriyi bir araya getirmektedir (Krausert, 1991; Ross, 1998; Deville ve Deville, 2003). Ayrıca, konuşma becerilerinin öz değerlendirmesine yönelik yürütülen bu az sayıdaki çalışma, özellikle konuşma yeteneklerinin öz değerlendirmesinin geçerliliği ve güvenilirliği ile çelişkili sonuçlar doğurmaktadır. Bu sebeple, öz değerlendirme araştırmalarında devam eden tartışmaya katkıda bulunmak için sadece konuşma becerilerinin öz değerlendirmesine dayalı bir çalışma yürütmek çok önemlidir.

Bir beceri olarak konuşmaktan söz etmek gerekirse, konuşmacıların fikirlerini, inançlarını ve duygularını sözel dilde ifade etmelerini ve bilgi aktarmalarını sağlayan verimli bir beceridir. Burns ve Joyce (1997) konuşmayı "bilginin üretilmesi, alınması ve işlemden geçirilmesini içeren bir anlam inşa etmenin etkileşimli bir süreci" olarak tanımlamaktadır. Böyle bir konuşmada, konuşmacı ve dinleyicinin rollerinin birbirinin yerine geçtiği ve sözel olarak da dahil olduğu iletişimsel bir süreç olduğunu göstermektedir. Konuşmanın bu özellikleri onu dil öğreniminin kaçınılmaz bir parçası haline getirir, dolayısıyla konuşma becerilerinin nasıl değerlendirileceğini bilmek önemlidir. Söyleşiler, bilgi boşluğu teknikleri, resim testleri, rol oyunları, cümle tamamlama testleri, yarı-doğrudan testler, bant tanımlayıcıları ve teknoloji gibi farklı konuşma becerisini değerlendirme yöntemleri vardır.

Nasıl test edileceği kadar, konuşma becerisinde neyin test edileceğini anlamak da önemlidir. Dilbilgisel doğruluk, kelime bilgisi, fikirlerin organizasyonu, akıcılık ve telaffuz yetenekleri gibi konuşma becerilerinin altında yatan belirli yetkinlikler değerlendirilir.

Konuşma aktiviteleri sırasında kişinin kendi performansını değerlendirmesi zor bir süreçtir. İyi kurulmuş bir öz değerlendirme sürecini gerçekleştirmek için öğretmenler tarafından izlenecek dört adım vardır. Her şeyden önce, öğretmenler, hem öğretmenlerin hem de öğrencilerin öz değerlendirme sürecine dâhil edilmesi hakkında bilgilendirilmeleri için bir takım kıstaslar oluşturmalıdır. İkinci olarak, öğrencilere, kendilerini gerginleştirebilecek not verme sürecini sormak yerine, onlardan başarılarını ve algıladıkları yetersizliklerini yansıtmaları istenir. Dahası, öğrencilerin kendilerine not vermelerini istemek de objektiflik ve güvenilirlik ile ilgili problemler yaratabilir. Üçüncü olarak, öz değerlendirme bir görevin tamamlanmasından hemen sonra gerçekleştirilmelidir, ancak ses ve video kaydı öğrencilerin performanslarını hatırlamalarına yardımcı olabilir. Son olarak, daha güvenilir bir değerlendirme aracı oluşturmak için, öz değerlendirme sonuçlarını öğretmenlerin gözlemlere dayanan bilgileriyle birleştirmek ve geri bildirimleri test etmek önemlidir.

Konuşma becerisinin öz değerlendirmesi sayesinde, öğrenciler ne kadar iyi konuştuklarını anlayabileceklerdir. Performanslarını değerlendirerek performanslarının güçlü ve zayıf yanlarının farkına varırlar. Örneğin, öğrenciler konuşma sürecini kaydedebilirlerse, performansı dinledikten veya izledikten sonra hataları bulabilir ve konuşma güçlerini anlayabilirler. (Yang ve Tseng, 2015). Bu nedenle, konuşma becerilerinin öz değerlendirmesi öğrencilerin konuşma becerilerini geliştirme konusunda daha fazla özerk olmalarına yardımcı olacaktır, çünkü öz değerlendirme öğrencilerin daha özerk olmasını sağlamış ve öğrenci merkezli dil öğrenimine katkıda bulunan bir temel nokta olmuştur.

Sayısız akademisyen, öz değerlendirme prosedürlerinin öğrenme ve değerlendirme sürecine katkılarını kabul etmesine rağmen, konuyla ilgili yapılan çalışmalar, dil becerilerinin geliştirilmesinde öz değerlendirmenin kullanımı için zengin bilgi sağlamamaktadır. Bu çalışmanın ana konusu olarak, konuşma becerilerinin öz değerlendirmesi hakkındaki araştırma çalışmaları azdır. Alanyazındaki mevcut çalışmaların, yabancı/ikinci dil olarak İngilizce öğrenme bağlamında öz değerlendirmenin kullanımıyla ilgili devam eden tartışmaya katkıda bulunduğu belirtilebilir. Genel olarak sonuçlar, konuşma becerilerinin öz değerlendirmesinin öğrencilerin sözel becerilerini geliştirdiğini ve hedef dilde konuşmaya yönelik

motivasyonlarını ve özgüvenlerini artırdığını göstermektedir. Ayrıca, öğrencilerin konuşma becerilerinin öz değerlendirmesine yönelik olumlu tutumları olduğu sonucuna varılabilir. Öğretmen ve öğrenci değerlendirmeleri dikkate alındığında, bazı çalışmalar öğretmenlerin ve öğrencilerin konuşma becerilerini birbirlerinden farklı değerlendirdiğini göstermektedir (Singh, 2015; Babaii ve diğ., 2015). Konuşma becerilerinin değerlendirilmesi hakkındaki başka bir çalışma, öğretmen ve öğrenci arasında anlamlı fark olmadığını göstermektedir. (Smith, 2015).

Türkiye'de dil becerilerinin öz değerlendirilmesi konusunda çok az çalışma bulunmaktadır. Bu çalışmalar genel olarak öz değerlendirme araçlarının ve öğrencilerin öz değerlendirme hakkındaki inançlarının güvenilirliğine odaklanmıştır. Öğrencilerin 2017 yılında, Burcu Karakaya tarafından tamamlanan İngilizce konuşma becerilerinin öz değerlendirilmesinde sadece bir çalışma vardır. Bu çalışma, sadece orta düzeyde (B1) öğrenciler ile yürütülmüştür ve amacı öz değerlendirmenin etkisini araştırmaktır.

Bu nedenle, yukarıda bahsedilen kaygılara odaklanan araştırma yapmak hem gerekli hem de önemlidir. Bu çalışma, anadili Türkçe olan ve İngilizce öğrenen öğrencilerin konuşma becerilerinin öz değerlendirilmesi ile öğretmenlerin öğrencilerin İngilizce konuşma becerilerini değerlendirmeleri arasında bir ayrım olup olmadığını belirlemeyi amaçlamaktadır. Araştırmanın diğer bir amacı, öğrencilerin İngilizce yeterlilik düzeylerinin, cinsiyetlerinin ve müfredat dışı etkinliklerle İngilizce'ye maruz kalma sürelerinin, öğrencilerin konuşma becerilerini öz değerlendirmeleri üzerinde bir etki yaratıp yaratmadığını da incelemektir.

Sonuç olarak, bu çalışma, öğrencilerin ve öğretmenlerin İngilizce konuşma becerilerini birbirlerinden farklı olarak değerlendirip değerlendirmediklerini ve öğrencilerin konuşma becerilerini genel olarak ve özellikle konuşma becerilerinin farklı bileşenleri için nasıl İngilizce olarak değerlendirdiklerini açıklığa kavuşturmayı amaçlamaktadır. Bu nedenle, bu çalışma, konuşma becerilerinin öz değerlendirmesi hakkında iç görü sağlayarak geçmiş araştırmalara katkıda bulunmayı amaçlamaktadır.

Bu çalışma TOBB Teknoloji ve Ekonomi Üniversitesi'nde (TOBB ETÜ) yürütülmektedir. TOBB ETÜ, 2003 yılında TOBB Vakfı tarafından kurulmuştur. Altı fakültesi, 21 bölümü ve yaklaşık 6000 öğrencisi vardır. Her yıl, yaklaşık 1000 öğrenci üniversitenin farklı bölümlerine kayıt yapar. Üniversitenin, Türkiye'deki diğer üniversitelerden farklı üç dönemlik bir eğitim yılı vardır.

Türk Dili ve Edebiyatı Bölümü ile Görsel İletişim ve Tasarım Bölümü dışındaki tüm bölümler %30 İngilizce ve %70 Türkçe ders işlemektedir. Bu nedenle, İngilizce eğitimli bölümlere kayıt yaptıran öğrencilerin İngilizce Hazırlık Programına katılmaları gerekmektedir. Diğer bir deyişle, Türk Dili ve Edebiyatı ve Görsel İletişim ve Tasarım bölümlerine kayıtlı olan öğrenciler, eğitimlerine doğrudan kendi bölümlerinde başlayabilirler. Ancak diğer öğrenciler İngilizce Hazırlık Programına başlamak zorundadırlar.

İngilizce Hazırlık Programı Yabancı Diller Bölümü tarafından uygulanır. Programda beş farklı seviye vardır: AF (başlangıç seviyesi), A (temel), B (Ön Orta), C (orta seviye) ve CR (İkinci kez programı tekrar eden orta düzey öğrenciler). Bu seviyeler, programa kayıtlı öğrencilerin girdikleri sınavlardan aldıkları puanlara göre belirlenir. Seviye belirleme sınavı, programın başında yapılan ilk sınavdır ve dinleme-anlama (60 soru), yapı (gramer ve kelime bilgisi 40 soru) ve okuduğunu anlamayı (25 soru) değerlendirir. Yeterlilik sınavı ikinci sınavdır ve ön sınavdan sonra uygulanır. Ön sınavdan 100 üzerinden 65 puan alan öğrenciler Yeterlilik sınavına girebilirler. Yeterlilik sınavı, dinlediğini anlama (50 soru), dilbilgisi (40 soru) ve okuduğunu anlama (50 soru) testini yapan TOEFL Kurumsal Test Programıdır (ITP). TOEFL ITP'den 677 üzerinden 500 puan alan öğrenciler hazırlık programından muaf olur ve bölüm derslerine devam edebilirler. 500'den düşük puan alan öğrenciler hazırlık programına başlamak zorundadırlar.

Bu çalışmanın katılımcıları, TOBB ETÜ İngilizce Hazırlık Programında okuyan ve 18-25 yaş aralığında olan 549 öğrencidir. Tüm katılımcılar, farklı seviyelerde İngilizce öğrenen ve anadili Türkçe olan öğrencilerdir. Hepsi, haftada 4 ila 8 saat İngilizce görmüş devlet ya da özel lise mezunlarıdır. Veri toplama döneminde katılımcılar TOBB ETÜ hazırlık okulundaydı. Katılımcıların dil yeterlilik düzeyi, 2016-2017 akademik yılı başında TOBB ETÜ'de Yabancı Diller Bölümü tarafından

verilen ön sınav ile belirlenmiştir. Bu ön sınav, yeni öğrencileri İngilizce seviyesindeki yeterlik seviyelerine göre farklı sınıflara yerleştirmek amacıyla yapılmıştır.

Bu çalışmada kullanılan veri toplama araçları arasında bir öz değerlendirme anketi, katılımcıların konuşma performanslarının öz değerlendirme puanları ve öğretmenlerin konuşma performansına ilişkin değerlendirme puanları bulunmaktadır. Öz değerlendirme anketi, katılımcıların demografik bilgileri, okul sonrası etkinliklerde İngilizce kullanımı ve konuşma becerilerindeki performansları ile ilgili sorulardan oluşmaktadır. Öğretmenlerin konuşma performansına ilişkin değerlendirme puanları, 2016-2017 bahar yarıyılı boyunca yapılan konuşma sınavında öğrencilere verilen değerlendirmelerden elde edilmiştir.

Öğrencilerin konuşma performanslarını nasıl değerlendirdikleri hakkında bilgi toplamak için geliştirilen öz değerlendirme anketi iki bölümden oluşmaktadır. İlk bölümde, katılımcıların demografik bilgileri ve günlük olarak İngilizce dışındaki günlük kullanımları hakkında 19 soru yer aldı. İkinci bölümde, katılımcıların gramer, kelime bilgisi, İngilizce akıcılık, telaffuz doğruluğu ve iletişim becerileri açısından İngilizce konuşma performanslarını değerlendirmelerini isteyen 25 soru vardır. Öğrencilerin ve öğretmenlerin konuşma performanslarını değerlendirmelerinde içerik tutarlılığını sağlamak için, öz değerlendirme anketi, katılımcıların değerlendirildiği konuşma sınavında öğretmenlerin kullandığı değerlendirme tablosuna göre hazırlanmıştır (bkz. Rubrik için Ek 2). Değerlendirme tablosunda yer alan maddeler, daha anlaşılabilir olmaları için "... -e/abilirim" ifadeleri şeklinde hazırlanmıştır. Değerlendirme listesi pilot olarak uygulanmış ve veriler uygun pilot uygulamadan iki hafta sonra toplandı. Katılımcılar, pilot süreyi tamamlamak için yaklaşık 20 dakikaya ihtiyaç duymuştur ve bu sürenin pilot uygulama deneyimlerine dayanarak yeterli olduğu anlaşılmıştır.

Diğer veri toplama aracı, TOBB ETÜ Hazırlık öğrencilerinin her akademik dönemde bir kez girdikleri konuşma sınavıydı. Sınav, öğrencilerin iki öğretmen tarafından değerlendirilmek üzere bireysel olarak davet edildiği bir mülakat olarak düzenlendi.

Bu çalışmadaki ilk araştırma sorusu, öğrencilerin öz değerlendirme ile öğretmenlerin öğrencilerin İngilizce konuşma becerilerini değerlendirmeleri arasında bir ilişki olup olmadığıdır. Sonuçlar, öğretmenlerin ve öğrencilerin konuşma becerilerini birbirinden önemli ölçüde farklı olarak değerlendirdiğini göstermektedir. Öğretmenlerin değerlendirmelerinin puan ortalaması 100 üzerinden 79,2 iken, öğrencilerin öz değerlendirme puan ortalaması 61,7'dir. Bu puanlar, öğretmenlerin, öğrencilerin konuşma becerilerini öğrencilerin kendilerinden daha yüksek puanla değerlendirdiği sonucuna varmamızı sağlar. Bunun nedeni, konuşma becerisinin, üretimsel bir beceri olduğu için değerlendirmenin en zor dil becerilerinden biri olmasından kaynaklanabilir (Harris, 1997), çünkü konuşma değerlendirmesi, birbiriyle ilişkilendirilemeyen yeteneklerin bir karışımını içerir.

Öğrencilerin öz değerlendirme ile öğretmenlerin değerlendirmesi arasındaki farkın bir başka nedeni, öğrencilerin konuşma yeteneklerine olan güven eksikliği ile ilişkili olabilir. Türkiye'de, İngilizce yabancı dil olarak öğretilir; öğrenciler, hazırlık sınıflarında yoğun bir şekilde hedef dile maruz kalmalarına rağmen, sınıf dışında iletişimsel amaçlar için İngilizce üretme konusunda yeterli fırsatlara sahip değildir. İngilizce kullanma fırsatının olmaması, konuşma becerilerinin gelişimini doğal olarak sınırlandırmaktadır. Dolayısıyla, öğrencilerin hedef dilde sözlü olarak konuşmada zorluk yaşayabilecekleri söylenebilir, sonuç olarak hedef dilde konuştuklarında kendilerine güvenemeyebilirler ve bunun sonucunda da konuşma yeteneklerini yeterince doğru değerlendiremezler.

Ayrıca, konuşma dilini bilmedikleri ve özellikle hedef dilde performans göstermedikleri için, anadillerinde konuşma şansları olduğu zaman, İngilizce konuşmamayı tercih edebilirler.

Özetlemek gerekirse, bu sonuçlar, öğrencilerin konuşma becerilerini öğretmenlerinin yaptığı gibi değerlendirdiklerini ifade edebilir. Bu çalışma önceki araştırmaların sonuçlarını doğrulamaktadır ve konuşma becerilerinin öz değerlendirmesinin uygun eğitim ve öğretmen gözetimi ile uygulanabileceği söylenebilir. Ayrıca, öğrenciler dil eğitiminde öz değerlendirme araçlarını kullanmaya alışık değildir. Türk eğitim sisteminde, değerlendirme sistemi öğretmen değerlendirmesine dayanmaktadır. Milli Eğitim Bakanlığı'nın Yabancı Dil

Sınavlarına İlişkin Kural ve Düzenlemelerine göre (2016 yılı resmi gazetesinde yayınlanmıştır), haftada kaç saat olursa olsun her ders için en az iki yazılı sınav gereklidir. Sınavların sayısı ve tarihleri her yarıyıl başında belirlenir ve okul müdürünün onayından sonra ders öğretmeni tarafından düzenlenir. Sınavlar için gerekli önlemler okul müdürlüğü tarafından alınır. Ayrıca yabancı dil dersleri sınavları, dinleme, konuşma, okuma ve yazma becerilerini ölçmek için yazılı ve uygulamalı olarak yürütülmektedir (MEB, 2016). Bu nedenle, öğrencilerin kendi konuşma becerilerini değerlendirmesinin, öz değerlendirme deneyiminin ve eğitiminin eksikliğinden etkilenmiş olabileceği söylenebilir.

Konuşma becerisi bileşenlerine bakıldığında, özet olarak, öğrenciler kelime, dilbilgisi, akıcılık, telaffuz ve iletişim becerilerini öğretmenlerinden daha düşük olarak değerlendirirler. Ek olarak, sonuçlar öğrencilerin bir bütün olarak konuşma becerileri hakkında ne kadar güvensiz olduklarını göstermektedir.

İkinci araştırma sorusu, konuşma becerilerinin öz değerlendirmesi ile öğrencilerin yeterlik düzeyleri arasındaki ilişkidir. Sonuçlar, öğrencilerin konuşma becerilerinin öz değerlendirmesinin öğrencilerin İngilizce yeterlilik düzeyine bağlı olarak değiştiğini göstermektedir. En yüksek yeterlilik düzeyine sahip olan grubun ortalama puanı 66.287 iken, en düşük düzeydeki grubun puan ortalaması ise 57.931'dir. Bu puanlar arasındaki fark istatistiksel olarak önemli ve anlamlıdır. Öğretmenlerin, öğrencilerin konuşma becerilerini değerlendirmesi, öğrencilerin seviyesine göre de değişir. En yüksek seviyedeki öğrencilerin puan ortalaması 80.856, en düşük puan ise 73.84'tür. Öğretmen değerlendirmesi ayrıca, öğrencilerin yeterlik düzeyleri açısından da önemli farklılıklar göstermektedir.

Öğretmen değerlendirmesinin toplam puan ortalaması dikkate alındığında (M = 79.269), daha yüksek yeterlik düzeyine sahip öğrencilerin, daha düşük yeterlik düzeyine sahip öğrencilere göre daha yakın bir ortalama puan (M = 66.287) olduğu söylenebilir. Sonuç olarak, daha yüksek yeterlilik seviyesine sahip öğrencilerin konuşma becerilerini değerlendirmede daha düşük yeterlilik seviyesine sahip öğrencilerden daha iyi oldukları söylenebilir. Bu gelişmekte olan dil farkındalığı ile ilgili olabilir. Yeterlik seviyesi ne kadar yüksekse, öğrenciler dil becerilerinin daha fazla farkında olurlar. Ne öğrendiklerinden ve ne kadar üretebildiklerinden daha

fazla haberdar olduklarından, öğrencilerin konuşma yeteneklerini değerlendirmesi daha kolaydır. Blue (1994) tarafından, dil farkındalığının yeterlikle ve kendi kendini değerlendirmenin doğruluğuyla ilişkili olduğu da öne sürülmektedir. Böylece, öğrenciler dil becerilerini, hedef dilde daha yetkin olduklarından, değerlendirmede ilerlemektedirler (Blue, 1994).

Bu çalışmanın diğer bir sorusu, cinsiyetin öğrencilerin ve öğretmenlerin öğrencilerin İngilizce konuşma becerilerini değerlendirmelerini nasıl etkilediğiyle ilgilidir. Bulgular, erkek öğrencilerin konuşma becerilerini kız öğrencilerden daha yüksek puanla değerlendirdiğini göstermiştir. Diğer bir deyişle, erkek öğrenciler (M = 64.9) kız öğrencilerden (M = 58,7) daha iyi İngilizce konuşma becerilerine sahip olduklarını düşünmektedir. Bu sonuç, erkek öğrencilerin kız öğrencilere göre konuşma becerilerine daha fazla güven duyduklarını göstermektedir. Bu, erkek öğrencilerin dil öğreniminde daha özerk ve kendinden emin olabilecekleri düşüncesiyle ilişkili olabilir. Öğretmenlerin değerlendirmesi ile öğrencilerin cinsiyetleri arasındaki ilişki söz konusu olduğunda, öğretmenlerin erkek ve kız öğrencilerin konuşma sınav puanlarını nasıl değerlendirdikleri arasında anlamlı bir fark yoktur. Benzer sonuçlar literatürde bildirilmiştir. Yazarlar, öğretmenlerin öğrencilerin konuşma yeteneklerini objektif ve adil bir şekilde değerlendirdiği sonucuna varmışlardır.

Son olarak bu çalışma, müfredat dışı etkinliklerin öğrencilerin konuşma becerilerini öz değerlendirme konusundaki etkisini de araştırmaktadır Bu müfredat dışı etkinlikler arasında İngilizce televizyon dizilerini izlemek, İngilizce şarkıları/radyo yayınlarını dinlemek, İngilizce video/çevrimiçi oyun oynamak, İngilizce web sitelerini kullanmak bulunmaktadır. Sonuçlar, yukarıda belirtilen faaliyetlerin hepsinin öğrencilerin öz değerlendirmesini etkilediğini göstermektedir. Yani, bu aktiviteleri diğerlerinden daha uzun sürelerle yapan öğrenciler konuşma becerilerini, bu aktiviteler için daha az zaman harcayan öğrencilerden daha iyi ve daha doğru olarak değerlendirirler. Her aktivitenin ortalama puanı, bu faaliyetlere harcanan saat arttıkça artar.

Dahası, öz değerlendirmenin, özerkliğe katkıda bulunduğu kadar, öğrencilerin öğrenme sürecine katılmalarına ve dolayısıyla motivasyonlarını artırmalarına

yardımcı olduğu öne sürülebilir. (Nunan, 1988; Oscarsson, 1989). Nurov'un (2000) belirttiği gibi, kendi öğrenim süreçleri ile ilgili yargılarda bulunmak, olumlu bir tutuma yol açabilir ve bu da, öğrenme sürecine yönelik olarak daha yüksek motivasyon ve güven oluşturmaya yol açabilir (Gardner, 1999; Nurov, 2000). Sonuç olarak, öğrenci motivasyonunun artması ve öz değerlendirmenin dil eğitimine katkıda bulunduğuna dair bir farkındalık, öğrencilerin dil becerilerini daha doğru bir şekilde değerlendirmelerine yardımcı olabilir.

# APPENDIX E: TEZ İZİN FORMU/THESIS PERMISSION FORM

**ENSTİTÜ** / INSTITUTE

**Fen Bilimleri Enstitüsü** / Graduate School of Natural and Applied Sciences ☐

**Sosyal Bilimler Enstitüsü** / Graduate School of Social Sciences ☒

**Uygulamalı Matematik Enstitüsü** / Graduate School of Applied Mathematics ☐

**Enformatik Enstitüsü** / Graduate School of Informatics ☐

**Deniz Bilimleri Enstitüsü** / Graduate School of Marine Sciences ☐

**YAZARIN /** AUTHOR

**Soyadı** / Surname     : BAŞAK
**Adı** / Name         : Helin
**Bölümü** / Department : İngiliz Dili Eğitimi

**TEZİN ADI / TITLE OF THE THESIS** (**İngilizce** / English) : SELF-ASSESSMENT OF SPEAKING

**TEZİN TÜRÜ /** DEGREE**:**  **Yüksek Lisans** / Master ☒    **Doktora** / PhD ☐

1. **Tezin tamamı dünya çapında erişime açılacaktır. /** Release the entire work immediately for access worldwide. ☐

2. **Tez iki yıl süreyle erişime kapalı olacaktır.** / Secure the entire work for patent and/or proprietary purposes for a period of **two year. \*** ☐

3. **Tez altı ay süreyle erişime kapalı olacaktır.** / Secure the entire work for period of **six months**. **\*** ☐

**\*** *Enstitü Yönetim Kurulu Kararının basılı kopyası tezle birlikte kütüphaneye teslim edilecektir.*
*A copy of the Decision of the Institute Administrative Committee will be delivered to the library together with the printed thesis.*

        **Yazarın imzası** / Signature   ..........................       **Tarih** / Date