BAŞKENT UNIVERSITY
INSTITUTE OF EDUCATIONAL SCIENCES
DEPARTMENT OF FOREIGN LANGUAGES
MASTER'S OF ENGLISH LANGUAGE TEACHING WITH THESIS

DEPICTING EFL INSTRUCTORS' LANGUAGE ASSESSMENT LITERACY IN ASSESSING PRODUCTIVE LANGUAGE SKILLS AND THEIR PERCEPTIONS AND PRACTICES ON FORMATIVE ASSESSMENT OF PRODUCTIVE LANGUAGE SKILLS

PREPARED BY

SELDA ARSLAN

MASTER THESIS

ANKARA-2022

BAŞKENT UNIVERSITY
INSTITUTE OF EDUCATIONAL SCIENCES
DEPARTMENT OF FOREIGN LANGUAGES
MASTER'S OF ENGLISH LANGUAGE TEACHING WITH THESIS

**DEPICTING EFL INSTRUCTORS' LANGUAGE ASSESSMENT
LITERACY IN ASSESSING PRODUCTIVE LANGUAGE SKILLS
AND THEIR PERCEPTIONS AND PRACTICES ON FORMATIVE
ASSESSMENT OF PRODUCTIVE LANGUAGE SKILLS**

PREPARED BY

SELDA ARSLAN

MASTER THESIS

THESIS ADVISOR

ASST. PROF. DR. SEVGİ ŞAHİN

ANKARA-2022

BAŞKENT ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Yabancı Diller Eğitimi Anabilim Dalı İngiliz Dili Öğretimi Tezli Yüksek Lisans Programı çerçevesinde Selda Arslan tarafından hazırlanan bu çalışma, aşağıdaki jüri tarafından Yüksek Lisans olarak kabul edilmiştir.

Tez Savunma Tarihi: 07/01/2022

Tez Adı: İngilizce Öğretmenlerinin Üretici Dil Becerilerinin Değerlendirilmesinde Ölçme ve Değerlendirme Okur-Yazarlık Seviyesi ve Üretici Dil Becerilerini Biçimlendirici Değerlendirme Yoluyla Değerlendirmeye Dair Bakış Açıları

Tez Jüri Üyeleri (Unvanı, Adı - Soyadı, Kurumu)   **İmza**

Prof. Dr. Çiler Hatipoğlu, Orta Doğu Teknik Üniversitesi   ..………………...

Asst. Prof. Dr. Ahmet Remzi Uluşan, Başkent Üniversitesi   …………………

Tez Danışmanı Asst. Prof. Dr. Sevgi Şahin, Başkent Üniversitesi   …………………

ONAY

Prof. Dr. Servet Özdemir

Eğitim Bilimleri Enstitüsü Müdürü

Tarih: …/…/2022

BAŞKENT ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
YÜKSEK LİSANS TEZ ÇALIŞMASI ORİJİNALLİK RAPORU

Tarih: 22/01/2022

Öğrencinin Adı, Soyadı: Selda Arslan

Öğrencinin Numarası: 21820012

Anabilim Dalı: Yabancı Diller Eğitimi

Programı: İngiliz Dili Öğretimi Tezli Yüksek Lisans Programı

Danışmanın Unvanı/Adı, Soyadı: Dr. Öğr. Üyesi Sevgi Şahin

Tez Başlığı: İngilizce Öğretmenlerinin Üretici Dil Becerilerinin Değerlendirilmesinde Ölçme ve Değerlendirme Okur-Yazarlık Seviyesi ve Üretici Dil Becerilerini Biçimlendirici Değerlendirme Yoluyla Değerlendirmeye Dair Bakış Açıları

Yukarıda başlığı belirtilen Yüksek Lisans tez çalışmamın; Giriş, Ana Bölümler ve Sonuç Bölümünden oluşan, toplam 174 sayfalık kısmına ilişkin, 22/01/2022 tarihinde tez danışmanım tarafından Turnitin adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı %17'dır. Uygulanan filtrelemeler:

1. Kaynakça hariç

2. Alıntılar hariç

3. Beş (5) kelimeden daha az örtüşme içeren metin kısımları hariç

"Başkent Üniversitesi Enstitüleri Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Usul ve Esaslarını" inceledim ve bu uygulama esaslarında belirtilen azami benzerlik oranlarına tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Öğrenci İmzası:…………………

ONAY

Tarih: 22/01/2022

Öğrenci Danışmanı Unvan, Ad, Soyad, İmza:

……………………………..

Dr. Öğr. Üyesi Sevgi Şahin

To my family who supported me throughout this journey.

ACKNOWLEDGMENTS

<div align="right">

Selda ARSLAN

Ankara 2022

</div>

ÖZET

SELDA ARSLAN

İngilizce Öğretmenlerinin Üretici Dil Becerilerinin Değerlendirilmesinde Ölçme ve Değerlendirme Okur-Yazarlık Seviyesi ve Üretici Dil Becerilerini Biçimlendirici Değerlendirme Yoluyla Değerlendirmeye Dair Bakış Açıları

Başkent Üniversitesi

Eğitim Bilimleri Enstitüsü

Yabancı Diller Eğitimi Anabilim Dalı

İngiliz Dili Öğretimi Tezli Yüksek Lisans Programı

2022

Bu çalışmanın amacı, Türkiye'deki üniversitelerin hazırlık bölümlerinde görev yapan İngilizce öğretim elemanlarının, üretici dil becerilerinin değerlendirilmesinde ölçme ve değerlendirme okur-yazarlık seviyesi ve üretici dil becerilerini biçimlendirici değerlendirme yoluyla değerlendirmeye dair bakış açılarını ve uygulamalarını incelemektir. Karma yöntem kullanılan bu çalışmada, veriler yarı yapılandırılmış görüşme ve "Dil Öğretiminde Ölçme ve Değerlendirme Okuryazarlığı" anketi ile toplanmıştır. Çalışmanın nicel kısmı için anket, Ölmezer-Öztürk'ün (2018) "Dil Değerlendirme ve Bilgi Ölçeğinden" adapte edilmiştir ve farklı üniversitelerin hazırlık birimlerinde çalışan 60 İngilizce öğretmeni bu anketi tamamlamıştır. Bu çalışmanın nitel kısmı içinse İngilizce öğretmenlerinin üretici dil becerilerini biçimlendirici değerlendirme yöntemiyle değerlendirmeye olan bakış açısını araştırmak amacıyla yarı yapılandırılmış görüşme soruları uygulanmıştır ve 11 gönüllü İngilizce öğretmeni görüşmelere katılmıştır. Anketten elde edilen veriler betimsel istatistik ile analiz edilmiştir. Nitel veriler Creswell (2014)' in sistematik içerik analizi kullanılarak analiz edilmiştir. Anketten elde edilen sonuçlar yüzdelikler açısından incelendiğinde, İngilizce öğretim elemanlarının üretici dil becerilerini değerlendirmede genel olarak orta seviyenin üzerinde bir ölçme ve değerlendirme okur-yazarlık bilgisi olduğunu göstermiştir. Fakat bu seviye hem "Konuşma becerisi değerlendirme maddeleri" hem de "Yazma becerisi değerlendirme maddeleri"nde yer alan "Değerlendirme araçları" kategorilerinde ortalama seviyenin altında çıkmıştır. Sonuçlar aynı zamanda, öğretim elemanlarının biçimlendirici değerlendirmeyi, çoğunlukla da gözlem aracını ve sınıf içi aktivitelerini, öğrencilerin üretici dil becerilerini değerlendirmek için sınıflarında uyguladıklarını ortaya çıkarmıştır. Bunun dışında, öğretmenlerin sınıfta üretici dil becerilerini değerlendirirken karşılaştıkları problemleri öğrenci tabanlı gösterirken, bu sorunlara çözümleri ise öğretmen tabanlı sunmuşlardır. Bu çalışmanın sonunda, pedagojik uygulamalar ve ileride yapılacak olan çalışmalar için öneriler sunulmuştur.

Anahtar Kelimeler: İngilizce öğretiminde ölçme ve değerlendirme, yabancı dil değerlendirme okur-yazarlığı, biçimlendirici değerlendirme, üretici dil becerilerini değerlendirme.

ABSTRACT

SELDA ARSLAN

Depicting EFL Instructors' Language Assessment Literacy in Assessing Productive Language Skills and Their Perceptions and Practices on Formative Assessment of Productive Language Skills

Başkent University
Institute of Educational Science
Department of Foreign Languages
Master's of English Language Teaching with Thesis

2022

The purpose of this study was to depict EFL instructors' language assessment literacy in assessing productive language skills and their perceptions on formative assessment of productive language skills in preparatory schools in Turkey. In this mixed-method research, the data were collected through semi-structured interview and language assessment knowledge questionnaire (LAKQ). For the quantitative part of the study, a questionnaire was adapted from Ölmezer-Öztürk (2018)'s "Language Assessment and Knowledge Scale" and 60 EFL instructors working in preparatory schools in different universities in Ankara completed the questionnaire. For the qualitative part of the study, a semi-structured interview was utilized to investigate the EFL instructors' perceptions on assessing productive language skills through formative assessment and 11 volunteer EFL instructors participated in the interview. The data obtained from the questionnaire were analysed using SPSS through descriptive statistics and the qualitative data were analysed using Creswell's (2014) systematic content analysis framework. The results revealed an average-above level of language assessment literacy (LAL) of the EFL instructors in assessing productive language skills. However, with respect to certain skills in LAL, this level was found below average in "Assessment Tools" parts of both Items for Assessing Writing and Speaking parts. The result also revealed that the instructors implemented formative assessment, mostly observation and in-class activities, in their classrooms to assess their learners' productive language skills. Moreover, the problems they encounter in the classrooms were mostly learner-based however, the solutions they offered were mostly instructor-based. The study also provided some pedagogical implications and suggestions for further studies.

Key Words: English Language assessment, language assessment literacy, formative assessment, assessing productive language skills

TABLE OF CONTENTS

APPENDIX 17: Research Approval (University of Turkish Aeronautical Association)

LIST OF TABLES

LIST OF FIGURES

ABBREVIATION LIST

EFL   English as a Foreign Language

AL   Assessment Literacy

LAL   Language Assessment Literacy

ELTE   English Language Testing and Evaluation

AFL   Assessment for Learning

AAL   Assessment as Learning

FA   Formative Assessment

SA   Summative Assessment

LTA   Language Testing and Assessment

ELT   English Language Teaching

ELL   English Language and Literature

ACL   American Culture and Literature

TI   Translation and Interpretation

CHAPTER 1

1. INTRODUCTION

This chapter presents and discusses the background of the study, statement of the problem, purpose of the study, the research questions, significance of the study, and definition of key terms.

1.1. Background of the Study

English, considered to be one of the most widely spoken languages in the world, is used for many different purposes including education, business, and trade. It is considered a lingua franca as it is "a language widely adopted for communication between two speakers whose native languages are different from each other's and where one or both speakers are using it as a "-second language-"" (Harmer, 2001, p. 1). Since English has appeared in every field of life, the demand for teaching and learning English has increased accordingly and this demand has brought with it different learning and teaching methods and approaches. As the world changes, the methods and approaches to teaching and learning English adapt themselves to those changes. Both teachers and learners shape these renewed methods and approaches for themselves. Testing and assessment is perhaps one of the most affected educational components by these changes, requiring adaptations and adjustments in how the language performances of learners are assessed (Şahin, 2019).

Testing and assessment can be seen as inseparable components in language education. Whenever there is teaching, there is also assessment and testing. Tosuncuoğlu (2018) states that "assessment includes information about student awareness, understanding, perception and attitude to learning" (p. 165), representing an attempt to meet learners' needs. It is implemented for various reasons, such as showing how many learners have accomplished their learning objectives, which learners have difficulties with their learning, and which techniques are better or more useful in teaching language (Tosuncuoğlu, 2018). Brown (2003), on the other hand, defines testing simply as "a method of measuring a person's ability, knowledge or performance in a given domain" (p. 3). In other words, a test, as a component of assessment, is applied to see what learners can achieve or have achieved. Hughes (2003), further, states that testing influences both teaching and learning. This impact known as "backwash" may be either harmful or beneficial (Hughes, 2003). Harmful backwash may happen when the contents of tests and techniques differ from the goals of the

course while beneficial backwash may happen when the goals of the course are in line with the contents of tests and techniques (Hughes, 2003). Assessment clearly does not only include tests as assessment tools; it also includes teacher observations, quizzes, projects, learner diaries, and portfolios (Hughes, 2003). Therefore, teachers do not only use tests to assess learners; they also use other different tools according to specific instructional goals and learning objectives.

There are different types of assessment according to different purposes. As one type of assessment, summative assessment is generally applied at the end of a teaching-learning procedure to understand what learners have learned and how much they have achieved. According to Dolin, Black, Harlen and Tiberghien (2018), "the aim of summative assessment is generally to report on students' level of learning at a particular time, rather than to impact on ongoing learning, as in the case of formative assessment" (p. 61). As another type of assessment, formative assessment is carried out by teachers to understand the improvement of learners, how much they have learned, and what they still need to learn, and teachers use that information to change their teaching plans with a strong emphasis on providing feedback to learners (Hughes, 2003). Brown (2003) explains formative assessment as follows: "most of our classroom assessment is formative assessment: evaluating students in the process of "forming" their competencies and skills with the goal of helping them to continue that growth process" (p. 6). Formative assessment which is an ongoing process aims to facilitate both learning and teaching processes and requires feedback. Examples include portfolios, dairies, teacher observations, project works, self-assessment and peer-assessment (Hughes, 2003). This need for assessment also raises the issue of the assessment literacy of teachers.

Assessment literacy (AL) is a general term referring to skills and knowledge regarding testing and assessment, how to prepare assessment instruments in a reliable and valid way, how to apply and score tests and the usage of assessment techniques applied in the classroom. "Although there is not currently a universally agreed upon definition of the term, assessment literacy can be defined as the creation and use of the spectrum of assessment techniques and instruments as part of the teaching-and-learning process" (Gareis, & Grant, 2015, p. 8). Jannati (2015) states that "teachers' knowledge of assessment or assessment literacy has a great effect on the quality of education" (p. 27). Being assessment literate can provide beneficial backwash. According to Elshawa, Heng, Abdullah, and Rashid (2016), when learners and teachers know the test and its goals together with its format, beneficial

backwash can take place and in order to ensure the development of learners' learning, in line with positive backwash, teachers should have AL. In other words, a teacher's AL level may have an influence on learners' success and learning in a course. When beneficial backwash takes place in a classroom, the teaching and learning processes might be more advantageous for both teachers and learners, which might result in incremented success levels of learners. AL is a general term applicable in every field related to assessment, but there is one specific term for the area of language learning and teaching: "language assessment literacy".

"Language assessment literacy (LAL), as a term specific to the field of language teaching and learning refers to the knowledge, skills and principles that stakeholders involved in assessment activities are required to master in order to perform assessment tasks" (Inbar-Lourie, 2017, p. 257). Scarino (2013) explains that in second language education, teachers' knowledge and training experiences related to assessment are very important. The more teachers know about how to assess learners, the better the outcomes may be in the classrooms, because a teacher with LAL may know what to assess and how to assess what learners know and can do. This may also help to decrease the level of anxiety and stress among learners, which may result in increased success levels of learners both in the classroom and on tests. Based on the results of successful assessment practices, teachers can identify common mistakes and not-internalized language topics and do remedial teaching. LAL also encompasses skills and knowledge regarding appropriate and accurate assessment tools for assessing different language skills because it is important to use the appropriate assessment tools for specific language skills. LAL provides teachers with the necessary information about what kind of assessment tools should be used to assess different language skills and gain accurate and reliable results from learners for each language skill after the implementation of those assessment tools (Hughes, 2003).

When educational systems are taken into consideration, it is known that the Turkish educational system is highly exam-oriented (Hatipoğlu, 2016) with various high stake tests with different purposes in different phases of education (e.g., High School Entrance Exam, University Entrance Exam- Basic Proficiency Test, Field Proficiency Test, Foreign Language Test, and Proficiency exams for preparatory schools). In the context of LAL, such tests generally include multiple-choice questions intended to assess different language skills. However, traditional assessment methods "such as multiple choice, matching, fill-in the blank and true-false questions and classical examinations with few open-ended questions" do not really assess learners' "higher order cognitive skills, informative and affective

behaviors within a short time frame" (Birgin & Baki, 2009, p. 681). In the Turkish context, there are universities (state and private) with preparatory schools that mainly offer intense English courses to learners with the aim of introducing grammar and vocabulary together with listening, reading, speaking, and listening skills (Çetinavcı & Topkaya, 2012). This aim is generally met with the help of separate courses and sometimes it is met by combining some skills in one course "such as speaking –listening"" (Çetinavcı & Topkaya, 2012, p. 83). However, these schools generally focus most on teaching grammar and vocabulary at the expense of eventually neglecting productive language skills (Çetinavcı & Topkaya, 2012) -because of high emphasis on grammar and vocabulary that might be related to teachers' and learners' educational background and beliefs on language learning- which signals that studies related to assessing productive language skills need to be carried out with the aim of demonstrating how important they are and why they should not be neglected. This is one of the reasons why assessing productive language skills appropriately and accurately is perhaps one of most the important issues in foreign language education. There are a lot of assumptions made about how to assess these skills. For instance, Luoma (2004) says that "from a testing perspective, speaking is special because of its interactive nature" (p. 170). Speaking as a language skill may be considered different from other language skills because it is interactive, requires turn-taking, and indispensably encompasses listening and other dynamics of spoken interaction. Luoma (2004) further adds that speaking is typically tested live, where the discourse of the test may not be totally certain, just as even if speakers have the same roles and goals and the conversation is about the same topics, the conversations are never exactly the same. For the skill of writing, according to Kroll (1998), "…the underlying concern with L2 students in our English-speaking institutions is to find ways to insure that their proficiency in English is adequate to meet their needs as degree candidates" (p. 232) but she further notes that "skill in writing is clearly only one aspect of such proficiency" (p. 232). Weigle (2002) affirms that "in choosing or designing a writing test, the logical place to begin is considering what we plan to use the test for" (p. 40). Crusan, Plakans, and Gebril (2016) support the idea that good assessment practices are crucial for writing. Both of these productive language skill types require different assessment techniques and tools to be implemented in classrooms. In some universities, there are testing units that prepare tests for learners, and teachers do not need to prepare their own tests. However, tests are not the only tools that teachers can use to assess learners in classrooms. Assessment may not be implemented through tests or quizzes, but observing learners in the classrooms can also be considered as a type of assessment allowing teachers to understand how learners cope with

these productive language skills. Therefore, learning processes for productive language skills may be provided smoothly with the aid of teachers' high levels of LAL, which will eventually provide internal knowledge about the instruments to be used, the techniques to be implemented, how to implement formative assessment tools inside and outside the classroom, and how to interpret the scores.

With all this information in mind, the purpose of this study was to depict the LAL of instructors of English as a foreign language (EFL) in assessing productive language skills and their perceptions of the assessment of productive language skills through formative assessment.

1.2. Statement of the Problem

Teachers are one of the key components of education and they are responsible for delivering instruction efficiently and effectively and implementing assessment to understand whether given instruction is delivered successfully to learners or not. With this aim, throughout teaching and learning processes, assessment is applied by teachers to see both how the learning process continues and how much learners gained in those processes. In order for teachers to be able to integrate assessment into their teaching properly, they should be assessment-literate since assessment and teaching are intertwined. When foreign language teaching is taken into consideration, language teachers should be aware of the importance LAL, which has been highlighted by many scholars (e.g., Fulcher, 2012; Inbar-Lourie, 2017; Inbar-Lourie, 2008a; Khadijeh & Amir, 2015; Tosuncuoğlu, 2018; Scarino, 2013). It is crucial in both teaching and learning English because the more important assessment is, the more important LAL is, as well. However, a great number of researchers have demonstrated that the LAL levels among EFL instructors are insufficient (Ölmezer-Öztürk & Aydın, 2019; Ölmezer-Öztürk, 2018; Tsagari & Vogt, 2017; Vogt & Tsagari, 2014). Research carried out on LAL with pre-service EFL teachers revealed low LAL levels as well (Çetin-Argün, 2020; Sarıyıldız, 2018; Şahin, 2019; Tamerer, 2019). In addition to LAL, formative assessment, which is an important component of LAL, has been depicted as crucial by many scholars (Black, Harrison, Lee, Marshall, & Wiliam, 2004; Black & Wiliam, 2010; Cauley & McMillan, 2010). When integrated with teaching, it can enhance a powerful teaching and learning environment and provide different opportunities for both learners and teachers to see their strengths and weaknesses with the help of mutual feedback.

As described in the previous section, Turkey is an exam-oriented country (Hatipoğlu, 2016). Teachers are responsible for the assessment of their learners and they need to apply different assessment tools in that process. From primary schools to universities, teachers are expected to prepare and design their own assessment tools to understand the situations of learners. However, in some universities, there are testing development units (TDU) that prepare tests for learners so that teachers do not need to design their own. Tests, however, are not the only tools that a teacher can utilize to understand how proficient a learner is in English and various assessment tools to be implemented in classrooms depending on variables such as the needs of learners. For teachers to know about these assessment tools, they require LAL regardless of whether they have to prepare their own tests or not. When the related literature is reviewed, it is seen that the LAL levels of EFL instructors are generally low and they need related training related in this regard. In particular, assessing productive language skills can be one of the most demanding works in language teaching because it takes long time due to fact that they fall under the category of subjective tests, and classrooms are often crowded. Also, teachers may not provide enough opportunities for learners to practice those skills in the classrooms for similar reasons such as time and practicality. While assessing productive language skills, there are many components that a teacher must take into consideration, such as, fluency, accuracy, and the use of different vocabulary items, for the skill of speaking or language use, punctuation, and context for the skill of writing. The problem, however, lies in the fact that EFL teachers receive insufficient training in assessment and testing to perform valid and reliable assessments of these types (Stiggins, 2002; Vogt & Tsagari, 2014; Tsagari & Vogt, 2017) which ultimately leads to deficiencies in AL (Brookhart, 2001).

The aforementioned studies address both pre-service and in-service teachers' lacks and needs in terms of LAL and formative assessment. These studies have investigated language skills as a whole instead of focusing on one or two specific language skills. In other words, in the related literature, there are several studies about LAL and formative assessment including mostly learners and pre-service EFL teachers (e.g., for LAL: Çetin-Argün, 2020; Hatipoğlu, 2015; Şahin, 2019; Tamerer, 2019; for FA: Büyükkarcı, 2010; Sarı, 2019; Sönmez, 2013; Yurtdakal, 2015) however, to the best of the researcher's knowledge, in Turkey, there remains a need for a study investigating the LAL levels of EFL instructors' who work in preparatory schools in terms of assessment of productive language skills and their perceptions of assessing productive language skills through formative assessment. This

study attempts to fill in this niche to contribute to the field of foreign language teacher education and language testing and assessment.

1.3. Purpose of the Study

The purpose of this study is to depict EFL instructors' language assessment literacy in assessing productive language skills and their perceptions on assessing productive language skills through formative assessment in preparatory schools.

1.3.1. Research Questions

The following main research questions have been formulated in line with the research purpose:

1. What is the EFL instructors' level of language assessment literacy in assessing students' productive language skills in English?

2. What are the EFL instructors' perceptions on assessing productive language skills through formative assessment?

3. How do the EFL instructors implement formative assessment tools in their classrooms?

4. What problems and challenges do the EFL instructors encounter while assessing learners' productive language skills in English?

5. What do the ELF instructors suggest as solutions for the stated problem/s?

1.4. Significance of the Study

There is an increasing demand for both teaching and learning English and this demand reflects itself in the need for assessment and testing. Since teachers are one the key components of the process of education, they should be aware that they need to improve themselves in assessment and testing to be more confident in the classroom while assessing learners' language skills (Şahin, 2019). They need to have knowledge about what to assess and how to assess in a reliable and valid way. For that, they need to improve their LAL, which allows teachers to be knowledgeable about the implementation of different assessment tools, which assessment tools meet the needs of the learners, and how teaching and learning processes advance.

Preparatory schools in Turkey offer mostly English courses to learners which provides opportunities for them to improve academically. Since the main focus of these schools is to teach English, they need to focus on language skills together with their sub-components such as grammar and vocabulary. However, as Çetinavcı and Topkaya (2012) stated, these schools mostly focus on grammar and vocabulary and might neglect productive language skills. That can be considered the reason why this present study focused on EFL instructors in preparatory schools and their LAL levels in assessing productive language skills together with perceptions application of formative assessment. This research can be considered important in several ways. First, it aims to depict EFL instructors' LAL in assessing productive language skills together with their perceptions of assessing productive language skills through formative assessment. It is believed that the study will accordingly contribute to the understanding of EFL instructors' perceptions of assessing productive language skills through formative assessment and how they implement formative assessment tools in classrooms. Second, it is believed that this study will shed light upon the problems and challenges EFL instructors' encounter while assessing productive language skills and their suggestions and possible solutions for those problems. Third, the findings of this study may increase the awareness of how important it is to develop and improve the LAL of EFL instructors and teachers. Finally, this study may provide insights for teachers, policy makers, and learners about how to implement formative assessment in classrooms to assess productive language skills.

1.5. The Definition of Key Terms

Research related to language testing and assessment shows that there are different terms that may be used interchangeably or conceptualized differently by various scholars. Followings are the key terms and their definitions as conceptualized by the researcher herself.

Assessment: According to Brown (2003), "assessment is an ongoing process that encompasses a much wider domain. Whenever a student responds to a question, offers a comment, or tries out a new word or structure, the teacher subconsciously makes an assessment of the student's performance" (p. 4).

Measurement: "measurement in the social sciences is the process of quantifying the characteristics of persons according to explicit procedures and rules" (Bachman, 1990, p. 18).

Testing: Testing measures "a person's ability, knowledge, or performance in a given domain" via tests" (Brown, 2003, p. 3).

Evaluation: "the fundamental use of testing in an educational program is to provide information for making decisions, that is, for evaluation" (Bachman, 1990, p. 54).

Assessment Literacy (AL): Stiggins (1995) explains that assessment literate teachers know "what they are assessing, why they are doing so, how best to assess the achievement of interest, how to generate sound samples of performance, what can go wrong, and how to prevent those problems before they occur" (p. 240).

Language Assessment Literacy (LAL): "language assessment literacy (LAL) refers to the knowledge skills and principles that stakeholders involved in assessment activities are required to master in order to perform assessment tasks" (Inbar-Lourie, 2017, p. 257).

Formative Assessment (FA): Black and Wiliam (2010) defines formative assessment as "…activities that provide information to be used as feedback to modify teaching and learning activities. Such assessment becomes formative assessment when the evidence is actually used to adapt the teaching to meet student needs" (p. 82).

Summative Assessment (SA): Taras (2005) explains summative assessment as "the process of assessment leads to summative assessment, that is, a judgement which encapsulates all the evidence up to a given point" (p. 468).

CHAPTER 2

2. LITERATURE REVIEW

In this chapter, a review of literature on language assessment literacy and formative assessment for productive language skills is presented in detail.

2.1. Assessment

Teaching is a general term that concerns every field of education. According to Azis (2012), "teaching is complex; it involves elements such as curriculum, subject matter and epistemology, teaching and learning, and also assessment and evaluation" (p. 41). Assessment is an umbrella term and an expansive field in language teaching. It is considered as "one of the cornerstones of the educational process but perceptions and needs related to assessment are evolving and changing and might be context dependent" (Hatipoğlu, 2015, p.111). This massive field includes testing and testing types as well. As Can (2017) explains:

> "As an ongoing process, assessment is a comprehensive field of study, which is closely related to such terms as testing, evaluation, measurement and so on; accordingly, there have always been various definitions of the term, assessment, with regard to different perspectives and contexts" (p. 6).

Testing and assessment are understood to be two different terms. Bachman (1990) defines a test as a measurement tool that is supposed to be used to understand one's attitude. Andhika (2011), on the other hand, states that testing is defined "as the narrowest in scope. It is one of the techniques for collecting the data or scores" (p. 94). Norris (2006) defines assessment as "the systematic gathering of information about student learning in support of teaching and learning" (p. 579). Ezir (2013) depicts the difference between testing and assessment by stating that assessment is a general term that encapsulates information about what students can achieve and how, covering "many ways and methods of information gathering, formal and informal, at different times and in different contexts" (p. 38), while testing, on the other hand, is an element of assessment by way of collecting information about students and it focuses on "finding the norm" (p. 38). They may be different terms however, they are both inseparable parts of education and they have a strong relationship. Without one, the other would not survive on its own.

Together with changes in the educational system, the types of assessment have also changed. While grammar-translation method was popular, assessment was applied by that

method's rules. According to Larsen-Freeman and Anderson (2013), "written tests in which students are asked to translate from their native language into the target language or vice versa are often used. Questions about the target culture or questions that ask students to apply grammar rules are also common" (p. 20). As a result of some factors that led people to question and reject the grammar-translation method, the method changed (Richards & Rodgers, 2014). The "direct method", in opposition to the grammar-translation method, was adopted, and assessment practices also changed. In the direct method, students use the target language with the help of oral and written skills (Larsen-Freeman & Anderson, 2013). Since today the new era is called as "communicative approach" era, assessment has changed one more time. Richards and Rogers (2014) state that "the Communicative Approach in language teaching starts from a functional theory of language- one that focuses on language as a means of communication" (p. 87). While traditional assessment, based on pen and paper exams, was once popular, today alternative assessment methods can be more preferred. Together with the "communicative approach", a student-centered teaching approach is applied in classrooms more often. One of the positive outcomes of teaching in a student-centered classroom is that students have much more control of their learning than they do in teacher-centered classrooms. With this change, teachers have transitioned new techniques. As Kaya (2020) explains, "teachers' role in the classroom is to make use of assessment practices in order to improve learning by adapting existing techniques in assessment to meet the needs of their students" (p. 14).

While it is important to assess learners, it is even more important for teachers to know what to consider while assessing learners. According to Harris and McCann (1994), there are five elements to consider while performing assessments. First, assessments should be carried out "constructively" with a focus on the achievements of students rather than failure (p. 4). The second element is "reliability," which focuses on consistency and fairness, while the third element, "validity," focuses on establishing certain objectives and reaching them through assessment (p. 4). The fourth element is "practicality," which supports the idea that assessment tools should not be too time-consuming, and the final element is "accountability," signalling that "as professionals, teachers should be able to provide learners, parents, institutions and society in general, with clear indications of what progress has been made and if it has not, why that is so"" (p. 4). These five elements should be taken into consideration when assessment takes place in the classrooms in terms of realizing the objectives of the course and gaining better results from students. Assessment is an important

component that enables teachers to have an idea about what learners are capable of doing and provides feedback. According to the feedback that teachers receive, the learning environment and assessment tools may be changed to maintain a better atmosphere for learners to internalize the target language. As Ballıdağ (2020) summarizes, "an accurate assessment will motivate students in return and facilitate the learning environment for both students and teachers" (p. 5).

All things considered, assessment and teaching are complementary elements in foreign language education. Assessment is a broad term covering tests as measurement instruments and the importance of assessment in education is not negotiable since it is crucial in efforts of learners to learn the target language and it plays an important role "in the process of learning, and connects students to new knowledge using their current abilities"" (Tosuncuoğlu, 2018, p.166).

2.2. Assessment Types

There are many different assessment types, including formative assessment, summative assessment, self-assessment, and peer assessment. In this study, formative assessment (FA) and its differences from summative assessment (SA) when necessary will be considered. The basic difference between SA and FA is that their focuses are on different things. SA is mainly applied at the end of the term, while FA is typically applied throughout the term consistently and continuously.

2.2.1. Summative Assessment

SA is one of the assessment types generally applied at the end of the term to see what students can achieve after a certain period of time. According to Harlen (2007), "summative assessment is carried out for the purpose of reporting the achievement of individual students at a particular time" (p. 16). SA is associated with more learning aims that may be reached with time and it may be applied via testing (Harlen, 2007). Dixson and Worrell (2016) explain that "summative assessments are almost always graded, are typically less frequent, and occur at the end of segments of instruction" (p. 156). They further add that "examples of summative assessments are final exams, state tests, college entrance exams (e.g., GRE, SAT, & LSAT), final performances, and term papers" (p. 156). Taras (2005), on the other hand, explains that "the process of assessment leads to summative assessment, that is, a judgement which encapsulates all the evidence up to a given point" (p. 468). According to Brookhart (2001), "for grading and other times of accountability, teachers do indeed collect

and use summative assessment information" (p. 157). Dolin, Black, Harlen, and Tiberghien (2018) state that SA "involves collecting, interpreting and reporting evidence of learning" and that "interpretation of evidence is in relation to the goals that students are intended to have achieved at a certain point, such as the end of a year, semester or stage" (p. 62). There are different assumptions made about SA; it may be positive and beneficial for both teachers and students or it may have some negative effects. According to Solgun-Günel (2014), SA is an instrument used by teachers after the education process and students are expected to respond to the learning aims. It also helps teachers understand what learners have learned from their instructions throughout the semester.

Yüksel and Gündüz (2017) emphasize the usage of SA, stating that regardless of how important it is to ensure the continuity of the teaching-learning process in a better way, SA is still applied most of the time. Biggs (1998) further argues that SA gives rise to backwash and this backwash is generally believed to be negative:

> "The backwash from SA is generally agreed to be negative, generating ego-related and other non-task priorities that in turn create a surface approach, where the learner uses lower cognitive level activities than those actually required, so that performance is low level and fragmentary" (p. 106).

SA which is generally applied at the end of the term may cause problems for learners such as decreased motivation on the grounds that it is applied only once as an opportunity for learners to show what they have learned throughout the term. This may be good for them, or it may cause stress and anxiety. According to Perera-Diltz and Moe (2014), on the other hand, SA has an advantage of holistic evaluation and integration occurring one time. They also note the negative side of SA; if a learner cannot do his or her best with a selected SA type, an accurate assessment of the learners' learning is not possible and learners' enthusiasm and eagerness for the process of learning might be reduced.

2.2.2. Formative Assessment

Black and Wiliam (1998) state that FA does not have a widely accepted definition but maybe understood as "encompassing all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged" (pp. 7-8). Büyükkarcı (2010) says that "in general terms, formative assessment, also known as assessment for learning, on-going assessment, or dynamic assessment, is concerned with helping pupils to improve their learning" (p. 21). Clark (2012) states that "formative assessment is connected by two

contiguous assessment objectives: assessment for learning (AfL) and assessment as learning (AaL)" (p. 5). Cauley and McMillan (2010) state that "formative assessment is a process through which assessment-elicited evidence of student learning is gathered and instruction is modified in response to feedback" (p. 1). FA, unlike SA, is not applied at the end of the term. Instead, it is applied consistently and continuously throughout the term. According to Black (1993), FA entails close relationships between instructors and learners. It is a significant means of making learning possible, and especially deeper learning (Rushton, 2005). Sadler (1989) affirms that "formative assessment is concerned with how judgments about the quality of student responses (performances, pieces, or works) can be used to shape and improve the student's competence by short-circuiting the randomness and inefficiency of trial-and-error learning" (p. 120). FA is generally used to reach learning goals and improve learners' learning. According to Ökten (2009), FA is used to understand the exact situation of the students and what should be done for improvement, and it provides both teachers and students with information on what and how much students are learning and how well the learning takes place. Teachers use FA for different purposes. Stiggins and DuFour (2009) explain the main reasons why teachers prefer FA in classrooms as follows:

> "Teachers and schools can use formative assessment to identify student understanding, clarify what comes next in their learning, trigger and become part of an effective system of intervention for struggling students, inform and improve the instructional practice of individual teachers or teams, help students track their own progress toward attainment of standards, motivate students by building confidence in themselves as learners, fuel continuous improvement processes across faculties, and, thus, drive a school's transformation" (p.640).

FA has several advantages. One of the advantages FA facilitates is that it allows learners to control and be aware of what they learn. In other words, it facilitates self-regulated learning. According to Afitska (2014), FA can allows students to evaluate and observe their own improvement and what they do. Uysal-Kurtulmuş (2018) claims that "formative assessment or assessment for learning has become of a great value in education with the increase in learner-centered and goal oriented learning" (p. 13). Moreover, Stull, Varnum, Ducette and Schiller (2011) state that "in formative assessment, students become active participants with their instructors, sharing learning goals and understanding how their learning is progressing, what steps they need to take and how to take them" (p. 30). As learners are active participants, they may be more eager to learn or reach learning goals. Teachers' attitudes and constructive feedback lead learners to find their own way in this process, and thus learn how to deal with the obstacles they face while trying to reach learning

goals. As grades are not the primary concern of FA, learners may feel less stress and more control at the same time.

The second advantage of FA is that, unlike SA which focuses on results, there is a strong emphasis on the process, which gives learners opportunities to correct or improve themselves. They may have a chance to observe themselves in learning and teaching processes. According to Doğru (2020), FA may be used to provide a process for learning and teaching and it may also aid students in their learning processes; instead of teachers, students should be the focal point here. Kaya (2020), on the other hand, states that "Both students and the teacher have an active role in formative assessment" (p. 14). Boston (2002) affirms that if "teachers know how students are progressing and where they are having trouble, they can use this information to make necessary instructional adjustments, such as reteaching, trying alternative instructional approaches, or offering more opportunities for practice" (p. 1). This may result in increased levels of success among the students. Wiliam (2006) argues that for assessment to be formative, it does not matter how long the feedback process is, where the assessment takes place, who applies it or who is responsible for it; what is important is to obtain the information, clarify it in terms of needs of learning, and use it to make improvements so as to meet the needs of learning better.

The third major advantage of FA is that throughout the whole teaching-learning process, feedback is given by teachers and/or learners. This helps learners see how much they have improved. Wiggins (2012) defines feedback as a term "used to describe all kinds of comments made after the fact, including advice, praise, and evaluation" (p. 1). Feedback should certainly be provided however, it must be conveyed in a meaningful context and in an appropriate format. Wiggins (2012) suggests that effective types of feedback are "goal-referenced; tangible and transparent; actionable; user-friendly (specific and personalized); timely; ongoing; and consistent" (p. 2). Gedye (2010), on the other hand, states that:

> "The least restrictive way of viewing formative assessment is that it is assessment which provides the learner with information that allows them to improve their learning and performance. In this sense, an end-of-module, graded assignment may be formative if the student receives good quality feedback on how they might improve their work, whilst a mid-semester, ungraded assessment may not be formative if all the feedback says is, "good work, well done"" (p. 40).

As noted by Gedye (2010), for feedback to be effective, while applying FA, it is important to provide the feedback in an appropriate form. Providing learners with only one type of feedback may not completely help teachers to reach the learning goals that they set.

Irons (2008) defines FA as, "any task or activity which creates feedback (or feedforward) for students about their learning. Formative assessment does not carry a grade which is subsequently used in a summative judgement" (p. 7). Irons (2008) further emphasizes that feedback must be comprehensible and conveyed in a way that allows learners to apply the feedback in order to achieve "the learning outcomes or reaching the required standard"" (p. 23). Moreover, feedback should be presented to encourage learners to start seeking find any learning problems (Irons, 2008). Tunstall and Gsipps (1996) also state that "feedback from teachers to children, in the process of formative assessment, is a prime requirement for progress in learning"" (p. 389). They clarify that FA is a procedure whereby teachers assess or criticise the performances or productions of learners to construct and advance their capabilities. It is clear that feedback is beneficial and should be provided by teachers to improve learners' understanding and learning. Nicol and Macfarlane-Dick (2006) refer to seven benefits of good feedback:

- "helps clarify what good performance is (goals, criteria, expected standards);
- facilitates the development of self-assessment (reflection) in learning;
- delivers high quality information to students about their learning;
- encourages teacher and peer dialogue around learning;
- encourages positive motivational beliefs and self-esteem;
- provides opportunities to close the gap between current and desired performance;
- provides information to teachers that can be used to help shape teaching" (p. 205).

Even though there are many opportunities provided by feedback, it should also be noted that giving detailed feedback may become highly time-consuming as Wingate (2010) cautions. Teachers should be careful while giving feedback in terms of finding an appropriate form and considering the time they can devote to it per learner. So as to avoid problems in the classroom, teachers should clarify the goals and those goals should encourage all learners, even those with low-motivation, to pursue the learning process (Wingate, 2010).

All things considered, FA is crucial with regard to enabling learning and providing feedback as one of its main components (Rushton, 2005). Moreover, when teachers recognize how learners improve themselves and where they have problems, they can utilize that information to make required alterations such as remedial teaching, the adoption of different instructional approaches or methods, and the offering more chances, which might improve the success of learners (Boston, 2002).

2.2.3. Distinctions Between FA and SA

FA and SA are different from each other as described by Harlen and James (1997). While SA "takes place at certain intervals when achievement has to be reported" (p. 372), for FA, they state that:

> "it has to take into account several instances in which certain skills and ideas are used and there will be inconsistencies as well as patterns in behaviour; such inconsistencies would be 'error' in summative evaluation, but in formative evaluation they provide diagnostic information" (p. 372).

According to Brookhart (2001), while FA is exclusive and focuses on learners' needs, SA is an answer to "external pressures and constraints, and the need for accountability" (p.157). Hattie (2003) refers to the difference between FA and SA by stating that "…the timing of the interpretations has major implications for the quality of the interpretations from both" (p. 6). It is clear that the main distinction between these two assessment types lies in which parts of the teaching-learning process they emphasizes. While FA focuses on processes, SA generally focuses on results. Glazer (2014) explains this as follows:

> "Formative assessment is any task that provides feedback to students on their learning achievements during the learning process. Summative assessment is used for evaluation, in which there is limited or no feedback beyond the achievement report, and is usually a numerical or letter grade score" (p. 277).

Hughes (2003) describes the difference between FA and SA by touching upon the points to those that Glazer (2014) mentions. Hughes (2003) clarifies that the focal point of FA is the process of learning and it is applied by teachers to understand how learners deal with the learning procedure and how much they have improved, and allowing that to be used to alter teaching plans if necessary. He also explains that the feedback given throughout the whole process, observations, and portfolios, as well as informal tests and quizzes, are used. SA, on the other hand, is applied at the end of teaching-learning process to understand what learners have achieved so far and teachers use formal tests to measure learners' performances (Hughes, 2003). According to Brown (2003), the aim of FA is to supervise the learning process and provide appropriate feedback in terms of learners' performances while the aim of SA is to evaluate what a learner has achieved which generally happens at the end of a term. According to Dixson and Worrel (2016), the characteristics of formative and summative assessments differ in the following ways:

Table 2.1. Characteristics of Formative and Summative Assessments

Characteristics	Formative Assessment	Summative Assessment
Purpose	To improve teaching and learning To diagnose student difficulties	Evaluation of learning outcomes Placement, promotion decisions
Formality	Usually informal	Usually formal
Timing of Administration	Ongoing, before and during instruction	Cumulative, after instruction
Developers	Classroom teachers to test publishers	Classroom teachers to test publishers
Level of stakes	Low-stakes	High-stakes
Psychometric Rigor	Low to high	Moderate to High
Types of questions asked	What is working What needs to be improved How can it be improved	Does student understand the material Is the student prepared for next level of activity
Examples	Observations Homework Question and answer sessions Self-evaluations Reflections on performance Curriculum-based measures	Projects Performance assessments Portfolios Papers In-class examinations State and national tests

Note. Adopted from "Formative and Summative Assessment in the Classroom, by Dante D. Dixson & Frank C. Worrell, 2016, *Theory Into Practice*, 55(2), p.153-159. Copyright 2010 by The College of Education and Human Ecology, The Ohio State University.

According to Dolin; et al. (2018), there are certain steps to be taken in both FA and SA which generally differ:

Table 2.2. Dimensions of assessment purposes and practices

	Formative ⟷ Summative			
	Informal Formative	Formal Formative	Informal Summative	Formal Summative
Major Focus	What are the next steps in learning?		What has been achieved to date?	
Purpose	To inform next steps in teaching and learning	To inform next steps in teaching and learning	To monitor progress against plans	To record achievements of individuals
How evidence collected	As normal part of class work	Introduced into normal class work	Introduced into normal class work	Separate task or test
Basis of judgement	Student- and criterion-referenced	Student and criterion-referenced	Criterion and student-referenced	Criterion-referenced
Judged by	Student and teacher	Teacher and student	Teacher	Teacher or external examiner
Action taken	Feedback to students and teacher	Feedback to students and into teaching plans	Feedback to students and into teaching plans	Report to student, parent, other teachers, etc.
Epithet	Assessment for learning	Matching	Dip stick	Assessment of learning
Examples of feedback modes	Verbal feedback on-the-fly	Written feedback on classroom work	Response to informal test or quiz	Synoptic report on achievement of course goals

Note. Adopted from "Exploring relations between formative and summative assessment", by Dolin, Black, Harlen, & Tiberghien, 2018, In *Transforming assessment* (pp. 53-80). Springer, Cham.

When the two figures above are taken into consideration, the differences between FA and SA can be clearly seen in detail. In Table 2.1., the characteristic differences between FA and SA described by Dixson and Worrell (2016) are shown. They identify eight different evaluation criteria and each criterion takes shape according to the type of assessment. There is a difficulty level for each assessment type and some example activities are specified for both FA and SA.

Table 2.2 provides a deeper analysis of the differences between FA and SA as indicated by Dolin et al. (2018). In contrast to Table 2.1., there is an "epithet" line in Table 2.2 that shows the different names used for each assessment type and there are four categories divided among formal formative/summative and informal formative/summative headings. In this schema, Dolin, et al. (2018) highlight how teachers can collect information for each assessment type, and it is clearly understood that in FA, feedback is given to the learners. In SA, in contrast, reports are given not only to learners but also to other teachers and parents. The purposes of FA and SA as clarified in Tables 2.1 and 2.2 are clarified also reflect the differences observed by Hughes (2003), as FA focuses on the learning process while SA focuses on results.

Another difference between these two assessment types is specified by Wiliam and Black (1996), who state that "…all assessments can be summative (i.e. have the potential to serve a summative function), but only some have the additional capability of serving formative functions" (p. 544). Therefore, it is important for teachers to specialize their teaching goals and materials. They need to determine teaching and learning goals first and then they need to find a way to use FA and SA throughout the semester with appropriate materials and feedback considered as important parts of those assessment types.

2.3. Assessment Literacy (AL)

Assessment takes place in every field of education. Brown (2003) defines assessment as "an ongoing process that encompasses a much wider domain" (p.4). Moreover, he states (2003) that teachers are assessing learners' performances, sometimes subconsciously, while they attempt to answer questions or form new sentences. If this is the case, then teachers are required to make use of different sorts of assessment materials to evaluate learners' knowledge and harmonize different kinds of assessment instruments to both strengthen their own guidance and evaluate learners' improvement; therefore, teachers need to advance and sustain an accurate awareness of assessment exercises and ideologies to ensure that

assessments boost and evaluate learners' knowledge while maintaining effective feedback (DeLuca & Klinger, 2010). Brown (2003) highlights the relationship among tests, assessment, and teaching (See Figure 2.1.). This importance of assessment raises the issue of AL.

Figure 2.1. Tests, assessment, and teaching
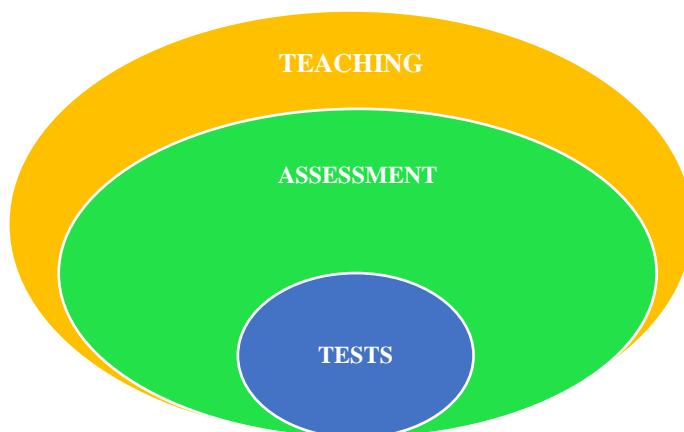
Note. Adapted from "*Language Assessment: Principles and Classroom Practises*" by Brown, D.H., 2003, p. 5. Pearson ESL.

Webb (2002) defines AL as "the knowledge of means for assessing what students know and can do, how to interpret the results from these assessments, and how to apply these results to improve student learning and program effectiveness" (p. 1). When assessment takes place appropriately, it helps teachers, learners and testers have information about learners' performances and to what extent learning goals are achieved in classrooms; therefore, it should be harmonized with teaching, building a relationship wherein assessment serves to advise and develop teaching and this mutual relationship cannot develop if teachers do not have sufficient background or education in assessment to improve, choose, and make use of tests and understand their results (Malone, 2013). This demonstrates that AL is as important as assessment itself.

Together with the educational changes occurring in the 21st century, the responsibilities that both teachers and learners have are also changing. According to Fulcher (2012), the 21st century is an era in which teachers have duties for testing and assessment and, "the range and number of stakeholders who require a level of assessment literacy has grown" (p. 115). Therefore, teachers are expected to be aware of what is changing in the educational system and how those changes affect testing and assessment because each

change brings something new to the field and new responsibilities for both teachers and learners. Popham (2011) states that "AL consists of an individual's understandings of the fundamental assessment concepts and procedures deemed likely to influence educational decisions" (p. 267). For teachers, Tsagari and Vogt (2017) say that "they are expected to exploit the increased variety of assessment procedures such as alternative forms of assessment that have become part of "mainstream" practices in many educational contexts" (p. 42). Yastıbaş (2018), on the other hand, explains that assessment-literate teachers are required to be careful in their efforts to cope with both theoretic and applied dimensions of assessment efficaciously. It is important that be assessment-literate in regards to both having theoretic knowledge and applying that knowledge in the classroom with the help of the necessary assessment tools.

Stiggins (1991a) discusses two different groups while explaining assessment literacy: assessment-illiterate and assessment-literate individuals. He claims that assessment illiterate teachers do not have an understanding of what it takes to develop "high quality achievement data" and do not assess the data they have. (p. 535). Moreover, they acknowledge "data at face value and can easily be intimidated and by a complicated presentation of test scores" (p. 535). Briefly, assessment-illiterate teachers are deficient in the means of being crucial users of data for assessments (Stiggins, 1991a). On the other hand, Stiggins (1991a) explains that assessment-literate teachers "have a basic understating of the meaning of high and low quality assessment and are able to apply that knowledge to various measures of student achievement" (Stiggins, 1991a, p. 535). Stiggins (1995) also highlights the fact that people who are conscious of the possible negative effects of false assessment do not allow learners to be in a position in which their achievements may be measured wrong. He suggests five standards to complete understanding of the concept of AL and sound assessment: "starting with clear purposes, focusing on achievement targets, selecting proper assessment methods, sampling student achievement and avoiding bias and distortion" (Stiggins, 1995, pp. 240-242). The American Federation of Teachers, the National Council on Measurement in Education, and the National Education Association (1990) jointly clarified several standards for teachers' responsibility for proficient assessments:

- deciding appropriate assessment methods for their own teaching plans
- having an ability to improve appropriate assessment methods for their own teaching plans;

22

- having an ability administer, mark, and read the results of the assessment methods created by themselves or other people;

- making use of the results of the assessment methods while drawing a conclusion about learners, designing their teaching, improving the syllabus, and schools

- having an ability to create valid procedures for grading learners;

- interacting with learners, parents, and educators about the results of assessments applied; and

- identifying improper assessment methods and the usage of assessment knowledge.

Considering those standards, it is clear that teachers are responsible for both theoretical and practical knowledge related to testing and assessment instruments, preparing appropriate tests, and administering and assessing them.

The importance of AL lies in the answers to the questions that teachers should ask themselves. These questions are highlighted by Stiggins (1991a) as follows:

1. "What does this assessment tell students about the achievement outcome we value?
2. And what is likely to be the effect of this assessment on students?" (p. 535).

The answers obtained from these questions reflect the knowledge of assessment-literate teachers. In other words, answering these questions reveals whether or not teachers are aware of sound assessments and how they apply and evaluate them. Assessment-literate teachers have an awareness of what to do and what not to do together with why and how to do it. According to Siegel and Wissehr (2011), "a teacher with AL can gather information, communicate with students about the status of their understanding, use this information to set immediate and long-term goals, and adjust instruction based upon the results of classroom assessments" (p. 374). Stiggins (1999) similarly explains that assessment-literate teachers can "connect assessments to clear purposes, clarify achievement expectations, apply proper assessment methods, develop quality assessment exercises and scoring criteria and sampling appropriately, avoid bias in assessment, communicate effectively about student achievement, and use assessment as an instructional intervention" (pp. 25-26). Pastore and Andrade (2019) claim that:

> "an assessment literate teacher understands and differentiates the aims of the assessment, and articulates a sound, cyclical process of collection, interpretation, use evidence, and communication of feedback. In this way, assessment information can be used to make inferences about student learning, inform decisions about curriculum, and adjust instruction" (p. 135).

There is another perspective that teachers should take care of their own assessment education in terms of helping both themselves and learners go further. Popham (2009) thinks that "it seems that AL is a commodity needed by teachers for their own long-term well-being, and for the educational well-being of their students" (p. 11). In short, AL entails understanding assessment, assessment tools and how to use them to improve learners' learning and teachers own development.

2.4. Language Assessment Literacy (LAL)

Assessment plays a crucial role in every field of education. In English language education, the teaching and learning processes might be challenging for both teachers and learners since they require intensive knowledge in different areas including culture, syntax, semantics, and pragmatics. Both teachers and learners are expected to have knowledge about those areas to transfer for teaching and learning to be realized successfully. AL, on the other hand, is a more general term with LAL. In other words, LAL is just one branch of AL and Giraldo (2018) explains that "language as a construct for assessment is what differentiates LAL from its generic term" (p. 183). He adds that "knowledge, skills, and principles in language assessment coexist with teachers' ways of thinking and acting upon the act of assessment" (p. 185). According to Kremmel and Harding (2020):

> "Given the widespread use of language assessments for decision-making across an increasing number of social domains (education, immigration and citizenship, professional certification), it has become vital to raise awareness and knowledge of good practice in language assessment for a wide range of stakeholder groups" (p. 100).

LAL is related to the way in which teachers think about assessment and what they do to assess their learners' language performances. It is also noted by Berry, Sheehan and Munro (2019) that since assessment can affect and improve learning processes, it is important for teachers to be able to gather various teaching, learning, and assessment exercises. Accordingly, it is important that teachers be aware of what to do and what not to do when it comes to language assessment. Inbar-Lourie (2017) explains that "LAL has drawn considerably from the literature and research on general AL, while attempting to set itself apart as a knowledge base that incorporates unique aspects inherent in theorizing and assessing language-related performance" (p. 259). LAL refers to

> "the knowledge, skills and abilities required to design, develop, maintain or evaluate, large-scale standardized and/or classroom based tests, familiarity with test processes, and awareness of principles and concepts that guide and underpin practice, including ethics and codes of practice. The

ability to place knowledge, skills, processes, principles and concepts within wider historical, social, political and philosophical frameworks in order to understand why practices have arisen as they have, and to evaluate the role and impact of testing on society, institutions, and individuals" (Fulcher, 2012, p.125).

Inbar-Lourie (2008a) explains that "the language assessment knowledge base in fact comprises layers of assessment literacy skills combined with language specific competencies, forming a distinct entity that can be referred to as 'language assessment literacy' (henceforth, LAL)" (pp. 389-390).

As can clearly be understood from these definitions proposed by different researchers, LAL is a major concept including sub-components such as testing and classroom-based assessment (Şahin, 2019). It is important for teachers to be aware of how they can make use of practices of classroom based assessment to advance learning "such as assigning grades, providing feedbacks to learners and modifying teachers' teaching and learners' learning are all intended to renew, keep track and record learners' performance in the classroom so as to enable teachers to enhance learners' learning" (Lan & Fan, 2019, p. 120). Pill and Harding (2013), on the other hand, have a different perspective on this subject and they discuss the existence of different types of literacies such as "media literacy, and health literacy," while LAL symbolises a particular type of literacy including skills of reading and writing (p. 382). In short, "LAL thus refers to stakeholders' ability to understand, interpret, and use the language testing concepts within their specific professional context" (Deygers & Malone, 2019, p. 2). Taylor (2013) asserts that due to the multifaceted nature of effects of testing in education, teachers have varying LAL levels and the ways in which they deal with assessment issues vary as well. Recognizing the range of proper stakeholders and assessing their particular requirements in relation to meanings given to test scores in their own contexts and; how those scores might be used have become important in settings where assessment has a major role (Taylor, 2013). As Taylor (2013) further explains:

> "AL/LAL development activities may need to be integrated within professional development programmes or briefing sessions for other stakeholder groups, such as civil servants or politicians. In each case, the primary content relevant to the needs of that stakeholder group (e.g., will it be measurement theory? practical know-how? ethical principles?) has to be extracted and translated into a language that particular group can access and understand" (p. 407).

Figure 2.2. Levels of AL/LAL differentiated according to stakeholders' constituency
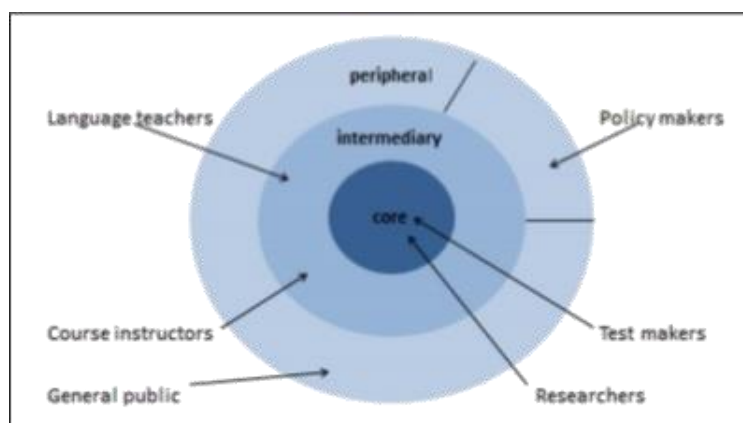
Note. Adopted from "Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections" by Taylor, L., 2013, *Language testing*, *30*(3), p. 403-412. Copyright 2013 by The Author(s).

According to O'Loughlin (2013), "its relevance is not restricted to language testing specialists and language teachers but also to educational policy developers and test users, such as university admissions staff" (pp. 363-364). As can be understood from Figure 2.2., there are three levels of LAL and AL, the core, intermediary, and peripheral levels which is consistent with the position taken by O'Loughlin (2013). The core of this model represents researchers' and test makers' requirement for training in language testing and assessment. At the intermediary level, language teachers and course instructors are located and at the peripheral level are the general public and policy makers.

Yastıbaş (2018) describes the importance of LAL for teachers, providing four main reasons: "teachers as the agents of assessment, language assessment training, assessment and testing cultures and educational and political reforms" (p. 23). Teachers have many responsibilities in terms of ensuring that their instructions are in line with learning procedures which is an important aspect of language assessment. While making decisions about the content of lessons, materials to be used, and learning tasks, teachers can consider the strengths and weakness of the options that are accessible to them; they may proceed according to experience, their understanding of learning, the development of language, and the proficiency of language with what they believe to be the most convenient beneficial option and as a part of their profession, teachers observe learners all the time, which gives rise to the improvement of their understanding of learners' development and their intuitions

26

about particular learning outcomes and general performance (Rea-Dickins, 2004). Here it may be considered that teachers are at the centre of the language learning procedure, similarly to be more general teaching procedures. Being in the centre of the learning procedure carries with it many responsibilities for teachers, such as preparing appropriate assessment tools, improving them, applying them, grading work, and evaluating results. Moreover, they need to understand the effects of the assessment they carry out in order to see the difference between appropriate assessments and poor quality assessment. Language teachers are expected to carry out assessments carefully and efficiently and this may be why they are seen as "agents of assessment" (Rea-Dickins, 2004).

Yastıbaş (2018) further states that language assessment training is a significant part of assessment practice. According to Yan, Zhang, and Fan (2018), "assessment includes all activities related to the development, validation, and use of language assessments for various purposes" (p. 158). Jeong (2011), on the other hand, advocates that "as more educational systems place greater assessment demands on teachers, the more support should be provided, and the first step is to provide teachers with appropriate professional development courses" (p. 19). The importance of language assessment training is emphasized by various researchers (Stiggins, 1991b, 1999; Jeong, 2013; Lam, 2015). Stiggins (1991b) stresses that for teachers to identify learners' needs, give sound grades, and evaluate the effect of remedial instruction with the help of sound assessments, they need assessment training. He (1999) states that "productive classroom assessment training promotes an understanding of, an ability to apply, and a commitment to meeting standards of valid and reliable classroom assessments" (p. 24), which is an aim for language classes and teachers to achieve. Jeong (2013) asserts that as teachers should be assessment-literate in classroom practices, they should be given proper teacher training in assessment. Vogt, Tsagari and Spanoudis (2020) suggest that when providing the LAL training with the aim of improving LAL levels of instructors, "various contextual factors, characteristics, needs and traditions" (p. 18) should be taken into consideration. Lam (2015), on the other hand, approaches this topic as follows:

> "language assessment training is not an only means to foster LAL; however, it is indispensable for equipping pre-service teachers with fundamental knowledge, skills, and principles in handling both large-scale and classroom-based assessments efficiently and mastering how to utilize these assessment outcomes to promote student learning through continuous professional learning (i.e., a community of practice)" (p. 191).

Language assessment training is given as a course for both pre-service and in-service teachers. However, in spite of its importance, Yastıbaş (2018) argues that the training being given is not satisfactory. Little time is allocated for the theory and practice of assessment in language education programs at the graduate level; even though there are many books about language testing and assessment, most of them are considered either highly theoretical or overly specific for language teachers who are trying to comprehend the fundamental principles and practices of assessment (Taylor, 2009).

According to Yastıbaş (2018), "educational theories affect language assessment and evaluation directly by creating two different cultures related to language assessment and evaluation: assessment culture and testing culture" (p. 25). Teachers' roles are obvious in the testing culture: they are powerful and they create tests (or use already prepared tests), they observe the classroom to prevent cheating, and they keep the records of the grades of learners and moreover, in this culture, there is an important rule that requires only teachers to know both the questions and answers that is of what are referred to as secure tests (Kleinsasser, 1995). In Table 2.3, both learners and teachers' roles in assessment culture are outlined.

Table 2.3. Comparing a Culture of Testing with a Culture of Assessment

Comparing a Culture of Testing with a Culture of Assessment	
Culture of Testing	**Culture of Assessment**
Roles	**Roles**
Teacher (test giver) is sole authority.	Teacher (assessor) and learner are in a conversation about learning.
Teacher's job is to keep test uncontaminated.	Student has an active role in own learning and assessment.
Teacher is a judge, evaluator, and scorekeeper in an economy of grades.	Student and teacher collaborate to build assessments, create connections.
Student is powerless, mystified by the testing process.	Students are self-appraisers, co-appraisers with teachers and other assessors.
Rules	**Rules**
Student goes it alone.	Teaming, cooperation, collaboration are valued.
Paper-pencil tests predominate.	Worthwhile assessment takes many forms: paper-pencil tests, performance-based assessments, demonstrations, exhibitions.
Teacher is the sole audience.	Assessment has multiple audiences, e.g., self, peers, teachers, parents, others.
Teacher owns the test.	
Rituals	**Rituals**
Testing episodes interrupt learning. Learning is for the test.	Distinctions between learning and assessment are blurred.
Testing is mysterious, an ordeal sometimes requiring extra sleep and nutrition, a tradition that is seldom questioned or challenge.	Worthwhile assessments are celebrations. Assessments are demystified; student knowledge about assessments and the assessment process increases.

Note. Adapted from "Assessment culture and national" by Kleinsasser, A. M., 1995, *The Clearing House*: *A Journal of Educational Strategies, Issues and Ideas,* 68(4), p.205-210.

As seen in the Table 2.3., Kleinsasser (1995) categorizes some features within the categories of culture of testing and culture of assessment revealing some important differences between these two cultures. According to this categorization, there are three main features of roles, rules, and rituals and each feature provides clear information. While there is a weak image of learners and a very strong image of teachers in testing culture, in assessment culture, this image is completely different in terms of giving learners a more powerful image and distributing responsibilities to both teachers and students (Kleinsasser, 1995). Inbar-Lourie (2008b) adds that the term assessment culture was coined as a reaction to the term testing culture and it refers to a broader and more theoretical structure for assessing knowledge: "assessment culture refers to educational evaluation practices that are compatible with current ideologies, social expectations, attitudes and values" (p. 285). In contrast to the traditions of testing culture, by seeing "external and internal assessment as dichotomous and detached, assessment cultures acknowledge the value of broadening the assessment construct by augmenting assessment from different sources and from different informers" (Inbar-Lourie, 2008b, p. 288). It is important that teachers be aware of what was done before and what is currently being done in terms of assessment and testing.

Education does not take place only to help improve students' understanding, knowledge, abilities, and attributes in quickly changing world and another important aim of education is to improve the capacities that support learning, thus allowing learners to make use of the advantages of several new forms of learning opportunities (Broadfoot, 2005). According to Brindley (2008), educational reforms have "origins in the identification of aspects of an educational system that are viewed as problematic such as:

- falling standards of students achievement
- poor performance by students in international comparisons
- lack of national standardization in assessment
- lack of comparability of outcomes across different educational systems
- lack of transparent accountability reporting
- public concerns about teacher competence" (p. 365).

Furthermore, while describing the role of testing in educational reform, several points should be taken into consideration:

- "the social and political context of reform
- the process by which it is implemented and adopted
- its effect on teaching and learning" (Brindley, 2008, p. 365).

Therefore, it might be said that educational reforms cover the whole concept of testing and assessment to make education better. Language education, on the other hand, requires that well-established specific goals be set by teachers and that those goals be realized accordingly. Since educational reforms make use of testing and assessment to realize their aims (Yastıbaş, 2018), language teachers are also responsible for setting and obtaining some goals aimed at increasing learners' success. Educational reforms may be required if there is a decrease in learners' achievements. Therefore, it is important to balance the success levels of the learners and the aims of teachers. In other words, teachers' aims should be directed toward increasing success levels.

With these different definitions and claims in mind, it can be said that instructors and other stakeholders generally try to adapt the concept of LAL without altering its the basic features so as to have both theoretical knowledge and practice in the classroom and help learners improve themselves in terms of language learning.

2.5. Assessing Productive Language Skills

For any language, there are two types of language skills: receptive language skills and productive language skills. While receptive language skills include reading and listening skills, productive language skills include writing and speaking skills. It is important to acknowledge that receptive language skills do not require learners to produce anything but rather only to understand what they read or hear. This is in contrast to productive language skills; productive language skills, which require learners to both understand a situation and produce something reasonably in line with that situation. In other words, with receptive language skills, learners receive knowledge passively and then perhaps try to produce spoken language in dialogues; the same procedure may be applied for productive language skills, which are also referred to as active language skills, but in this case, learners are expected to exert more energy while producing something either orally or in writing; but, these two categories of language skills are both indispensable parts of teaching and learning processes (Golkova & Hubackova, 2014). Golkova and Hubackova (2014) says:

> "When one starts learning a foreign language, he surely and subconsciously is exposed to both categories of language skill. As mentioned before, productive skills - also called active skills - mean the transmission of information that a language user produces in either spoken or written form" (p. 478).

According to Matthews (1990), when assessing productive skills, all responsibility is given to assessors. She adds that "this is especially the case where the assessor plays a dual or even triple role, acting as a plausible, sympathetic interlocutor and an efficient administrator, as well as assessor" (p. 117).

Assessing productive language skills may require both teachers and learners to be critical and careful about the procedure of assessment on the grounds that productive language skills may not involve only a single valid answer, as is generally the case for receptive language skills. While assessing receptive language skills, there is typically one correct answer for any question asked; however, in the assessment of productive language skills, answers may vary according to the questions, which may require learners to change their points of view or offer teachers a new perspective.

2.5.1. Assessing Speaking Skill

Speaking is one type of productive language skill and it is an indispensable part of a language since it allows people to communicate orally. In language education, this skill is one of the important skills to teach and learn. However, while assessing speaking, some problems may arise regarding reliability. Learners' speaking skills are generally evaluated in face-to-face interactions between an assessor and a speaker and moreover, the assessment may simultaneously depend on speech features such as fluency and accuracy, and the gender of the learner, and the language level and characteristics of the assessor; in addition, during these interactions, the specific tasks and questions directed to learners might also have effects on their performances (Ahmed & Alamin, 2014). According to Luoma (2004), although "assessment developers are the key players in the speaking assessment cycle, the examinees, interlocutors, raters and score users also have a role to play in the activities" (p. 4). Luoma (2004) further explains that speaking is one of the crucial skills in language teaching curricula which makes it a significant item to be assessed even though such assessments are difficult since there are various factors that affect teachers' ideas about how well a person can speak the target language, and since teachers expect results to be correct and convenient for their goals. Knight (1992) states that "we recognize the importance of relevant and reliable assessment for providing vital information to the students and teachers about the progress made and the work to be done" (p. 294). Therefore, speaking might be one of the most difficult language skills to assess (Hughes, 2003; Ginther, 2012). Correct evaluations of speaking skills might be difficult and it takes a great amount of time and effort to get valid and reliable outcomes, but that might be considered a necessity if there is

backwash or if a test is very important, and moreover, it should also be considered that the selection of the contents, the levels of the rubrics, and the techniques applied in testing speaking may all depend on the needs of institutions (Hughes, 2003).

The aim of teaching spoken language is to improve the skill of communicating in the target language successfully requiring understanding and production at the same time and moreover, at the beginning of spoken interactions in the classroom, teachers do not need any formal tests; rather informal observations will provide diagnostic information for them (Hughes, 2003). Roever and Kasper (2018), on the other hand, highlight the relationship between speaking and listening and state that it is not possible to separate speaking and listening with the understanding that listening helps speakers to produce meaningful and appropriate sentences. While assessing speaking, it is important to take listening into consideration so as to determine whether learners comprehend enough of what they hear to respond to the teacher appropriately. Brown (2003) suggests that speaking as a productive skill might be monitored directly but that such monitoring can be influenced by the accuracy and impact of learner's listening skills, which weakens the reliability and validity of speaking tests. Therefore, the relationship between speaking and listening cannot be ignored and the fact that listening has an effect on speaking should be considered at all times. It is also important to acknowledge that the effect of listening might be negative. The correlation between these skills will be either good or bad depending on a given learner's listening skill. However, in some situations learners' listening skills are not included in assessments of their speaking skills, such as in cases of monologues or reading aloud, because these activities do not require an interlocutor. This creates an environment for learners in which they do not need to use their listening skills to be able to produce something oral. Hughes (2003) states that there are three procedures for teachers in the process of assessing speaking skills. First, teachers prepare some tasks that provide a representative model of the speaking tasks that they expect learners to achieve; second, the tasks that the teachers prepared should produce behaviours that symbolize the learners' skill and finally, the scores given to those behaviours might be reliable and valid (Hughes, 2003).

The basic format of speaking in the classrooms is question and answer. Teachers ask some questions and expect students to provide some answers. Brown (2003) states that there are five different speaking types that teachers can utilize to assess learners' speaking skills: "imitative, intensive, responsive, interactive and extensive" (pp. 141-142). Imitative tasks generally focus on repetition and may involve repetition tasks that require learners to repeat

a sentence or a structure that the teacher has uttered (Brown, 2003). Scoring, according to Brown (2003), is generally based on accuracy in pronunciation and the intonation of learners. While learners repeat words or sentences, teachers generally focus on how the learners pronounce the words, and how good their intonation is. Figure 2.3 illustrates Brown's (2003) scoring scale for repetition tasks demonstrating how teachers evaluate learners' speech.

Figure 2.3. Scoring scale for repetition tasks

2	Acceptable pronunciation
1	Comprehensible, partially correct pronunciation
0	Silence, seriously incorrect pronunciation

Note. Adopted from "*Language Assessment: Principles and Classroom Practises*" by Brown, D.H., 2003, p. 145. Pearson ESL.

As seen in Figure 2.3., there are three potential scores when considering how learners pronounce words or sentences. It is important for learners to be understandable while speaking. The intensive type of speaking, on the other hand, involves short sentence structures (Brown, 2003). According to Brown (2003), in this case, learners are supposed to produce only one short sentence to demonstrate their linguistic abilities. The intensive type of speaking encompasses tasks that include reading aloud, sentence or dialogue completion, and picture-cued tasks, and what is important in the assessment of this type of speaking is comprehension and whether learners are able utilize the targeted grammar or vocabulary correctly (Brown, 2003). The responsive type of speaking requires interaction between an interlocutor and a speaker (Brown, 2003). In this case, tasks that teachers can apply to assess learners' speaking ability include question and answer, and paraphrasing tasks. In contrast to the intensive type of speaking that requires the production of only one sentence, learners may produce more than one sentence, but the amount will still generally be limited (Brown, 2003). The fourth type of speaking identified by Brown (2003) is interactive speaking, also called "interpersonal" (p. 167). This type of speaking requires learners to speak longer. Assessment tasks for interactive speaking consist of interviews, role-plays, games, and discussions (Brown, 2003). The final type of speaking is extensive speaking; it requires learners to be able to produce more complex sentences and the assessment tasks include presentations, story-telling with pictures, retelling stories, and translation (Brown, 2003).

Thornbury (2005) explains that "testing, both informally and formally, takes place at the beginning and at the end of most language courses, as well as at various times during the course itself" (p. 124). Thornbury (2005) also mentions some types of spoken tests, such as "interviews, live monologues, recorded monologues, role-plays and collaborative tasks and discussions" (pp. 125-126) which are similar to the tasks described by Brown (2003). These types of tests may be used at the beginning of the term to understand what learners know and before the learning process takes place, and what they are able to do can be observed. They may also be used at the end of the term to assess what learners have learned and the extent to which they can reflect their learnings orally. These types of tests may also be used in the classroom throughout the term to observe learners. Thus, teachers can apply at least one of these test types to assess learners' speaking performances. While assessing speaking, teachers evaluate specific criteria such as accuracy, fluency, grammar and vocabulary. In the assessment of speaking, there are two types of scoring: "holistic" and "analytic". Holistic scoring can be considered as less detailed but more time-saving. It is used to assess speaking skill as a whole and it is generally applied when the evaluation needs to be done immediately; however, it may not provide learners with feedback and it is generally used for individual improvement (Ulker, 2017; Aleksandrzak, 2011). Analytic scoring, on the other hand, can be considered as more time-consuming and more detailed. Analytic scoring gives information about what a learner has accomplished and it can also provide feedback for both learners and teachers thus learners can identify their weaknesses in the target language and teachers can gain opportunities for remedial teaching (Mukminatien, 2015). Backlund, Brown, Gurry, and Jandt (1982), on the other hand, suggest that:

- Assessment of speaking should be realized in a natural environment instead of a planned environment.
- Assessment of speaking should also include observations by teachers.
- The assessment tool should be a scale that includes criteria measuring how learners use the language, how they organize their ideas, how they create utterances to respond in a relevant way, and how they use their voices (intonation, pitch, accent, etc.). The assessment should be carried out by teachers who are trained in using, grading, and explaining the scale.
- While the tool is being improved, assessment ought to be confirmed.
- A confirmation study should be carried out when the tool is first improved.

With this information in mind, it can be said that assessing speaking may be difficult since there might be problems of reliability and validity issues as speaking is a relatively subjective skill. While teachers try to assess learners' speaking skills, they need to be very careful and consider what elements can affect the assessment environment. Different speaking assessment tasks for different purposes should be conducted in the classroom. Anything that may compromise reliability and validity should be eliminated.

2.5.2. Assessing Writing Skill

Writing is another productive language skill. Students are expected to convey their messages or what they know about a topic in writing. Writing has gained popularity in terms of providing communication throughout the whole world and it is one of the most important components of a language since it renders communication possible even when oral interactions are impossible. Therefore, it is vital to teach and learn this productive language skill since it helps both learners and teachers communicate around the world.

There are three different types of writing which are classified as "academic writing, job-related writing, and personal writing" (Brown, 2003, p. 219). Academic writing consists of "essays, compositions, academically focused journals, theses and dissertations" (Brown, 2003, p. 219). According to Brown (2003), while job-related writing includes "letters, e-mails, schedules, advertisements, and annuals", personal writing includes "letters, e-mails, messages, diaries, personal journals, and fictions" (p. 219). According to Fahimi and Rahimi (2015), writing, among other language skills, has drawn attention owing to the communication needs arising in different places in the world. They further state that writing should be taught early and that skill should be assessed and marked. Brown (2003) explains that there are four types of writing performance including "imitative, intensive (controlled), responsive, and extensive" (p. 220). These types of writing performances help teachers determine which assessment type is most appropriate for learners. While imitative tasks, which generally focus on assessing learners' skills in producing written symbols in a language, include "copying, listening cloze selection tasks, picture-cued tasks, form completion tasks, and converting numbers and abbreviations to words" (Brown, 2003, pp. 222-223), intensive (controlled) tasks mainly include "dictation and grammatical transformation tasks" (pp. 225-226). Responsive and extensive tasks, on the other hand, include "paraphrasing, guided question and answer, and paragraph construction tasks" (Brown, 2003, pp. 234-235). As can be understood clearly from the assessment types that each performance task has, imitative writing generally focuses on how students are able to

produce letters and symbols writing. Intensive (controlled) writing focuses on grammar and correct usage of structures that have been taught. Responsive and extensive writings, on the other hand, focus on more intense types of writing where learners are supposed to create longer sentences and even paragraphs.

It is vital for teachers to devote time to writing tasks in classrooms. However, it may sometimes be hard for teachers to carry out classroom writing tasks because of time limitations. As a result, written tasks may often be ignored. Dempsey, PytlikZillig and Bruning (2009) claim that even though writing is one of the best known skills in the world, it is often ignored since it requires a complex and varied set of abilities and the learning of those abilities requires teachers who are qualified and have the self-confidence to support learners' improvement toward literacy while giving accurate instruction and guiding those learners. Writing tasks may seem easy to apply, but teachers need to give clear instructions and determine what to assess and how to assess it. This also requires that teachers be more careful and have an ability to include writing tasks in the classroom with well-prepared instructions and the integration of different skills. Javed, Juan and Nazli (2013), on the other hand, claim that writing achievement is used for various goals and that ensuring guidance for learners both inside and outside the classroom, giving grades, matching students and tasks appropriately, and evaluating programs are regarded as significant points in the assessment of writing. While assessing writing, first, teachers should determine the goal and the assessment type that is convenient for the learners. The goal clarifies which criteria teachers should apply in a writing task. Bachman and Palmer (1996) explain that goal setting means making decisions about what to do and involves three mean features:

- determining the type of the task(s), whether language use tasks or test tasks
- selecting one or more tasks from a group of probable tasks; and
- making a decision about accomplishing the task(s) or not.

Together with the types of writing and writing performances (Brown, 2003), setting a goal has become an indispensable part of this procedure. Teachers, after deciding which types of writing and performance would be better for learners, should determine the goal, which is an important part of assessment. However, some problems might occur throughout the writing assessment procedure. According to Zheng and Yu (2019), problems related to writing assessments have altered and become more complex, or perhaps the complex problems "…behind writing assessment have finally become evident, after a century in which they have been concealed by special interests and technical problems" (p. 1). White

(1994) also supports the idea that the problems occurring in the assessment of writing have altered and become more complex. Moreover, White (1994) asserts that teachers grow disappointed and angry when assessment, which they may consider an unnecessary political tool, is very time-consuming and encroaches upon their teaching time. People occasionally consider writing as only a part of teaching and learning grammar and structure which eventually minimizes the significance of writing skills and hampers their improvement (Fareed, Ashraf & Bilal, 2016). However, writing not only includes grammatical rules and structures; it includes vocabulary, punctuation, meaning, and organization as well.

In the assessment of writing, two types of approaches are prominent: direct and indirect assessment (Breland, 1983; Stiggins, 1982). Direct assessment in writing is associated with directly observing learners' writing procedures and it is generally evaluated by English teachers who are trained in making evaluations about writing skills while indirect assessment in writing is associated with indirectly observing learners' writing procedures and generally includes multiple-choice questions, which are intended to comprehend learners' writing skills through grammar and sentence structures (Breland, 1983). Breland (1983) describes some of the direct assessment types for writing tasks as "letters, narrative, descriptive, argumentative, expressive, role-playing, and diary entry" (pp. 2-3). Holistic and analytic scoring can also be applied to evaluate written tasks. Stiggins (1982) accounts for this by clarifying the differences between these types of assessments and in the context of writing assessment. He states that even though these two types differ in form and provide different types of information for written tasks, they may both be very beneficial in terms of providing necessary information for making decisions in educational contexts such as instructional management and selection. According to Stiggins (1982), direct assessment requires learners to actually write in response to a given prompt, and in a later stage, the composition skills of the learners are evaluated by certain performance-based criteria while indirect assessment, on the other hand, offers learners multiple-choice questions, which lead teachers to understand whether learners can write appropriately or not through grammar and sentence structure; actual writing is not needed.

Writing may be assessed by both FA and SA. FA has gained more popularity in English language classrooms recently since SA is believed to be insufficient owing to the fact that it fails to utilise the formative capability of assessments to improve learning; on the other hand, formative approach focuses on analysis and the opportunities that assessment offers to advance both teaching and learning (Guadu & Boersma, 2018). Lee (2011),

explains that FA might take place directly in writing classrooms or through peer evaluation, feedback through portfolios, and conferencing. The chief goal of FA is to guide and advance learners' learning by giving them information relevant to the gap between their present performance and the aimed performance (Wingate, 2010). To close that gap, feedback should be offered in the classroom, but the feedback given by teachers might only be as good as the information that has been established thus, the quality of the prompt that a teacher utilizes to decide the actual situations of learners is a crucial part of a successful feedback circle (Furtak & Ruiz-Primo, 2008). Feedback, as an indispensable part of FA, can provide many opportunities for both learners and teachers as long as it takes place within teaching and learning procedures. One of these opportunities is remedial teaching. Teachers might make use of FA to identify their learners' exact stage of learning advancement in writing. Therefore, it can be easier for both learners and teachers to detect problematic subjects and, if possible, remedial teaching might take place.

Hughes (2003) suggests that there are several formative writing tasks that teachers can implement in classrooms, such as portfolios, diaries, teachers' observations, projects, and quizzes. Apart from formal tests applied to evaluate learners' writing skills, informal tests or tasks can be applied regularly to see how learners deal with the language. While it is important to set tasks, scoring is also an important part of this process. Once again, two approaches influence scoring: holistic scoring and analytic scoring. The first thing that teachers should do is set tasks that can be reliably scored. According to the tasks, one of the two approaches can be chosen for scoring. Holistic scoring is a fast, qualitative process for either sorting or scoring the writing tasks of learners and it is not constructed to adjust or alter the products or to identify deficiencies; instead, it is a procedure intended to give a value to a writing task according to predetermined criteria (Charney, 1984). It is hard to develop a holistic rubric but this facilitates easier and quicker assessments since it entails the assignment of a single score for the whole written product (Beyreli & Ari, 2009). Moreover, rubrics for holistic scoring are generally short, and do not contain particularized criteria for evaluation, and enable a written task to be scored after reading it once therefore, it serves the universities well in the context of cost/benefit analyses; moreover, these rubrics are commonly applied for assessing written performance in major assessment contexts, which makes holistic scoring popular for written performance in tests such as the Test of English as a Foreign Language (Ghalib & Al-Hattami, 2015). Analytic scoring, on the other hand, is believed to provide more detailed information about learners' performances on written tasks

(Ahour & Mukundan, 2009) and an analytic scoring rubric generally contains some writing elements relevant to learners' verbal, grammatical, and discourse competence (Ghalib & Al-Hattami, 2015).

With this information in mind, it can be said that assessing writing might be difficult since it is a subjective skill and may give rise to problems of reliability and validity. Teachers should be careful while choosing tasks to meet learners' needs and find a correct type of scoring rubric. Moreover, anything that might be an obstacle to reliability and validity should be eliminated.

2.6. Studies Related to LAL Conducted in Turkey and Throughout the World

The concept of LAL is gaining popularity among researchers; "…researchers have turned their attention to this concept having realized the significance of effective and efficient testing and assessment practices for better and quality teaching and learning" (Şahin, 2019, p. 26). Several studies about LAL have been conducted in Turkey (Çetin-Argün, 2020; Hatipoğlu, 2015; Kaya, 2020; Ölmezer-Öztürk, 2018; Sarıyıldız, 2018; Sevilen-Yılmaz, 2021; Şahin, 2019; Tamerer, 2019; Yastıbaş, 2018).

In her master's thesis, Çetin-Argün (2020) aimed to explore the language assessment knowledge of 154 participating pre-service English teachers. Her results demonstrated that pre-service English teachers have average levels of knowledge about LAL and require more detailed and inclusive training about it. Upon analysing skill-based, it was found that reading and speaking skills were above average while writing and listening skills were below average. In another study, Hatipoğlu (2015) investigated what pre-service EFL teachers in English Language Teaching (ELT) departments knew about language testing and what they thought their ELTE course should include with regard to content. The study was conducted with 124 pre-service English teachers and the data were obtained with the help of a questionnaire and interviews. The findings of her research were similar to results to those of Çetin-Argün (2020) in that the pre-service English teachers had limited knowledge about testing and assessment both in general and in English.

In her master's thesis, Kaya (2020) investigated the LAL of EFL instructors in language preparatory programs and examined ELF instructors' viewpoints on language assessment and its reflections in their teaching practice. She conducted the study with 195 EFL instructors in different universities in Turkey and collected the data with the Language

Assessment Knowledge Scale (LAKS) to measure the instructors' language assessment knowledge levels. She also conducted interviews with 17 instructors to determine their perceptions of language assessment. According to the findings of the questionnaire, the EFL instructors' language assessment knowledge was higher than average and they were able to apply their knowledge on their practices. Another finding of her study also revealed that factors including demographic features did not have an influence on the instructors' language assessment knowledge.

In another study, Ölmezer-Öztürk (2018) aimed to determine the language assessment knowledge of EFL instructors working in higher education in Turkey. In order to see the general picture of the situation of LAL among EFL instructors in Turkey, she developed a scale including different items for each language skill (in total 60 items- 15 per). The LAKS was completed by 542 EFL instructors, and for the qualitative part of the study, 11 instructors answered the questions in detail to provide in-depth data. The quantitative data, each item was scored either 1 or 0 thus, the highest score that an instructor could obtain was 60 while the lowest was 0. Each correct answer was scored as 1 while "Don't Know" was always scored 0. When her quantitative results were analysed, it was found that the general language assessment knowledge of the instructors was below average. Moreover, the instructors were more competent with reading skills and less competent with listening skills. The qualitative part of the study revealed that the education was not sufficient and the participants needed further training.

In her master's thesis, Sarıyıldız (2018) investigated LAL levels of pre-service EFL teachers in Turkey, trying identify their perceptions of the theoretical and practical training for assessment that they had in their teacher education programs and whether they needed more training in language assessment or not. Moreover, she aimed to determine how English language teachers evaluated school experience course and English Language Testing and Evaluation course (ELTEC). She used a mixed research design and carried out the study with 101 fourth year pre-service teachers in the ELT department of Middle East Technical University (METU). She conducted interviews with 25 participants; quantitative data were gathered with the Language Assessment Literacy Survey, and qualitative data were gathered with semi-structured interviews. According to the results of this study, the pre-service EFL teachers had received training but still needed more basic training. Second, according to qualitative data, the participants could not apply their theoretical knowledge of language assessment and testing in real-life practice and neither the theory nor practice of assessment

was encompassed in the school experience course. Finally, the results revealed that participants were knowledgeable about the importance and contributions of ELTE course. In her master's thesis, Sevilen-Yılmaz (2021) focused on examining and comparing the LAL levels of EFL instructors working in preparatory schools in both foundation and state universities. 336 EFL instructors were asked to participate in the study and fill out the questionnaire with the purpose of examining and comparing their LAL levels. In addition, 10 preparatory students were interviewed to investigate their perceptions of assessment. The findings of this study revealed that the language assessment knowledge of the instructors was generally high, but the language assessment knowledge of the instructors working in state schools was higher than that of instructors working in foundation schools. The qualitative results, on the other hand, revealed that preparatory school students that participated in the interviews generally knew about the assessment criteria and they were aware of the importance of assessment together with its impact on their achievements.

In her PhD dissertation, Şahin (2019) investigated the current situation of the ELTEC in English Language Teaching Education Programs while addressing LAL for potential English teachers. She applied a convergent parallel mixed methods approach and data were obtained with the help of various data collection instruments. The results revealed that one Language Testing and Assessment (LTA) course was not enough to impart the necessary knowledge, including theoretical and practical knowledge together with skills, for the improvement of LAL. Moreover, the findings of her study highlighted that the emphasis of the considered ELTEC was mainly on theoretical aspects rather than practical aspects and that pre-service English teachers described their LAL training in the ELTEC was as satisfactory, yet nevertheless demonstrated less success with regard to FA and the assessment of productive language skills.

In another study, Tamerer (2019) aimed to investigate the LAL levels and training needs of pre-service EFL teachers in her master's thesis. She applied a questionnaire to 30 pre-service EFL teachers and semi-structured interviews to 10 participants and found that the LAL levels of the participants were low and they needed training in this specific area. In his dissertation, Yastıbaş (2018) aimed to explain how LAL is applied in language classes; the study was carried out with eight Turkish EFL instructors. Yastıbaş used a multiple-case study research design and collected the data with the help of different data collection instruments such as interviews, observations, document analysis, think-aloud protocols and focus group discussions. His research revealed that participants had a criticial attitude

towards both assessment and evaluation. It also revealed that the pre-service training was efficient in terms of assessment and evaluation practices of those who graduated from faculty of education while this training was found inadequate of those who graduated from different faculties.

When research conducted throughout the world is taken into consideration, there are numerous studies about LAL (Jannati, 2015; Hakim, 2015; Mellati & Khademi, 2018; Sultana, 2019; Vogt & Tsagari, 2014). Jannati (2015), for example, investigated Iranian ELT teachers' perceptions and practices related to assessment. She conducted the study with 18 Iranian ELT teachers and divided them into three different categories of low, medium and high according to their teaching experience. Interviews were conducted and data were processed with content analysis. Her findings demonstrated that the teachers were familiar with the concepts and there was no significant relationship between teaching experience and the ways that they perceived assessment. However, she found that even though the teachers were assessment-literate, that literacy was not demonstrated in their practices.

In another study, Hakim (2015) aimed to explore instructors' assessment awareness levels and how they applied assessment tools for better learning of learners in Saudi Arabia. A questionnaire was administered to 30 EFL instructors and it was found that all participants had knowledge about how to use assessment tools. Moreover, although instructors' perceptions of assessment tools were least influenced by duration of teaching experience, instructors with more years of teaching experience were found to have clearer understandings of the components and tools of assessment.

Mellati and Khademi (2018) tried to evaluate instructors' levels of AL and its effect on their current assessment practices and students' writing outcomes. To do so, they applied various data collection instruments. The research was carried out with 10 male EFL instructors and 75 male college students selected from four language schools in Iran. The findings revealed that instructors' AL levels had an important influence on learners' writing success.

In another study, Sultana (2019) focused on determining whether English teachers in Bangladesh were prepared to carry out a variety of tasks academically and professionally and how these teachers perceived LAL in their teaching. To answer these questions, the researcher conducted semi-structured interviews with 10 teachers. According to the findings, the teachers' knowledge was insufficient and they generally utilised traditional tests.

Accordingly, it was deduced that teachers' assessment knowledge was inadequate in this Bangladeshi context.

Vogt and Tsagari (2014) investigated the current level of LTA literacy among foreign language teachers and tried to identify their relevant training needs. They made use of a mixed methods approach administering a questionnaires to 853 teachers and conducting interviews with 63 teachers in selected European countries. Their findings revealed that 26.3% of the teachers stated that they had no training and 39.9% of the teachers stated that they had a little training while 33.8% of the teachers stated that they had advanced training. Moreover, 47.7% of the teachers stated that they needed more advanced training in the field of LTA.

When these studies are taken into consideration, it can be concluded that EFL instructors need training about LAL to compensate their lack in this field. Moreover, it can be seen that while some studies on this topic have been conducted, studies on the LAL levels of EFL instructors in higher education context especially in preparatory schools, where learners are prepared to pass a language proficiency test including all four language skills, regarding the assessment of productive language skills are still limited in the Turkish context. Therefore, the present study was designed to shed light on LAL levels of EFL instructors in assessing productive language skills as well as their perceptions of assessing productive language skills through the application of FA.

2.7. Studies on formative assessment conducted in Turkey and Throughout the World

Despite the significance of FA and its critical role in language learning and teaching, there is a lack of research on it in the Turkish context. While some studies have investigated FA more generally (Arda-Özkan, 2020; Büyükkarcı, 2010; Gökçe, 2014; Ökten, 2009; Sarı, 2019; Sönmez, 2015; Solgun-Günel, 2014; Uysal-Kurtulmuş, 2018); only two studies have been identified that investigated EFL instructors' perception of FA (Gökçe, 2014; Arda-Özkan, 2020). The rest of the aforementioned works mostly investigated the effects of FA on learners.

Arda-Özkan (2020) conducted a study to determine EFL instructors' attitudes and practices regarding FA in English preparatory schools of universities in İstanbul. With the participation of 144 EFL instructors working in both public and foundation universities, the researcher found that all participating teachers acknowledged the importance of FA for

improving learner success and involvement in educational processes. These teachers claimed that "they have inspirational mentalities, view and objectives with respect to FA (p.92)". Despite their practical intentions, however, when the issue was the implementation of these assessment techniques in the classroom, not all teachers were certain about them. Some participating teachers reported that they could not take advantage of FA in their practices, but most stated that they utilized FA in the classroom while teaching and assessing. Furthermore teachers from both foundation and public universities had positive thoughts about FA and no significant difference was found between these two groups. There are three noteworthy conclusions that can be drawn from the results of Arda-Özkan's study (2020). The first regards the implementation of FA; even though most of the EFL teachers made use of FA in their classrooms, there were still some who did not. The second regards the EFL teachers' changing attitudes and practices regarding FA based on factors such as age, undergraduate department, and professional development activities. The third regards the teachers' levels of FA implementation in relation to their undergraduate departments as. EFL teachers graduating from ELT departments were more inclined to implement FA than EFL teachers graduating from English Language and Literature (ELL) departments.

In another study, Büyükkarcı (2010) aimed to identify the effects of FA on learners' test anxiety and their preferences for assessment. The participants were chosen from ELT departments. The researcher applied a mixed methods approach with one experimental group (n= 38) and one control group (n= 48). The Test Anxiety Inventory was administered to the participants and the findings demonstrated that FA implementation decreased the experimental group's anxiety levels while the control group's anxiety levels remained the same.

In one of the theses on in-service EFL instructors' perceptions of FA, Gökçe (2014) investigated whether there was a significant difference between the FA perceptions of EFL teachers working in public and private schools in the Turkish context. Questionnaires were administered to 100 EFL teachers from 39 different schools including 22 private and 17 public schools in Turkey to determine their perceptions of and attitudes toward assessment. The researcher found that most of the participants understood the concept of assessment in the same way; however, EFL teachers working in private and public schools were dissimilar in terms of their classroom practices.

In her master's thesis, Ökten (2009) investigated the effects of the implementation of FA on learners' motivation and awareness of their strengths and weaknesses together with their beliefs about language learning and assessment. She also aimed to determine the effects of the implementation of FA on learners' proficiency, particularly for lower-achieving learners. The research was designed as a case study and conducted with 10 learners using both quantitative and qualitative data collection tools. According to the results of the qualitative research conducted with the help of interviews and reflection papers, it was found that most of the learners were only evaluated with written exams and they were not very pleased about it. With the implementation of FA, learners indicated that their motivation increased and they gained awareness of their strengths and weaknesses. It was also found that the success level of learners increased after the implementation of FA.

In another study, Sarı (2019) investigated to the extent to which EFL students' anxiety about writing skills altered with the application of FA. The study was conducted with 30 sophomore students as an action research study. The data were collected with the help of an evaluation process survey, interviews, and field notes. The findings revealed that the implementation of FA decreased the anxiety levels of the participating EFL students in terms of writing skills. Similarly, Sönmez (2015) examined the effects of FA on the autonomy of Turkish EFL students in an English preparatory school at a specific university. The research was conducted as a case study and FA was implemented in the course of research. As data collection tools, Sönmez (2015) utilized quantitative and qualitative methods. She administered the Autonomy Learner Questionnaire and Assessment Preference Scale with pre-tests and post-tests. It was found that FA had a positive effect on Turkish EFL learners' autonomy. In another study, Solgun-Günel (2014) investigated the effects of FA on learners' participation in an English language program. Semi-structured interviews, a teacher journal, teacher notes, and weekly reflection sheets were used to obtain the data. The findings revealed that FA had positive effects on learners' involvement in the classrooms.

Uysal-Kurtulmuş (2018) explored perceptions of FA among Turkish and international adult EFL learners in English classes with regard to gender and nationality at a foundation university in Turkey. It was additionally aimed to investigate the ways in which those perceptions influenced these learners' language learning processes and how instructors perceived these learners' perceptions. According to the findings of the study, learners held a positive perception of FA but a few differences were observed with regard to nationality and gender, which affected their language learning processes to some extent. The study also

revealed that differences between the instructors' perceptions and learners' perceptions of FA and its effect on learners' learning which might be linked to the lack of knowledge about FA.

Turning to studies conducted throughout the world, are taken into consideration, some additional research on FA has been conducted (Dixson & Worrel, 2016; Wiliam, 2011; Xiao & Yang, 2019). Dixson and Worrel (2016) explained the importance of FA in their study, stating that FA should be implemented in the course of instruction and throughout the whole learning process to help to learners learn the material, while SA could be implemented at the end of a chapter or a unit to evaluate how much learning has occurred. Finally, they highlighted that instructors should be aware of the aims of their assessment and how they will use the outcomes of the assessment results to select the best instruments to implement for accomplishing their aims. In another study, Wiliam (2011) reviewed recent improvements and attempts to understand FA and AfL in different ways, particularly with regard to classroom strategies and techniques that instructors can use to develop the quality of the outcomes of their instructional plans. He stated that even though there are many components that still need to be developed further, there is strong theoretical work suggesting that integrating assessment into teaching might increase learner engagement and develop learning outcomes. In another study, Xiao and Yang (2019) investigated how FA can promote the self-regulated learning of secondary-learners' in the process of learning English. They conducted their study with 16 learners and 2 instructors in a foreign language secondary school and revealed that, with the help of the guidance of their instructors, the participants engaged in FA and behaved like self-regulated learners. The participants perceived the FA activities implemented in the classroom and the feedback that they obtained as being helpful for their improvement in understanding and self-regulation processes in English language learning.

When the studies presented here are considered, it can be concluded that few studies have specifically focused on EFL instructors' perceptions of FA. Moreover, the studies conducted on this topic to date have largely involved learners, and the number of the studies investigating the FA perceptions of EFL instructors' in higher education setting in Turkey regarding the assessment productive language skills is limited. Therefore, the present study aims to shed light upon EFL instructors' perceptions of assessing productive language skills through FA.

CHAPTER 3

3. METHODOLOGY

In this chapter, the research design, the participants, the research context of the study, the data collection tool construction process and the pilot studies, the data collection instruments, and the data analysis procedures are presented and discussed in detail.

3.1. The Research Design

This study has been designed as explanatory sequential mixed methods research study (Creswell, 2014). Mixed methods study consists of both quantitative and qualitative data in a study and while qualitative data are generally open-ended and do not require a preparation for answers, quantitative data generally contain close-ended answers which can be discovered in a questionnaire (Creswell, 2014). The logic behind the mixed methods study is based on the fact that quantitative and qualitative methods, alone, are not enough to understand the trends and elaboration of a condition (Ivankova, Creswell, & Stick, 2006). In other words, mixing both of the methods allows researchers to utilise the features of both methods and gain an understanding of the situation better. Explanatory sequential mixed methods design, on the other hand, is one of the types of mixed methods (Creswell, 2014). This type of mixed method requires researchers to implement the quantitative research and analyse the results first and use the results to explain them in detail with qualitative research (Creswell, 2014).

Since the designed study was explanatory mixed method research study, first, the quantitative data collection instrument was conducted (See Figure 3.1.). To gather quantitative data, there are several different instruments such as questionnaires and experiments. According to Dörnyei (2003), questionnaires are efficient in terms of time, effort and, financially making it possible to reach as many participants as possible to gather data in a short time. Therefore, to obtain more quantitative data in a short time, a questionnaire as a quantitative research instrument was preferred. As a quantitative data collection tool, a questionnaire was decided and it was adapted from Ölmezer-Öztürk's (2018) "Language Assessment Knowledge Questionnaire" to examine the level of LAL of EFL instructors in assessing productive language skills.

Qualitative research, on the other hand, includes several different data collection tools such as interview, observation and asks researchers to involve in the study for long-term so as to have a natural situation and in this type of research, the researcher listens, observes, or asks some questions to people (Stainback & Stainback, 1989). Qualitative data provide researchers with great details and context and reflects the contextual setting and what lies there (Libarkin & Kurdziel, 2002). Thus, as a second data collection tool, semi-structured interviews designed by the researcher herself were conducted and carried out with EFL instructors to elaborate the results and to collect the qualitative data to uncover EFL instructors' perspectives on formative assessment in assessing productive language skills.

Stages of Explanatory Sequential Mixed Methods Design

Figure 3.1. Three Basic Mixed Methods Design
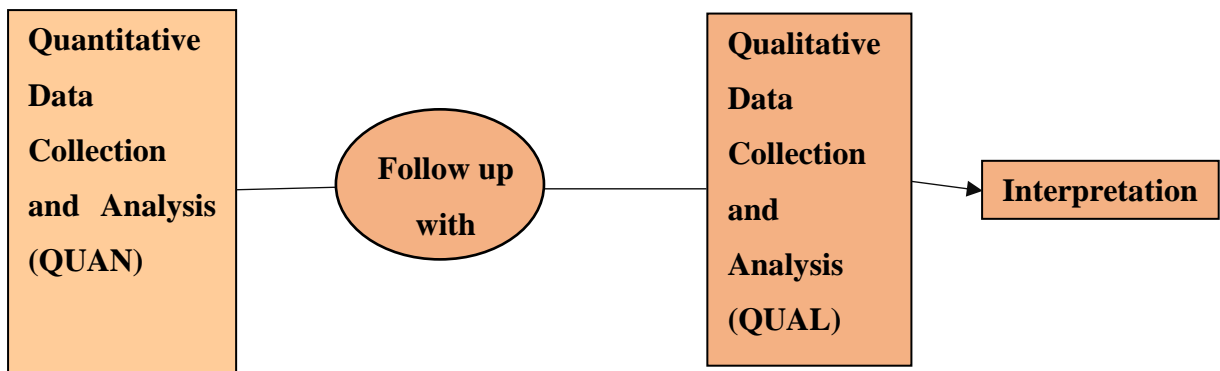
Note. Adapted from "*Research Design: Qualitative, Quantitative and Mixed Methods Approaches (4th Edition)*" by Creswell, J. W., 2014, p.220. SAGE Publications.

3.2. The Participants

The data for this study were collected from only one group: English language instructors working at preparatory schools in private and state universities in Ankara. The EFL instructors working at preparatory schools at the universities in Ankara were requested to fill out the questionnaire several times in three months during 2020-2021 academic year.

In total, 60 EFL instructors from 9 different universities (49 female and 11 male) participated in the study. The age of the participants ranged from 24 to 60 years old, with a mean age of 34.76 and had between 1 week to 31 years of teaching experience. The

participants graduated from different departments such as ELT (n.24), ELL (n.21), ACL (n.7), TI (n.5), Linguistics (n.2) and Others (Biology, n.1).

Table 3.1. Demographic information of the EFL instructors

| Name of the University | Gender of the Participant | | | |
	Female	Male	Total	
Ankara University	2	1	3	State
Atılım University	10	2	12	Private
Başkent University	16	2	18	Private
Gazi University	2	1	3	State
Hacettepe University	5	0	5	State
Middle East Technical University	2	0	2	State
Ostim Technical University	5	1	6	Private
TED University	5	3	8	Private
University of Turkish Aeronautical Association	2	1	3	Private
Total	**49**	**11**	**60**	

In the second part of the study, a semi-structured interview was carried out with the EFL instructors. The main aim of the interview was to reveal what instructors thought about formative assessment in assessing productive language skills. Among the EFL instructors who filled out the questionnaire, 11 EFL instructors agreed for an interview (9 female, 2 male). These instructors were chosen by looking at the number of correct answers that they supplied (high, average and low). The age of the participants ranged from 24 to 55 with a mean age of 32,54. Each interview lasted about 25 minutes, thus the researcher completed the whole interview session in 5 hours 19 minutes. The interviewees were informed about the general purpose of the study and that anonymity would be maintained. Moreover, their consents were taken to record the whole interview.

The data collection procedure was carried out online because of the COVID-19 pandemic. The researcher utilised Google forms to collect the quantitative data and arranged Zoom meetings to conduct interviews.

3.3. Data Collection Instruments

This study was a mixed research study, and the data were gathered through two different data collection instruments: Language Assessment Knowledge Questionnaire (LAKQ) and semi-structured interviews.

3.3.1. Language Assessment Knowledge Questionnaire (LAKQ)

As this study was an explanatory sequential mixed study, first, the questionnaire was administered to the EFL instructors. Language Assessment Knowledge Questionnaire was adapted from the questionnaire that Ölmezer-Öztürk (2018) used in her study and before implementing the questionnaire, the necessary permission was taken (See Appendix 1). The questionnaire that the researcher used in her study consisted of 60 items with 0.91 reliability. For each language skill, there were 15 questions. However, since the focus of this study was on productive language skills, reading and listening skills parts were removed from the questionnaire. At first, in total, there were 30 questions, 15 questions for each language skill. However, the study was too specific, therefore the researcher added more questions for each skill in line with the research questions to obtain more valid and reliable outcomes. First, the researcher added 17 questions to items for assessing writing skill part and then she added 25 questions to items for assessing speaking skill part. The original questionnaire includes Likert-scale items and has three different sections that each aims different purposes.

The first part of the questionnaire was demographic information part (See Appendix 2) which has 23 questions aiming to reveal EFL instructors' background information such as age, gender, experience etc., and their educational background. The second part of the questionnaire included items for assessing writing skill part which has 32 questions aiming to reveal how much EFL instructors are literate about assessing writing skills of the learners and the last part of the questionnaire included items for assessing speaking skill which has 40 questions aiming to reveal how much EFL instructors are literate about assessing learners' speaking skills. The statements had only one correct answer. Therefore, "Don't Know" and "Not Sure" answers were not assumed as correct. The in-service EFL teachers were asked to tick "True", "False", "Don't Know" and "Not Sure" for the statements related to these two productive language skills.

3.3.2. Semi-Structured Interviews

Semi-structured interviews with EFL instructors working at preparatory schools in the universities in Turkey, Ankara were carried out and recorded. Before recording the interview, permission was taken from each participant. To prepare the interview questions, the researcher took Hughes (2003) and Brown (2003)'s books as examples and constituted the questions according to their books. The semi-structured interview data depicted the qualitative part of this particular study. Semi-structured interview provides an opportunity to interviewees to elaborate on the issues asked by the interviewer; namely, the interviewer guides and leads together with follow up questions and allows interviewees to elaborate on issues (Dörnyei, 2007).

The interviews were carried out in English. Each interview was recorded, and they lasted approximately 26.36 minutes. During the interviews, the participants were asked questions about their educational background, their work and field experiences, how they implemented assessment in language classes, how they implemented formative assessment, their opinions on formative assessment, the problems and challenges they encountered when assessing productive language skills, and if they had possible solutions for the stated problem (s) (see Appendix 3 for the interview questions).

3.4. Data Collection Procedure

In the process of a research study, it is crucial to evaluate the data collection instruments before implementing the actual study. Especially, in social sciences, this evaluation carries great importance since human beings are the focal points in those studies. To evaluate the data collection instruments, pilot studies are carried out. Pilot study might be defined as a study which is carried out as pre-testing of a specific data collection tool such as a questionnaire and interview (Van Teijlingen & Hundley, 2002). The aim of the pilot studies is to figure out whether there are any problems in the questionnaires or interviews before the participants carry out them and the researcher collects the main data. Moreover, one of the advantages of administering a pilot study is that it may allow researchers to see where the actual study may be unsuccessful and if the chosen instruments and methods are improper or very complex (Van Teijlingen & Hundley, 2002). On the other hand, pilot studies try to detect and clarify any ambiguous parts which are presented in the data collection tools (Şahin, 2011). In the contrary case, the desired answers may not be obtained

from the instruments. Therefore, the aim of the first pilot study was to obtain information about the length, format, and the wording of the questions and if there were any ambiguous statements in the questionnaire. With this aim in mind, the researcher first developed a Likert-scale questionnaire to depict EFL instructors' LAL level in assessing productive language skills.

3.4.1. Constructing the LAKQ before the pilot study

The four stages followed after deciding the statements which would take place in LAKQ are listed below:

Stage 1: First, the necessary permission to implement the questionnaire was taken from Ölmezer-Öztürk (See Appendix 1). Before the LAKQ was administered to the pilot group, the questionnaire was specialised in line with the aim of the study. Normally, there were 60 statements (15 statements for each skill), however, since this study only included productive language skills, only 30 statements were included (writing and speaking skills statements).

Stage 2: Before the LAKQ was administered to the pilot group, a demographic part was added by the researcher (see Table 3.2.) and since 15 statements for each skill might not provide sufficient and satisfactory results, some statements for each skill were added by being utilized from different researchers (i.e. Backlund, Brown, Gurry & Jandt, 1982; Hughes, 2003; Luoma, 2004; Stiggins, 1982; Weigle, 2002) by the researcher herself.

Stage 3: Before the LAKQ was administered to the pilot group, one expert opinion was asked to evaluate the statements in the questionnaire to ensure that they were not ambiguous and unclear. A Turkish-English bilingual (who has a PhD in language assessment and testing and works on testing and assessment and has 12 years of experience teaching to non-native speakers of English) evaluated the statements in terms of the language use and clarity. Based on the feedback received from the expert, some questions were added to the demographic part, some wording and order changes were made in the statements and some of the statements were removed since they did not contribute to the whole aim of the study.

Table 3.2. An example of a question included in the demographic part of the questionnaire before the pilot study

Demographic Part of the Questionnaire
Age:

Table 3.3. An example for the changes made on the questions asked in the demographic part of the questionnaire before the pilot study

Demographic Part of the Questionnaire
7. How long have you been teaching English? (former version)

0-5 years	☐
6-10 years	☐
11-15 years	☐
16-20 years	☐
21-25 years	☐
More than 25 years	☐
(Former version).	

7. How long have you been teaching English? (Latter version)

This change was made (See Table 3.3.), based on the feedback received, since the range of the years would not be necessary for the study and the mean of the question could not be possible to be calculated. However, what was important for the study was to obtain the actual year/years. Therefore, it would be easier to see the exact year of experience and calculating the mean of the question was possible.

Table 3.4. An example for statement included in the questionnaire before the pilot study

Items for Assessing Writing	True	False	Don't Know
19. Writing exams that only include language use is enough to understand learners' writing ability.	☐	☐	☐

Table 3.5. An example for statement included in the questionnaire before the pilot study

Items for Assessing Speaking	True	False	Don't Know
30. Speaking skill can be assessed through an indirect test such as close tests.	☐	☐	☐

Table 3.6. An example for the changes made on the statements before the pilot study

Items for Assessing Writing	True	False	Don't Know
24. Feedback may lack the skill of providing learners with opportunities to close the gap between their current and desired performance (Former version).	☐	☐	☐
24. Feedback is necessary to provide learners with opportunities to close the gap between their present and desired performances (Latter version).	☐	☐	☐

Table 3.7. An example for the changes made on the statements before the pilot study

Items for Assessing Speaking	True	False	Don't Know
9. It is significant if interviewers or raters are not trained (Former version).	☐	☐	☐
9. Whether the interviewers or raters are trained for the speaking exams or not does not affect the reliability of the scoring. (Latter version).	☐	☐	☐

As it is seen clearly, some new statements were included to obtain a detailed answer for each skill and some necessary changes were made on the wording and the order to sustain more clear and unambiguous statements (See Table 3.4., 3.5., 3.6., 3.7. respectively). With the additional statements, for items for assessing writing skill, there were 32 statements while for items for assessing speaking skill, there were 40 statements. In total, there were 72 statements and when the questions in the demographic part were added, there were 81 questions and 1 e-mail request before it was implemented to the pilot group (see Appendix 4).

3.4.2. Piloting LAKQ with the EFL Instructors

After finalizing the procedure of the questionnaire, it was sent to 10 English language instructors working at preparatory schools in three private universities through an online platform for the pilot study. First, the researcher received permission from each participant to record each session on Zoom. Then, each participant was asked to fill out the questionnaire. When they finished the questionnaire, the researchers obtained instructors' ideas and opinions about the statements, wording, format, length, and the questionnaire as a whole unit by asking some questions. The questionnaire took approximately 20 minutes of the participants.

As the questionnaire consisted of 3 main categories (demographic part, items for assessing writing part and items for assessing speaking part), the outcomes of the pilot study were analysed under these 3 categories.

3.4.2.1. Demographic Part

After the participants submitted the questionnaire, the researcher asked some questions and obtained the participants ideas. The first question the researcher asked was if the participant had any questions or any suggestions for demographic part of the questionnaire. Most of the answers were positive and moreover, they stated that the questions were all necessary and logical. Some of the participants suggested some ideas about the questions. The questions were reorganized and some new questions were added in line with the needs and suggestions indicated by the participants (See Table 3.8. and 3.9. below to see the changes in the questions in the demographic part of the questionnaire).

> ***Instructor 2:*** "In the demographic part, the education part which asks if we have taken courses related to testing and assessment in MA and PhD, since I did not take it, I left it blank. You can reorganize it. You can give an instruction in parentheses like "If you haven't taken any courses, write this". (…) Moreover, since I have not graduated yet (MA), I was indecisive about writing the name of the university and the department. Maybe you can reorganize it like "graduate from/ are you currently studying". (…) Other than that, the rest of the questions were necessary for you. I think they were logical".

Table 3.8. A question change in the demographic part made after the pilot study

Demographic Part
6. Which university and department did you graduate from (MA)? (Former version)
6. Which university and department did you graduate from/are you currently studying (MA)? (If you do not hold an MA degree or if you are not studying, please leave it blank). (Latter version).
8. Which university and department did you graduate from (PhD)? (Former version)
8. Which university and department did you graduate from/are you currently studying (PhD)? (If you do not hold a PhD degree or if you are not studying, please leave it blank). (Latter version).

Table 3.9. Questions in the demographic part added after the pilot study

Demographic Part
14. Have you received any courses about testing and assessment English or testing and assessment in general through formation, CELTA and/or DELTA?
15. If yes, how many courses have you received?

3.4.2.2. Items for Assessing Writing Part

The researcher asked the participants if there were any ambiguous or unclear sentences in this part. Most of the participants stated that terminology was difficult but other than that the statements were necessary and helped them realize how important assessment was. Some stated that a few questions were unclear, created ambiguity and including some strong adjectives which led the participants into choosing wrong answer (See Table 3.10. below to see the statement including strong adjective in the items for assessing writing part of the questionnaire). (See Appendix 6 for the changes in the statements).

> ***Instructor 2:*** "In the writing part, I only noted technical stuff. However, there was a question including "best way" and it made me think on it more. (…) I realized that after that question, some of the questions included "best" while some of them did not. As I told you, that term was stuck in my head because when you say "best", it becomes something different. (…) For example, you say "can be assessed" for some of the questions, yes I approach it positively however, for that question, even

though I approached positively, I picked False since it includes "best way". I don't know how I can say it, and I don't know if you have some certain lines in your study and in your questionnaire or not, if you don't, you can alter it "a good way" or "one of the good ways". If you do it like that, I definitely choose True because you know, every techniques has its pros and cons. (…) So, you can form a standardization according to what you want in your questions. Either make them include "best" or make them include "one of the good ways". But I don't know if you especially assess that like if it is the best or not".

Table 3.10. Statement including strong adjective

Items for Assessing Writing	True	False	Don't Know
7. Classroom assessment of learning in terms of writing is best served through analytic scoring rather than holistic scoring (original version).	☐	☐	☐

3.4.2.3. Items for Assessing Speaking Part

The researcher asked the participants if there were any ambiguous or unclear sentences in items for assessing speaking part. Most of the participants stated that this part was easier than the other part. Some stated that a few questions were unclear, created ambiguity and including some strong adjectives which led the participants into choosing wrong answer (See Table 3.11. below to see the changes in the statements in the items for assessing speaking part of the questionnaire).

Instructor 10: "About the speaking part, there is item 4. Here, it was about interlocutor, right? So, interlocutor is there and the student is there. It means this, right? I just want to be clear. It is ended at the time the task doesn't continue because the student couldn't give the correct answer but there are some parts they should be going on and you stop it. It is now clear when you explain it because I was thinking that it could be something like this there are some specific answers and students couldn't reach this level and then you just end the task. 12th question… There is "contrived". I don't know. I wasn't used to that term. I did not know. Other than that, yes, it was nice. The questions were clear I guess and there were only 3 options and that was nice. I don't remember the numbers of the questions. (…) I felt writing

questions are more because I took so much time but the speaking was better. I think it was nice, there was no problem. Thank you for your effort, that was well-designed (As a suggestion) Maybe the other teachers who are away from the field maybe the literature graduates, they could have some problems with "contrived" for example. But other than that, it was nice. We know that how to assess speaking and writing and we could answer the questions".

Table 3.11. Statement change in the items for assessing speaking part made after the pilot study

Items for Assessing Speaking Skill	True	False	Don't Know
11. The results of the speaking assessment should be communicated to learners and parents (former version).	☐	☐	☐
11. The score of the speaking assessment with feedback should be communicated to learners to provide them with feedback (latter version).	☐	☐	☐
15. Feedback should be adapted according to the audience (former version).	☐	☐	☐
15. Feedback should be adapted according to the learners (latter version).	☐	☐	☐
26. Oral presentation, when it is recorded, may decrease the stress level of learners. (Former version)	☐	☐	☐
26. Oral presentation, when it is recorded during the lesson, may decrease the stress level of learners (Latter version).	☐	☐	☐

In the pilot study, in total, there were 10 instructors working at three different private universities in Ankara. There were 2 male and 8 female instructors. According to the feedback received by each instructor, necessary changes were made on the statements in the questionnaire and added some questions. Through the addition of some questions, in demographic part, 23 questions were constituted, in items for assessing writing part, 32 questions were constituted and in items for assessing speaking part, 40 questions were constituted. In total, there were 94 questions and 1 e-mail request. The final version of the questionnaire can be seen in Appendix 5.

3.4.3. Constructing the semi-structured interview questions for the EFL instructors

Four stages that were followed after deciding the questions which would take place in the semi-structured questionnaire are listed below:

Stage 1: First, while preparing the questions, Brown (2003) and Hughes (2003)' books were taken as example and depending on the research questions of this thesis, some certain questions were added. The researcher added demographic questions first to reveal EFL instructors' background information such as age, gender, experience etc., their educational background in general and their educational background related to language testing and assessment and formative assessment.

Stage 2: Before the semi-structured interview was administered, two expert opinions were asked to evaluate the questions in the semi-structured interview to ensure that they were not ambiguous and unclear and whether they achieved the aim of this thesis or not. A Turkish-English bilingual (who has a PhD in language assessment and testing and works on testing and assessment and has 12 years of experience teaching to non-native speakers of English) and a Turkish-English bilingual (who studies on language assessment and testing and has a 3 years of experience teaching to non-native speakers of English) evaluated the questions in terms of language use, clarity, and comprehensibility. Based on the feedback received from the experts, some questions were added to eight categories of the interview, some wording and order changes were made in the questions and some of the questions were removed since they did not contribute to the whole aim of the study (See Table 3.12. for examples for included questions in different categories and see Table 3.13. for examples for changes in the questions in different categories).

Table 3.12. Examples for question included in the interview before implementing

1st Category: Educational Background

- Did you take any courses about testing and assessment in English or any other courses related testing and assessment? **If yes, when and how many?**

2nd Category: Work and Field Experience

- Have you worked in a testing unit before? If yes, how long have you worked? **And in what kind of assessment practices have you involved in?**

3rd Category: Application of Assessment in Language Classes

- **How do you assess your students' speaking skills in English?**
- **What are your assessment criteria in assessing students' speaking skills?**
- **How do you assess your students' writing skills in English?**
- **What are your assessment criteria in assessing students' writing skills?**

4th Category: Application of Formative Assessment

- **What types of materials and in-class activities do you implement to teach and practice writing skills?**

5th Category: Opinions about Formative Assessment

- **What is the place of "providing feedback" in the assessment procedure?**
- **Do you provide feedback for your students speaking and writing productions during in-class activities? If yes how?**

8th Category: Others

- **Do you have any other comments or suggestions to add?**

Table 3.13. An example for a change made on the question in the interview before implementing

4th Category: Application of Formative Assessment

- What types of materials do you use in the classroom to allow speaking to take place? (former version)

- What types of materials and in-class activities do you implement and practice speaking skills? (Final version)

As it is seen clearly, according to the feedback received from the experts, some new questions were included to obtain a detailed answer for each skill and some necessary

changes were made on the wording and the order to sustain more clear and unambiguous questions in the interview. With the additional questions, in total, there were 30 questions, which consist of sub-questions as well, and 8 categories (See Appendix 3).

3.5.   Data Analysis Procedure

This thesis has been designed as explanatory sequential mixed methods research. More specifically, two different data were obtained from the EFL instructors at different times. First, the researcher implemented the questionnaire to obtain quantitative data. The data obtained from the questionnaire were analysed through SPSS (Statistical Package for the Social Sciences). Second, the researcher implemented semi-structured interviews to obtain qualitative data. The data obtained from the semi-structured interviews were analysed with the help of content analysis.

3.5.1. Quantitative Data Analysis

As this research has been designed as explanatory sequential mixed method research, two different data were analysed differently. First, the researcher implemented questionnaire to obtain quantitative data from EFL instructors. The obtained data were analysed through SPSS by running descriptive analysis. In the questionnaire, there are three different parts; demographic part, items for assessing writing skill part and items for assessing speaking skill part. Participants' demographic information such as gender and educational background and their LAL level were analysed through descriptive statistics. The aim of the questionnaire was to examine the level of LAL of EFL instructors in assessing productive language skills. For this purpose, the LAKQ questionnaire designed by Ölmezer-Öztürk (2018) was adapted. Two parts (reading and listening skills) from the scale were removed since they were not related to the main aim of this study; however, some certain statements were added later for writing and speaking skill parts. 32 items were determined for assessing writing skill part while 40 items for determined for assessing speaking skill part after an elaborated process. The items were designed in "true, false, don't know" format but after the pilot study, the format changed and "not sure" choice was added.

3.5.2. Qualitative Data Analysis

A fine qualitative research utilizes an organized and rigid approach targeting to reply questions that deal with people's thoughts and feelings about an event, and moreover, it might focus on the reason why an event has happened the way it has (Seers, 2012). According to Seers (2012), qualitative data generally focus on texts and may contain images as well. Qualitative data analysis includes a process of alteration of the qualitative data into more clear, comprehensible, perceptive and reliable analysis through some certain analytic procedures (Gibbs, 2007). For this study, the qualitative data were gathered through semi-structured interviews and the obtained data were analysed by content analysis. With this in mind, qualitative data analysis of this study (semi-structured interviews) was carried out using qualitative content analysis merging the steps suggested by Creswell (2014, p.197-200) as demonstrated in Figure 3.2.

Figure 3.2. Data Analysis in Qualitative Research
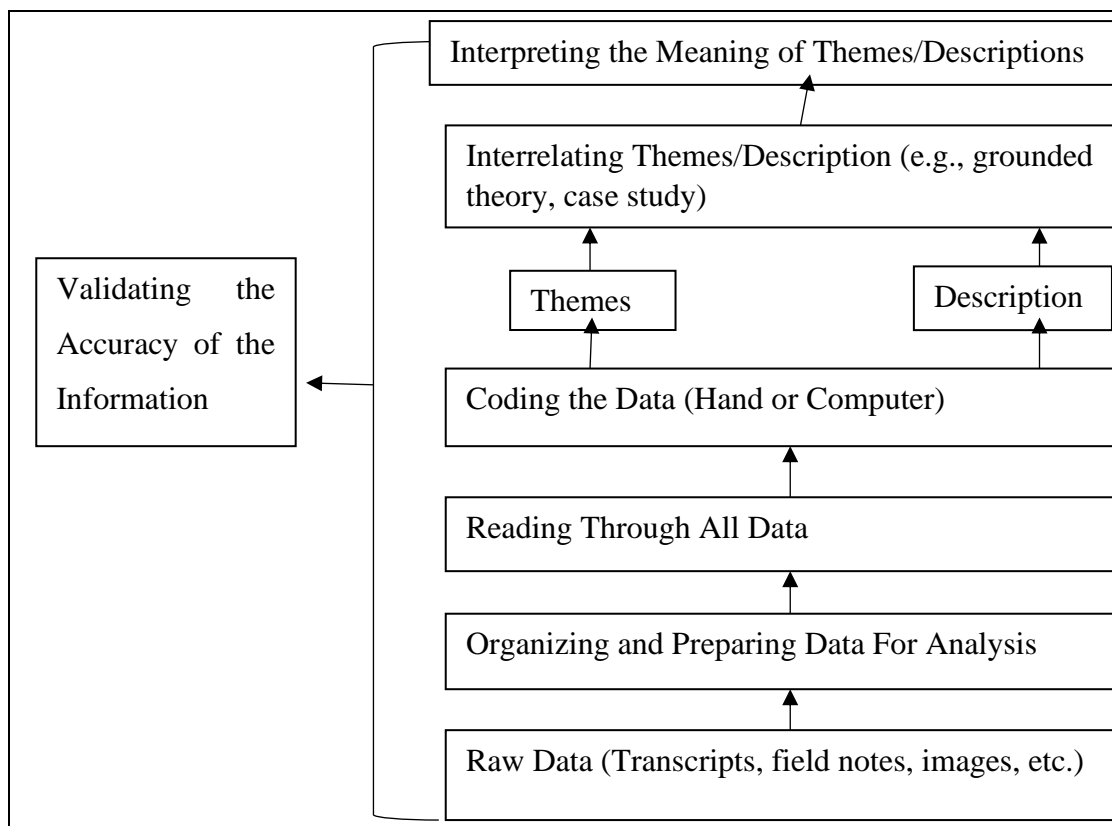
Note. Adapted from *"Research Design: Qualitative, Quantitative and Mixed Methods Approaches (4th Edition)"* by Creswell, J. W., 2014, p. 197. SAGE Publications.

The qualitative data analysis of this research compromised these steps:

The Step 1 was "Organize and prepare the data for analysis" (Creswell, 2014, p.197). First, the researcher transcribed the interviews recorded via Zoom after each interview and the responses were written. Each participant tried to contribute to the research.

The Step 2 was "Read or look at all the data" (Creswell, 2014, p.197). After finalizing transcribing the raw data, the researcher watched each video and checked transcriptions so as not to lose a datum. Thereupon, all of the answers given for each interview question were read three times before coding the data. After the third reading of the dataset, in the fourth cycle, with regard to the research questions of the study, the researcher jotted down some ideas and notes in the process.

The Step 3 was "Start coding all of the data" (Creswell, 2014, p.197). In this step, the researcher coded the data with a word or sentence representing each category in the responses. While the researcher analysed the interview data, she tried to combine a general premise coding based on the interview questions. The answers of the participants were analysed in line with these coded categories. In order to establish a valid and reliable coding, the researcher went over the obtained data, data analysis and the interpretation of the results. After finalizing the codes and categories, the researcher analysed the structured interviews one by one and tried to create a reasonable information patterns from these categories and codes. Then the researcher looked at the frequency of the codes to expand the interpretations that encompass the consistency and find out the components underlying the research questions that were formulated for this research. So as to ensure the reliability of the results, data analyses were checked by one coder (i.e. an academic specialized in LTA). In the end, the presentation of the categories and the codes, tables and figures were developed for a detailed discussion.

CHAPTER 4

4. FINDINGS AND DISCUSSIONS

The findings of this study together with discussion are presented under each research question of this study. The research questions and the data collection tools that the researcher utilized to answer each are given below in Table 4.1.

Table 4.1. The Research Questions and the data collection tools used to explore the answers

Research Questions	Data Collection Tools
1- What is the EFL instructors' level of language assessment literacy in assessing students' productive language skills in English?	Language Assessment Knowledge Questionnaire (LAKQ)
2- What are the EFL instructors' perceptions on assessing productive language skills through formative assessment?	Semi-Structured Interview (Questions – Part 5; 1, 2, 2.1, 2.1.1, 3, 4)
3- How do the EFL instructors implement formative assessment tools in their classrooms?	Semi-Structured Interview (Questions- Part 3; 1, 2, 3, 4, 5, 6, 7, 8 – Part 4; 1, 2, 3, 3.1)
4- What problems and challenges do the EFL instructors encounter while assessing learners' productive language skills in English?	Semi-Structured Interview (Questions Part 6; 1, 2)
5- What do the ELF instructors suggest as solutions for the stated problem/s?	Semi-Structured Interview (Questions – Part 7; 1)

4.1. Research Question 1: What is the EFL instructors' level of language assessment literacy in assessing students' productive language skills in English?

The first research question of this study is "What is the EFL instructors' level of language assessment literacy in assessing students' productive language skills in English?" In order to be able to answer this question, descriptive analysis was implemented. The questionnaire was filled out by 60 EFL instructors working in preparatory schools in Ankara and the findings that were derived from their answers are presented in tables below respectively and the bold ones refer to the correct answers.

Table 4.2. The Frequency of the Answers provided by the EFL instructors for Items for Assessing Writing Part: Category Scoring

CATEGORY 1: SCORING		n	%
1. Giving two options to learners and asking them to write about one ensure reliable and valid scoring.	True	32	53,3%
	False	**8**	**13,3%**
	Don't Know	8	13,3%
	Not Sure	12	20,0%
2. Analytic scoring is used to see the strengths and weaknesses of learners.	**True**	**49**	**81,7%**
	False	2	3,3%
	Don't Know	7	11,7%
	Not Sure	2	3,3%
3. The parts of a scoring scale and the scores in each part do not change for different levels of learners.	True	12	20%
	False	**40**	**66,7%**
	Don't Know	6	10%
	Not Sure	2	3,3%
4. When there is a huge difference between the scores of the two raters, the same raters score the written work again.	True	24	40,0%
	False	**30**	**50,0%**
	Don't Know	2	3,3%
	Not Sure	4	6,7%
5. Holistic scoring is used to see whether the learner is proficient or not at the end of the term.	**True**	**36**	**60,0%**
	False	8	13,3%
	Don't Know	9	15%
	Not Sure	7	11,7%
6. Analytic scoring leads to greater reliability than holistic scoring in writing.	**True**	**25**	**41,7%**
	False	14	23,3%
	Don't Know	12	20,0%
	Not Sure	9	15,0%
8. Mechanical errors (e.g. spelling and punctuation) are dealt with in the assessment of later stages of a written work.	**True**	**25**	**41,7%**
	False	28	46,7%
	Don't Know	5	8,3%
	Not Sure	2	3,3%

9. While scoring a writing test, the expectation of raters does not have an influence.	True	16	26,7%
	False	**40**	**66,7%**
	Don't Know	2	3,3%
	Not Sure	2	3,3%
10. The complex relation among writing task variables, rater variables and scoring criteria does not have any effects on the scoring the writing tests.	True	5	8,3%
	False	**47**	**78,3%**
	Don't Know	6	10,0%
	Not Sure	2	3,3%
11. While planning a scoring rubric for writing, factors including who will use the rubric and what aspects of writing are crucial are not taken into consideration.	True	8	13,3%
	False	**49**	**81,7%**
	Don't Know	2	3,3%
	Not Sure	1	1,7%
15. Irrelevant ideas are ignored in the assessment of initial stages of a written work in process writing.	True	17	28,3%
	False	**33**	**55,0%**
	Don't Know	4	6,7%
	Not Sure	6	10,0%

Items for assessing writing part were divided into categories. The first category is "Scoring" and there are 11 questions (See Table 4.2.). In the 1st statement, it was seen that the instructors could not differentiate the items that are valid for valid scoring because only 8 (13.3%) EFL instructors answered it correctly which also demonstrated that 52 (86.6%) EFL instructors could not choose the correct option. In the 2nd statement, it was found out that the awareness of the EFL instructors of knowing that the analytic scoring is implemented to see the learners' weaknesses and strengths was high with a percentage of 81.7%. In the 3rd item, it was found out that the EFL instructors' awareness of the fact that the parts of the scoring scale and scores in each part can change according to learners at different levels was average with a percentage of 66.7%. In the 4th statement, it was found out that only 50% of the EFL instructors knew that when it comes to score a written work, if there is a disagreement between the raters, there needs to be another rater to score the written work in terms of providing reliability and validity. In the 5th statements, it was found out that while 60% of the instructors knew that holistic scoring is implemented to see if learners are proficient at the end of the term or not, 40% of the instructors did not know. In the 6th statement, it was found out that only 41.7% of the EFL instructors were aware of the difference between analytic and holistic scoring in terms of providing reliability which was higher in analytic scoring; however, the majority of the instructors (58.3%) were not aware of this difference. In the 8th statement, it was found out that only 41.7% of the EFL instructors knew that mechanical parts are taken into consideration at the later stages of assessment of

a written work while the majority of the instructors (58.3%) did not know this. In the 9th statement, it was found out that 66.7% of the EFL instructors knew that when assessing a written work, the expectation of the rater has an effect. In the 10th statement, it was found out that 78.3% of the EFL instructors knew that the complex relation among writing task variables, rater variables and scoring criteria have an effect on scoring. In the 11th statement, it was found out that 81.7% EFL instructors, which is a high percentage, knew that when planning a scoring rubric for writing, there are some factors that should be taken into consideration such as the person who will use the rubric and what aspects of writing are crucial. In the 15th statement, it was found out that 55% of the EFL instructors knew that in process writing, irrelevant ideas are not ignored in the assessment of initial stages of a written work.

Table 4.3. The Frequency of the Answers provided by the EFL instructors for Items for Assessing Writing Part: Category Formative Assessment

CATEGORY 2: FORMATIVE ASSESSMENT		**n**	**%**
7. Classroom assessment of learning in terms of writing is carried out in a more reliable way through analytic scoring rather than holistic scoring.	**True**	**31**	**51,7%**
	False	12	20,0%
	Don't Know	11	18,3%
	Not Sure	6	10,0%
16. In the assessment of writing, the assessment method should be chosen according to instructional decisions.	**True**	**44**	**73,3%**
	False	8	13,3%
	Don't Know	5	8,3%
	Not Sure	3	5,0%
17. Utilising the results of writing assessment, while making decisions about learners and teaching plans, carries great importance.	**True**	**56**	**93,3%**
	False	2	3,3%
	Don't Know	1	1,7%
	Not Sure	1	1,7%
24. Feedback is necessary to provide learners with opportunities to close the gap between their present and desired performances.	**True**	**59**	**98,3%**
	False	1	1,7%
	Don't Know	0	0%
	Not Sure	0	0%
25. Providing feedback after carrying out the writing tasks not only helps teachers collect information about learners but also helps them carry out remedial teaching if necessary.	**True**	**58**	**96,7%**
	False	2	3,3%
	Don't Know	0	0%
	Not Sure	0	0%
30. Portfolio is considered as one of the good instruments for writing assessment.	**True**	**55**	**91,7%**
	False	1	1,7%
	Don't Know	2	3,3%
	Not Sure	2	3,3%

31. Self-assessment is a required tool to assess learners' writing skills.	**True**	**46**	**76,7%**
	False	9	15,0%
	Don't Know	0	0,0%
	Not Sure	5	8,3%

For this category, which is about formative assessment, there were 7 statements in total (See Table 4.3.). In the 7th item, it was found out that 51.7% of the EFL instructors knew that analytic scoring is a more reliable way than holistic scoring in terms of classroom assessment of learning with regards to writing implementation. In the 16th statement, it was found out that 73.3% of the EFL instructors knew that when assessing writing, the assessment method should be opted for according to instructional decisions. In the 17th statement, it was found out that 93.3% of the EFL instructors, which is a high percentage, knew that it is important to utilise the results of the writing assessment when making decisions about learners and teaching plans. In the 24th statement, it was found out that 98.3% of the EFL instructors, which is a high percentage, knew that it is necessary to provide learners with feedback to help them close the gap between their present and desired performances. In the 25th statement, it was found out that 96.7% of the EFL instructors knew that after implementing writing tasks, providing feedback help teachers collect information about learners and do remedial teaching if necessary. In the 30th statement, it was found out that 91.7% of the EFL instructors knew that portfolio is one of the good instruments for writing assessment. In the 31st statement, it was found out that 76.7% of the EFL instructors knew that self-assessment is a necessary instrument to assess learners' writing skills.

Table 4.4. The Frequency of the Answers provided by the EFL instructors for Items for Assessing Writing Part: Category Reliability and Validity

CATEGORY 3: RELIABILITY AND VALIDITY

		n	**%**
12. Learners are required to write about at least 2 tasks in the exam rather than 1 task.	**True**	**8**	**13,3%**
	False	33	55,0%
	Don't Know	12	20,0%
	Not Sure	7	11,7%
13. Giving restrictive prompt/guidelines for the writing task is avoided.	True	23	38,3%
	False	**32**	**53,3%**
	Don't Know	4	6,7%
	Not Sure	1	1,7%
18. Valid and reliable assessment of writing can be best ensured when it is integrated with other language skills.	**True**	**40**	**66,7%**
	False	13	21,7%
	Don't Know	3	5,0%
	Not Sure	4	6,7%
19. Writing exams that only include language use is enough to understand learners' writing ability.	True	3	5,0%
	False	**54**	**90%**
	Don't Know	2	3,3%
	Not Sure	1	1,7%
20. Clarity in task demand is a required factor in the writing tasks for learners.	**True**	**58**	**96,7%**
	False	1	1,7%
	Don't Know	1	1,7%
	Not Sure	0	0%
21. Writing tasks should be prepared as much authentic as possible.	**True**	**48**	**80%**
	False	8	13,3%
	Don't Know	1	1,7%
	Not Sure	3	5,0%
22. Validity may not be sustained if writing tasks measure learners' creativity and imagination.	**True**	**25**	**41,7%**
	False	26	43,3%
	Don't Know	4	6,7%
	Not Sure	5	8,3%
23. In writing tasks, the prompts that guide learners about what to do in the task should include full sentences so that they can understand how to carry out the task.	True	37	61,7%
	False	**13**	**21,7%**
	Don't Know	3	5,0%
	Not Sure	7	11,7%
26. Giving learners an opinion or a topic and asking them to discuss that opinion is a valid way of assessing their writing skills.	True	37	61,7%
	False	**13**	**21,7%**
	Don't Know	2	3,3%
	Not Sure	8	13,3%

In this category, which is about reliability and validity, there were 9 questions (See Table 4.4). In the 12th statement, it was found out that 86.7% of the EFL instructors did not

know that giving 2 writing tasks in exam is necessary to provide reliability and validity while only 13.3% of the EFL instructors chose the correct answer. In the 13th statement, it was found out that 53.3% of the EFL instructors knew that giving limited prompts or guidelines in writing tasks is necessary to provide reliability and validity. In the 18th statement, it was found out that 66.7% of the EFL instructors knew that integrating other language skills to assess writing can be a good way. In the 19th statement, it was found out that 90% of the EFL instructors knew that language use is not enough to comprehend learners' writing skills in writing exams when reliability and validity is taken into consideration. In the 20th statement, it was found out that 96.7% of the EFL instructors knew that clarity in writing tasks is important with regards to reliability and validity. In the 21st statement, it was found out that 80% of the EFL instructors knew that writing tasks should be authentic. In the 22nd statement, it was found out that only 41.7% of the EFL instructors knew that writing tasks do not provide validity if they assess creativity and imagination of the learners while a majority of the instructors (58.3%) did not know it. In the 23rd statement, it was found out that only 21.7% of the EFL instructors knew that writing tasks should not include prompts with full sentences that guide learners about what to do in the tasks in terms of providing reliability and validity; however, a majority of the instructors (78.4%) did not know it. In the 26th statement, it was found out that only 21.7% of the EFL instructors knew that giving an opinion or a topic to learners and asking them to discuss it is not a valid way of assessing their writing skills; however a majority of the instructors (78.4%) did not know it.

Table 4.5. The Frequency of the Answers provided by the EFL instructors for Items for Assessing Writing Part: Category Assessment Tools

CATEGORY 4: ASSESSMENT TOOLS		n	%
14. In controlled writing, learners have the chance to convey new information.	True	20	33,3%
	False	**29**	**48,3%**
	Don't Know	5	8,3%
	Not Sure	6	10,0%
27. Using visuals which guide learners for writing poses a problem.	True	8	13,3%
	False	**34**	**56,7%**
	Don't Know	7	11,7%
	Not Sure	11	18,3%
28. Writing skill might be assessed indirectly through multiple choice questions.	**True**	**23**	**38,3%**
	False	30	50,0%
	Don't Know	3	5,0%
	Not Sure	4	6,7%
29. Providing a reading text is a way of assessing writing skills.	**True**	**30**	**50,0%**
	False	17	28,3%
	Don't Know	6	10,0%
	Not Sure	7	11,7%
32. Non-interactive writing tasks (e.g. changing verbs in a paragraph from present to past tense) require metacognitive skills (e.g. understanding the goal, what they need to do to complete the task etc.) while interactive tasks (e.g. brainstorming) do necessarily require.	True	17	28,3%
	False	**26**	**43,3%**
	Don't Know	7	11,7%
	Not Sure	10	16,7%

When "assessment tools category" is taken into consideration, it is seen that there are 5 statements (See Table 4.5.). In the 14th statement, it was found out that only 48.3% of the EFL instructors knew that learners do not provide new information in controlled writing while 51.6% of the instructors did not know it. In the 27th statement, it was found out that 56.7% of the EFL instructors knew that making use of visuals can help learners for writing while 43.3% of the instructors did not know it. In the 28th statement, it was found out that only 38.3% of the instructors knew that multiple choice questions are indirect assessment tools for writing assessment however, a majority of the instructors (61.7%) were not aware of this fact. In the 29th statement, it was found out that 50% of the EFL instructors knew that a reading text can be used as a writing assessment tool. In the 32nd statement, it was found out that only 43.3% of the instructors knew the difference between non-interactive tasks and interactive tasks while 56.7% of the instructors did not know it.

Table 4.6. The Frequency of the Answers provided by the EFL instructors for Items for Assessing Speaking Part: Category Reliability and Validity

CATEGORY 1: RELIABILITY AND VALIDITY

		n	%
1. When the interlocutor does not understand the learner, giving that feeling or saying it poses a problem.	True	42	70,0%
	False	**12**	**20,0%**
	Don't Know	6	10,0%
	Not Sure	0	0,0%
2. Giving learners one task is enough to assess speaking skills.	True	1	1,7%
	False	**53**	**88,3%**
	Don't Know	1	1,7%
	Not Sure	5	8,3%
3. Interlocutors' showing interest by verbal and nonverbal signals poses a problem.	True	23	38,3%
	False	**28**	**46,7%**
	Don't Know	3	5,0%
	Not Sure	6	10,0%
5. In interlocutor-learner interviews, the teacher has the chance to adapt the questions being asked.	**True**	**24**	**40%**
	False	26	43,3%
	Don't Know	4	6,7%
	Not Sure	6	10,0%
8. Carrying out the interview in a quiet room with an acoustic may not help learners focus on the questions.	True	21	35%
	False	**23**	**38,3%**
	Don't Know	5	8,3%
	Not Sure	11	18,3%
9. Whether the interviewers or raters are trained for the speaking exams or not does not affect the reliability of the scoring.	True	5	8,3%
	False	**54**	**90%**
	Don't Know	0	0%
	Not Sure	1	1,7%
10. In productive performance, the oral stimulus does not have to be particular so that learners can give diverse answers.	True	28	46,7%
	False	**18**	**30,0%**
	Don't Know	7	11,7%
	Not Sure	7	11,7%
12. Speaking assessment of learners should take place in naturalistic rather than planned situations.	**True**	**22**	**36,7%**
	False	22	36,7%
	Don't Know	5	8,3%
	Not Sure	11	18,3%
13. While the instrument for assessing speaking is being improved, assessment should be verified in terms of providing inter-rater reliability.	**True**	**44**	**73,3%**
	False	6	10,0%
	Don't Know	5	8,3%
	Not Sure	5	8,3%
14. A confirmation study is not necessary to be conducted when the instrument is first being developed.	True	6	10,0%
	False	**39**	**65,0%**
	Don't Know	8	13,3%
	Not Sure	7	11,7%

33. The interlocutor gives the score when the learner is in the exam room.	True	11	18,3%
	False	**43**	**71,7%**
	Don't Know	2	3,3%
	Not Sure	4	6,7%
36. Only teachers who have been trained in practising, scoring, and interpreting the rating scale should be responsible for the assessment of speaking skills.	**True**	**37**	**61,7%**
	False	15	25,0%
	Don't Know	2	3,3%
	Not Sure	6	10,0%

Items for assessing speaking part were divided into categories. The first category is "Reliability and Validity" and according to this category, there are 12 questions (See Table 4.6.). In the 1st statement, it was found out that only 20% of the EFL instructors knew that giving a feeling or saying that they do not understand the learner does not pose a problem with regards to reliability and validity in speaking assessment while majority of the instructors (80%) did not know it. When the 2nd statement was analysed, it was found out that 88.3% of the EFL instructors knew that giving one speaking task is not enough in terms of maintaining reliability and validity. In the 3rd statement, it was found out that only 46.7% of the EFL instructors knew that showing interest verbally or non-verbally does not pose a problem in terms of reliability and validity of speaking assessment. When the 5th statement was analysed, it was found out that 40% of the EFL instructors knew that they have a chance to adapt the questions in the interview while 60% of the instructors did not know this. In the 8th statement, it was found out that only 38.3% of the EFL instructors knew that interview in a quiet room with an acoustic might help learners focus on the questions while majority of the instructors (61.6%) did not know it. When the 9th statement was analysed, it was found out that 90% of the EFL instructors knew that whether or not interviewers or raters are trained for speaking exams affect the reliability of the scoring. In the 10th statement, it was found out that only 30% of the instructors knew that oral stimulus should be particular for learners while 70% of the instructors did not know the answer. In the 12th statement, it was found out that only 36.7% of the EFL instructors knew that assessment of speaking should take place in natural situations rather than planned situations while 63.3% of the instructors did not know it. When the 13th statement was analysed, it was found out that 73.3% of the EFL instructors knew that when the instrument for speaking assessment is improved, assessment should be verified in terms of maintaining inter-rater reliability. In the 14th statement, it was found out that 65% of the EFL instructors knew that confirmation study is necessary to be implemented when the instrument is being developed. In the 33rd statement,

it was found out that 71.7% of the EFL instructors knew that score should be given after the learners leave the exam room. In the 36th statement, it was found out that 61.7% of the EFL instructors knew that only the teachers trained in practicing, scoring and interpreting the rating scale should be responsible for the speaking assessment in terms of reliability and validity.

Table 4.7. The Frequency of the Answers provided by the EFL instructors for Items for Assessing Speaking Part: Category Summative Assessment

CATEGORY 2: SUMMATIVE ASSESSMENT

		n	**%**
1. When it becomes apparent that the learner cannot reach the criterion level during the speaking exam, the task is ended.	**True**	**22**	**36,7%**
	False	28	46,7%
	Don't Know	1	1,7%
	Not Sure	9	15,0%
7. In a speaking exam, production and comprehension are assessed together.	**True**	**51**	**85,0%**
	False	7	11,7%
	Don't Know	1	1,7%
	Not Sure	1	1,7%
28. Paired interaction tasks in speaking exams may result in domination of one learner over another.	**True**	**48**	**80%**
	False	5	8,3%
	Don't Know	4	6,7%
	Not Sure	3	5,0%

When summative assessment category is taken into consideration, it is seen that there are 3 questions (See Table 4.7.). In the 4th statement, it was found out that only 36.7% of the EFL instructors knew that when it becomes obvious that learners cannot reach the criterion level throughout speaking exams, the task is ended while majority of them (63.4%) did not know it. When the 7th statement was analysed, it was found out that 85% of the EFL instructors knew that in speaking exams, comprehension and production are assessed together. In the 28th statement, it was found out that 80% of the EFL instructors knew that in speaking exams, paired interaction tasks might result in domination of one learner over another.

Table 4.8. The Frequency of the Answers provided by the EFL instructors for Items for Assessing Speaking Part: Category Assessment Tools

CATEGORY 3: ASSESSMENT TOOLS

		n	**%**
6. In interactive tasks, more than two learners pose a problem.	**True**	**17**	**28,3%**
	False	22	36,7%
	Don't Know	9	15,0%
	Not Sure	12	20,0%
17. Asking learners to repeat a word, phrase or a sentence is a way of assessing speaking skills.	**True**	**20**	**33,3%**
	False	34	56,7%
	Don't Know	1	1,7%
	Not Sure	5	8,3%
20. In peer interaction, random matching is avoided.	**True**	**9**	**15,0%**
	False	40	69%
	Don't Know	3	5,0%
	Not Sure	8	13,3%
21. Structured speaking tasks are generally implemented to assess linguistic features of a language, especially pronunciation and grammar.	**True**	**49**	**81,7%**
	False	1	1,7%
	Don't Know	5	8,3%
	Not Sure	5	8,3%
30. Speaking skill can be assessed through an indirect test such as cloze tests.	**True**	**3**	**5,0%**
	False	44	73,3%
	Don't Know	7	11,7%
	Not Sure	6	10,0%
31. Phoneme discrimination (a kind of activity carried out to show the difference between the sounds- e.g., minimal pair practice (ship-sheep)) is an example of direct testing for assessing speaking.	True	20	33,3%
	False	**22**	**36,7%**
	Don't Know	8	13,3%
	Not Sure	10	16,7%
39. Grammatical competence (understanding and using grammar, vocabulary, spelling etc. accurately) does not contribute to learners' fluency.	True	10	16,7%
	False	**48**	**80,0%**
	Don't Know	0	0%
	Not Sure	2	3,3%
40. Strategic competence (appropriate use of communication strategies to compensate lack of knowledge in L2) helps learners use the language appropriately and effectively.	**True**	**57**	**95,0%**
	False	2	3,3%
	Don't Know	0	0%
	Not Sure	1	1,7%

When "assessment tool category" table is examined, there are 8 questions (See Table 4.8.). In the 6th statement, it was found out that only 28.3% of the EFL instructors knew that more than two learners in an interactive task pose a problem while majority of the instructors (71.7%) thought the opposite. When the 17th statement was analysed, it was found out that only 33.3% of the EFL instructors knew that repetition is a way of assessing speaking while

62.7% of them did not know it. In the 20th statement, it was found out that only 15% of the EFL instructors knew that random matching is avoided in peer interaction while 87.3% of them did not know it. When the 21st statement was analysed, it was found out that 81.7% of the EFL instructors knew that structured speaking tasks are generally carried out to assess linguistic features. In 30th statement, it was found out that only 5% of the EFL instructors knew that indirect tests can be utilised such as cloze tests to assess speaking skills while majority of the instructors (95%) did not know this. When the 31st statement was analysed, it was found out that only 36.7% of the EFL instructors knew that phoneme discrimination is an example of indirect testing for assessing speaking. In the 39th statement, it was found out that 80% of the EFL instructors knew that grammatical competence contributes to learners' fluency. When the 40th statement was analysed, it was found out that 95% of the EFL instructors had satisfactory information about strategic competence.

Table 4.9. The Frequency of the Answers provided by the EFL instructors for Items for Assessing Speaking Part: Category Formative Assessment

CATEGORY 4: FORMATIVE ASSESSMENT

		n	%
11. The scores of the speaking assessment with feedback should be communicated to learners to provide them with feedback.	**True**	**45**	**75,0%**
	False	5	8,3%
	Don't Know	4	6,7%
	Not Sure	6	10,0%
15. Feedback should be adapted according to the learners.	**True**	**49**	**81,7%**
	False	7	11,7%
	Don't Know	2	3,3%
	Not Sure	2	3,3%
16. Reading aloud is a technique used to assess speaking skills.	**True**	**23**	**38,3%**
	False	30	50,0%
	Don't Know	2	3,3%
	Not Sure	5	8,3%
18. Discussion among learners is a way of assessing speaking skills.	**True**	**52**	**86,7%**
	False	4	6,7%
	Don't Know	2	3,3%
	Not Sure	2	3,3%
19. When the focus is to assess speaking skills, role plays are used.	**True**	**45**	**75,0%**
	False	7	11,7%
	Don't Know	2	3,3%
	Not Sure	6	10,0%
22. Short answer questions try to measure learners' comprehension of the questions asked and ability to provide relevant information in response.	**True**	**45**	**75,0%**
	False	5	8,3%
	Don't Know	5	8,3%
	Not Sure	5	8,3%

23. In the assessment of speaking, appropriate speaking assessment methods should be chosen for instructional plans.	**True**	**53**	**88,3%**
	False	3	5,0%
	Don't Know	1	1,7%
	Not Sure	3	5,0%
24. Speaking assessment should include teachers' observations of the learners' oral language productions in the classes.	**True**	**43**	**71,7%**
	False	7	11,7%
	Don't Know	1	1,7%
	Not Sure	9	15,0%
25. Mimicry (repetition of series sentences after the teacher or examiner) is one of the techniques used to assess speaking skills of learners.	**True**	**22**	**36,7%**
	False	25	41,7%
	Don't Know	5	8,3%
	Not Sure	8	13,3%
26. Oral presentation, when it is recorded in the lesson, may decrease the stress level of learners.	True	22	36,7%
	False	**22**	**36,7%**
	Don't Know	4	6,7%
	Not Sure	12	20,0%
27. Self-assessment is reliable even if it is used in the early stages.	True	20	33,3%
	False	**21**	**35%**
	Don't Know	6	10,0%
	Not Sure	13	21,7%
29. Peer-assessment is a way of assessing learners' speaking skills.	**True**	**42**	**70,0%**
	False	10	16,7%
	Don't Know	3	5,0%
	Not Sure	5	8,3%
38. Analytic scoring does not provide feedback for learners.	True	4	6,7%
	False	**40**	**66,7%**
	Don't Know	10	16,7%
	Not Sure	6	10,0%

When this "formative assessment category" table is examined, there are 13 questions (See Table 4.9.). When the 11th statement was analysed, it was found out that 75% of the EFL instructors knew that the scores of the speaking assessment should be communicated to learners to provide feedback. In the 15th statement, it was found out that 81.7% of the EFL instructors knew that feedback should be adapted according to learners. In the 16th statement, it was found out that only 38.3% of the EFL instructors knew that reading aloud is a technique implemented to assess speaking skills of the learners while a great number of the instructors (61.6%) did not know this. When the 18th statement was analysed, it was found out that 86.7% of the EFL instructors knew that discussion is a way of assessing speaking skills. In the 19th statement, it was found out that 75% of the EFL instructors knew that role plays are used when the focus is to assess speaking skills. In the 22nd statement, it was found

out that 75% of the EFL instructors knew that short answer questions try to measure learners' comprehension of the questions asked and if they can provide relevant information in response. When the 23rd statement was analysed, it was found out that 88.3% of the EFL instructors knew that when assessing speaking, appropriate speaking assessment methods should be chosen for instructional plans. In the 24th statement, it was found out that 71.7% of the EFL instructors knew that observation should take place in speaking assessment of learners. In the 25th statement, it was found out that only 36.7% of the EFL instructors knew that mimicry is one of the techniques implemented to assess speaking skills of learners while 63.3% of the instructors did not know this technique. When the 26th statement was analysed, it was found out that only 36.7% of the EFL instructors knew that when oral presentation is recorded in the lesson, it may increase the stress level of learners while majority of learners thought the opposite. In the 27th statement, it was found out that only 35% of the EFL instructors knew that self-assessment in not reliable if it is used in early stages. In the 29th statement, it was found out that 70% of the EFL instructors knew that peer assessment is one of the ways of assessing learners' speaking skills. When the 38th statement was analysed, it was found out that 66.7% of the EFL instructors knew that analytic scoring provides feedback for learners.

Table 4.10. The Frequency of the Answers provided by the EFL instructors for Items for Assessing Speaking Part: Category Scoring

CATEGORY 5: SCORING		n	%
32. Using holistic and analytic scales at the same time poses a problem.	True	21	35,0%
	False	**17**	**28,3%**
	Don't Know	12	20,0%
	Not Sure	10	16,7%
34. A checklist is a means of scoring oral presentations in in-class assessment.	**True**	**51**	**85,0%**
	False	3	5,0%
	Don't Know	5	8,3%
	Not Sure	1	1,7%
35. Analytic rubrics provide detailed, focused, and precise assessment of speaking skills.	**True**	**53**	**88,3%**
	False	4	6,7%
	Don't Know	2	3,3%
	Not Sure	1	1,7%
37. Analytic scoring provides teachers with more beneficial diagnostic information about learners' speaking skills.	**True**	**42**	**70,0%**
	False	5	8,3%
	Don't Know	5	8,3%
	Not Sure	8	13,3%

When this "scoring category" is taken into consideration, it is seen that there are 4 questions (See Table 4.10.). In the 32nd statement, it was found out that only 28.3% of the EFL instructors have knowledge about scoring. When the 34th statement was analysed, it was found out that 85% of the EFL instructors knew that checklist is a means of scoring oral presentations in in-class assessment. In the 35th statement, it was found out that 88.3% of the EFL instructors knew that analytic scoring provides detailed, precise and focused assessment of speaking skills. When the 36th statement was analysed, it was found out that 70% of the EFL instructors knew that analytic scoring provides instructors with beneficial diagnostic information about learners' speaking skills.

The answers of the participants were analysed through descriptive statistics and the percentages for each statement together with the correct answers are presented on the tables. When analysed thoroughly, it can be seen that some statements were answered correctly and had a high percentage while some statements were answered incorrectly including the wrong answer together with Don't Know and Not Sure answers and had a low percentage. When the categories for "Items for Assessing Writing" part are taken into consideration, it can be seen that out of 32 items, it can be said that "Items for Assessing Writing" part level was average. When the categories for "Items for Assessing Speaking" are taken into consideration, it can be seen that the level is average as well. When each item is analysed based on the percentages, it can be concluded that the overall LAL level of instructors in assessing productive language skills is above average (50%). However, this level was found below the average in some specific categories (e.g. Items for Assessing Writing: Assessment Tools category indicates a low level: 47,32%- Items for Assessing Speaking: Assessment Tools Category demonstrates low level: 46,87%). This finding shows that even though the overall items show average percentages which is in line with the level of LAL of the EFL instructors, some specific parts show low percentages therefore, instructors may need more trainings related to these categories. This finding is in line with the Kaya's findings (2020). Kaya (2020) conducted a study with a purpose of revealing the LAL knowledge of the EFL instructors working in preparatory schools in Turkish context. 195 EFL instructors from different universities participated Kaya's study and a questionnaire was implemented to obtain the data. According to the findings of the study, it was found out that instructors participating the study had extensive knowledge of language assessment but there is only one contradict about this study. In Kaya' study (2020), speaking skill was found to be the lowest while in this study, the situation is the opposite. In a similar study, Sevilen-Yılmaz

(2021) administered a study with EFL instructors working in preparatory schools. 336 EFL instructors working in state and foundation universities participated in the study and filled out the questionnaire. The findings of her study revealed that instructors had high level of LAL. When analysed skill-based LAL, it was also found that LAL level in assessing writing skill was not significantly higher while it was found that LAL level in assessing speaking skill was significantly higher. In another study, Çetin-Argün (2020), on the other hand, conducted a study with pre-service EFL instructors aiming to find out language assessment knowledge of them. In her study, it was found out that LAL level is below the average and when analysed in terms of productive language skills, it was depicted that LAL level in terms of writing skill was below the average while in speaking, it was above the average. However, when the literature related to this very topic is analysed, it can be understood that a great number of studies found low levels of LAL. First, Mede and Atay (2017) conducted a study with a purpose of investigating the assessment literacy of EFL instructors working in preparatory schools in Turkey and it was found out that Turkish EFL teachers were lack training on classroom focused testing and assessment and this eventually turned into a need for training in Turkish EFL context related to testing and assessment which also depicted that EFL instructors' knowledge on LTA was limited. In another study, Ölmezer-Öztürk (2018) conducted a study with EFL instructors working in higher education system in Turkey aiming to find out the general LAL level of the participants. In her study, it was found out that the general LAL level of the EFL instructors was low. When productive language skills parts were analysed, it was found out that both writing and speaking parts were below the average. Hatipoğlu (2015) conducted a study with pre-service EFL instructors aiming to find out what pre-service EFL instructors in ELT departments knew about language testing and what they thought their ELTE course should include with regards to content. The findings of her study demonstrated that the pre-service EFL instructors in ELT departments had limited knowledge regarding this field. It can be understood that the findings of this study were at variance with the findings of the studies in the related literature. There might be several reasons for this situation. This study consists of 60 participants working in preparatory school in Ankara. However, Mede and Atay (2017) conducted their study with 350 participants working in preparatory schools in universities in Turkey and all these participants graduated from ELT departments. The reason of the contradict might derive from the fact that this study's population was limited and narrowed down to Ankara context only while Mede and Atay (2017) spread their study throughout Turkey. When Ölmezer-Öztürk' study (2018) was taken into consideration, it can be seen that the researcher also

conducted her study in universities in Turkey and the number of the participants was 542 which is a lot higher than this study' participant number. The reasons behind this variance might be the difference between the number of the participants in the studies, different contexts where the studies were conducted.

4.2. Research Question 2: What are the EFL instructors' perceptions on assessing productive language skills through formative assessment?

The second research question of this study is "What are the EFL instructors' perceptions on assessing productive language skills through formative assessment?" The first interview question related to this research question was "In your opinion, what is formative assessment, how can you define it?". The aim of this question was to find out the idea/ideas that EFL instructors hold for formative assessment and whether they are aware of it or not. The findings demonstrated that most of the instructors had an idea what formative assessment is and what kind of tools it incudes.

Table 4.11. EFL Instructors' Definitions for Formative Assessment

Category	Codes	Frequency
Definitions made based on the characteristics of formative assessment	Process-based	6
	Consistency	2
	Feedback oriented	3
	Increasing motivation of learners	1
	No necessity of giving grades	1
	Assessment for learning	2
Others	No idea	1

When the Table 4.11. is taken into consideration, the first category is "definitions made based on the characteristics of formative assessment". Out of 11 EFL instructors, 10 defined formative assessment according to its characteristics and the most stated definition was based on its "process-based" feature. To illustrate this, a quotation was given as an example:

> **Instructor 1:** *"Formative assessment, **I can say that these are the things which we apply in process. (...) So, they can be accepted as formative assessment because they are not at the end or in the beginning (…)**".*

The second category is "others". Out of 11 EFL instructors, 1 stated that they had no idea about formative assessment as stated in the quotation below:

Instructor 4: *"Actually, I've looked that up. But since as I said I've taken my testing class in Turkish, formative assessment means nothing to me right now and I am also tired (…)".*

The other interview question related to this research question was "Do you provide feedback for your students speaking and writing productions during in-class activities? If yes, how?" (See Table 4.12.). The aim of this question was to find out whether EFL instructors give feedback or not for learners' speaking and writing productions during in-class activities and if they do, how they deliver it. The findings illustrated that most of the instructors provide feedback when it comes to assess speaking and writing skills of learners.

Table 4.12. EFL Instructors' Perceptions on Providing Feedback for Students' Speaking and Writing Productions during In-Class Activities

Category	Codes	Frequency
Feedback for Speaking Skill	Immediate feedback	4
	Delayed feedback	9
	Repeating the mistake	1
	Direct feedback	1
	By giving scores	1
	Hinting	1
Feedback for Writing Skill	Writing error codes	3
	Indirect feedback	1
	Delayed feedback	4
	Immediate feedback	2
	By giving scores	1
Others	Coxtext specific reaction for giving feedback	1

The first category is "feedback for speaking skill". Out of 11 instructors, 10 stated that they made use of different types of feedback when it came to learners' language productions in English. The most stated type of feedback given for speaking skill was "delayed feedback" which includes mostly "writing the mistake on the board". To illustrate this, a quotation was given as an example;

Instructor 1: *"(…) I write the common mistakes on the board or sometimes I write the common mistakes on my notebook. Then in the end, I share the common mistakes*

on the board and I want them to correct them so I try to raise their self-awareness of their mistakes. That was for speaking (...)".

The second category is "feedback for writing skills". Out of 11 EFL instructors, 10 mentioned different feedback types that they gave in the classroom. The most stated feedback type was "delayed feedback" as clearly stated in the quotation:

Instructor 11: *"While in class, while I was monitoring the class, in activities, I always monitor class and I have a notebook. I take notes.* **(...) I just give them delayed feedback at the last 5 minutes of the lesson (...)".**

The last category for this question is "Others". Out of 11 EFL instructors, 1 stated that they implemented feedback according to the type activity and the number of people in the classroom. To illustrate this, a quotation was given as an example;

Instructor 3: **"It depends on the activity and the number of people in the class (…)".**

The other interview question related to this research question was "How do your students tend to react to the feedback you provide for their oral and written productions?" (See Table 4.13.). The aim of this question was to find out the reaction of the learners when feedback is given for their speaking and writing productions during in- class activities. The findings indicated that learners mostly have a positive attitude towards given feedback for their oral language productions.

Table 4.13. Learners' Attitudes Towards Given Feedback

Category	Codes	Frequency
Positive	High motivation to learn from feedback	5
	Asking for feedback for their productions	2
	Afraid of making mistakes	1
Negative	Disinterested in feedback	1
	Decrease in motivation in students' engagement	1
Others	Students' characteristics	3
	Depending on the feedback type	1

As clearly seen in the table 4.13., the EFL instructors stated different answers that were collected under three categories. The instructors generally stated that they obtain positive reactions, which takes place in the first category, from their learners such as "high motivation to learn from feedback, asking for feedback for their productions and afraid of making mistakes". The most stated one was "high motivation to learn from feedback" which includes "keeping feedback for later use, eager to learn and giving importance to feedback" as clearly stated in the quotation below:

> **Instructor 2:** *"(...) **Even if you do not provide any feedback, they ask "Teacher, where is my feedback?" or "Are you going to give me feedback?" because they are responsible for these assignments in their exams so there is a concern for grades. And for speaking, actually, I have not received any negative reaction. (...)"***.

Moreover, the instructors touched upon some negative reactions that they obtain from their learners such as "disinterested in feedback" and "decrease in motivation in students' engagement". To illustrate this, a quotation is given as an example;

> **Instructor 11:** *"(…) **Sometimes, strong ones, they really don't like it. They feel like they are better than the rest of the class like they are the best. (…)"***.

Out of 11 EFL instructors, 4 stated that the reactions obtained from their learners change in line with the characteristics of the learners, the type of the given feedback as clearly stated in the quotation below:

> **Instructor 9:** *"Actually it depends on correction. We have many techniques as I said. (…) **It depends on the students' characteristics firstly.** If he is a or she is a shy student, we should determine the correction way very carefully. Because when you correct this learner directly, you can lose your learners. **That is why it depends on my learners' characteristics as I said and the level"**.*

The other interview question related to this research question was "What is the place of "providing feedback" for students' oral language productions in teaching?" (See Table 4.14.). The aim of this question was to find out the perceptions of the EFL instructors on giving feedback for their learners' oral language productions in teaching. The findings exhibited that most of the EFL instructors consider feedback as necessary for oral language productions.

Table 4.14. EFL Instructors' Perceptions on "Providing Feedback" for Students' Oral Language Productions in Teaching

Category	Codes	Frequency
Necessary	Preventing fossilization	3
	Improving learners	3
	Stimulating learning	2
	Engaging learners in a topic	1
Unnecessary	To avoid speaking anxiety	1

When the instructors were asked about their perception on providing feedback for their learners' oral productions and out of 11 EFL instructors, 10 stated that providing feedback for learners' oral production is necessary. The most stated answers were "preventing fossilization and improving learners" as stated in the following quotation:

> **Instructor 3:** *"Well, I think it is very important. Because if we don't give feedback, they can't improve or what's the role of the teacher? To give feedback (…)".*

Regardless of the high number of instructors stating that providing feedback is necessary for learners' oral productions, there was only one instructor who was at variance with these claims by stating that giving too much feedback may disturb the learners and minimise the amount of their eagerness to speak which was linked with avoiding speaking anxiety as clearly stated in the quotation below:

> **Instructor 4:** *"(…) **I try not to disturb my students with the feedback for oral productions. I don't know if I make any sense".***

The final interview question related to this research question was "What is the place of "providing feedback" for students' written language productions in teaching?" (See Table 4.15.). The aim of this question was to find out the perceptions of the EFL instructors on giving feedback for their learners' written language productions in teaching. The findings revealed that all EFL instructors participating the interview consider feedback as necessary for written language productions.

Table 4.15. EFL Instructors' Perceptions on "Providing Feedback" for Students' Written Language Productions in Teaching

Category	Codes	Frequency
Necessary	Preventing fossilization	2
	Improving learners	3
	Stimulating learning	4
	Preparing learners for the exam	1

When the instructors were asked about their perception on providing feedback for their learners' written productions and out of 11 EFL instructors, 11 stated that providing feedback for learners' written production is necessary as stated in the following quotation:

> **Instructor 5:** *"(...) We need to provide written feedback. Not just correcting their mistakes but maybe letting them see their mistakes together with the teacher. Maybe trying to discover what's wrong. (...) So, it would be like a guided correcting I would say maybe. We need to guide them, make them discover and then help them at times".*

4.3. Research Question 3: How do the EFL instructors implement formative assessment tools in their classrooms?

The third research question of this study is "How do the EFL instructors implement formative assessment tools in their classrooms?". The interview question related to this research question was "How do you assess your students' language skills in English?" (See Table 4.16.). The aim of this question was to find out how EFL instructors implement formative assessment tools in their classrooms. The findings demonstrated that formative assessment tools are implemented more in the classrooms with the aim of assessing learners' language skills.

Table 4.16. EFL Instructors' perceptions on how to assess their students' language skills in English

Category	Codes	Frequency
TDU materials	TDU prepares	3
Summative assessment tools	proficiency exams	9
	formative assessment test	1
	speaking exam	3
Formative assessment tools	Observation	5
	quizzes	3
	writing activities	2
	speaking activities	2
	homework	3
	portfolio	2
	grammar tests	1
	ICQs & CCQs	1
	in-class activities	2
	brainstorming	1
	weekly writing assignments	1

As seen in table (Table 4.16.), formative assessment was implemented more than summative assessment. Out of 11 EFL instructors, 3 of them stated that TDU (Testing Development Unit) prepared the materials used in the classroom as one of the instructors clearly stated in the quotation:

> **Instructor 1:** "***Actually, I do not prepare the testing materials individually. Our testing unit prepares it*** *and we have two achievement exams in one course and there are two writing quizzes during 8 week course but this has changed after the pandemic (...)".*

In that vein, summative assessment no matter how less it was mentioned took place in the interviewees' answers. There were different answers as seen in the table 4.16 but most occurred summative assessment types were proficiency exams including final exams, mid-term exams, progress tests and achievement tests, and with speaking exam as one of the instructors clearly stated in the following quotation:

> **Instructor 3:** *"At school, okay. **At school, we have final exam and mid-term exam and there are written parts and before each and every exam, we have standardization sessions where we come together and marked exam papers (…)**.*

When it comes to speaking assessment, again we have standardization sessions before each and every mid-term and final and before proficiency exam because speaking and writing the students are awarded 25 points (…)".

For formative assessment, according to the interview, no matter how the answers varied, observation was found to be the most implemented tool in the classrooms, followed by quizzes and homework, as can be seen in one of the instructors answer:

Instructor 4: *"(…) When teaching young learners, even a look or even an "okay" means fine. (…) You can understand it from their vocabulary usage, grammatical usage if it is a written form if it is not, you can also see it in the confidence, in the position that they are taking in the class. (…). That's why I always observe my students and I try to see their improvement in the process".*

In addition, the instructors were specifically asked "How do you assess your students' speaking skills in English?" According to the answers of the instructors, the researcher constituted three different categories and tried to make a deeper analysis to find out the codes (See Table 4.17.). The findings indicated that the instructors implement formative assessment tools more.

Table 4.17. EFL Instructors' preferences on how to assess their students' speaking skills in English

Category	Codes	Frequency
Summative Assessment tools	Speaking exams	4
Formative Assessment tools	Observation	3
	In-class activities	9
	Presentation	1
	Listening activities	1
	Question-answer sessions	4
Others	Dependency on the course book	2

When instructors' preferences on how to assess speaking skills of learners were asked, they stated various answers and these answers formed three different categories including several different codes. The first category was summative assessment category. As can be

seen in the table, speaking exams were implemented in the classroom to assess learners' speaking skills as summative assessment tool. Out of 11 EFL instructors, 4 stated that they implemented speaking when assessing learners' speaking skills as one of the instructors clearly stated in the quotations:

> **Instructor 1:** *"Via our speaking assessment at the end of the term and at the end of the half of the term. They are given some tasks and they talk about these tasks. Sometimes they are given some pictures and they are supposed to explain what is going on on the photo. That kind of activities. Actually, I do not assess their speaking skills in classroom. We observe but we do not grade (…)".*

For formative assessment, in-class activities were mentioned the most. The instructors utilized the one of the forms of formative assessment, which was in-class activities including brainstorming, speaking activities, picture description, repetition, group work and peer work, for this particular case in the classroom as one of the instructors clearly stated in the quotation:

> **Instructor 7:** *"(…)For speaking, again, at the beginning, short conversations with their peers or maybe by group work if they are not ready to speak if they don't trust themselves enough (…)".*

The last category for this question was "Others" and here, the code was dependency on the course book. Instructors stated that the book has lots of activities so that they have implemented them in their classrooms that can be seen in the following quotation:

> **Instructor 6:** *"Well, most of the time I try to use all kinds of speaking activities in the course book. I never skip them and the other thing even from the very beginning of the semester, I take our speaking exam as a model and I try to apply it as many times as possible (…)".*

In addition, the instructors were asked "What are your assessment criteria in assessing students' speaking skills?" Most of the instructors stated that they implemented rubric based on different parts while assessing their learners' speaking skills in English (Table 4.18).

Table 4.18. EFL Instructors' perspectives on their assessment criteria in assessing learners' speaking skills

Category	Codes	Frequency
Rubric based on Language Use	Accuracy	17
	Fluency	5
Rubric based on Content and meaning	Content	3
	Task achievement	1
Observation	Observation	3

As seen in the table 4.18., there are three different categories formed according to the answers of the interviewees: Rubric based on Language Use, Rubric based on Content and Meaning and observation. Accuracy was the most implemented criterion which includes grammar, vocabulary usage and pronunciation followed by fluency in the first category which is rubric based on language use. In rubric based on content and meaning, the most frequently stated answer was content. In addition, observation as an assessment criterion was stated by 3 EFL instructors as can be seen clearly in the give quotation:

> **Instructor 5:** *"Okay, what is the criteria? **Firstly, self-confidence I would say. A student can be making a lot of grammatical mistakes but using his or her body language a lot and if I can understand that student, that means a lot more than knowing a lot of grammar (...)"**.*

The other interview question related to the research question was "How do you assess your students' writing skills in English?". The aim of this question was to understand EFL instructors' perceptions on assessing writing skills in English. The findings revealed that the instructors implement formative assessment tools when it comes to writing skills of the learners (See Table 4.19.).

Table 4.19. EFL Instructors' perspectives on how to assess their students' writing skills in English

Category	Codes	Frequency
Summative Assessment	Achievement tests	6
Formative Assessment	Quizzes	3
	In-class writing activities	14
	Homework	2
Others	Dependency on the course book	1

The interviewees stated several different answers for that particular question. The researcher collected the answers under three categories which are summative assessment, formative assessment and others. This tables demonstrates that EFL instructors utilize different tools when it comes to assessing writing skills of learners. The most stated tool was in-class writing activities which include writing essays, writing paragraphs, writing activities and writing e-mail. As it can be understood from the table, the instructors implement not only formative assessment tools but they also implement summative assessment tools such as achievement tests including mid-term exams and finals. Moreover, it was stated by some instructors that they made use of course book. An answer for each category is given in the quotations below respectively;

> **Instructor 11:** *"We have weekly homework. **We expect them to write creative writing, opinion essays, argumentative essays and problem solution essays sometimes balanced opinion**. We have **achievement tests, mid-terms and quizzes and homework. That's all"**.*

> **Instructor 5:** *"Anything about writing, okay. If I've given them assigned piece of writing mark. I check it according to how the student has organized his or her writing, **essay or paragraph.** Maybe originality. "Is the text original?" or like just the ones similar to every student can write? Does it have original points that get attention? So organization, originality, use of grammar yes to some extent (…)".*

> **Instructor 9:** *"Writing skills. As I said, I am a skilled teacher. **I am using Skillful and in the Skillful book**, there are many writing tasks. First, we prepare our learners with a model writing (…)".*

The other interview question related to this research question was "What are your assessment criteria in assessing students' writing skills?". The aim of this question was to

understand EFL instructors' criteria in assessing writing skills in English. The findings showed that the instructors make use of rubrics based on different characteristics when it comes to assess writing skills of learners (See Table 4.20.).

Table 4.20. EFL Instructors' Perspectives on their assessment criteria in assessing students' writing skills

Category	Codes	Frequency
Rubric based on Language Use	Grammar	11
	Vocabulary	8
	Mechanics	3
Rubric based on content and meaning	Originality	2
	Organization	6
	Content	6

The interviewees stated that they implemented rubrics when it came to assess writing skills of the learners. The researcher formed two different categories named as rubric based on language use and rubric based on content and meaning. This table demonstrates the types of rubric that the instructors utilized.

> **Instructor 8:** *"Okay, how can I say? Firstly we need to look at the rubric, right? This is so important and according to the rubric, we need to look at the* ***grammar mistakes and also some other mistake word or something like that.*** *We need to pay attention to this in writing assessment, okay?"*.

> **Instructor 9:** *"In assessment criteria, we use generally 4 different parts. These are* ***organization, content, vocabulary and grammar and originality.*** *(…) Especially in online classes, students can plagiarize the other writers easily. That's why we also added "originality". In these days, we assess our learners based on these four I can say. But for me, let me say this one as well, the most important parts are organization and content"*.

The other interview question related to this research question was "Which of the following do you utilize for assessing productive language skills? Portfolio/ presentation/ informal tests/ quizzes/ observation/ diary/ self-assessment/ peer-assessment/ project works?". The aim of this question was to understand what type of formative assessment tool/tools EFL instructors implemented to assess writing and speaking skills in English. The

findings indicated that the instructors made use of almost every tool that was mentioned in the question (See Table 4.21.). Apart from these tools, some instructors gave a different tool that was not mentioned in the question as a preference (The frequency of the implementation of these tools can be seen in Appendix 7).

Table 4.21. EFL Instructors' Preferences on Formative Assessment Tools

Category	Codes	Frequency
Formative Assessment	Portfolio	7
	Presentation	3
	Informal Test	4
	Quizzes	6
	Observation	6
	Learner Diary	2
	Self-Assessment	4
	Peer-Assessment	3
	Project works	4
	Role-play	1
Online Tools	Google Sheets	1
	Padlet	2
	Kahoot	1
	Word Wall	1

For this particular question, the interviewees opted for different formative assessment tools that they implement in their classrooms to assess writing and speaking skills of their learners. According to the table, the most mentioned tool was portfolio followed by quizzes and observation while the least mentioned tool was diary followed by presentation and peer-assessment. In addition, out of 11 EFL instructors, 4 mentioned that together with the given tools, they also applied online tools to assess learners' productive language skills in English as clearly stated in the following quotations:

> **Instructor 2:** *"**Portfolios** and exams".*

> **Instructor 11:** *"Well, sometimes like every 2 weeks I can say, I prepare some online education tools like when it comes to productive language skills, it is hard but for receptive skills, I prepare **Kahoot, Word Wall,** maybe **Padlet** for productive skills. I try to use online tools to assess them. Formatively of course. If I have time".*

The other interview question related to this research question was "What types of materials and in-class activities do you implement to teach and practice speaking skills?". The aim of this question was to find out the materials and activities EFL instructors implemented to teach and practice speaking skills in English. The findings demonstrated that most of the instructors implement in-class activities to teach and practice speaking skills (See Table 4.22.).

Table 4.22. Materials and In-Class Activities Implemented by EFL Instructors to Teach and Practice Speaking Skills

Category	Codes	Frequency
CDU materials	Materials provided by school	2
Online Tools	Flipgrid	1
	Quizlet	1
	Telegraph	1
Games	Taboo	2
	Games	1
In-Class Activities	Mechanical exercises	1
	Guided speaking activities	6
	Acting out/Role play	2
	Discussion	1
	Picture description	1
	Question-answer	2
	Brainstorming	1
	Scaffolding	1
	Self-assessment	1
Authentic Materials	Talking about magazine news	1
	Talking about political news	1
	Talking about daily life	2
	Talking about films	1
	Watching YouTube videos and talking about them	2
Others	Dependency on Course Book	4

11 EFL instructors were asked what types of materials and in-class activities they implement to teach and practice speaking skill. Out of 11 EFL instructors, 2 stated that together with the course book, they implemented materials provided by the school as clearly seen in the quotation below:

Instructor 1: *"I actually follow the course book and we cannot use any other thing except our course book and* **the material which is provided by the school to be**

standard. Because there are lots of classrooms in the school so we have to be standard. I just follow the course book. But sometimes I adapt the materials or I change some parts of it".

The second category is "online tools"; out of 11 EFL instructors, only 2 stated that they made use of online tools as stated in the quotation:

Instructor 6: *"Well I have my own portfolio of questions and pictures for both Program 1 and Program 2 students. Apart from the school's materials, I always use them. I ask the questions. Sometimes I also use an application I don't remember the name right now. Flipgrid maybe. You have heard about it. I send the questions and pictures to the students and they record their voice and they send it to me and I give feedback to them. So, they like it. Sometimes in the class and sometimes through these application I can assess their speaking skills".*

Out of 11 EFL instructors, 3 mentioned that they made use of games while teaching and practising speaking skill as can be seen in the quotation below:

Instructor 8: *"Actually when we look at the speaking skills, for example brainstorming is important in speaking. They can share everything with brainstorming and also some **games like Taboo (…)**".*

The other category is "in-class activities". Most of the interviewees gave their answers in this category. The most stated answers were "guided speaking activities", "role play" and "question-answer" as clearly stated in the quotation below:

Instructor 5: *"Okay, generally in speaking classes, I may come up with a material but not all the time. What I may come up with would be giving situations and just acting out those. First myself showing that I can do it and I have confidence and I can act like a big person or an actress or an actor. Then I involve students to my personal area so I act out with students (…)".Yes, usually acting out, repetition, videos, talking about films, talking about daily life(…)".*

The fourth category is "authentic material"; the instructors implemented real-life materials in their classrooms to teach and practice speaking skills of their learners. The instructors mentioned that they bring magazine news, political news, daily life, films and YouTube videos in the classroom and they talk about them.

Instructor 4: *"Actually, I try to use funny or hot topics to make the room feel more energetic. Sometimes, it might be something political if it is not going too far actually. Sometimes it might be about magazine news. Sometimes it might be about their future or any kind of topic or even using my day personal day and I am giving them a story about my life and I want to know what they think about it or if they have gone through something similar lately or general (…)".*

The last category is "others" and the most mentioned tool was "course book". Out of 11 instructors, 4 stated that together with the other tools, their activities and materials were mainly based on course book as stated in the quotation:

Instructor 2: *"We mainly use our course books, the activities included in our course books. But if I do not like the activity, I can adapt it. I can bring some new ideas or topics or through the activities in the course book".*

As can be understood from the results for that question, the instructors implemented different types of tools when it comes to assess speaking skill of learners.

The other interview question related to this research question was "What types of materials and in-class activities do you implement to teach and practice writing skills?". The aim of this question was to find out the materials and activities EFL instructors implemented to teach and practice writing skills in English. The findings revealed that most of the instructors implement in-class activities to teach and practice writing skills (See Table 4.23.).

Table 4.23. Materials and In-Class Activities Implemented by EFL Instructors to Teach and Practice Writing Skills

Category	Codes	Frequency
CDU materials	Materials provided by school	2
In-Class Activities	Homework	1
	Free writing activities	9
	Brainstorming	5
	Online tools	1
	Guided Writing Activities	8
Authentic Materials	Newspapers	2
	Articles	2
Others	Course Book	4

11 EFL instructors were asked what types of materials and in-class activities they implement to teach and practice writing skill. The first category of the question is "CDU materials". Out of 11 EFL instructors, 2 stated that they implemented materials provided by the school as clearly seen in the quotation below:

> **Instructor 2:** *"We have actually extra materials that were prepared by the material development unit for writing so we give the topics for the essay types and students for us to write these essays so we can say extra materials".*

The second category is "in-class activities". The most stated answer given for this category was free writing activities" followed by "guided writing activities". To illustrate this, the quotation below is given as an example;

> **Instructor 1:** *"I can say the same thing. This is the same for writing as well but as I said before, after teaching* **writing a paragraph or writing an essay,** *I try to implement process writing like they just first write and they bring, they get some feedback and they edit it. That is all".*

The third category is "authentic materials". Out of 11 instructors, 2 stated that they brought newspapers and articles to the classroom as stated in the quotation below:

> **Instructor 7:** ***"For the writing skills, I think generally I provide some pictures again for the lower levels, they can compare and write about it or I show some slides only made with pictures and I want them to write about it. (…) For upper levels, I can just give them a piece of newspaper, article. They can read and write about it (…)".***

The last category is "Others" which includes "use of course book". Use of course book together with other materials was mentioned four times as can be seen in the quotation below:

> **Instructor 6:** *"Well, the book has lots of activities so I don't have any other time to do other writing activities. So, I try to apply all the writing activities in the book. I read the instruction and I apply the writing activity".*

The other interview question related to this research question was "Do you integrate assessment into your teaching? If yes, how and why?". The aim of this question was to understand whether EFL instructors implemented assessment into their teaching and in what

way they realized this implementation. The findings displayed that most of the instructors integrate assessment into their teaching with different tools (See Table 4.24.).

Table 4.24. EFL Instructors' Perceptions on Integrating Assessment into Teaching

Category	Codes	Frequency
Yes	Through observation	8
	Through homework	1
	Through feedback	2
	Through writing activities	1
	Through giving assignments	1
	Through worksheet	1
No		2
Others	Depending on the type of the activity	1

The first category is "Yes". Instructors gave several answers and the most stated answer was integrating assessment through observation. Out of 11 instructors, 8 of them integrated assessment into their teaching and use of observation was high when the instructors assessed their learners while teaching. To illustrate this, a quotation was given as an example;

> **Instructor 1:** *"Actually, I integrate assessment to my teaching but I do not grade them. Through the observations and the interactions between me and my students. They are all kinds of assessment I guess. When you say the name of one of my students, I can tell anything and everything about him or her. Because all the time, **I observe them and I have some information related their learning. That is all I can say"**.*

The second category is "No". Out of 11 EFL instructors, 2 stated that they did not integrate assessment into their teaching as stated in the quotation below:

> **Instructor 6:** *"(...) Because you know our system I mean it is a central system. So, the Testing Unit, they describe everything, they determine everything and we apply it. So, I think we apply what they determine. **I think I can say "No" for this question".***

The last category formed for this question is "Others". Out of 11 EFL instructors, 1 stated that they integrated assessment according to the activity type/types. To illustrate this statement, a quotation was given as an example:

Instructor 4: *"Sometimes when it comes to pair work or group work yes, I do. But generally not. I like my observations and I trust my guts actually".*

4.4. Research Question 4: What problems and challenges do the EFL instructors encounter while assessing learners' productive language skills in English?

This research questions aimed to find out the problems and challenges the EFL instructors encountered while assessing writing and speaking skills of learners through interview questions. The problems and challenged the EFL instructors encountered while assessing productive language skills of learners lied in the questions 6.1 and 6.2. The researcher jotted down the answers of each instructor studiously and constituted two categories and codes. The findings presented that most of the instructors attached the occurring problems while assessing productive language skills to learners (See Table 4.25.).

Table 4.25. The problems and challenges EFL Instructors Encountered While Assessing Speaking Skills of Learners

Category	Codes	Frequency
Learner based problems	Lack of motivation to participate	9
	Underestimating the importance of the skill	1
	Affective Domain Problems	9
	Students' personal matters	1
	Higher students' over dominating	1
	Attendance problems	1
	L1 interference	1
Others	Crowded Classrooms	1
	Technological problems	1

As seen in the Table 4.25., there are different types of problems and challenges that EFL instructors encountered while assessing speaking skills. According to the answers of the instructors, two different categories were constituted named as "Learner based problems" and "Others". When the first category was taken into consideration, it was seen that most of the problems and challenges that the instructors stated evolved out of learners while only

two of the answers which are the number of students in a classroom and technological problems evolved out of other reasons. The most encountered problems were "students' lack of motivation to participate", which include reluctant to speak and reluctant to participate, and "affective domain problems", which include feeling shy, low self-confidence, feeling anxious, feeling insecure and afraid of making mistake, as one of the instructors stated in the following quotation:

> **Instructor 6:** *"Well, I think some students are very reluctant to participate in speaking activities so, that might be a problem (…)".*

The second most stated answer was "affective domain problems" when it came to speaking in the classroom as one of the instructors stated clearly in the following quotation:

> **Instructor 2:** *"That's a good question. Because some students shy away speaking. They just give short answers like yes or no or they do not want to produce sentences. They can be shy, or they think that their level is low so there can be lots of reasons behind that (…)".*

As seen in table, out of 11 EFL instructors, only 2 stated different reasons that did not evolve out of the learners. To illustrate this, a quotation was given as an example;

> **Instructor 3:** *"The class number is a problem. There must be 10 or 12 learners. But unfortunately we don't have 12 students (…)".*

Table 4.26. The problems and challenges EFL Instructors Encountered While Assessing Writing Skills of Learners

Category	Code	Frequency
Learner based problems	L1 interference	1
	Lack of writing ability in L1	1
	Unable to understand the instruction	3
	Students' showing no progress	2
	Lack of student engagement	2
	Copying from friends	1
	Using translation	1
	Lack of motivation to write	1
Instructor based problems	Indecisiveness about correcting all the mistakes	1
	Time management	2
	Teachers' lack of knowledge on rubrics	1
Others	Crowded classes	1

As seen in the table 4.26., there are different types of problems and challenges that the EFL instructors encountered while assessing writing skills. Three different categories were constituted in line with the instructors' answers named as "Learner based problems", "Instructor based problems" and "Others". When categories are taken into consideration, it is clearly seen that the problems and challenges that the instructors encountered evolve out of different factors. For the first category, the most encountered problems were "students not being able to understand the instruction" and "lack of student engagement" as two of the instructors stated in the following quotations relatively;

> **Instructor 8:** *"Generally they create their writing... They are reading a lot of topics right?* ***They cannot understand the instructions and they cannot understand the topic, the main topic.*** *Maybe words limit. That is so important. For example, I am saying "You need to write 50-70 words but they are writing 100 or 20. That's a big problem I think".*

> **Instructor 10:** *"****They don't want to write in class. When I give this task to them because I want to guide them while they are writing, I want to correct them at the time of their writing (...)".***

For the second category, the most encountered problem was "time management" as one of the instructors stated in the following quotation:

Instructor 9: *"Time management. Time management is the most challenging part for teachers who are giving feedback to learners (…)".*

For the last category, out of 11 EFL instructors, 1 stated that the problem that they encountered was "crowded classrooms" as clearly stated in the following quotation:

Instructor 5: *"(…) But the student does not show any progress. At that point, it is a problem. You need to go and really help, deal with that student individually. Otherwise, in a crowded class, it is sometimes hard. You can't always work with students individually even if they need. That's an issue, crowded classes (…)".*

When the problems and challenges encountered while assessing speaking skill of learners were taken into consideration, it was seen that out of 11 EFL instructors, 7 focused on problems that were related to learners and the most mentioned problem was learners' lack of motivation to participate which was mentioned 9 times. In addition, when the problems and challenges encountered while assessing writing skill of learners were taken into consideration, it was seen that the problems and challenges encountered while assessing speaking skill were similar to the ones encountered while assessing writing skill of learners. Out of 11 EFL instructors, 3 instructors focused on the problems given rise to by instructor and the most mentioned problem was "time management". Instructors stated that they had difficulty in managing the time especially when they had more than 20 learners in the classroom. 1 instructor focused on "crowded classes" as a problem and claimed that it was difficult to deal with learners individually which eventually gave rise to not to assess learners' writing skills. Having analysed the data, the researcher extrapolated that most of the problems and challenges encountered and stated by EFL instructors while assessing productive language skills were in line with the learners' attitudes and perceptions on language learning.

4.5. Research Question 5: What do the ELF instructors suggest as solutions for the stated problem/s?

This research questions aimed to find out the possible solutions for the problems and challenges that the EFL instructors stated while assessing writing and speaking skills of learners through interview questions. The possible solutions for the problems and challenges

that the EFL instructors stated lied in the question 7. The researcher jotted down the answers of each instructor studiously and constituted three categories and codes. The findings indicated that most of the instructors offered solutions attached to the instructors (See Table 4.27.).

Table 4.27. The possible solutions offered by EFL Instructors for the stated problems

Category	Codes	Frequency
Instructor based solutions	Increasing learner awareness	2
	Diversity of the activities	4
	Softening the language of the given feedback	1
	Focusing on feedback more	1
	Knowing students better	1
	Sitting plan	1
	Teachers' involvement in the activities	1
	Balancing the difficulty of the tasks	1
	Starting with writing first	1
	Giving more homework	1
	Giving limited time	1
	Giving understandable instructions	1
	Engaging learners more	1
	Using hard copy	1
Others	Change in curriculum	3
	Less crowded classrooms	1

As seen in the table 4.27., there are several solutions suggested by EFL instructors based on different factors. Two different categories were constituted in line with the instructors' answers named as "Instructor based solutions" and "Others". It was also clearly seen that, unlike the problems and challenges encountered, solutions were mainly circled around the instructors as stated in the following quotation:

> **Instructor 5:** *"(…) Teachers are the best source of solutions I think. Everything is up to the teacher actually especially when speaking. Teacher should be involved everything. If they are playing a game, being involved in that game, if acting out, being involved in acting out. So, a teacher should be like a friend and a part of students (…)".*

For the last category, out of 11 EFL instructors, 4 stated that the solution might be about institutional based as clearly stated in the following quotation:

Instructor 9: *"What might be the solutions?* **We should have more writing classes I believe. Thanks to this, we can easily manage the time and we can give much more detailed feedback for our learners. What about speaking? We should also have speaking classes I believe because in speaking, we have many activities (…)".**

The instructors mentioned different solutions for the stated problems and each datum was analysed studiously. When the categories of the stated solutions were taken into consideration, it was seen that out of 11 EFL instructors, 7 focused on solutions based on instructors and, 4 stated solutions based on the institution. Even though the problems mentioned were almost similar to one another, the solutions were completely different. Having analysed the data, the researcher inferred that the interviewees' foci depended on instructors when it came to solutions.

The answers of the participants were analysed through content analysis suggested by Creswell (2014) and the categories and codes found for each statement are presented on the tables. When analysed thoroughly, it can be seen even though the study was about formative assessment and the participants were informed about it, they still gave answers including summative assessment which might show that even though a high number of the instructors (10 out of 11) made a self-made definition for formative assessment, they could not differentiate it with summative assessment. This might also indicate that they do not get a good grasp of formative assessment. The second research question aimed to find out the EFL instructors' perceptions on assessing productive language skills through formative assessment. According to the findings, most of the interviewees had an idea about formative assessment and what tools it includes. They also stated that they give feedback for both of these productive language skills which indicates an important aspect of formative assessment is mostly applied in the classroom. When the interviewees were asked to define formative assessment, they mostly focused on its "process-based" aspect. Only 1 of the instructors stated that she had no idea what formative assessment is. When they were asked their perceptions on providing feedback for students' speaking and writing productions during in-class activities, they mostly stated that they implemented delayed feedback for both of these skills.

The findings demonstrated that most of the instructors had an idea what formative assessment is and what kind of tools it incudes. Moreover, when they were asked how learners react to the given feedback, they stated that there are some positive and negative reactions together with context specific. For positive attitudes, the most stated answer was "high motivation to learn from feedback" while for negative attitudes, the most stated answers were "disinterested in feedback" and "decrease in motivation in student engagement". Out of 11 EFL instructors, 4 of them also stated that the reaction depends on "students' characteristics" and "the feedback type". When the instructors were asked their perceptions on providing feedback for students' oral language productions in teaching and providing feedback for students' written language productions in teaching, they all stated that it is necessary to provide feedback for students' written language productions because it improves learners and stimulate learning while 10 instructors stated that it is necessary to provide feedback for students' oral language productions in terms of preventing fossilization and improving learners and only 1 of the instructors stated that it is unnecessary to provide feedback for students' oral language productions in terms of avoiding speaking anxiety. Overall, it can be stated that most of the instructors knew about formative assessment and provide feedback for their learners' productive language skills. This finding was in line with the Guadu and Boersma (2018)'s study. They (2018) found out that teachers teaching writing skills have a positive belief about the significance of the FA and they practice it averagely; furthermore, it was concluded that their beliefs and practices have a positive and significant correlation. Moreover, Gökçe (2014) conducted a study with English teachers working in private and state schools mostly in Central Black Sea region aiming to find out how EFL instructors perceive FA and if there is significant difference between their perceptions in terms of the type of the school they work in. One of the findings of his study found out that the instructors perceive assessment as necessary for students to see their weaknesses and strengths which is in line with the finding of this study. In another study, Duy and Vien (2020) conducted a study with 40 EFL teachers to explore their comprehension of FA, how they implement it in their classes together with the challenges they encounter when implementing this type of assessment in their classes. They found out that most of the EFL instructors were aware of FA and how important and necessary it is.

The third research question aimed to find out how the EFL instructors implement formative assessment tools in their classes with the aim of assessing productive language skills. According to the findings, most of the interviewees implemented formative

assessment tools to assess learners' productive language skills and they stated that they utilise rubrics as assessment criteria. However, it was also found out that even though the instructors were specifically asked about formative assessment, some of them also stated summative assessment which ultimately shows that they have difficulty in differentiating formative assessment and summative assessment. When the interviewees were asked about their perceptions on how to assess their learners' language skills in English, they mostly stated that they made use of formative assessment tools such as in-class activities for both writing and speaking skills, which is the most frequent one, observation, which is one of the most frequent ones, and quizzes while there are some interviewees stated that they made use of summative assessment when it comes to assess learners' language skills in English such as achievement exams which exhibits that the instructors may not know the difference between SA and FA and they may not have enough training on these issues. When analysed skill-based in terms of implementing FA, for assessing writing skills, it can be said that the instructors assess their learners' written productions mostly through in-class writing activities which include writing essays, writing paragraphs and writing e-mail together with quizzes, homework and course book (as an SA tool, they stated that they implemented achievement tests) and when they were asked about their assessment criteria in assessing writing skills, they stated that they mostly focus on grammar together with vocabulary, content and organization. When it comes to assess speaking skills of the learners, the interviewees stated that they mostly implement in-class activities including picture description, repetition, and brainstorming and when they were asked about the assessment criteria that they implement for speaking assessment, they mostly stated that accuracy is taken into consideration which includes grammar, vocabulary use and pronunciation and it is followed by fluency, content, observation and task achievement. Overall, it can be stated that EFL instructors implement formative assessment tools mostly as in-class activities to assess productive language skills of learners and moreover, they utilize online tools as well. This finding was in line with the Elmahdi, Al-Hattami and Fawzi, (2018)'s study. They collected the data from 166 learners in University of Bahrain in order to comprehend the effectiveness of an online tool which is Plickers. The result of their study demonstrated that the learners believe the importance of FA and immediate feedback after the implementation of the online tool. Moreover, it was found out that this online tool helped the learning process when used as formative assessment tool, improved learners and saved time.

The fourth research question aimed to find out the problems and challenges the EFL instructors encounter while assessing productive language skills of learners. According to the findings, it was found out that the EFL instructors generally circled the problems around learners. When the problems and challenges were analysed separately for writing and speaking skills, it can be concluded that there are different type of problems. When problems and challenges encountered while assessing speaking skills of learners are taken into consideration, the instructors stated two types of problems: problems circled around learners and other. A great number of the instructors mentioned that problems circle around learners. These problems circled around learners include "lack of motivation to participate consisting of reluctant to speak and reluctant to participate, affective domain problems including feeling shy, anxious, insecure, afraid of making mistakes and low self-confidence, underestimating the importance of the skill, students' personal matters, high students' over dominating, attendance problems and L1 interference". Among these problems, the most stated ones are lack of motivation to participate and affective domain problems. Out of 11 EFL instructors, only two of them stated different problems including crowded classes and technological problems. When problems and challenges encountered while assessing writing skills of learners, the instructors stated three types of problems: problems circled around learners, problems circled around instructors and others including crowded classes. The problems circled around learners include unable to understand the instruction, students' showing no progress, lack of students engagement, L1 interference, lack of writing ability in L1, copying from friends, using translation and lack of motivation to write while the problems circled around instructors include indecisiveness about correcting all the mistakes, time management and teachers' lack of knowledge on rubric. The last category which is others includes crowded classes. Overall, it can be stated that most of the problems that the EFL instructors encountered were mostly circled around learners while they also stated some of the problems circled around instructors and others. This finding was in line with the Ozsevik's findings (2010). He (2010) also found out that the problems that instructors encounter were "large classrooms, teachers' heavy workload, heavily-loaded program, curriculum/assessment mismatch, students' poor communicative abilities and students' low motivation" (p.67). In another study, Hossain (2015) conducted a study which focuses on the challenges on the way of effective implementation of productive skills in Bangladeshi context with learners and instructors. In his study, he found out that all the instructors participating the interview stated that the main problems that they encounter was time management and large classes with less classroom durations and his finding was in line with

this study's finding. In another study, Karademir and Gorgoz (2019) conducted a study aiming to exhibit teachers' opinions on the problems that they face in the process of teaching receptive and productive skills. They conducted interviews with English teachers and they found out that, for speaking skill, the problems were circled around the learners generally. They also stated that in the interviews, the teachers mentioned affective domain problems including students' shyness, anxiety, fear, low self-confidence and shame. Moreover, they found out that, for writing skill, the teachers stated that problems evolved out of learners and the teachers stated that these problems were linked to learners' lack of motivation to write which is in line with the findings of this study. In Koosha and Yakhabi (2013) conducted a study aiming to examine the challenges that might give rise to the failure of communicative language teaching in EFL context by touching upon the difficulties that teachers encounter in the process of speaking and offer some solutions for the problems. The problem that they mentioned in their study was the low motivation of learners to speak foreign language together with crowded classes which was in line with the findings of this study.

The fifth research question aimed to find out the solutions offered by the EFL instructors for the stated problems and challenges encountered while assessing productive language skills of learners. According to the findings, it was found out that the EFL instructors generally offered solutions circled around instructors even though they stated that the problems evolved out of learners. When solutions were analysed, it can be concluded that there are different type of solutions. The instructors circled their solutions around two categories: instructor based solutions and others. Instructor based solutions include a great number of codes but the most stated ones are diversity of the activities and increasing learner awareness. Others category, on the other hand, includes change in curriculum such as having separate speaking and writing classes, and less crowded classes. Overall, it can be stated that most of the solutions that the EFL instructors suggested were mostly circled around instructors while they also stated some of the suggestions circled around others. When the literature is reviewed, there are several studies about this issue. In one study, Koosha and Yakhabi (2013) investigated some problems related to communicative language teaching in EFL context by touching upon the difficulties that teachers encounter in the process of speaking and offer some solutions for the problems. They offer some solutions and these include various aspects. The solutions, which are similar to the findings of this study, they offer include those: training opportunities should be offered to teachers, computer mediated communication might be brought to the classes so that learners can have an opportunity to

use the target language outside the class, there can be an improvement about teaching methods which are appropriate to their countries EFL context, and setting some more natural and real needs for learners.

CHAPTER 5

5. CONCLUSION

In this chapter, first the overview of the study is presented with its methodology. Then, the findings and the implications are presented. Finally, it presents limitations and the suggestions for further studies.

5.1. Summary of the Study

This study has depicted LAL level of the EFL instructors in assessing productive language skills and their perceptions on assessing productive language skills through formative assessment. Initially, it has depicted LAL level of the EFL instructors in assessing productive language skills, through a questionnaire that was adapted from Ölmezer-Öztürk (2018). Then it has investigated the instructors' perceptions on formative assessment in assessing productive language skills with a semi-structured interview. This study has been designed as explanatory sequential mixed methods research (Creswell, 2014). The data were collected from 60 EFL instructors working in 9 different universities in Ankara via utilizing a LAL questionnaire; 11 EFL instructors working in different universities in Ankara through semi-structured interview. Quantitative data were analysed with the help of IBM SPSS Statistics Program to run descriptive analyses and frequency analyses to depict LAL of the EFL instructors in assessing productive language skills. Qualitative data analysis procedure of this study was implemented making use of content analysis that integrated the steps Creswell (2014, p.197) suggested.

5.2. The Findings of the Study

Based on the related literature, LAL has an important role in teaching and assessing English. According to Öz and Atay (2017), assessment has numerous advantages for education and it shows teachers' success, what is valid or not and which parts need development; moreover, with the aid of assessment, teachers can comprehend and follow the teaching and learning procedure and implement necessary changes so as to develop learners' learning. In Sevilen-Yılmaz's study (2021), it was found out that the general level of language assessment knowledge of the instructors was high and moreover in Kaya's study (2020) the level of language assessment knowledge of the instructors was sufficient.

However, in Ballıdağ's study (2020), it was found out that the LAL knowledge of the EFL teachers working in different school structures was insufficient.

Based on the literature, formative assessment, defined as assessment for learning, helps learning when it gives information to instructors and learners can utilise as feedback while assessing themselves and it also helps instructors adapt teaching to meet the needs of learners (Black, Harrison, Lee, Marshall, & Wiliam, 2004). In Gökçe's study (2014), it was found out that most of the instructors see formative assessment in the same way but the EFL instructors who work in public and private schools hold opposite views about classroom practices. There are numerous studies that investigate language assessment literacy level of EFL instructors and formative assessment separately. Most of these studies are related to pre-service teachers and learners.

The findings of each research question for this study provided discussion and particular evidence to complement one another. The first research question of this study aimed to explore the LAL level of the EFL instructors in assessing productive language skills. In order to answer this question, a questionnaire was implemented and to analyse the data, descriptive statistics was utilized. The findings of the quantitative data revealed that EFL instructors LAL level in assessing productive language skills was above average however, in some specific parts this level was below average ("Assessment Tools" Parts of Items for Assessing Writing and Items for Assessing Speaking Parts). Since the scope of this study was LAL and FA in assessing productive language skills, when the questionnaire was categorized, it was also found out that the FA categories in both parts (Items for Assessing Writing and Items for Assessing Speaking) had high level.

The second research question of this study aimed to investigate the EFL instructors' perceptions on assessing productive language skills through formative assessment. The answers of the participants were analysed through content analysis suggested by Creswell (2014). When analysed thoroughly, it can be seen even though the study was about formative assessment and the participants were informed about it, they still gave answers including summative assessment which might show that they could not differentiate it with summative assessment. This might also indicate that they do not get a good grasp of formative assessment. Moreover, it was understood that the instructors, mostly, touched upon the characteristics of formative assessment including important definite procedures such as its process-based feature together with focusing on its feedback oriented feature. In addition to

this, the second conclusion that was drawn with the help of the responds obtained from the instructors was that they provide feedback, which is one of the important aspect of formative assessment, for their learners' productive language skills and a great number of the instructors considered feedback as a necessary process for learners' learning procedure. When specifically asked what type of feedback, the instructors generally provide for their learners' oral and written productions during in-class activities, they stated that they give feedback for both of these productive language skills which indicates an important aspect of formative assessment is mostly applied in the classroom. While 10 instructors stated that it is necessary to provide feedback for students' oral language productions in terms of preventing fossilization and improving learners and only 1 of the instructors stated that it is unnecessary to provide feedback for students' oral language productions in terms of avoiding speaking anxiety, they all claimed that it is important to provide feedback for students' written language productions because it improves learners and stimulate learning. Moreover, when they are asked how learners react to the given feedback, they stated that there are some positive and negative reactions together with context specific. Overall, it can be stated that most of the instructors knew about formative assessment and provide feedback for their learners' productive language skills.

The third research question of this study aimed to explore how EFL instructors' implement formative assessment tools in their classrooms to productive language skills. Many important conclusions were drawn in line with the responses of the interviewees. The first conclusion that was drawn based on the responses of the interviewees was that they made use of different assessment tools when it comes to assessing their learners' productive language skills. The most implemented formative assessment tools were "observation" together with "in-class activities" stated by most of the interviewees. According to the findings, the interviewees mostly engage learners to the assessment procedure by providing various assessment tools including games and role-plays as well as their observation. There were also some interviewees stated that they made use of summative assessment when it comes to assess learners' language skills in English such as achievement exams which exhibits that the instructors may not differentiate SA and FA and they may not have enough training on these issues. Overall, it can be stated that EFL instructors implement formative assessment tools mostly as in-class activities to assess productive language skills of learners and moreover, they utilize online tools as well.

The fourth research question of this study aimed to explore the problems and challenges the EFL instructors encounter while assessing productive language skills in English. Numerous important conclusions were drawn in accordance with the responses of the interviewees for these particular interview questions to discover the answers for the fourth research question. The first conclusion that was drawn based on the responses of the interviewees was that the problems and challenges that the instructors encountered while assessing productive language skills generally circled around learners. The instructors touched upon various problems and challenges that they encountered while assessing productive language skills. Since the study focused on productive language skills, this question was divided into two categories in the interview in order to obtain absolute answers; moreover, they were analysed separately. The instructors' foci was learners and the problems evolved out of them. When problems and challenges encountered while assessing speaking skills of learners are taken into consideration, the instructors stated two types of problems: problems circled around learners and other. A great number of the instructors mentioned that problems circle around learners. These problems circled around learners include "lack of motivation to participate consisting of reluctant to speak and reluctant to participate, affective domain problems including feeling shy, anxious, insecure, afraid of making mistakes and low self-confidence, underestimating the importance of the skill, students' personal matters, high students' over dominating, attendance problems and L1 interference". Among these problems, the most stated ones are lack of motivation to participate and affective domain problems. Out of 11 EFL instructors, only two of them stated different problems including crowded classes and technological problems. When problems and challenges encountered while assessing writing skills of learners, the instructors stated three types of problems: problems circled around learners, problems circled around instructors and others including crowded classes. The problems circled around learners include unable to understand the instruction, students' showing no progress, lack of students engagement, L1 interference, lack of writing ability in L1, copying from friends, using translation and lack of motivation to write while the problems circled around instructors include indecisiveness about correcting all the mistakes, time management and teachers' lack of knowledge on rubric. The last category which is others includes crowded classes. As can be seen, the instructors not only focused on problems and challenges evolved out of learners but they focused on the problems and challenges evolved out of other elements as well. Moreover, the researcher deduced that not only problems and challenges related to learners affected the

assessment procedure of productive language skills but also the problems and challenges related to instructors and institutions affected this procedure.

The fifth research question of this study aimed to explore the solutions suggested by EFL instructors for the stated problems. An important conclusion drawn in line with the responses of the interviewees for these particular interview questions to explore the answers for the fifth research question. The first conclusion that was drawn based on the responses of the interviewees was that the solutions mostly circled around the instructors on the contrary to problems and challenges encountered. Instructor based solutions include a great number of codes but the most stated ones are diversity of the activities and increasing learner awareness. Others category, on the other hand, includes change in curriculum such as having separate speaking and writing classes, and less crowded classes. According to the interviewees, the instructors can handle the problems that occur in the classroom environment. Moreover, they can have a chance to change the material or they can create a better learning environment for learners.

All things considered, this study revealed that the LAL level of EFL instructors in assessing productive language skills was found average. However, it was also realized that the instructors needed further training in this particular field. Moreover, even though the instructors had a slight idea on what FA was according to the self-definitions they made, throughout out the interview, their answers included SA even when they were asked about FA. This might indicate that the instructors might not know the difference between FA and SA which might be solved with the help of advanced training. In addition, the instructors generally implemented in-class activities for productive language skills together with observation. The reason behind using observation as an instrument might be based on the fact that instructors do not need to prepare themselves and they might have problems with the time so observation is used when practically is taken into consideration. They mostly made use of feedback, which is an important component of FA, to both learn things about learners and inform them about their success together with implementing rubrics which are generally based on accuracy. Moreover, the problems and challenges the instructors encounter in classrooms were mostly circled around learners by touching upon affective filter domain and their lack of motivation. Crowded classrooms were also seen as a big problem. Finally, as solutions, they generally focused on instructor based solutions. In addition, suggestions related to separate writing and speaking classrooms were offered as solutions together with less crowded classrooms.

5.3. Implications of the Study

Testing and assessment is one of the important components of foreign language education. The concepts of LAL and formative assessment, on the other hand, have gained importance and popularity among researchers. There are different studies carried out on these concepts; however, this particular study might be considered as sui generis because there have been no researches that have focused on depicting the EFL instructors LAL in assessing productive language skills and investigate their opinions on implementing formative assessment to assess productive language skills, by formulating five research questions that enable the researcher to complete the study by including one of the key stakeholders in language teaching and assessment that is EFL instructors working in preparatory schools. Thus, the results of the study indicate the depiction of the EFL instructors LAL in assessing productive language skills and how they implement formative assessment in their teaching, what sort of problems they encounter while assessing productive language skills and what they suggest as solutions for the problems they stated moreover, they imply how effective teaching and learning can be in the classroom with the help of assessment and present opinions for instructors about the problems, the assessment tools and the possible solutions.

First, it was found out that the LAL level of EFL instructors in assessing productive language skills is average. However, in some parts of the questionnaire this level was found low. As a conclusion, it can be better if more ELTE courses are offered to instructors or it can be better to investigate the effectiveness of the courses that are already offered in universities. Moreover, participating in workshops or trainings about testing and assessment may be beneficial and may contribute to their proficiency in this particular area.

Second, the result of the qualitative data revealed that a great number of the instructors implemented formative assessment, mostly in-class activities together with observation, in their classrooms to assess their learners' productive language skills. Moreover, the problems they encounter in the classrooms were mostly occurred because of learners however, the solution they offered were mostly instructor-based. As a result, increasing the awareness of formative assessment and how to implement it into classrooms with different assessment tools might be effective in terms of improving the quality of the lessons and removing the possible problems encountered on the way of teaching and learning.

As a conclusion, when all the results of the study are taken into consideration, it can be said that more ELTE courses can be beneficial for instructors in order to cope with the difficulties of the assessment procedure. Moreover, it might be better if not only ELT departments but also the other related departments can take ELTE courses to be competent enough to understand and implement assessment appropriately and effectively into their teaching. There might be more workshops and training opportunities for instructors to improve themselves in the sense of testing and assessment. Briefly, the findings of this particular study shed light upon the depiction of the EFL instructors LAL in assessing productive language skills and how they implement formative assessment in their teaching, what sort of problems they encounter while assessing productive language skills and what they suggest as solutions for the problems they stated. The implications mentioned above may help EFL instructors improve testing and assessment area with the aid of LAL.

5.4. Limitations of the Study and Suggestions for Further Studies

This study has two limitations. The first limitation is the number of the participants. This research includes 60 participants. Even though the number is enough for this study, the research could be conducted in more universities and with more participants to elaborate the study and results. The other limitation is the narrow-focused qualitative data. As this research includes only instructors' perceptions as qualitative data, learners' perceptions could also be taken into consideration to see two different points of view. The same methodology and the data collection tools might be used with different participants in different institutions. Moreover, the perspective might be altered from instructors to learners completely or one of the tools might be implemented to learners to see the difference between the perceptions.

REFERENCES

Afitska, O. (2014). Use of Formative Assessment, Self-and Peer-Assessment in the Classrooms: Some Insights from Recent Language Testing and Assessment (LTA) Research. *Journal on English Language Teaching*, *4*(1), 29-39. Retrieved from: https://files.eric.ed.gov/fulltext/EJ1068443.pdf

Ahmed, S., & Alamin, A. (2014). Assessing speaking ability in academic context for fourth year Taif University students. *International Journal of English Linguistics*, *4*(6), 97. DOI: 10.5539/ijel.v4n6p97

Ahour, T., & Mukundan, J. (2009). Analytic assessment of writing: Diagnosing areas of strength and weakness in the writing of TESL undergraduate students. *Iranian Journal of Language Studies*, *3*(2).

Aleksandrzak, M. (2011). Problems and challenges in teaching and learning speaking at advanced level. Retrieved from: https://repozytorium.amu.edu.pl/bitstream/10593/1680/1/Aleksandrzak.pdf

American Federation of Teachers, National Council on Measurement in Education, and National Education Association (1990). Standards for teacher competence in educational assessment of students, Educational Measurement: Issues and Practice, 9(4), 1990, 30-32. Retrieved from: http://www.unl.edu/buros/bimm/html/article3.html.

Andhika, T. R. W. (2011). Assessment in ELT: from Theory to Practice. *JURNAL ILMIAH BINA BAHASA*, *4*(2), 91-100.

Arda Özkan, S. (2020). *Exploring the Attitudes, Intentions and Practices of EFL Preparatory School Instructors Regarding Formative Assessment in A Turkish Context.* [Unpublished Master's Thesis]. Istanbul Sabahattin Zaim University, Istanbul.

Azis, A. (2012). Teachers' conceptions and use of assessment in student learning. *Indonesian Journal of Applied Linguistics*, *2*(1), 41-51.

Bachman, L. F. (1990). *Fundamental considerations in language testing (1ˢᵗ Edition).* Oxford, England: Oxford university press.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford University Press.

Backlund, P. M., Brown, K. L., Gurry, J., & Jandt, F. (1982). Recommendations for assessing speaking and listening skills. *Communication Education*, *31*(1), 9-17. Doi: https://doi.org/10.1080/03634528209384655

Ballıdağ, S. (2020). *Exploring the language assessment literacy of Turkish in-service EFL teachers.* [Unpublished Master's Thesis]. Kocaeli University, Kocaeli.

Berry, V., Sheehan, S., & Munro, S. (2019). What does language assessment literacy mean to teachers?. *ELT Journal*, *73*(2), 113-123. Doi: https://doi.org/10.1093/elt/ccy055

Beyreli, L., & Ari, G. (2009). The Use of Analytic Rubric in the Assessment of Writing Performance--Inter-Rater Concordance Study. *Educational Sciences: Theory and Practice*, *9*(1), 105-125.

Biggs, J. (1998). Assessment and classroom learning: a role for summative assessment?. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 103-110. Doi: https://doi.org/10.1080/0969595980050106

Birgin, O., & Baki, A. (2009). An investigation of primary school teachers' proficiency perceptions about measurement and assessment methods: The case of Turkey. *Procedia-Social and Behavioral Sciences*, *1*(1), 681-685.

Black, P. J. (1993). Formative and summative assessment by teachers. *Studies in Science Education*, 21(1), 49-97. Doi: 10.1080/0305726930856001

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, *5*(1), 7-74. Doi: http://dx.doi.org/10.1080/0969595980050102

Black, P., & Wiliam, D. (2010). Inside the Black Box: Raising Standards through Classroom Assessment. *Phi Delta Kappan, 92(1), 81–90.* Doi:10.1177/003172171009200119

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi delta kappan*, *86*(1), 8-21. Doi:

Boston, C. (2002). The concept of formative assessment. *Practical Assessment, Research, and Evaluation*, *8*(1), 9. Doi: https://doi.org/10.7275/kmcq-dj31

Breland, H. M. (1983). The Direct Assessment of Writing Skill: A Measurement Review. College Board Report No. 83-6. Retrieved from: https://files.eric.ed.gov/fulltext/ED242756.pdf

Brindley, G. (2008). Educational Reform and Language Testing. In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopedia of language and education* (2nd ed.) (pp. 365-378). Springer Science+Business Media LLC.

Broadfoot, P. M. (2005). Dark alleys and blind bends: Testing the language of learning. *Language Testing*, *22*(2), 123-141.

Brookhart, S. M. (2001). *The "Standards" and Classroom Assessment Research.* Paper presented at the Annual Meeting of the American Association of Colleges for Teacher Education, Dallas, TX. (ERIC Document Reproduction Service No. ED451189).

Brown, D.H. (2003). *Language Assessment: Principles and Classroom Practises*. Pearson ESL.

Büyükkarcı, K. (2010). *The Effect of Formative Assessment on Learners' Test Anxiety and Assessment Preferences in EFL Context*. [Unpublished Doctoral Dissertation]. Çukurova University, Adana.

Can, E. (2017). *English teachers' classroom assessment practices and their views about the ethics of classroom assessment practices.* [Unpublished Master's Thesis]. Çağ University, Mersin.

Cauley, K. M., & McMillan, J. H. (2010). Formative assessment techniques to support student motivation and achievement. *The clearing house: A journal of educational strategies, issues and ideas*, *83*(1), 1-6. Doi: https://doi.org/10.1080/00098650903267784

Charney, D. (1984). The validity of using holistic scoring to evaluate writing: A critical overview. *Research in the Teaching of English*, 65-81.

Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review*, *24*(2), 205-249. Doi: DOI 10.1007/s10648-011-9191-6

Creswell, J. W., (2014). *Research Design: Qualitative, Quantitative and Mixed Methods Approaches (4ᵗʰ Edition).* SAGE Publications.

Crusan, D., Plakans, L., & Gebril, A. (2016). Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices. *Assessing writing*, *28*, 43-56. Doi: http://dx.doi.org./10.1016/j.asw.2016.03.001

Çetinavcı, U. R., & Topkaya, E. Z. (2012). A Contrastive Qualitative Evaluation of Two Different Sequential Programs Launched at the School of Foreign Languages of a Turkish University. *Turkish Online Journal of Qualitative Inquiry*, *3*(3).

Çetin Argün, B. (2020). *Language Assessment Knowledge of Preservice Teachers of English as a Foreign Language*. [Unpublished Master's thesis]. Çağ University, Mersin.

DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice*, *17*(4), 419-438. Doi: https://doi.org/10.1080/0969594X.2010.516643

Dempsey, M. S., PytlikZillig, L. M., & Bruning, R. H. (2009). Helping preservice teachers learn to assess writing: Practice and feedback in a Web-based environment. *Assessing writing*, *14*(1), 38-61. Doi: https://doi.org/10.1016/j.asw.2008.12.003

Deygers, B., & Malone, M. E. (2019). Language assessment literacy in university admission policies, or the dialogue that isn't. *Language Testing*, *36*(3), 347-368. Doi: https://doi.org/10.1177/0265532219826390

Dixson, D. D., & Worrell, F. C. (2016). Formative and summative assessment in the classroom. *Theory into practice*, *55*(2), 153-159. Doi: 10.1080/00405841.2016.1148989

Doğru, D. (2020). *Assessment Literacy Conceptions and Practices of EFL Teachers in Higher Education.* [Unpublished Master's Thesis]. Bahçeşehir University, İstanbul.

Dolin, J., Black, P., Harlen, W., & Tiberghien, A. (2018). Exploring relations between formative and summative assessment. In *Transforming assessment* (pp. 53-80). Springer, Cham. Doi: 10.1007/978-3-319-63248-3_3

Dörnyei, Z. (2003). *Questionnaires in second language research: Construction, administration, and processing*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Dörnyei, Z. (2007). *Research Methods in Applied Linguistics (1ˢᵗ Edition).* Oxford University Press.

Duy, P. C. M., & Vien, T. (2020). An Investigation into EFL Teachers' perceptions and Practice of Formative Assessment at Some Colleges in Thua Thien Hue Province. *Journal of Science, Hue University of Education, ISSN 1859-1612, No. 04(56) A/2020, 7-20.* Retrieved from: http://tckhgd.huce.vn/DOC_BAIBAO/51_775_PhanCanhMyDuy,TruongVien_03_t ieng%20anh_phan%20canh%20my%20duy_ppdh.pdf

Elmahdi, I., Al-Hattami, A., & Fawzi, H. (2018). Using Technology for Formative Assessment to Improve Students' Learning. *Turkish Online Journal of Educational Technology-TOJET*, *17*(2), 182-188. Retrieved from: https://files.eric.ed.gov/fulltext/EJ1176157.pdf

Elshawa, N., Heng, C. S., Abdullah, A. N., & Rashid, S. (2016). Teachers' assessment literacy and washback effect of assessment. *Teachers' Assessment Literacy and Washback Effect of Assessment*, *5*(4).

Ezir, E. (2013). Assessment and Testing in ELT: the Difference Between Assessment and Testing. *Jurnal SAINTIKOM Vol*, *12*(1). Retrieved from: https://prpm.trigunadharma.ac.id/public/fileJurnal/hpRKJurnal%2012-1-2013_5%20Eli%20Eziir%20OK.pdf

Fahimi, Z., & Rahimi, A. (2015). On the impact of self-assessment practice on writing skill. *Procedia-Social and Behavioral Sciences*, *192*, 730-736. Doi: doi: 10.1016/j.sbspro.2015.06.082

Fareed, M., Ashraf, A., & Bilal, M. (2016). ESL learners' writing skills: Problems, factors and suggestions. *Journal of Education and Social Sciences*, *4*(2), 81-92.

Fulcher, G. (2012). Assessment Literacy for the Language Classroom. *Language Assessment Quarterly*, *9*(2), 113-132. Doi: http://dx.doi.org/10.1080/15434303.2011.642041

Furtak, E. M., & Ruiz-Primo, M. A. (2008). Making students' thinking explicit in writing and discussion: An analysis of formative assessment prompts. *Science Education*, *92*(5), 799-824.

Gareis, C. R., & Grant, L. W. (2015). Assessment Literacy for Teacher Candidates: A Focused Approach. *Teacher Educators' Journal*, 4-21.

Gedye, S. (2010). Formative assessment and feedback: A review. *Planet*, *23*(1), 40-45.

Ghalib, T. K., & Al-Hattami, A. A. (2015). Holistic versus Analytic Evaluation of EFL Writing: A Case Study. *English Language Teaching*, *8*(7), 225-236.

Gibbs, G.R. (2007). *Analyzing qualitative data.* (In Flick, U., ed. The SAGE qualitative research kit. Volume 1-8 Thousand Oaks, California: SAGE. p. 1-176)

Ginther, A. (2012). Assessment of speaking. *The Encyclopedia of applied linguistics*. Retrieved from: https://www.researchgate.net/profile/April_Ginther/publication/277707664_Assessment_of_Speaking/links/5a107383a6fdccc2d79995b9/Assessment-of-Speaking.pdf

Giraldo, F. (2018). Language assessment literacy: Implications for language teachers. *Profile Issues in Teachers Professional Development*, *20*(1), 179-195. Doi: http://dx.doi.org/10.15446/profile.v20n1.62089

Glazer, N. (2014). Formative plus Summative Assessment in Large Undergraduate Courses: Why Both?. *International Journal of Teaching and Learning in Higher Education*, *26*(2), 276-286.

Golkova, D., & Hubackova, S. (2014). Productive skills in second language learning. *Procedia-Social and Behavioral Sciences*, *143*, 477-481.

Gökçe, Ö. F. (2014). *A Comparison of EFL Teachers' Perception of Formative Assessment in Public and Private Schools.* [Unpublished Master's Thesis]. Çağ University, Mersin.

Guadu, Z. B., & Boersma, E. J. (2018). EFL instructors' beliefs and practices of formative assessment in teaching writing. *Journal of Language Teaching and Research*, *9*(1), 42-50.

Hakim, B. (2015). English language teachers' ideology of ELT assessment literacy. *International Journal of Education and Literacy Studies*, *3*(4), 42-48.

Harlen, W. (2007). *Assessment of learning*. Sage.

Harlen, W., & James, M. (1997). Assessment and learning: differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy & Practice*, *4*(3), 365-379.

Harmer, J. (2001). *The Practise of English Language Teaching (3rd Edition)*. England: Pearson Education Limited, Longman.

Harris, M., & McCann. P. (1994). *Assessment (Handbooks for the English Classroom) (1st Edition).* Macmillan Education.

Hatipoğlu, Ç. (2015). English language testing and evaluation (ELTE) training in Turkey: Expectations and needs of pre-service English language teachers. *ELT Research Journal*, *4*(2), 111-128.

Hatipoğlu, Ç. (2016). The impact of the university entrance exam on EFL education in Turkey: Pre-service English language teachers' perspective. *Procedia-Social and Behavioral Sciences*, *232*, 136-144.

Hattie, J. (2003). Formative and summative interpretations of assessment information. *Retrieved September*, *28*, 2011.

Hossain, M. I. (2015). *Teaching productive skills to the students: a secondary level scenario.* (Unpublished Master's Thesis). BRAC University, Bangladesh.

Hughes, A. (2003). *Testing for Language Teachers (2nd Edition).* England: Cambridge University Press.

Inbar-Lourie, O. (2008a). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, *25*(3), 385-402.

Inbar-Lourie, O. (2008b). Language assessment culture. In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopedia of language and education* (2nd ed.) (pp. 285-299). Springer Science+Business Media LLC.

Inbar-Lourie, O. (2017). Language Assessment Literacy. *Language Testing and Assessment*, 257-270. doi: 10.1007/978-3-319-02261-1_19

Irons, A. (2008). *Enhancing learning through formative assessment and feedback.* Routledge.

Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using mixed-methods sequential explanatory design: From theory to practice. *Field methods*, *18*(1), 3-20.

Jannati, S. (2015). ELT teachers' language assessment literacy: Perceptions and practices. *The International Journal of Research in Teacher Education*, *6*(2), 26-3. Retrieved from: http://ijrte.eab.org.tr/media/volume6/issue2/s_jannati.pdf

Javed, M., Juan, W. X., & Nazli, S. (2013). A Study of Students' Assessment in Writing Skills of the English Language. *Online submission*, *6*(2), 129-144.

Jeong, H. (2011). *Perspectives of language assessment training for teachers and testing professionals* (Doctoral dissertation, University of Illinois at Urbana-Champaign). Retrieved from: https://www.ideals.illinois.edu/bitstream/handle/2142/26071/Jeong_Heejeong.pdf?sequence=1&isAllowed=y

Jeong, H. (2013). Defining assessment literacy: Is it different for language testers and non-language testers?. *Language Testing*, *30*(3), 345-362.

Karademir, C. A., & Gorgoz, S. (2019). English Teachers' Problems Encountered in Teaching Four Basic Language Skills. *International Education Studies*, 12(4), 118-127.

Kaya, T. (2020). *Exploring Language Assessment Literacy of EFL Instructors in Language Preparatory Programs*. [Unpublished Master's Thesis]. Bahçeşehir University, İstanbul.

Khadijeh, B., & Amir, R. (2015). Importance of teachers' assessment literacy. *International Journal of English Language Education*, *3*(1), 139-146.

Kleinsasser, A. M. (1995). Assessment culture and national testing. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *68*(4), 205-210.

Knight, B. (1992). Assessing speaking skills: a workshop for teacher development. *ELT journal*, *46*(3), 294-302.

Koosha, M., & Yakhabi, M. (2013). Problems associated with the use of communicative language teaching in EFL contexts and possible solutions. *International Journal of Foreign Language teaching and research,* 1(2), 77-90.

Kremmel, B., & Harding, L. (2020). Towards a comprehensive, empirical model of language assessment literacy across stakeholder groups: Developing the Language Assessment Literacy Survey. *Language Assessment Quarterly*, *17*(1), 100-120.

Kroll, B. (1998). Assessing writing abilities. *Annual Review of Applied Linguistics*, *18*, 219-240. Doi: https://doi.org/10.1017/S0267190500003561

Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*, *32*(2), 169-197.

Lan, C., & Fan, S. (2019). Developing classroom-based language assessment literacy for in-service EFL teachers: The gaps. *Studies in Educational Evaluation*, *61*, 112-122.

Larsen-Freeman, D., & Anderson, M. (2013). *Techniques and principles in language teaching (3rd edition-Oxford handbooks for language teachers)*. Oxford university press.

Lee, I. (2011). Formative assessment in EFL writing: An exploratory case study. *Changing English*, *18*(1), 99-111.

Libarkin, J. C., & Kurdziel, J. P. (2002). Research methodologies in science education: The qualitative-quantitative debate. *Journal of Geoscience Education*, *50*(1), 78-86.

Luoma, S. (2004). *Assessing Speaking (1st Edition)*. Cambridge, England: Cambridge University Press.

Malone, M. E. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing*, *30*(3), 329-344.

Matthews, M. (1990). The measurement of productive skills: doubts concerning the assessment criteria of certain public examinations. *ELT Journal, 44(2), 117–121.* doi:10.1093/elt/44.2.117

Mede, E., & Atay, D. (2017). English language teachers' assessment literacy: The Turkish context. *Dil Dergisi*, 168(1), 43-60.

Mellati, M., & Khademi, M. (2018). Exploring teachers' assessment literacy: Impact on learners' writing achievements and implications for teacher development. *Australian Journal of Teacher Education (Online),* 43(6), 1-18.

Mukminatien, N. (2015). The advantages of using an analytic scoring procedure in speaking assessment. *TEFLIN Journal*, *11*(1), 35-46.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in higher education*, *31*(2), 199-218.

Norris, J. M. (2006). The issue: The why (and how) of assessing student learning outcomes in college foreign language programs. *The Modern Language Journal*, *90*(4), 576-583.

O'Loughlin, K. (2013). Developing the assessment literacy of university proficiency test users. *Language Testing*, *30*(3), 363-380.

Ozsevik, Z. (2010). *The use of communicative language teaching (CLT): Turkish EFL teachers' perceived difficulties in implementing CLT in Turkey*. [Unpublished Master's Thesis]. University of Illinois at Urbana-Champaign, The USA. Retrieved from:https://www.ideals.illinois.edu/bitstream/handle/2142/16211/Ozsevik_Zekariy a.pdf?sequ

Ökten, A. (2009). *Effects of Formative Assessment Application on Students' Language Proficiency and Beliefs in Language Learning in E.F.L Context: A Case Study*. [Unpublished Master's Thesis]. Çukurova University, Adana.

Ölmezer-Öztürk, E. (2018). *Developing and Validating Language Assessment Knowledge Scale and Exploring the Assessment Knowledge of EFL Teachers*. [Unpublished Doctoral Dissertation]. Anadolu University, Eskişehir.

Ölmezer-Öztürk, E., & Aydın, B. (2019). Investigating language assessment knowledge of EFL teachers. *Hacettepe University Journal of Education*, 34(3), 602-620. doi: 10.16986/HUJE.2018043465

Öz, S., & Atay, D. (2017). Turkish EFL instructors' in-class language assessment literacy: Perceptions and practices. *ELT Research Journal*, 6(1), 25-44.

Pastore, S., & Andrade, H. L. (2019). Teacher assessment literacy: A three-dimensional model. *Teaching and Teacher Education*, *84*, 128-138.

Perera-Diltz, D. M., & Moe, J. L. (2014). Formative and summative assessment in online education. *Journal of research in innovative teaching*, *7*(1).

Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing*, *30*(3), 381-402. Doi: https://doi.org/10.1177/0265532213480337

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental?. *Theory into practice*, *48*(1), 4-11.

Popham, W. J. (2011). Assessment literacy overlooked: A teacher educator's confession. *The Teacher Educator*, *46*(4), 265-273.

Rea-Dickins, P. (2004). Understanding teachers as agents of assessment. *Language Testing, 21(3),* 249-258.

Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching*. Cambridge university press.

Roever, C., & Kasper, G. (2018). Speaking in turns and sequences: Interactional competence as a target construct in testing speaking. *Language Testing*, *35*(3), 331-355.

Rushton, A. (2005). Formative assessment: a key to deep learning?. *Medical teacher*, *27*(6), 509-513.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional science*, *18*(2), 119-144.

Sarı, H. (2019). *The Impact of the Use of Formative Assessment on EFL Students' Writing Anxiety: An Action Research Study.* [Unpublished Master's Thesis]. Çağ University, Mersin.

Sarıyıldız, G. (2018). *A Study into Language Assessment Literacy of Preservice English as a Foreign Language Teachers in Turkish Context.* [Unpublished Master's Thesis]. Hacettepe University, Ankara.

Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing*, *30*(3), 309-327. Doi: https://doi.org/10.1177/0265532213480128

Seers, K. (2012). Qualitative data analysis. *Evidence-based nursing*, *15*(1), 2-2.

Sevilen Yılmaz, Ç. (2021). *Exploring the Language Assessment Literacy of EFL Instructors and Students' Perceptions about Assessment in English Preparatory Program.* [Unpublished Master's Thesis]. Bahçeşehir University, İstanbul.

Siegel, M. A., & Wissehr, C. (2011). Preparing for the plunge: Preservice teachers' assessment literacy. *Journal of Science Teacher Education*, *22*(4), 371-391.

Solgun Günel, A. (2014). *The Effects of Formative Assessment on Students' Participation in an English Language Program.* [Unpublished Master's Thesis]. Çağ University, Mersin.

Sönmez, T. (2015). *The Effects of Formative Assessment on Learner Autonomy of Turkish Adult EFL Learners.* [Unpublished Master's Thesis]. Necmettin Erbakan University, Konya.

Stainback, W., & Stainback, S. (1989). Using qualitative data collection procedures to investigate supported education issues. *Journal of the Association for Persons with Severe Handicaps*, *14*(4), 271-277.

Stiggins, R. J. (1982). A comparison of direct and indirect writing assessment methods. *Research in the Teaching of English*, 101-114.

Stiggins, R. J. (1991a). Assessment literacy. *Phi Delta Kappan*, *72*(7), 534-39.

Stiggins, R. J. (1991b). Relevant classroom assessment training for teachers. *Educational Measurement: Issues and Practice*, *10*(1), 7-12.

Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, *77*(3), 238.

Stiggins, R. J. (1999). Evaluating classroom assessment training in teacher education programs. *Educational measurement: Issues and practice*, *18*(1), 23-27.

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, *83*(10), 758-765.

Stiggins, R., & DuFour, R. (2009). Maximizing the power of formative assessments. *Phi Delta Kappan*, *90*(9), 640-644.

Stull, J., Varnum, S. J., Ducette, J., & Schiller, J. (2011). The many faces of formative assessment. *International Journal of Teaching and Learning in Higher Education*, *23*(1), 30-39.

Sultana, N. (2019). Language assessment literacy: An uncharted area for the English language teachers in Bangladesh. *Language Testing in Asia*, *9*(1), 1-14.

Şahin, S. (2011). *American English, Turkish and Interlanguage Refusals: A Cross-Cultural Communication and Interlanguage Pragmatics Study.* [Unpublished Master's Thesis]. Middle East Technical University, Ankara.

Şahin, S. (2019). *An Analysis of English Language Testing and Evaluation Course in English Language Teacher Education Programs in Turkey: Developing Language Assessment Literacy of Pre-Service EFL Teachers.* [Unpublished Doctoral Dissertation]. Middle East Technical University, Ankara.

Tamerer, B.R. (2019). *An investigation of Turkish pre-service EFL teachers' language assessment literacy.* [Unpublished Master's Thesis]. Kocaeli University, Kocaeli.

Taras, M. (2005). Assessment–summative and formative–some theoretical reflections. *British journal of educational studies*, *53*(4), 466-478.

Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics*, *29*, 21-36. Doi: 10.1017/S0267190509090035

Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language testing*, *30*(3), 403-412.

Thornbury, S. (2005). *How to Teach Speaking*. Harlow, England: Pearson-Longman.

Tosuncuoğlu, I. (2018). Importance of Assessment in ELT. *Journal of Education and Training Studies*, *6*(9), 163-167. Doi: 10.11114/jets.v6i9.3443

Tsagari, D., & Vogt, K. (2017). Assessment literacy of foreign language teachers around Europe: Research, challenges and future prospects. *Papers in Language Testing and Assessment*, *6*(1), 41-63.

Tunstall, P., & Gsipps, C. (1996). Teacher feedback to young children in formative assessment: A typology. *British educational research journal*, *22*(4), 389-404.

Ulker, V. (2017). The Design and Use of Speaking Assessment Rubrics. *Journal of Education and Practice*, *8*(32).

Uysal Kurtulmuş, N. (2018). *A Comparative Study of Turkish and International Adult EFL Students' Perceptions of Formative Assessment*. [Unpublished Master's Thesis]. Bahçeşehir University, İstanbul.

Van Teijlingen, E., & Hundley, V. (2002). The importance of pilot studies. *Nursing Standard (through 2013)*, *16*(40), 33.

Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, *11*(4), 374-402.

Vogt, K., Tsagari, D., & Spanoudis, G. (2020). What Do Teachers Think They Want? A Comparative Study of In-Service Language Teachers' Beliefs on LAL Training Needs. *Language Assessment Quarterly*, 1-24.

Webb, N. (2002, April). Assessment literacy in a standards-based urban education setting. In *annual meeting of the American Educational Research Association, New Orleans*.

White, E. M. (1994). Issues and problems in writing assessment. *Assessing Writing*, *1*(1), 11-27.

Weigle, S. C. (2002). *Assessing Writing (1ˢᵗ Edition)*. England: Cambridge University Press.

Wiggins, G. (2012). Seven keys to effective feedback. *Feedback*, *70*(1), 10-16.

Wiliam, D., & Black, P. (1996). Meanings and consequences: a basis for distinguishing formative and summative functions of assessment?. *British educational research journal*, *22*(5), 537-548.

Wiliam, D. (2006). Formative assessment: Getting the focus right. *Educational assessment*, *11*(3-4), 283-289.

Wiliam, D. (2011). What is assessment for learning?. *Studies in educational evaluation*, *37*(1), 3-14.

Wingate, U. (2010). The impact of formative feedback on the development of academic writing. *Assessment & Evaluation in Higher Education*, *35*(5), 519-533.

Xiao, Y., & Yang, M. (2019). Formative assessment and self-regulated learning: How formative assessment supports students' self-regulation in English language learning. *System*, *81*, 39-49.

Yan, X., Zhang, C., & Fan, J. J. (2018). "Assessment knowledge is important, but…": How contextual and experiential factors mediate assessment practice and training needs of language teachers. *System*, *74*, 158-168.

Yastıbaş, A., E. (2018). *Language Assessment Literacy of Turkish EFL Instructors: A Multiple-Case Study*. [Unpublished Doctoral Dissertation]. Atatürk University, Erzurum.

Yurtdakal, H. D. (2015). *The Effects of Formative Assessment on EFL Learners' Vocabulary Enrichment.* [Unpublished Master's Thesis]. İstanbul Sabahattin Zaim Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.

Yüksel, H. S., & Gündüz, N. (2017). Formative and summative assessment in higher education: opinions and practices of instructors. *European Journal of Education Studies*.

Zheng, Y., & Yu, S. (2019). What has been assessed in writing and how? Empirical evidence from Assessing Writing (2000–2018). *Assessing Writing*, 100421. Doi: https://doi.org/10.1016/j.asw.2019.100421

APPENDICES

APPENDIX 1: THE PERMISSION TO IMPLEMENT THE QUESTIONNAIRE

Selda Arslan
Alıcı: elcinolmezerozturk ▾

13 Tem 2020 13:12 ☆ ↰ ⋮

Merhabalar Elçin Hocam,

Ben Selda Arslan. Başkent Üniversitesi Eğitim Bilimleri Enstitüsü İngiliz Dili Eğitimi Tezli Yüksek Lisans öğrencisiyim. Tez danışmanım Dr. Öğr. Üyesi Sevgi Şahin hocamla Language Assessment Literacy Development üzerine çalışacağız. Daha önce konum üzerine çalışırken "Investigating Language Assessment Knowledge of EFL Teachers" başlıklı makalenizi okudum ve geliştirdiğiniz araştırma ölçeğini inceleme fırsatı buldum ayrıca ölçeği geliştirdiğiniz doktora tezinizi de inceledim. Geliştirdiğiniz ölçeğin, tez çalışmam için uygun olduğunu düşünüyorum. Etik kurallar çerçevesinde ilerlemek amacıyla, geliştirdiğiniz ölçeği uyarlayıp veri toplama araçlarımdan biri olarak kullanabilmek için izninizi isteyecektim.

Şimdiden çok teşekkür ederim,

Saygılarımla,

Selda ARSLAN

Elçin ÖLMEZER ÖZTÜRK
Alıcı: ben ▾

14 Tem 2020 13:26 ☆ ↰ ⋮

Merhabalar,
Tezinizde anketimi kullanabilirsiniz
Kolaylıklar ve başarılar dilerim

132

APPENDIX 2: THE DEMOGRAPHIC QUESTIONS

1. Gender: Female ☐   Male: ☐
2. Age:
3. The university you are currently working:
3.   Which university did you graduate from (BA):
4.   Which department did you graduate from (BA)? English Language Teaching   ☐ English Language and Literature ☐ American Culture and Literature ☐ Translation and Interpretation   ☐ Linguistics   ☐ Others (Please specify):   ☐
5.   Which university did you graduate from / are you currently studying (MA)? (If you do not hold an MA degree or if you are not studying, please leave it blank.)
6.   Which department did you graduate from/are you currently studying (MA)? (If you do not hold an MA degree or if you are not studying, please leave it blank) English Language Teaching   ☐ English Language and Literature ☐ American Culture and Literature ☐ Translation and Interpretation   ☐ Linguistics   ☐ Others (Please specify):   ☐ I have not graduated yet   ☐ (If you have not graduated yet, please specify your department by selecting one of the choices together with selecting this choice)
7.   Which university did you graduate from/are you currently studying (PhD)? (If you do not hold a PhD degree or if you are not studying, please leave it blank)

8.	Which department did you graduate from/are you currently studying (PhD)? (If you do not hold a PhD degree or if you are not studying, please leave it blank) English Language Teaching ☐ English Language and Literature ☐ American Culture and Literature ☐ Translation and Interpretation ☐ Linguistics ☐ Others (Please specify): ☐ I have not graduated yet ☐ (If you have not graduated yet, please specify your department by selecting one of the choices together with selecting this choice)
9.	How long have you been teaching English:
10.	Have you received any courses about testing and assessment English or testing and assessment in general in BA? Yes ☐ No ☐
11.	The total number of the courses about testing and assessment English or testing and assessment in general you have received in BA (If you have not received any courses, then please write "0"):
12.	Have you received any courses about testing and assessment English or testing and assessment in general in MA? Yes ☐ No ☐
13.	The total number of the courses about testing and assessment English or testing and assessment in general you have received in MA (If you have not received any courses, then please write "0"):
14.	Have you received any courses about testing and assessment English or testing and assessment in general in PhD? Yes ☐ No ☐
15.	The total number of the courses about testing and assessment English or testing and assessment in general you have received in PhD (If you have not received any courses, then please write "0"):

16.	Have you received any courses about testing and assessment English or testing and assessment in general through pedagogic formation, CELTA and/or DELTA? Yes □ No □
17.	If yes, which one/s offered a course/s about testing and assessment English or testing and assessment in general? □ Pedagogic Formation □ CELTA □ DELTA
18.	The total number of the courses about testing and assessment English or testing and assessment in general you have received through pedagogic formation (If you have not received any courses, then please write "0")
19.	The total number of the courses about testing and assessment English or testing and assessment in general you have received in CELTA (If you have not received any courses, then please write "0")
20.	The total number of the courses about testing and assessment English or testing and assessment in general you have received in DELTA (If you have not received any courses, then please write "0")
21.	Have you worked in one of these units in the preparatory school? CDU (Curriculum Development Unit) □ TDU (Testing Development Unit) □ TTDU (Teacher Training Development Unit) □ No, I haven't. □
22.	Please write your institutional e-mail address:

INFORMED CONSENT FORM

This present study is conducted for a master thesis work titled "EFL Instructors' Language Assessment Literacy Level in Assessing Productive Language Skills and Their Perceptions on Assessing Productive Language Skills through Formative Assessment" by Selda Arslan and her thesis advisor Assist. Prof. Sevgi Şahin. For this study, the ethics committee approval from Başkent University has been received. The study aims to measure the EFL instructors' perceived level of language assessment literacy in assessing productive language skills and to obtain their perceptions on assessing productive language skills through formative assessment. While the first part of the study focuses on measuring the perceived level of language assessment literacy of EFL instructors in assessing productive language skills via a questionnaire, this will be the second part of the study which aims to obtain EFL instructors' perceptions on assessing productive language skills through formative assessment via semi-structured interview. Participation in the study must be on a voluntary basis. No personal identification information is required in the interview questions. Your verbal responses will be kept anonymous and your answers will be kept strictly confidential and evaluated only by the researchers; the obtained responses will be used for scientific purposes only. We expect you to read the consent form carefully and put a tick to the box to participate the study. The interview does not contain any questions that may give rise to discomfort in the participants. However, during participation, for any reason, if you feel uncomfortable, you are free to quit at any time and if you do so, please inform the researcher. We would like to thank you in advance for your contribution to this study.

I am told what the study is about and I am participating in this study totally on my own will and I am aware that I can quit participating at any time I want. I also know that my verbal responses will be kept anonymous and used for only the study. I give my consent for the use of the information I provide for scientific purposes.

☐ I agree to participate this study.

CATEGORIES	QUESTIONS
1. Educational Background	1. Which university and department did you graduate from (BA, MA, PhD)? 2. Did you take any courses about testing and assessment in English or any other courses related testing and assessment? 2.1. If yes, when and how many?
2. Work and Field Experience	1. Where are you currently working? 2. How long have you been teaching English? 3. Have you worked in a testing unit before? 3.1. If yes, how long have you worked? 4. And in what kind of assessment practices have you involved in?
3. Application of Assessment in Language Classes	1. How do you assess your students' language skills in English? 2. How do you assess your students' speaking skills in English? 3. What are your assessment criteria in assessing students' speaking skills? 4. How do you assess your students' writing skills in English? 5. What are your assessment criteria in assessing students' writing skills? 6. Which of the following do you utilize for assessing productive language skills? ▪ Portfolio ▪ Presentation ▪ Informal tests ▪ Quizzes ▪ Observation ▪ Diary ▪ Self-assessment ▪ Peer-assessment

		▪ Project works
	7.	How often do you use them?
	8.	Do you utilise other assessment instruments differently from the ones mentioned above?
4. Application of Formative Assessment	1.	What types of materials and in-class activities do you implement to teach and practice speaking skills?
	2.	What types of materials and in-class activities do you implement to teach and practice writing skills?
	3.	Do you integrate assessment into your teaching?
		3.1. If yes, how and why?
5. Opinions about Formative Assessment	1.	In your opinion, what is formative assessment, how can you define it?
	2.	Do you provide feedback for your students speaking and writing productions during in-class activities?
		2.1. If yes, how?
		2.1.1. How do your students tend to react to the feedback you provide for their oral and written productions?
	3.	What is the place of "providing feedback" for students' oral language productions in teaching?
	4.	What is the place of "providing feedback" for students' written language productions in teaching?
6. The problems and challenges encountered when assessing productive language skills	1.	What sort of problems do you encounter while assessing speaking skills in the classroom?
	2.	What sort of problems do you encounter while assessing writing in the classroom?
7. Possible solutions for the stated problem (s)	1.	What might be the solutions for the problems you mentioned as far as you are concerned?
8. Others	1.	Do you have any other comments or suggestions to add?

PART 1: DEMOGRAPHIC INFORMATION

1. Gender: Female ☐ Male ☐
2. Age:
3. The university you are currently working: Atılım University ☐   Başkent University ☐   TED University ☐
4. Which university and department did you graduate from (BA)?   **The name of the university:**   **Department:**   English Language Teaching ☐   English Language and Literature ☐   American Culture and Literature ☐   Translation and Interpretation ☐   Linguistics ☐   Others (Please specify):
5. Which university and department did you graduate from (MA)?   **The name of the university:**   **Department:**   English Language Teaching ☐   English Language and Literature ☐   American Culture and Literature ☐

Translation and Interpretation   ☐
Linguistics   ☐
Others (Please specify):
I have not graduated yet   ☐
(Please specify your department by selecting one of the choices together with selecting this choice)

6. Which university and department did you graduate from (PhD)?

The name of the university:

Department:

English Language Teaching   ☐

English Language and Literature   ☐

American Culture and Literature   ☐

Translation and Interpretation   ☐

Linguistics   ☐

Others (Please specify):

I have not graduated yet   ☐

(Please specify your department by selecting one of the choices together with selecting this choice)

7. How long have you been teaching English:

8. Have you received any courses about testing and assessment English or testing and assessment in general in BA, MA or PhD? If yes, how many?

For BA:

Yes   ☐   No   ☐

The total number of the courses you have received:

For MA:

Yes ☐ No ☐
The total number of the courses you have received:
For PhD:
Yes ☐ No ☐
The total number of the courses you have received:

9. Have you worked in one of these units in the preparatory school?

CDU (Curriculum Development Unit) ☐

TDU (Testing Development Unit) ☐

TTDU (Teacher Training Development Unit) ☐

No, I haven't. ☐

Please write your institutional e-mail address:

ITEMS FOR ASSESSING WRITING	TRUE	FALSE	DON'T KNOW
1. Giving two options to learners and asking them to write about one ensure reliable and valid scoring.	☐	☐	☐
2. Analytic scoring is used to see the strengths and weaknesses of learners.	☐	☐	☐
3. The parts of a scoring scale and the scores in each part do not change for different levels of learners.	☐	☐	☐
4. When there is a disagreement between the scores of the two raters, they score the written work again.	☐	☐	☐
5. Holistic scoring is used to see whether the learner is proficient or not at the end of the term.	☐	☐	☐
6. Analytic scoring leads to greater reliability than holistic scoring in writing.	☐	☐	☐
7. Classroom assessment of learning in terms of writing is best served through analytic scoring rather than holistic scoring.	☐	☐	☐
8. Mechanical errors (e.g. spelling and punctuation) are dealt with in the assessment of later stages of a written work.	☐	☐	☐
9. While scoring a writing test, the expectation of raters does not have an influence.	☐	☐	☐
10. The complex relation among writing task variables, rater variables and scoring criteria does not have any effects on the scoring the writing tests.	☐	☐	☐
11. While planning a scoring rubric for writing, factors including who will use the rubric and what aspects of writing are crucial are not taken into consideration.	☐	☐	☐
12. Learners are required to write about at least 2 tasks in the exam rather than 1 task.	☐	☐	☐

13. Giving restrictive prompt/guidelines for the writing task is avoided.	☐	☐	☐
14. In controlled writing, learners have the chance to convey new information.	☐	☐	☐
15. Irrelevant ideas are ignored in the assessment of initial stages of a written work in process writing.	☐	☐	☐
16. In the assessment of writing, the assessment method should be chosen according to instructional decisions.	☐	☐	☐
17. Utilising the results of writing assessment, while making decisions about learners and teaching plans, carries great importance.	☐	☐	☐
18. Writing skill is best assessed when it is integrated with other language skills.	☐	☐	☐

ITEMS FOR ASSESSING WRITING	TRUE	FALSE	DON'T KNOW
19. Writing exams that only include language use is enough to understand learners' writing ability.	☐	☐	☐
20. Clarity in task demand is a required factor in the writing tasks for learners.	☐	☐	☐
21. Writing tasks should be prepared as much authentic as possible.	☐	☐	☐
22. Validity may not be sustained if writing tasks measure learners' creativity and imagination.	☐	☐	☐
23. In writing tasks, the notes that guide learners about what to do in the task must include full sentences so that they can understand how to carry out the task.	☐	☐	☐
24. Feedback is necessary to provide learners with opportunities to close the gap between their present and desired performances.	☐	☐	☐
25. Providing feedback after carrying out the writing tasks not only helps teachers collect information about learners but also helps them carry out remedial teaching if necessary.	☐	☐	☐
26. Giving learners an opinion and asking them to discuss it is a valid way of assessing their writing skills.	☐	☐	☐
27. Using visuals which guide learners for writing poses a problem.	☐	☐	☐
28. Writing skill might be assessed indirectly through multiple choice questions.	☐	☐	☐
29. Providing a reading text is a way of assessing writing skills.	☐	☐	☐
30. Portfolio is considered as one of the good instruments for writing assessment.	☐	☐	☐
31. Self-assessment is a required tool to assess learners' writing skills.	☐	☐	☐
32. Non-interactive writing tasks require metacognitive skills (e.g. understanding the goal, what they need to do to complete the task etc.) while interactive tasks do not necessarily require.	☐	☐	☐

ITEMS FOR ASSESSING SPEAKING

		TRUE	FALSE	DON'T KNOW
1.	When the interlocutor does not understand the learner, giving that feeling or saying it poses a problem.	☐	☐	☐
2.	Giving learners one task is enough to assess speaking skills.	☐	☐	☐
3.	Interlocutors' showing interest by verbal and nonverbal signals poses a problem.	☐	☐	☐
4.	When it becomes apparent that the learner cannot reach the criterion level during the speaking exam, the task is ended.	☐	☐	☐
5.	In interlocutor-learner interviews, the teacher has the chance to adapt the questions being asked.	☐	☐	☐
6.	In interactive tasks, more than two learners pose a problem.	☐	☐	☐
7.	In a speaking exam, production and comprehension are assessed together.	☐	☐	☐
8.	Carrying out the interview in a quiet room with an acoustic may not help learners focus on the questions.	☐	☐	☐
9.	Whether the interviewers or raters are trained for the speaking exams or not does not affect the reliability of the scoring.	☐	☐	☐
10.	In productive performance, the oral stimulus does not have to be particular so that learners can give diverse answers.	☐	☐	☐
11.	The results of the speaking assessment should be communicated to learners and parents.	☐	☐	☐
12.	Speaking assessment of learners should take place in naturalistic rather than contrived situations.	☐	☐	☐
13.	While the instrument for assessing speaking is being improved, assessment should be verified in terms of providing inter-rater reliability.	☐	☐	☐

14. A confirmation study is not necessary to be conducted when the instrument is first being developed.	☐	☐	☐
15. Feedback should be adapted according to the audience.	☐	☐	☐
16. Reading aloud is a technique used to assess speaking skills.	☐	☐	☐
17. Asking learners to repeat a word, phrase or a sentence is a way of assessing speaking skills.	☐	☐	☐
18. Discussion among learners is a way of assessing speaking skills.	☐	☐	☐
19. When the focus is to assess speaking skills, role plays are used.	☐	☐	☐
20. In peer interaction, random matching is avoided.	☐	☐	☐
21. Structured speaking tasks are generally implemented to assess linguistic features of a language, especially pronunciation and grammar.	☐	☐	☐

ITEMS FOR ASSESSING SPEAKING

	TRUE	FALSE	DON'T KNOW
22. Short answer questions try to measure learners' comprehension of the questions asked and ability to provide relevant information in response.	☐	☐	☐
23. In the assessment of speaking, appropriate speaking assessment methods should be chosen for instructional plans.	☐	☐	☐
24. Speaking assessment should include teachers' observations of the learners' oral language productions in the classes.	☐	☐	☐
25. Mimicry is one of the techniques used to assess speaking skills of learners.	☐	☐	☐
26. Oral presentation, when it is recorded, may decrease the stress level of learners.	☐	☐	☐
27. Self-assessment is reliable even if it is used in the early stages.	☐	☐	☐
28. Paired interaction tasks in speaking exams may result in domination of one learner over another.	☐	☐	☐
29. Peer-assessment is a way of assessing learners' speaking skills.	☐	☐	☐
30. Speaking skill can be assessed through an indirect test such as close tests.	☐	☐	☐
31. Phoneme discrimination is an example of direct testing for assessing speaking.	☐	☐	☐
32. Using holistic and analytic scales at the same time poses a problem.	☐	☐	☐
33. The interlocutor gives the score when the learner is in the exam room.	☐	☐	☐
34. A checklist is a means of scoring oral presentations in in-class assessment.	☐	☐	☐

35. Analytic rubrics provide detailed, focused, and precise assessment of speaking skill.	☐	☐	☐
36. Only teachers who have been trained in practising, scoring, and interpreting the rating scale should be responsible for assessment of speaking.	☐	☐	☐
37. Analytic scoring provides teachers with more beneficial diagnostic information about learners' speaking skills.	☐	☐	☐
38. Analytic scoring does not provide feedback for learners.	☐	☐	☐
39. Grammatical competence does not contribute to learners' fluency.	☐	☐	☐
40. Strategic competence helps learners use the language appropriately and effectively.	☐	☐	☐

APPENDIX 5: THE FINAL VERSION OF THE QUESTIONNAIRE AFTER THE PILOT STUDY

LANGUAGE ASSESSMENT KNOWLEDGE SCALE INFORMED CONSENT FORM

This present study is conducted for a master thesis work titled "EFL Instructors' Language Assessment Literacy Level in Assessing Productive Language Skills and Their Perceptions on Assessing Productive Language Skills through Formative Assessment" by Selda Arslan and her thesis advisor Assist. Prof. Sevgi Şahin. For this study, the ethics committee approval from Başkent University has been received. The study aims to measure the EFL instructors' perceived level of language assessment literacy in assessing productive language skills and to obtain their perceptions on assessing productive language skills through formative assessment. This is the first part of the study and it focuses on measuring the perceived level of language assessment literacy of EFL instructors in assessing productive language skills via a questionnaire. Participation in the study must be on a voluntary basis. No personal identification information is required in the questionnaire only institutional e-mail addresses will be collected for the second part of the study. Your answers will be kept anonymous and strictly confidential and they will be evaluated only by the researchers; the obtained data will be used for scientific purposes only. We expect you to read the questions and statements carefully and give an answer. It will take you approximately 20 minutes to complete the questionnaire. The questionnaire does not contain any questions that may give rise to discomfort in the participants. However, during participation, for any reason, if you feel uncomfortable, you are free to quit at any time and if you do so, please inform the researcher. We would like to thank you in advance for your contribution to this study.

I am told what the study is about and I am participating in this study totally on my own will and I am aware that I can quit participating at any time I want. I also know that my responses will be kept anonymous and used for only the study. I give my consent for the use of the information I provide for scientific purposes.

☐ I agree to participate this study.

LANGUAGE ASSESSMENT KNOWLEDGE SCALE

PART 1: DEMOGRAPHIC INFORMATION

In this part of the questionnaire, first, the demographic information (gender, age, your educational background, the teaching experience and so) will be gathered from you as participants.

1. Gender: Female ☐ Male: ☐
2. Age:
3. The university you are currently working:
4. Which university did you graduate from (BA):
5. Which department did you graduate from (BA)?   English Language Teaching ☐   English Language and Literature ☐   American Culture and Literature ☐   Translation and Interpretation ☐   Linguistics ☐   Others (Please specify): ☐
11. Which university did you graduate from / are you currently studying (MA)? (If you do not hold an MA degree or if you are not studying, please leave it blank.)
12. Which department did you graduate from/are you currently studying (MA)? (If you do not hold an MA degree or if you are not studying, please leave it blank)   English Language Teaching ☐   English Language and Literature ☐   American Culture and Literature ☐   Translation and Interpretation ☐   Linguistics ☐   Others (Please specify): ☐   I have not graduated yet ☐   (If you have not graduated yet, please specify your department by selecting one of the choices together with selecting this choice)

13. Which university did you graduate from/are you currently studying (PhD)? (If you do not hold a PhD degree or if you are not studying, please leave it blank)

14. Which department did you graduate from/are you currently studying (PhD)? (If you do not hold a PhD degree or if you are not studying, please leave it blank)

English Language Teaching   □

English Language and Literature   □

American Culture and Literature   □

Translation and Interpretation   □

Linguistics   □

Others (Please specify):   □

I have not graduated yet   □

(If you have not graduated yet, please specify your department by selecting one of the choices together with selecting this choice)

15. How long have you been teaching English:

16. Have you received any courses about testing and assessment English or testing and assessment in general in BA?

     Yes □   No □

17. The total number of the courses about testing and assessment English or testing and assessment in general you have received in BA (If you have not received any courses, then please write "0"):

18. Have you received any courses about testing and assessment English or testing and assessment in general in MA?

     Yes □   No □

19. The total number of the courses about testing and assessment English or testing and assessment in general you have received in MA (If you have not received any courses, then please write "0"):

20. Have you received any courses about testing and assessment English or testing and assessment in general in PhD?

     Yes □   No □

21. The total number of the courses about testing and assessment English or testing and assessment in general you have received in PhD (If you have not received any courses, then please write "0"):

22. Have you received any courses about testing and assessment English or testing and assessment in general through pedagogic formation, CELTA and/or DELTA?   Yes ☐   No ☐
23. If yes, which one/s offered a course/s about testing and assessment English or testing and assessment in general?   ☐ Pedagogic Formation   ☐ CELTA   ☐ DELTA
24. The total number of the courses about testing and assessment English or testing and assessment in general you have received through pedagogic formation (If you have not received any courses, then please write "0")
25. The total number of the courses about testing and assessment English or testing and assessment in general you have received in CELTA (If you have not received any courses, then please write "0")
26. The total number of the courses about testing and assessment English or testing and assessment in general you have received in DELTA (If you have not received any courses, then please write "0")
27. Have you worked in one of these units in the preparatory school?   CDU (Curriculum Development Unit)   ☐   TDU (Testing Development Unit)   ☐   TTDU (Teacher Training Development Unit)   ☐   No, I haven't.   ☐
28. Please write your institutional e-mail address:

LANGUAGE ASSESSMENT KNOWLEDGE SCALE (LAKS)

This scale is intended to measure the perceived level of language assessment literacy - in assessing productive language skills- of English language instructors working at preparatory schools. In the scale, there are both True and False statements. Your names and information will be kept anonymous. Thank you for sparing your valuable time.

PART 2: ITEMS FOR ASSESSING WRITING: This part of the questionnaire includes items for assessing writing skill. In the questionnaire, there are both True and False statements. Thank you for sparing your valuable time to fill out the questionnaire.

ITEMS FOR ASSESSING WRITING	TRUE	FALSE	DON'T KNOW	NOT SURE
1. Giving two options to learners and asking them to write about one ensure reliable and valid scoring.	☐	☐	☐	☐
2. Analytic scoring is used to see the strengths and weaknesses of learners.	☐	☐	☐	☐
3. The parts of a scoring scale and the scores in each part do not change for different levels of learners.	☐	☐	☐	☐
4. When there is a huge difference between the scores of the two raters, the same raters score the written work again.	☐	☐	☐	☐
5. Holistic scoring is used to see whether the learner is proficient or not at the end of the term.	☐	☐	☐	☐
6. Analytic scoring leads to greater reliability than holistic scoring in writing.	☐	☐	☐	☐
7. Classroom assessment of learning in terms of writing is carried out in a more reliable way through analytic scoring rather than holistic scoring.	☐	☐	☐	☐
8. Mechanical errors (e.g. spelling and punctuation) are dealt with in the assessment of later stages of a written work.	☐	☐	☐	☐
9. While scoring a writing test, the expectation of raters does not have an influence.	☐	☐	☐	☐
10. The complex relation among writing task variables, rater variables and scoring criteria does not have any effects on the scoring the writing tests.	☐	☐	☐	☐
11. While planning a scoring rubric for writing, factors including who will use the rubric and what aspects	☐	☐	☐	☐

of writing are crucial are not taken into consideration.				
12. Learners are required to write about at least 2 tasks in the exam rather than 1 task.	☐	☐	☐	☐
13. Giving restrictive prompt/guidelines for the writing task is avoided.	☐	☐	☐	☐
14. In controlled writing, learners have the chance to convey new information.	☐	☐	☐	☐
15. Irrelevant ideas are ignored in the assessment of initial stages of a written work in process writing.	☐	☐	☐	☐
16. In the assessment of writing, the assessment method should be chosen according to instructional decisions.	☐	☐	☐	☐
17. Utilising the results of writing assessment, while making decisions about learners and teaching plans, carries great importance.	☐	☐	☐	☐
18. Valid and reliable assessment of writing can be best ensured when it is integrated with other language skills.	☐	☐	☐	☐
19. Writing exams that only include language use is enough to understand learners' writing ability.	☐	☐	☐	☐
20. Clarity in task demand is a required factor in the writing tasks for learners.	☐	☐	☐	☐
21. Writing tasks should be prepared as much authentic as possible.	☐	☐	☐	☐
22. Validity may not be sustained if writing tasks measure learners' creativity and imagination.	☐	☐	☐	☐
23. In writing tasks, the prompts that guide learners about what to do in the task should include full sentences so that they can understand how to carry out the task.	☐	☐	☐	☐

24. Feedback is necessary to provide learners with opportunities to close the gap between their present and desired performances.	☐	☐	☐	☐
25. Providing feedback after carrying out the writing tasks not only helps teachers collect information about learners but also helps them carry out remedial teaching if necessary.	☐	☐	☐	☐
26. Giving learners an opinion or a topic and asking them to discuss that opinion is a valid way of assessing their writing skills.	☐	☐	☐	☐
27. Using visuals which guide learners for writing poses a problem.	☐	☐	☐	☐
28. Writing skill might be assessed indirectly through multiple choice questions.	☐	☐	☐	☐
29. Providing a reading text is a way of assessing writing skills.	☐	☐	☐	☐
30. Portfolio is considered as one of the good instruments for writing assessment.	☐	☐	☐	☐
31. Self-assessment is a required tool to assess learners' writing skills.	☐	☐	☐	☐
32. Non-interactive writing tasks (e.g. changing verbs in a paragraph from present to past tense) require metacognitive skills (e.g. understanding the goal, what they need to do to complete the task etc.) while interactive tasks (e.g. brainstorming) do not necessarily require.	☐	☐	☐	☐

LANGUAGE ASSESSMENT KNOWLEDGE SCALE (LAKS)

This scale is intended to measure the perceived level of language assessment literacy - in assessing productive language skills- of English language instructors working at preparatory schools. In the scale, there are both True and False statements. Your names and information will be kept anonymous. Thank you for sparing your valuable time.

PART 3: ITEMS FOR ASSESSING SPEAKING SKILL: This part of the questionnaire includes items for assessing speaking skill. In the questionnaire, there are both True and False statements. Thank you for sparing your valuable time to fill out the questionnaire.

ITEMS FOR ASSESSING SPEAKING	TRUE	FALSE	DON'T KNOW	NOT SURE
1. When the interlocutor does not understand the learner, giving that feeling or saying it poses a problem.	☐	☐	☐	☐
2. Giving learners one task is enough to assess speaking skills.	☐	☐	☐	☐
3. Interlocutors' showing interest by verbal and nonverbal signals poses a problem.	☐	☐	☐	☐
4. When it becomes apparent that the learner cannot reach the criterion level during the speaking exam, the task is ended.	☐	☐	☐	☐
5. In interlocutor-learner interviews, the teacher has the chance to adapt the questions being asked.	☐	☐	☐	☐
6. In interactive tasks, more than two learners pose a problem.	☐	☐	☐	☐
7. In a speaking exam, production and comprehension are assessed together.	☐	☐	☐	☐
8. Carrying out the interview in a quiet room with an acoustic may not help learners focus on the questions.	☐	☐	☐	☐
9. Whether the interviewers or raters are trained for the speaking exams or not does not affect the reliability of the scoring.	☐	☐	☐	☐
10. In productive performance, the oral stimulus does not have to be particular so that learners can give diverse answers.	☐	☐	☐	☐

11. The scores of the speaking assessment with feedback should be communicated to learners to provide them with feedback.	☐	☐	☐	☐
12. Speaking assessment of learners should take place in naturalistic rather than planned situations.	☐	☐	☐	☐
13. While the instrument for assessing speaking is being improved, assessment should be verified in terms of providing inter-rater reliability.	☐	☐	☐	☐
14. A confirmation study is not necessary to be conducted when the instrument is first being developed.	☐	☐	☐	☐
15. Feedback should be adapted according to the learners.	☐	☐	☐	☐
16. Reading aloud is a technique used to assess speaking skills.	☐	☐	☐	☐
17. Asking learners to repeat a word, phrase or a sentence is a way of assessing speaking skills.	☐	☐	☐	☐
18. Discussion among learners is a way of assessing speaking skills.	☐	☐	☐	☐
19. When the focus is to assess speaking skills, role plays are used.	☐	☐	☐	☐
20. In peer interaction, random matching is avoided.	☐	☐	☐	☐
21. Structured speaking tasks are generally implemented to assess linguistic features of a language, especially pronunciation and grammar.	☐	☐	☐	☐
22. Short answer questions try to measure learners' comprehension of the questions	☐	☐	☐	☐

asked and ability to provide relevant information in response.				
23. In the assessment of speaking, appropriate speaking assessment methods should be chosen for instructional plans.	☐	☐	☐	☐
24. Speaking assessment should include teachers' observations of the learners' oral language productions in the classes.	☐	☐	☐	☐
25. Mimicry (repetition of series sentences after the teacher or examiner) is one of the techniques used to assess speaking skills of learners.	☐	☐	☐	☐
26. Oral presentation, when it is recorded in the lesson, may decrease the stress level of learners.	☐	☐	☐	☐
27. Self-assessment is reliable even if it is used in the early stages.	☐	☐	☐	☐
28. Paired interaction tasks in speaking exams may result in domination of one learner over another.	☐	☐	☐	☐
29. Peer-assessment is a way of assessing learners' speaking skills.	☐	☐	☐	☐
30. Speaking skill can be assessed through an indirect test such as cloze tests.	☐	☐	☐	☐
31. Phoneme discrimination (a kind of activity carried out to show the difference between the sounds e.g., minimal pair practice (ship-sheep)) is an example of direct testing for assessing speaking.	☐	☐	☐	☐
32. Using holistic and analytic scales at the same time poses a problem.	☐	☐	☐	☐
33. The interlocutor gives the score when the learner is in the exam room.	☐	☐	☐	☐

34. A checklist is a means of scoring oral presentations in in-class assessment.	☐	☐	☐	☐
35. Analytic rubrics provide detailed, focused, and precise assessment of speaking skills.	☐	☐	☐	☐
36. Only teachers who have been trained in practising, scoring, and interpreting the rating scale should be responsible for the assessment of speaking skills.	☐	☐	☐	☐
37. Analytic scoring provides teachers with more beneficial diagnostic information about learners' speaking skills.	☐	☐	☐	☐
38. Analytic scoring does not provide feedback for learners.	☐	☐	☐	☐
39. Grammatical competence (understanding and using grammar, vocabulary, spelling etc. accurately) does not contribute to learners' fluency.	☐	☐	☐	☐
40. Strategic competence (appropriate use of communication strategies to compensate lack of knowledge in L2) helps learners use the language appropriately and effectively.	☐	☐	☐	☐

APPENDIX 6: STATEMENTS' CHANGES MADE IN THE ITEMS FOR ASSESSING WRITING PART AFTER THE PILOT STUDY

Items for Assessing Writing	True	False	Don't Know
4. When there is a disagreement between the scores of the two raters, they score the written work again (Former version)	☐	☐	☐
4. When there is a huge difference between the scores of the two raters, the same raters score the written work again. (Latter version).	☐	☐	☐
7. Classroom assessment of learning in terms of writing can be served better through analytic scoring rather than holistic scoring (latter version).	☐	☐	☐
18. Writing skill is best assessed when it is integrated with other language skills (former version).	☐	☐	☐
18. Writing skill can be assessed better when it is integrated with other language skills (latter version).	☐	☐	☐
23. In writing tasks, the notes that guide learners about what to do in the task must include full sentences so that they can understand how to carry out the task (former version).	☐	☐	☐
23. In writing tasks, the notes that guide learners about what to do in the task should include full sentences so that they can understand how to carry out the task (latter version).	☐	☐	☐
26. Giving learners an opinion and asking them to discuss it is a valid way of assessing their writing skills (Former version).	☐	☐	☐
26. Giving learners an opinion or a topic and asking them to discuss that opinion is a valid way of assessing their writing skills (latter version).	☐	☐	☐

APPENDIX 7: THE FREQUENCY OF THESE ASSESSMENT TOOLS IN THE CLASSROOM

Category	Codes	Frequency (Times mentioned)
Always	Every lesson	2
	Weekly	6
	Daily	1
	Every minute	1
Frequently	2-3 times a month	1
	Once in 2 units	1
Rare	3-4 times in a semester	1
	2-3 times in a semester	1

APPENDIX 8: ETHICS COMMITTEE APPROVAL

1993

BAŞKENT ÜNİVERSİTESİ
Akademik Değerlendirme Koordinatörlüğü

TS-EN-ISO 9001
KALİTE SİSTEM BELGESİ

BEND39822

Sayı : 62310886-302.14.03/
Konu : Tez Önerisi (Selda Arslan)

06/07/2020

EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

İlgi : 23/06/2020 tarih ve 17384 sayılı yazınız.

Enstitünüz İngiliz Dili Öğretimi Tezli Yüksek Lisans Programı öğrencisi Selda Arslan'ın, Dr. Öğretim Üyesi Sevgi Şahin danışmanlığında yürüteceği "İngilizce Öğretmenlerinin Üretici Dil Becerilerinin Değerlendirilmesinde Okur-Yazarlık Seviyesi ve Üretici Dil Becerilerini Biçimlendirici Değerlendirme Yoluyla Değerlendirmeye Dair Bakış Açıları" başlıklı tez önerisi değerlendirilmiş ve bilgilerinize ekte sunulmuştur.

e-imzalıdır
Prof. Dr. M. Abdülkadir VAROĞLU
Kurul Başkanı

Ek : Değerlendirme Formu

İlgili Makama

Üniversitemiz Eğitim Bilimleri Enstitüsü İngiliz Dili Öğretimi Tezli Yüksek Lisans Programı öğrencisi Selda Arslan'ın, Dr. Öğretim Üyesi Sevgi Şahin danışmanlığında yürüteceği, "İngilizce Öğretmenlerinin Üretici Dil Becerilerinin Değerlendirilmesinde Okur-Yazarlık Seviyesi ve Üretici Dil Becerilerini Biçimlendirici Değerlendirme Yoluyla Değerlendirmeye Dair Bakış Açıları" başlıklı tez önerisi değerlendirilmiş ve yapılmasında bir sakınca olmadığı tespit edilmiştir.
Bilgilerinize saygılarımızla sunarız.

Başkent Üniversitesi Sosyal ve Beşeri Bilimler ve Sanat Araştırma Kurulu

Ad, Soyad	Değerlendirme	İmza
Prof. Dr. M. Abdülkadir Varoğlu	Olumlu/~~Olumsuz~~	
Prof. Dr. Kudret Güven	Olumlu/Olumsuz	
Prof. Ali Sevgi	Olumlu/Olumsuz	
Prof. Dr. Işıl Bulut	Olumlu/Olumsuz	
Prof. Dr. Sadegül Akbaba Altun	Olumlu/~~Olumsuz~~	
Prof. Dr. Can Mehmet Hersek	Olumlu/~~Olumsuz~~	
Prof. Dr. Özcan Yağcı	Olumlu/Olumsuz	

Prof. Dr. Sadegül Akbaba Altun, Eğitim Bilimleri Enstitüsü İngiliz Dili Öğretimi Tezli Yüksek Lisans Programı öğrencisi Selda Arslan'ın, Dr. Öğretim Üyesi Sevgi Şahin danışmanlığında yürüteceği, "İngilizce Öğretmenlerinin Üretici Dil Becerilerinin Değerlendirilmesinde Okur-Yazarlık Seviyesi ve Üretici Dil Becerilerini Biçimlendirici Değerlendirme Yoluyla Değerlendirmeye Dair Bakış Açıları" başlıklı tezi yapabileceğini; ancak, araştırmacının uyarlayarak kullanacakları ölçeğin sahibinden izin alması gerektiği görüşündeler.

Prof. Dr. Özcan Yağcı, Eğitim Bilimleri Enstitüsü İngiliz Dili Öğretimi Tezli Yüksek Lisans Programı öğrencisi Selda Arslan'ın, Dr. Öğretim Üyesi Sevgi Şahin danışmanlığında yürüteceği, "İngilizce Öğretmenlerinin Üretici Dil Becerilerinin Değerlendirilmesinde Okur-Yazarlık Seviyesi ve Üretici Dil Becerilerini Biçimlendirici Değerlendirme Yoluyla Değerlendirmeye Dair Bakış Açıları" başlıklı tez önerisinin uygun olduğu kanaatlerini iletmişlerdir.

APPENDIX 9: RESEARCH APPROVAL (MIDDLE EAST TECHNICAL UNIVERSITY)
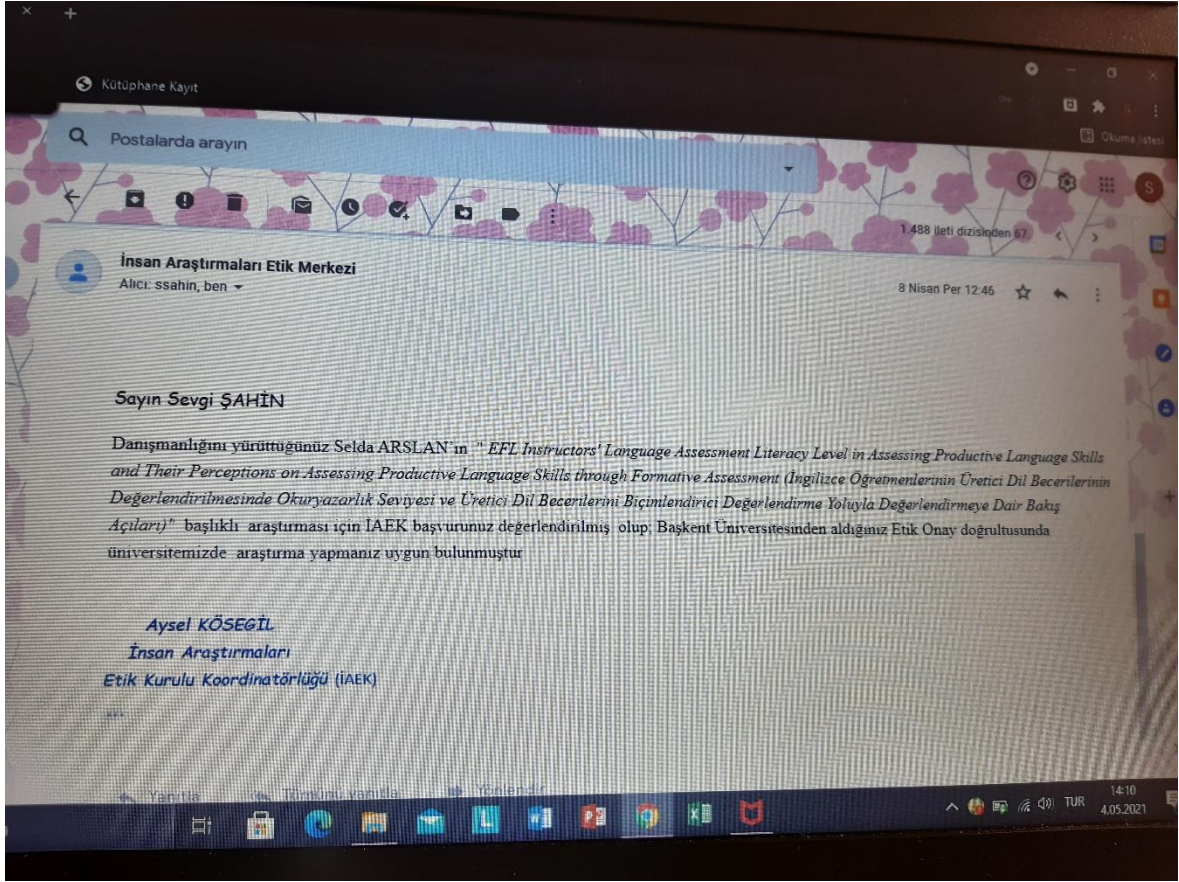
APPENDIX 10: RESEARCH APPROVAL (TED UNIVERSITY)

Suzan Kotan   8 Eki 2021 Cum 12:37 ☆ ↩ ⋮

Alıcı: ben, els ▾

Merhaba Selda Hocam,

çalışmanızı yapmanız uygundur. Size ne tür bir destek sağlamamızı istediğinizi iletirseniz, size elimizden gelen desteği verebiliriz.

İyi çalışmalar dilerim

Suzan Kotan

Director of English Language School

TED University

TED UNIVERSITY

...

APPENDIX 11: RESEARCH APPROVAL (BAŞKENT UNIVERSITY)

1993

BAŞKENT ÜNİVERSİTESİ
Yabancı Diller Yüksekokulu Müdürlüğü

05.02.2021

Sayı   : E-44447083-605.01-8984
Konu : Araştırma İzni (Selda Arslan)

EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

İlgi   :   02.02.2021 tarih ve 7443 sayılı yazınız.

İlgi yazınızda belirtilen Enstitünüz İngiliz Dili Öğretimi Tezli Yüksek Lisans programı öğrencisi Selda Arslan'ın "EFL Instructors' Language Assessment Literacy Level in Assessing Productive Language Skills and Their Perceptions on Assessing Productive Language Skills through Formative Assessment" adlı teziyle ilgili Yüksekokulumuzda görevli öğretim görevlileri ile çalışma yapması uygundur.
Bilgilerinize arz ederim.

Prof. Dr. Uğur GÜNGÖR
Müdür

Bu belge, güvenli elektronik imza ile imzalanmıştır.

APPENDIX 12: RESEARCH APPROVAL (ATILIM UNIVERSITY)

T.C.
ATILIM ÜNİVERSİTESİ REKTÖRLÜĞÜ

Sayı : E-59394181-605.01-2538
Konu : Araştırma İzni (Selda Arslan)

BAŞKENT ÜNİVERSİTESİ REKTÖRLÜĞÜNE
Eğitim Bilimleri Enstitüsü Müdürlüğü

Üniversiteniz İngiliz Dili Öğretimi Tezli Yüksek Lisans programı öğrencisi Selda Arslan'ın, Dr. Öğr. Üyesi Sevgi Şahin danışmanlığında hazırlamakta olduğu "EFL Instructors' Language Assessment Literacy Level in Assessing Productive Language Skills and Their Perceptions on Assessing Productive Language Skills through Formative Assessment (İngilizce Öğretmenlerinin Üretici Dil Becerilerinin Değerlendirilmesinde Okur-Yazarlık Seviyesi ve Üretici Dil Becerilerini Biçimlendirici Değerlendirme Yoluyla Değerlendirmeye Dair Bakış Açıları)" başlıklı tez çalışması kapsamında yapılması planlanan anket çalışması Üniversitemiz ilgili akademik birimi ile paylaşılmış olup çalışmaya katkı vermeyi uygun gören öğretim elemanları tarafından katılım sağlanabilecektir.

Bilgilerinizi arz ederim.

Prof.Dr. Serkan ERYILMAZ
Rektör Yardımcısı

Bu belge, güvenli elektronik imza ile imzalanmıştır.

Bu belge 5070 sayılı Elektronik İmza Kanununun 5. Maddesi gereğince güvenli elektronik imza ile imzalanmıştır.

APPENDIX 13: RESEARCH APPROVAL (OSTIM TECHNICAL UNIVERSITY)

T.C.
OSTİM TEKNİK ÜNİVERSİTESİ
Yabancı Diller Bölüm Başkanlığı

Sayı : E-33791099-100-1585   29.04.2021
Konu : Araştırma İzni

Sayın Prof. Dr. M. Abdulkadir VAROĞLU
Başkent Üniversitesi Rektör Yrd.

Başkent Üniversitesine bağlı Enstitünüzde, İngiliz Dili Öğretimi Tezli Yüksek Lisans programı öğrencisi Selda Arslan'ın, Dr. Öğr. Üyesi Sevgi Şahin danışmanlığında "EFL Instructors' Language Assessment Literacy Level in Assessing Productive Language Skills and Their Perceptions on Assessing Productive Language Skills through Formative Assessment (İngilizce Öğretmenlerinin Üretici Dil Becerilerinin Değerlendirilmesinde OkurYazarlık Seviyesi ve Üretici Dil Becerilerini Biçimlendirici Değerlendirme Yoluyla Değerlendirmeye Dair Bakış Açıları)" adlı teziyle ilgili olarak yapması gereken çevrim içi anket çalışmasını ve yarı yapılandırılmış görüşme çalışmalarını Üniversitemizin İngilizce Hazırlık Bölümünde görev yapan öğretim görevlileriyle gerçekleştirme isteği Bölüm Başkanlığımız tarafından onaylanmıştır.

Saygılarımla arz ederim.

Dr. Öğr. Üyesi Nesrin ERUYSAL
Bölüm Başkanı V.

Bu belge, güvenli elektronik imza ile imzalanmıştır.

APPENDIX 14: RESEARCH APPROVAL (HACETTEPE UNIVERSITY)

T.C.
HACETTEPE ÜNİVERSİTESİ REKTÖRLÜĞÜ
Rektörlük

Sayı : E-35853172-605-00001561760   3.05.2021
Konu : Selda ARSLAN (Uygulama İzni)

BAŞKENT ÜNİVERSİTESİ REKTÖRLÜĞÜNE

(Eğitim Bilimleri Enstitüsü Müdürlüğü)

İlgi: 10.03.2021 tarihli ve 605.01- 17802 sayılı yazınız.

Üniversiteniz Eğitim Bilimleri Enstitüsü İngiliz Dili Öğretimi tezli yüksek lisans öğrencisi Selda ARSLAN'ın Dr. Öğr. Üyesi Sevgi ŞAHİN danışmanlığında yürüttüğü "İngilizce Öğretmenlerinin Üretici Dil Becerilerinin Değerlendirilmesinde Okur-Yazarlık Seviyesi ve Üretici Dil Becerilerini Biçimlendirici Değerlendirme Yoluyla Değerlendirmeye Dair Bakış Açıları" adlı tez çalışmasının Üniversitemiz Yabancı Diller Yüksekokulu İngilizce Hazırlık Biriminde yapılması uygun bulunmuş olup, söz konusu anket bağlantısı Yüksekokulumuz öğretim görevlileriyle paylaşılmıştır.

Bilgilerinize arz ederim.

Prof. Dr. Mehmet Cahit GÜRAN
Rektör

Bu belge güvenli elektronik imza ile imzalanmıştır.

Bu belge 5070 sayılı Elektronik İmza Kanununun 5. Maddesi gereğince güvenli elektronik imza ile imzalanmıştır.

APPENDIX 15: RESEARCH APPROVAL (ANKARA UNIVERSITY)

T.C.
ANKARA ÜNİVERSİTESİ REKTÖRLÜĞÜ
Öğrenci İşleri Daire Başkanlığı

Sayı   : E-14267719-302.14.01-141358   16.06.2021

Konu   : Selda ARSLAN Hk.

BAŞKENT ÜNİVERSİTESİ REKTÖRLÜĞÜNE

İlgi   : a) 10.03.2021 tarihli ve 67284360-605.01.02.01-E.17818 sayılı yazınız.
          b) Üniversitemiz Yabancı Diller Yüksekokulu Müdürlüğü'nün 30.04.2021 tarihli ve
             E-90530290-302.14.01-112141 sayılı yazısı.

Üniversiteniz Eğitim Bilimleri Enstitüsü İngiliz Dili Öğretimi tezli yüksek lisans programı öğrencisi Selda ARSLAN'ın "EFL Instructors' Language Assessment Literacy Level in Assessing Productive Language Skills and Their Perceptions on Assessing Productive Language Skills Through Formative Assessment (İngilizce Öğretmenlerinin Üretici Dil Becerilerinin Değerlendirilmesinde Okur-Yazarlık Seviyesi ve Üretici Dil Becerilerini Biçimlendirici Değerlendirme Yoluyla Değerlendirmeye Dair Bakış Açıları)" adlı tez çalışmasında kullanılmak üzere anket uygulaması talebi hakkında Üniversitemiz Yabancı Diller Yüksekokulu Müdürlüğü'nden alınan 30.04.2021 tarih ve E-90530290-302.14.01-112141 sayılı yazı örneği ilişikte sunulmuştur.

Bilgilerinize saygı ile arz ederim.

Prof. Dr. Ayşen APAYDIN
Rektör a.
Rektör Yardımcısı

Ek: Yazı örneği (1 Sayfa)

Bu belge, güvenli elektronik imza ile imzalanmıştır.

171

Sayı   : E-90530290-302.14.01-112141   30.04.2021
Konu   : Selda ARSLAN Hk.

ANKARA ÜNİVERSİTESİ REKTÖRLÜĞÜNE
(Öğrenci İşleri Daire Başkanlığı)

İlgi   : Ankara Üniversitesi Rektörlüğü (Öğrenci İşleri Daire Başkanlığı)'nün 15.03.2021
          tarihli ve 14267719-302.14.01-E.66547 sayılı yazısı.

Başkent Üniversitesi Eğitim Bilimleri Enstitüsü İngiliz Dili Öğretimi tezli yüksek
lisans programı öğrencisi Selda ARSLAN'ın "EFL Instructors'Language Assessment Literacy
Level in Assessing Productive Language Skills through Formative Assessment (İngilizce
Öğretmenlerinin Üretici Dil Becerilerinin Değerlendirilmesinde Okur-Yazarlık Seviyesi ve
Üretici Dil Becerilerini Biçimlendirici Değerlendirme Yoluyla Değerlendirmeye Dair Bakış
Açıları) adlı tez çalışmasında kullanılmak üzere anket uygulaması talebi hakkındaki ilgi
yazınız alınmıştır.

Selda ARSLAN'ın söz konusu tez çalışmasında kullanılmak üzere Yüksekokulumuzda
anket uygulaması Müdürlüğümüzce uygun görülmüştür.

Gereğini bilgilerinize saygılarımla arz ederim.

Doç. Dr. Zeynep Zeren ATAYURT FENGE
Müdür

Bu belge, güvenli elektronik imza ile imzalanmıştır.

172

APPENDIX 16: RESEARCH APPROVAL (GAZI UNIVERSITY)

T.C.
GAZİ ÜNİVERSİTESİ
Yabancı Diller Yüksekokulu Müdürlüğü

Sayı : E-98579800-044-63702
Konu : Anketler (Selda ARSLAN)

REKTÖRLÜK MAKAMINA
(Öğrenci İşleri Daire Başkanlığı)

İlgi :   12.03.2021 tarihli ve 17311665-044- 50884 sayılı yazı.

İlgi yazınız ve ekleri incelenmiş olup, Selda Arslan'ın bahsi geçen tez çalışması kapsamında çevirimiçi anket çalışmasını ve yarı yapılandırılmış görüşme çalışmalarını Yüksekokulumuzda yapması Müdürlüğümüzce uygun görülmüştür.

Bilgilerinizi ve gereğini arz ederim.

Doç. Dr. Celal Turgut KOÇ
Yüksekokul Müdürü

Bu belge, güvenli elektronik imza ile imzalanmıştır.

173

APPENDIX 17: RESEARCH APPROVAL (UNIVERSITY OF TURKISH AERONAUTICAL ASSOCIATION)

T.C.
TÜRK HAVA KURUMU ÜNİVERSİTESİ
Yabancı Diller Bölüm Başkanlığı

Sayı   : E-40524177-605.01-3805   29.04.2021
Konu : Araştırma İzni (Selda ARSLAN)

BAŞKENT ÜNİVERSİTESİ REKTÖRLÜĞÜNE

İlgi   : Başkent Üniversitesi Rektörlüğü| 10.03.2021 tarihli ve 67284360-605.01.02.01-E.17815 sayılı yazısı.

İlgi yazıya istinaden, Başkent Üniversitesi, Eğitim Bilimleri Enstitüsü, İngiliz Dili Öğretimi Tezli Yüksek Lisans programı öğrencisi Selda Arslan'ın, Dr. Öğr.Üyesi Sevgi Şahin danışmanlığında yürütülen "EFL Instructors' Language Assessment Literacy Level in Assessing Productive Language Skills and Their Perceptions on Assessing Productive Language Skills through Formative Assessment (İngilizce Öğretmenlerinin Üretici Dil Becerilerinin Değerlendirilmesinde Okur-Yazarlık Seviyesi ve Üretici Dil Becerilerini Biçimlendirici Değerlendirme Yoluyla Değerlendirmeye Dair Bakış Açıları)" adlı tezine ilişkin çalışmaları Üniversitemiz, Yabancı Diller Bölümünde görev yapan öğretim görevlilerinin gönüllülük esasına dayanan katılımı ile yürütmesinin uygun olacağı düşünülmektedir.

Söz konusu hususu tensiplerinize arz ederim.

Prof. Dr. Hasan ERBAY
Rektör V.

Bu belge, güvenli elektronik imza ile imzalanmıştır.

174