



**T.R.
ONDOKUZ MAYIS UNIVERSITY
INSTITUTE OF GRADUATE STUDIES
DEPARTMENT OF FOREIGN LANGUAGE EDUCATION
ENGLISH LANGUAGE TEACHING PROGRAM**

**TEACHERS' PERSPECTIVES ON ASSESSING ENGLISH
SPEAKING SKILLS:
A POST-NEW EXAM MODEL INVESTIGATION**

Master's Thesis

Merve Nur DİNÇER

Supervisor
Assoc. Prof. Dr. Betül BAL GEZEGİN

SAMSUN
2024

**T.R.
ONDOKUZ MAYIS UNIVERSITY
INSTITUTE OF GRADUATE STUDIES
DEPARTMENT OF FOREIGN LANGUAGE EDUCATION
ENGLISH LANGUAGE TEACHING PROGRAM**



**TEACHERS' PERSPECTIVES ON ASSESSING ENGLISH
SPEAKING SKILLS:
A POST-NEW EXAM MODEL INVESTIGATION**

Master's Thesis

Merve Nur DİNÇER

Supervisor

Assoc. Prof. Dr. Betül BAL GEZEGİN

SAMSUN
2024

.ACCEPTANCE AND APPROVAL OF THE THESIS

The study entitled “**TEACHERS' PERSPECTIVES ON ASSESSING ENGLISH SPEAKING SKILLS: A POST-NEW EXAM MODEL INVESTIGATION**” prepared by **Merve Nur DİNÇER**, and supervised by Assoc. Prof. Dr. **Betül Bal Gezegin**, was found successful and unanimously accepted by committee members as Master thesis, following the examination on the date 13.6.2024

	Title Name SURNAME University Department/Art	Final Decision
Chairman	Assoc. Prof. Dr. Betül BAL GEZEGİN Ondokuz Mayıs University Department of Foreign Language Education	☒ Accept ☐ Reject
Member	Assoc. Prof. Dr. Ceylan YANGIN ERSANLI Ondokuz Mayıs University Department of Foreign Language Education	☒ Accept ☐ Reject
Member	Assoc. Prof. Dr. Şeyda SELEN ÇİMEN Muğla Sıtkı Koçman University Department of Foreign Language Education	☒ Accept ☐ Reject

This thesis has been approved by the committee members that already stated above and determined by the Institute Executive Board.

Prof. Dr. Ahmet TABAK
Head of Institute of Graduate Studies

DECLARATION OF COMPLIANCE WITH SCIENTIFIC ETHIC

I hereby declare and undertake that I complied with scientific ethics and academic rules in all stages of my Master's Thesis, that I have referred to each quotation that I use directly or indirectly in the study, and that the works I have used consist of those shown in the sources, that it was written in accordance with the institute writing guide and that the situations stated in the article 3, section 9 of the Regulation for TÜBİTAK Research and Publication Ethics Board were not violated.

Is Ethics Committee Necessary?

Yes ☒ (If it necessary, please add appendices.)

No ☐

15 /05 / 2024

Merve Nur DİNÇER

DECLARATION OF THE THESIS STUDY ORIGINALITY REPORT

Thesis Title : TEACHERS' PERSPECTIVES ON ASSESSING ENGLISH
SPEAKING SKILLS: A POST-NEW EXAM MODEL INVESTIGATION

As a result of the originality report taken by me from the plagiarism detection program on 15 /05 /2024 for the thesis titled above;

Similarity ratio : % 13

Single resource rate : % 2 has been released.

15/05/ 2024

Assoc. Prof. Dr. Betül BAL GEZEGİN

ÖZET

İNGİLİZCE KONUŞMA BECERİLERİNİN DEĞERLENDİRİLMESİNE İLİŞKİN ÖĞRETMENLERİN BAKIŞ AÇILARI: YENİ SINAV MODELİ SONRASI ARAŞTIRMASI

Merve Nur DİNÇER

Ondokuz Mayıs Üniversitesi

Lisansüstü Eğitim Enstitüsü

Yabancı Diller Eğitimi Ana Bilim Dalı

İngiliz Dili Eğitimi Programı

Yüksek Lisans, Haziran/2024

Danışman: Doç. Dr. Betül Bal Gezeğin

Bu araştırma, Türkiye Cumhuriyeti Milli Eğitim Bakanlığı (MEB) tarafından 2023-2024 eğitim-öğretim yılında uygulamaya konulan yeni düzenlemeler sonrasında ortaokul İngilizce öğretmenlerinin İngilizce konuşma becerilerinin değerlendirilmesine yönelik bakış açılarını incelemektedir. Yeni düzenlemelerle birlikte Türkiye'deki ortaokullarda öğrenciler ilk kez İngilizce konuşma sınavlarına girmeye başlamıştır. Bu çalışma, İngilizce öğretmenlerinin bu adaptasyon sürecinde sınav öncesi, sınav sırası ve sınav sonrası durumlara ilişkin bakış açılarını, duygularını, deneyimlerini ve düşüncelerini keşfetmeyi amaçlamaktadır. Karma araştırma deseninin kullanıldığı bu çalışmada, 11 öğretmenden açık uçlu sorularla bire bir yapılandırılmış görüşmeler yoluyla nitel veri ve 217 öğretmenden araştırmacı tarafından oluşturulan Likert tipi bir anket yoluyla nicel veri toplanmıştır. Çalışmanın sonucunda, öğretmenlerin konuşma sınavlarını önemli, gerekli ve faydalı buldukları ortaya çıkmıştır. Düzenlemelerden sonra öğretmenler konuşma etkinlikleri ve derslerde İngilizce kullanımı konusunda daha motive olduklarını belirtmişlerdir. Öğretmenler, konuşma sınavlarının ortaokullarda uygulanmaya başlaması ve uyum sürecinde meslektaşlarıyla iş birliği yapma konusunda genel olarak olumlu duygulara sahip olsalar da değişimin beklenmedik zamanlaması nedeniyle hazırlıksızlık, stres, tükenmişlik, yorgunluk ve umutsuzluk gibi olumsuz duygular da hissettiklerini belirtmişlerdir. Bu çalışmada ayrıca uyum sorunları, verimsiz sınavlar, öğretmenlerin artan iş yükü, işlevsel olmayan değerlendirme ve e-okul sistemine notların girişi ve öğretmenlerin yoğun programları gibi sorunlar da tespit edilmiştir. Konuşma sınavlarının geleceğine ilişkin olarak öğretmenler, bu sınavların devam edeceği ancak geliştirilmesi ve iyileştirilmesi gerektiği görüşündedirler. Genel olarak öğretmenler, bu sınavlarla ilgili sorunların zaman içinde çözülebileceğine inanmaktadırlar. Bu çalışmanın bulguları, İngilizce ölçme ve değerlendirme, müfredat geliştirme, eğitim planlama ve programlama, materyal geliştirme, öğretmenler için hizmet içi eğitim düzenlemeleri ve ortaokullarda konuşma sınavlarına ilişkin İngilizce öğretmeni adaylarının eğitimi için bir kaynak olarak hizmet edebilir.

Anahtar Sözcükler: İngilizce konuşma sınavları, Konuşma becerileri, Konuşma becerilerinin değerlendirilmesi, Öğretmenlerin bakış açısı, İngilizce Öğretmenleri

ABSTRACT

TEACHERS' PERSPECTIVES ON ASSESSING ENGLISH SPEAKING SKILLS: A POST-NEW EXAM MODEL INVESTIGATION

Merve Nur DİNÇER

Ondokuz Mayıs University

Institute of Graduate Studies

Department of Foreign Language Education

English Language Teaching Program

Master, June/2024

Supervisor: Assoc. Prof. Dr. Betül BAL GEZEGİN

This study investigates the secondary school English as Foreign Language (EFL) teachers' perspectives on assessing English speaking skills after the new regulations introduced in the 2023-2024 academic year by the Turkish Ministry of National Education (MoNE). With the new regulations, students started to take English speaking exams for the first time in secondary schools in Türkiye. The focus of the study is on EFL teachers' perspectives, feelings, experiences, and thoughts regarding pre-, during, and post-exam situations in this adaptation process. In this study, which used a mixed research design, qualitative data was gathered from 11 teachers via one-on-one semi-structured interviews with open-ended questions and quantitative data was gathered from 217 teachers with a Likert-type questionnaire created by the researcher. The analysis of the data revealed that EFL teachers find the speaking exams significant, necessary, and useful. Teachers reported to have become more motivated about speaking activities and using English in the lessons after the regulations. Although teachers generally had positive feelings about the introduction of speaking exams in secondary schools and collaborating with colleagues in the adaptation process, they also reported negative feelings such as unpreparedness, stress, burnout, fatigue, and hopelessness due to the unexpected timing of the change. This study also identified problems such as adaptation, ineffective exams, increased workload of teachers, impractical grading in the e-school system, and teachers' busy schedules. For the future of speaking exams, teachers have the idea that these exams will continue but should be developed and improved. In general, teachers believe that the problems with these exams can be solved over time. The findings of this study can serve as a resource for English testing and evaluation, curriculum development, educational planning and programming, material development, in-service training arrangements for teachers, and the training of pre-service EFL teachers regarding speaking exams in secondary schools.

Keywords: English speaking exam, Speaking skills, Assessment of speaking skills, Teachers' perspectives, English language teachers

ACKNOWLEDGEMENT

I would like to thank everyone who supported and advised me during the development of this thesis. First of all, I am grateful to my dear advisor, Assoc. Prof. Dr. Betül Bal Gezeğin, for the light she shed on my path with her knowledge and experience throughout this academic journey. I learned a lot from her until the end of our work, and I am completing the master's program with the professional development I aimed for.

I would also like to express my gratitude to my family. I owe endless thanks to my mother, father, and little daughter, Dünya. I could not have completed this work without their great sacrifice, understanding, and help. In addition, my cats Ateş Efe and Yaşa Nur (who unfortunately could not see the completion of my thesis) warmed me with their pure love throughout my challenging process of academic growth and contributed greatly to this thesis.

Last but not least, I want to thank me for believing in me, doing all this hard work, having no days off, and never quitting.

Finally, I dedicate this thesis to my beloved daughter, Dünya. I hope one day you will be proud of your mother, who loves you as much as the world.

Merve Nur DİNÇER

CONTENTS

ACCEPTANCE AND APPROVAL OF THE THESIS	i
DECLARATION OF COMPLIANCE WITH SCIENTIFIC ETHIC.....	ii
DECLARATION OF THE THESIS STUDY ORIGINALITY REPORT	ii
ÖZET	iii
ABSTRACT	iv
ACKNOWLEDGEMENT	v
CONTENTS	vi
SYMBOLS AND ABBREVIATIONS	vii
FIGURES LEGENDS	viii
TABLES LEGENDS	ix
1. INTRODUCTION	1
1.1. Background of the Study	1
1.2. Statement of the Problem	4
1.3. Purpose of the Study	5
1.4. Significance of the Study	5
1.5. Research Questions	5
2. LITERATURE REVIEW	7
2.1. Assessment and Testing in English Language Teaching	7
2.1.1. Principles of Assessment	9
2.1.1.1. Reliability	10
2.1.1.2. Validity	10
2.1.1.3. Practicality	12
2.1.1.4. Authenticity	12
2.1.1.5. Washback	12
2.1.2. Types of Assessment	13
2.1.2.1. Formative and Summative Assessment	13
2.1.2.2. Traditional and Alternative Assessment	14
2.1.3. Types of Tests	15
2.1.3.1. Achievement Tests	16
2.1.3.2. Proficiency Tests	16
2.1.3.3. Aptitude Tests	17
2.1.3.4. Diagnostic Tests	17
2.1.3.5. Placement Tests	18
2.2. Speaking Skills in English Language Teaching	18
2.2.1. Assessment of Speaking Skills	19
2.2.1.1. Types of Oral Assessments	20
2.2.1.2. Discrete Point vs Integrated Tests	23
2.2.1.3. Feedback in Oral Assessment	24
2.2.1.4. Evaluating and Scoring Speaking Tasks	25
2.2.2. Communicative Language Testing and Speaking Skills	26
2.2.3. Speaking Skills and CEFR	30
2.3. The New Examination System, Regulations, and Curriculum by Turkish MoNE	33
2.4. Research on Speaking Skills in Other Countries	37
2.5. Research on Speaking Skills in Türkiye	39
3. METHODOLOGY	42
3.1. Research Design	42
3.2. Participants	43
3.3. Data Collections Instruments	46
3.3.1. Semi-Structured Interviews	46
3.3.2. Questionnaire	46
3.4. Data Analysis	47

3.5. Reliability and Validity.....	47
4. FINDINGS	49
4.1. The Interview Findings.....	49
4.1.1. Theme 1: Teachers' Overall Thoughts.....	51
4.1.1.1. Positive Thoughts Related to the Recent Speaking Exams	52
4.1.1.2. Negative Thoughts Related to the Recent Speaking Exams	58
4.1.2. Theme 2: Teachers' Feelings.....	63
4.1.2.1. Positive Feelings Related to the Recent Speaking Exams.....	63
4.1.2.2. Negative Feelings Related to the Recent Speaking Exams	65
4.1.3. Theme 3: Experiences and Observations	71
4.1.3.1. Problems and Limitations of the Recent Speaking Exams	71
4.1.3.2. Procedural Observations of the New Speaking Exams	77
4.1.4. Future Thoughts.....	80
4.1.4.1. Expectations, Beliefs, and Requests for Speaking Exams	80
4.1.4.2. Recommendations to Improve the Speaking Exams.....	83
4.2. The Questionnaire Findings	91
4.2.1. Part I.....	91
4.2.2. Part II	95
5. DISCUSSION	103
5.1. Teachers' Overall Thoughts	103
5.2. Teachers' Feelings.....	104
5.3. Experiences and Observations	105
5.4. Future Thoughts	107
6. CONCLUSION	109
6.1. Implications.....	109
6.2. Suggestions	110
6.3. Conclusion	110
REFERENCES.....	112
APPANDICES	117
CIRCULUM VITEA	127

SYMBOLS AND ABBREVIATIONS

CLT	: Communicative Language Teaching
CEFR	: Common European Framework of Reference for Languages
EAP	: English for Academic Purposes
EFL	: English as a Foreign Language
ELF	: English as Lingua Franca
ELT	: English Language Teaching
ESL	: English as a Second Language
FLE	: Foreign Language Education
IELTS	: International English Language Testing System
LF	: Lingua Franca
L2	: Second Language
MLAT	: Modern Language Aptitude Test
MoNE	: Ministry of National Education
PLAB	: The Pimsleur Language Aptitude Battery
TOEFL	: Test of English as a Foreign Language

FIGURES LEGENDS

Figure 1.1. The Effect of Exams on Avarage Grades.....	3
Figure 1.2. 4Cs of 21st Century Skills	4
Figure 2.1. Relationship Between Testing, Assesment, and Teaching	9
Figure 2.2. Types of Tasks in Speaking Assessment According to Bygate	20
Figure 2.3. Types of Tasks in Speaking Assessment According to Luoma	21
Figure 2.4. Task Types in Speaking Assessment According to Brown and Abeywickrama ..	22
Figure 2.5. Bachman's Communicative Proficiency Model	27
Figure 2.6. Oral Production and Interaction Activities	33
Figure 3.1. Graduate Degrees of the Respondents	44
Figure 3.2. Teaching Experiences of the Respondents.....	45
Figure 3.3. The Grades in Which the Respondents Teach.....	45
Figure 4.1. Four Themes Created from the Interview Data.....	49
Figure 4.2. The Themes and Categories Created from the Interview Data	50
Figure 4.3. The Codes for Positive Thoughts Related to the Recent Speaking Exams	53
Figure 4.4. The Codes for Negative Thoughts Related to the Recent Speaking Exams	59
Figure 4.5. The Codes for Positive Feelings Related to the Recent Speaking Exams	64
Figure 4.6. The Codes for Negative Feelings Related to the Recent Speaking Exams	65
Figure 4.7. The Codes for Problems and Limitations of the Recent Speaking Exams.....	72
Figure 4.8. The Codes for Procedural Observations of the New Speaking Exams.....	78
Figure 4.9. The Codes for Expectations, Beliefs, and Requests for Speaking Exams	81
Figure 4.10. The Codes for Recommendations to Improve the Speaking Exams.....	83
Figure 4.11. The Experience of Teachers Regarding Exam Preparation	92
Figure 4.12. The Experience of Teachers Regarding Exam Administration	92
Figure 4.13. The Percentages of Teachers' Choice Regarding Exam Techniques	93
Figure 4.14. The Percentage of Teachers Using a Rubric for the Recent Speaking Exams ..	94
Figure 4.15. The Percentage of Teachers' Criteria for Speaking Exams	95

TABLES LEGENDS

Table 2.1. Principles of Assessment.....	10
Table 2.2. Paradigm Shift in ELT	14
Table 2.3. Overall Oral Production	30
Table 2.4. Overall Oral Interaction	31
Table 2.5. Primary Education English Lesson Curricula Weekly Lesson Hours by Year	34
Table 2.6. The Comparison of Changes in the Examination System for English Courses....	35
Table 3.1. Demographic Information of the Participants in the Interview Part	43
Table 4.1. Codes, Categories, and Themes of the Interview Data	50
Table 4.2. The Questionnaire and Research Questions, the Research Themes	96



1. INTRODUCTION

“What is assessed becomes what is valued, which becomes what is taught.”

(McEwen, 1995, p. 42).

In this chapter, background of the study, statement of the problem, and purpose of the research will be explained in order. Then, significance of the study will be presented. The chapter will end with research questions.

1.1. Background of the Study

The needs and demands in education shift quickly with the fast-changing world. Expectations from children define national and international education policies and their application from preschool to twelfth grade (K-12) accordingly. In order to keep up with this rapid change, societies need to create more space for the 21st-century skills in K-12 education. In the last century, there has been a transition in educational approaches in language education from behaviorism to cognitivism and then to the communicative approach. This situation is reflected in language teaching as the need for learner autonomy (Holec, 1981), where students take responsibility for their own learning. Learner-centered education has gradually gained value in the last decade. For this reason, governments are constantly updating their education systems and trying to adapt the 21st-century skills to the system and make it more learner-centered. However, examining the improvements and innovations made over time is vital in ensuring that these processes are fit for purpose, efficient, and meaningful.

Türkiye has a sizeable young population directly affected by these innovations and developments. According to a report by Turkish Ministry of National Education (MoNE) in 2022, more than 19 million students receive formal education at the preschool, primary, and secondary education levels in Türkiye. The number of teachers working in schools within the scope of formal education is stated as 1 million 139 thousand in the 2022 academic year. When precisely calculated for secondary schools within the scope of this research, there are more than 5 million students and 376 thousand teachers in 18 thousand secondary schools in Türkiye (MoNE, 2022). While these enormous numbers keep growing, this report helps us understand how large the impact of the decisions taken and those that can be made on the masses.

Foreign Language Education (FLE) has been one of the most significant focal points in educating students since the first years of the Turkish Republic. Atatürk put importance on foreign language education to make use of it for our national benefit (Sarıçoban, 2012; Sarıçoban & Sarıçoban, 2012). The educational reforms carried out by the Turkish Ministry of National Education in the last quarter century have also enabled us to take steps in this direction so that English, which is accepted as the current lingua franca (LF) in the world, can be better taught to students throughout K-12 education. In this regard, teaching that covers all four language skills (reading, writing, listening, speaking) and the integration of communicative language teaching (CLT) into foreign language education has been tried to be carried out step by step over the years (Kırkgöz, 2007). However, while writing and reading skills were given importance for years, speaking and listening skills were neglected. While these skills were included in books and curriculum, the lack of exams to evaluate them caused them to remain in the background. However, according to CLT theory, all four skills are essential and need to be assessed and taught. Achieving these four main foreign language skills at a whole level by using the changing texture and learning opportunities of the age will pave the way for students to be in more dialogue, sharing, and development with the world at the tertiary level, to be able to access information and indeed to the progress of the country. In this context, Turkish MoNE has made arrangements and improvements in the English education curriculum at regular intervals so that English lessons at primary, secondary, and high school levels can be given at sufficient levels and hours. Over time, it has tried to change the exam systems as well as the curriculum.

At the beginning of the 2023-2024 academic year, secondary schools' English language course examination system was radically changed by MoNE. The examination system that teachers and students have been accustomed to for years, which is applied twice a term in secondary schools and assessing reading and writing skills, has been phased by additional examination sessions, including listening and speaking exams. Including all four basic skills in the evaluation phase constitutes one part of the exam changes. According to the new regulations, the main exam measuring reading and writing skills is calculated out of 100 points and affects the student's grade by 50%. The whole school must hold this exam with the participation of all classes on the same day and time, that is, as a grade-wide exam. With the new regulation, it has been decided to hold two different exam sessions apart from this exam: listening and

speaking. These exams are evaluated out of 100 points, but each affects the student's grade by 25% (Figure 1.1.). After the students complete all the exams, the average grade is calculated by the teachers and uploaded to the electronic grading system "e-school."

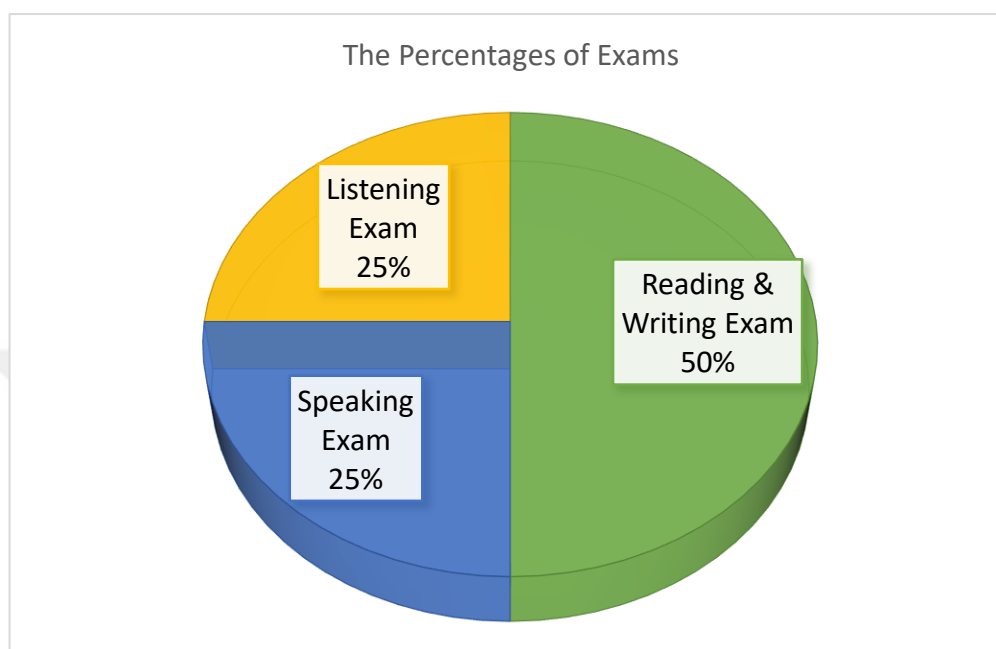


Figure 1.1. The Effect of Exams on Average Grades

With the new regulations, naturally, English department teachers started to prepare the exam questions by cooperating. In addition, the ministry has decided that teachers should eliminate multiple-choice questions in the exams and instead ask open-ended questions that do not contain options. The ministry shared sample exam scenarios with teachers for this transition. The ministry presented examples by showing the distribution of questions and functions by creating different scenarios with sample questions. With these arrangements, it is obvious that the 21st-century skills are underlined again.

21st-century skills are coded as the 4Cs of education: *communication, collaboration, critical thinking, and creativity* (Figure 1.2.). Acquiring these skills not only in English courses but also, in general, is the aim and expectation of the modern education system.

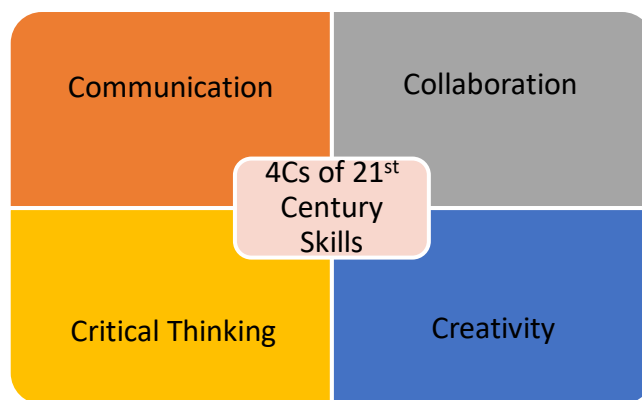


Figure 1.2. 4Cs of 21st Century Skills

Finally, the passing grade for the course is one of the changes made in this context. The course passing score, which was 45 out of 100 in previous years, was increased to 50 with the new regulations in 2023. These regulations in the examination system are new for teachers and students, so it is essential to examine them closely. For this purpose, this research aimed to make a detailed analysis of the change in the speaking exam model and exam content at the secondary school level from the teachers' perspective.

1.2. Statement of the Problem

The exam system change made by the Turkish MoNE at the beginning of the 2023-2024 academic year directly affected millions of secondary school students and thousands of English teachers. Obtaining comprehensive opinions about the exam from the teachers who prepare the speaking exams, which is a new dimension of the exam, can help obtain healthy data about these exam processes, reveal positive and negative situations about the state of the exam practices, the general perspectives of the teachers, and their motivation levels. Since these developments are new, EFL (English as Foreign Language) teachers' perspectives towards assessing speaking skills after the new exam model and regulations at secondary schools in Türkiye have not been examined in any previous study, to my knowledge. There is a need for studies in this field. Analyzing this exam, which concerns millions of students and thousands of EFL teachers, can positively contribute to the literature.

1.3. Purpose of the Study

The research aims to examine one part of the examination system introduced in the 2023-2024 academic year from the perspective of secondary school EFL teachers, focusing on "assessment and testing of speaking," one of the four language skills. Listening to the speaking examination system from teachers who closely observe every situation of the students and who prepare and administer the exams can help develop and better organize the systems for the following years by obtaining realistic data about these exams. For this reason, this study focuses on EFL teachers' perspectives toward assessing speaking skills after the new exam model and regulations at secondary schools in Türkiye and gives insight into teachers' perspectives.

1.4. Significance of the Study

This study was designed to evaluate the impact of the new regulations on teachers by obtaining answers from teachers about the new English exam system changes made in the 2023-2024 academic year. Although previous studies have been published in national and international sources on various subjects such as speaking exams, assessment of speaking skills, speaking anxiety, and building speaking skills, to my knowledge, the exact subject of this thesis has not been studied by any researcher before. The data obtained from this research can serve as a resource for English testing and evaluation, curriculum development, educational planning and programming, material development, in-service training arrangements for teachers, and the training of pre-service EFL teacher candidates. Thus, it can help improve in-service and pre-service EFL teachers' practices in assessing, preparing, and administering speaking exams and can be a guide for EFL teachers. Since this study deals with a necessary change that directly and indirectly affects a large part of society, it can be worth reference in the literature regarding the data it can provide.

1.5. Research Questions

In this study, where secondary school speaking exams and teachers' perspectives were centered, the answers to the following research questions were searched:

RQ1: What are the overall thoughts of EFL teachers towards the recent English speaking exams after the current regulations in the secondary school examination system in Türkiye?

RQ2: What are the feelings of EFL teachers related to the recent English speaking exams?

RQ3: What experiences were encountered by EFL teachers during the implementation of the recent English speaking exams?

RQ4: What are the future thoughts of EFL teachers related to the recent English speaking exams at secondary schools?

RQ1 is about teachers' overall thoughts related to the recent speaking exams at secondary schools. This research question examines what teachers think about the necessity of the new regulations and usefulness of the speaking exams, their positive or negative opinions towards them, and their readiness level for the changes. RQ2 deals more generally with teachers' and students' feelings. Motivation, satisfaction, self awareness, self-confidence, positive and negative feelings are the subjects of this question. RQ3 is about teachers' experiences after the new regulations based on the recent speaking exams. It examines problems, limitations, benefits, administration process, in-class situations, and observations related to students during the implementation of the new regulations. RQ4 deals with post-exam situations. It explores teachers' suggestions, recommendations, hopes, expectations, and predictions about upcoming exams. Teachers' belief in the continuity of the new system is also the subject of this question.

2. LITERATURE REVIEW

This chapter will introduce testing and assessment in English language teaching (ELT), the basic principles of assessment, and the types of tests. After that, speaking skills in ELT and their relation to the communicative approach, speaking skills in CEFR, and testing and assessment of speaking skills will be clarified. The new examination system, regulations, and curriculum by Turkish MoNE at secondary schools will be explained in detail. Lastly, to delve deeper into the literature on speaking skills, research on speaking skills in Türkiye and other countries will be provided.

2.1. Assessment and Testing in English Language Teaching

Assessment is a systematic approach to gathering information and making assumptions about the learners or the success of a teaching course based on different sources of evidence (Richards et al., 2002). It is undeniable that teaching and assessment are inseparable in education, even if the significance of assessment on behalf of language teaching has been known recently (Tosuncuoğlu, 2018). Assessment has two primary purposes: to help students learn and to report on what has been learned (Harlen, 2007). However, "the primary purpose of assessment is to collect relevant information to make inferences or decisions about individuals" (Tomlinson, 2005, p.42). Thus, assessment is vital in classifying and grading students and providing appropriate feedback while scaffolding new strategies to compensate for shortcomings in language teaching. With the help of assessment instruments, teachers can check learners' English proficiency and acquisition levels. By providing feedback about learners, assessment instruments help learners progress in return. Assessment allows teachers to make personal and introspective evaluations by asking how learning can be improved during the learning process. In this context, the assessment process affects both ELT teachers and learners mutually, and so it should be given significance.

Even if educators use them interchangeably, taking a closer look at the terms, assessment is different from evaluation. *Assessment* can be defined as a general component consisting of methods and techniques that inform teachers about the learners' ability, achievement, knowledge, and performance (Lambert & Lines, 2013). Assessment can be justified by specific weighted set goals (Taras, 2005). Harlen (2007) states that their processes are similar, but the kind of evidence and the basis of

judgment differ. Teachers' day-to-day classroom assessment is essential and can be improved to build responsible, engaged, and self-directed learners in an improving school environment (Chappuis & Stiggins, 2002). On the other hand, evaluation is a process that allows us to obtain information that determines whether a program achieves its purpose (Brown & Abeywickrama, 2021). Judgments related to courses, systems, programs, materials, course delivery, procedures, curriculum, or the process of making judgments are called evaluations, while judgments related to students' work are called assessments (Taras, 2005; Harlen, 2007).

Figure 2.1. indicates the relationship between assessment, testing, and teaching. Tests can serve pedagogical purposes, yet their primary function is to measure (Bachman & Palmer, 2015). Tests are not normally seen as facilities where new learning occurs, even if it is possible by receiving focused and constructive feedback during and after a test (Tomlinson, 2005). According to Tomlinson (2005), to ensure the learning validity of tests, authenticity, designing tests requiring communicative language ability tasks replicating real-life communication, positive washback, and self-assessment of learners are some highlights.

Analyzing the relationship between assessment and testing, the term "testing" is often associated with high-stakes standardized tests; meanwhile, the term "assessment" is used in a much broader sense. It is essential to examine the concepts individually to clarify them. First, there is a need to look at the meaning of testing and how it is defined in different sources. The Longman Dictionary of Language Teaching & Applied Linguistics identifies testing as "any procedure for measuring ability, knowledge, or performance" (Richards et al., 2002, p. 546), while Brown (1993) defines it as a method of measuring a person's ability or knowledge in a given area. He notes that testing is everywhere in our lives while emphasizing speaking skills:

Language learners are testing their newly acquired forms of language almost every time they speak. They devise hypotheses about how the language forms are structured and how certain functions are expressed in forms. On the basis of the feedback they receive, they make judgments and decisions. (p.252)

Even if this paragraph portrays informal tests of laypersons or individual learners, this study is about formal tests (newly introduced compulsory speaking

exams) in secondary schools and their reflections on teachers in the Turkish education system.

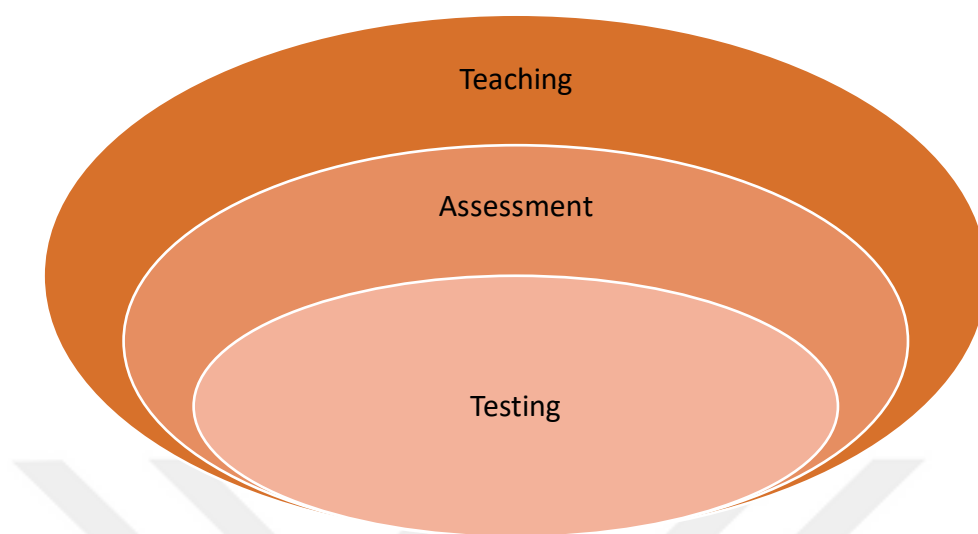


Figure 2.1. Relationship Between Testing, Assessment, and Teaching

Source. Adapted from “What does language assessment literacy mean to teachers?” by V. Berry, S. Sheehan, and S. Munro, 2019, *ELT Journal*, 73, p. 113–123.

The assessment literacy of Turkish EFL teachers is another dimension that should be examined after the new regulations by MoNE. Atay and Mede (2017) showcased that Turkish EFL teachers’ assessment literacy needed to be improved. They exemplified that teachers were not competent in preparing tests and giving feedback (informal, self, or peer feedback) on productive skills, especially in speaking. Even if they were comfortable with the microlinguistic aspect of ELT (grammar and vocabulary), which was possibly the result of the grammar instruction dominance in the Turkish ELT context, according to the research aforementioned, it was clear that teachers needed support in designing and managing test procedures.

2.1.1. Principles of Assessment

It is logical to get through criteria against which the test can be measured to evaluate the effectiveness of a test (Harmer, 2022). To ensure this, reliability, validity, practicality, authenticity, and washback can be analyzed under the characteristics of a good test. Reliability and validity are critical for tests and are referred to as essential measurement qualities that affect the usefulness of a test (Bachman & Palmer, 2015).

The criteria against which they are negotiated are required to evaluate the components of an assessment system. Table 2.1. briefly shows the criteria for evaluating student assessment procedures.

Table 2.1. Principles of Assessment

Principles	Associated Concepts
Reliability	Consistency
Validity	Assessment of intended objective
Practicality	Clear administrative details
Authenticity	Natural, real-world tasks
Washback	Positive, useful feedback to test taker

Source. Adapted from “*Teaching by principles: An interactive approach to language pedagogy*” by H. D. Brown, and H. Lee, 2015, Pearson Education.

2.1.1.1. Reliability

A good test needs to give consistent results, and this can be ensured by making the test instructions clear, providing restrictions for the scope of the answers, and, lastly, keeping test conditions constant (Harmer, 2022). “A reliable test is consistent and dependable” (Brown & Abeywickrama, 2021, p. 29). If you administer the same test to the same test-takers under the same conditions at two different times, the test should yield similar results (Brown, 1993; Harlen, 2007). As well as consistency from day to day, a reliable test also has clear directions for scoring or evaluation, has uniform rubrics, and encompasses unambiguous items (Brown & Abeywickrama, 2021). Students-related reliability, rater reliability, test administration reliability, and test reliability are types of considerations related to reliability. Sarıçoban (2011) states that Split-half, Kuder Richardson 20 and 21, and Cronbach’s can be used as the statistical methods to determine reliability.

2.1.1.2. Validity

Validity is the most essential quality of a test (Bachman, 1995). A test is considered valid if it tests what it is supposed to test (Harmer, 2022). A valid test measures what it is supposed to measure, and it gives the same results if it is repeated (Tosuncuoğlu, 2018), while “it does not measure irrelevant contaminating variables”

(Brown & Abeywickrama, 2021, p. 32). In other words, it includes everything students have learned, practiced, and trained in the class (Kadir et al., 2019). For example, testing general speaking ability with role-play activities where candidates have to perform tasks requiring specialist knowledge of English literature or science is invalid. Even if traditionally, it is classified into subcategories like content, criterion, and construct validity, validity is a unitary concept, and it always refers to the degree to which that evidence supports the inferences made from scores. “The estimation of reliability is an essential requisite of validation” (Bachman, 1995, p. 236).

- i. **Content Validity:** Sariçoban (2011) explains that the test should contain a representative sample of what is learned. For instance, if a teacher assesses a learner’s speaking ability in a conversational setting and asks the learner to answer a paper-pen test based on grammatical judgments, this lacks content validity (Brown & Abeywickrama, 2021).
- ii. **Construct Validity:** According to Bachman and Palmer (2015), it refers to the extent to which we can interpret a given test score as an indicator of the abilities or constructs we want to measure. Including irrelevant aspects is as threatening to construct validity as omitting relevant aspects in an assessment instrument (Harlen, 2007). For example, “the greater the relationship that can be demonstrated between a test of communicative competence in a language and the theory of communicative competence, the greater the construct validity of the test” (Richards & Schmidt, 2002, p.112).
- iii. **Face Validity:** Mousavi (2009) defines *face validity* as “the degree to which a test looks right and appears to measure the knowledge or abilities it claims to measure based on the subjective judgment of the examinees who take it, the administrative personnel who decide on its use, and other psychometrically unsophisticated observers” (p.247).
- iv. **Criterion Validity:** It is a “type of validity in regard to comparisons or correlations with an established external criterion measure” (Richards & Schmidt, 2002, p.133). To get criterion-related validity, we compare the results of the tests with the results of some criterion measures. These can be an existing

valid test, teacher's ratings, or independent assessments (Sarıçoban, 2011). "Concurrent validity and predictive validity" are two kinds of criterion-related validity.

2.1.1.3. Practicality

The practicality of a test is about the budget, time constraints, and human and material resources. A practical test must consider budget, timing, and clear directions and planning. It appropriately uses human and material resources (Brown & Abeywickrama, 2021). For instance, a proficiency test that lasts six hours is impractical, or a test that takes five minutes to administer but three hours to score.

2.1.1.4. Authenticity

Authenticity is the degree of correspondence between the characteristics of a given language test task and the features of a language task (Bachman & Palmer, 2015). An authentic test encompasses natural language as much as possible, has contextualized items, includes meaningful and relevant topics, and offers tasks that replicate real-world tasks (Brown & Abeywickrama, 2021). Authenticity creates interaction between the test-taker and the test task that essentially supports communicative language use or the negotiation of the meaning (Bachman, 1995). Using thematic organization in a contextual way with a natural language can help teachers prepare authentic tests (Brown, 1993).

2.1.1.5. Washback

Washback, also called backwash, in testing is "the impact of a test on classroom teaching or learning" (Richards et al., 2002, p.586). It is a complex phenomenon with intended and unintended effects, called positive, negative, or mixed washback (Alderson et al., 2017). Getting a grade, a numerical score, markings, percentages, and symbols is not necessarily constructive learning alone. After a test, learners should be able to get useful feedback that helps them create intrinsic motivation by arguing their strengths and weaknesses with the teacher (Brown, 1993; Chappuis & Stiggins, 2002). Even if it is challenging for overworked teachers to find time to give helpful feedback, especially in crowded classrooms, it is significant for learners to develop autonomy, self-confidence, and identity. As a result, teachers should comment generously and give diagnostic information about the exams so that test performance becomes a

fulfilling experience for learners (Brown & Lee, 2015). Furthermore, students must learn to self-assess to benefit from teacher feedback (Chappuis & Stiggins, 2002).

2.1.2. Types of Assessment

Different types of assessments can be distinguished according to their purpose, the period they cover, their content, and the way they are administered. The following paragraphs will revise formative, summative, traditional, and alternative assessment types.

2.1.2.1. Formative and Summative Assessment

Summative assessment is a summary of achievements at a particular time for keeping records and giving reports on progress to students, parents, and teachers, while formative assessment is about using evidence on students' achievements, the goals of learning, and the planning of the following steps to help learning and foster more profound engagement with it (Harlen, 2007). Taras (2005) clarifies the terms by asserting that formative assessment concentrates on assessing and using feedback, and summative assessment concentrates on the product. She points out that feedback is essential for an assessment to be formative, and learners make use of this information. However, she adds that a false separation has been created over time between summative and formative assessment because formative assessment, i.e., assessment for learning, is summative assessment plus feedback; they support each other in the classroom context, and there is no need for a dichotomy.

Even though learners improve with the help of constructive feedback during or after a test, they are still waiting for feedback from summative examinations, especially public summative examinations, which give little opportunity for useful learning and should provide more (Tomlinson, 2005). He underlines that formative tasks promote useful learning and do not create disadvantages for summative examinations. According to Chappuis and Richard (2002), classroom assessment can be a high-quality formative assessment that can boost student achievement. They support student involvement in assessment to take responsibility for and improve their learning. This way, students know where they are, communicate their status, make decisions, and plan their next step in their learning.

2.1.2.2. Traditional and Alternative Assessment

Assessment selection is an essential part of ELT's learning and teaching process. There is a paradigm shift from traditional to modern education systems worldwide (Table 2.2.). These new paradigms in education require more attention in choosing assessment format and preparing assessment procedures according to expectations from learners (Tosuncuoğlu, 2018).

Table 2.2. Paradigm Shift in ELT

Old teaching foreign language paradigm	New teaching foreign language paradigm
Focus on language	Focus on communication
Isolated skills	Integrated skills
Teacher-centered	Learner-centered
Emphasis on output	Emphasis on procedure
Tests only test	Tests to teach

Source. Adapted from “Importance of assessment in ELT” by I. Tosuncuoğlu, 2018, *Journal of Education and Training Studies*, 6, p. 163.

Table 2.2. clearly shows that language education is becoming more communication-focused, integrated skills are more in the spotlight, and classrooms are getting more learner-centered in language education. According to Table 2.2., the ongoing learning and teaching procedure is becoming critical, while tests are viewed as teaching instruments rather than tests only. This shift in applied linguistics and the shortcomings of traditional standardized paper-pen tests resulted in alternative assessments (Brown & Abeywickrama, 2021). The alternative assessment focuses on measuring learners' performance of the target language in real life in a specified period. They are continuous and long-term assessments. They require contextualized communicative tasks and are in harmony with 21st-century skills because learners need to be creative, communicative, and interactive. Alternative assessment can be applied in many different ways with a great variety of alternative assessment strategies. Projects, portfolios, concept maps, observation, drama, posters, interviews, and peer or self-assessments are some examples of alternative assessments. However, comparing traditional and alternative assessments, we cannot state that one is prior to the other. Choosing the appropriate assessment strategy and instrument is up to

teachers, considering many variables about learners, needs, budget, limitations, and facilities.

There can be some concerns about alternative assessment's validity, practicality, and effectiveness. Regarding this, Kadir, Zaim, and Refnaldi (2019) developed an instrument for evaluating the validity, practicality, and effectiveness of the authentic assessment of speaking skills at secondary schools, which can help teachers make better decisions and designs.

2.1.2. Types of Tests

A test measures a person's ability, knowledge, or performance in a given domain. Compared to the term assessment, tests are admitted as a subset of assessment (Brown & Abeywickrama, 2021). A test is a method and an instrument emphasizing the concept of method and measuring. Tests have identifiable scoring methods and administrative procedures applied systematically in specific periods (Brown & Lee, 2015).

One of the issues in testing is the dichotomy between norm-referenced and criterion-referenced tests. Norm-referenced tests interpret the scores of the testees concerning the other test takers, which is the mean (average score) (Richards et al., 2002). Norm-referenced tests place testees in rank order with a numerical score (e.g., 120 of 200). Standardized tests, such as the Test of English as a Foreign Language (TOEFL) and The International English Language Testing System (IELTS), are examples of norm-referenced tests. Besides, criterion-referenced tests measure testees' performance based on a particular standard or criteria.

In contrast to the norm-referenced tests, in criterion-referenced tests, testees' scores are interpreted concerning the criterion score that has already been agreed upon (Richards et al., 2002). Criterion-referenced tests are designed to provide testees with feedback. This feedback is in the form of grades according to the objectives of a specific curriculum, objectives, course, or lesson (Brown & Abeywickrama, 2021).

All in all, there are numerous tests with specific purposes and criteria for measuring. These tests are used in everyday practices in language teaching, certification, and gatekeeping (Chapelle et al., 2013).

2.1.2.1. Achievement Tests

An *achievement test* is an instrument designed to measure what a person has learned within or up to a given time (Davies, 2006). The results of achievement tests are used for decision-making processes at the end of a period, like a school term, a whole year, a unit, or an entire curriculum. Therefore, they are and should be limited to the particular material covered in the curriculum within a specific period (Brown & Abeywickrama, 2021). The primary concern of an achievement test is to decide whether or not the course objectives have been met. They are often summative because they are administered at the end of a specific period. However, they have a critical formative role in the learning process. Practical achievement tests provide quality feedback about the learners' performance (Brown & Lee, 2015).

2.1.2.2. Proficiency Tests

A proficiency test measures how much of a language has been learned by the testee. In this regard, it is about global competence in a language. Stansfield (2008) claims that the need to define language proficiency resulted from interest in testing communication as worldwide immigration and multicultural societies increased after the fall of the Soviet Union. A multilingual workforce worldwide, especially in Europe, North America, and Canada, has created a never-seen language learning interest, resulting in a need for convenient, reliable, and valid proficiency tests in multiple languages.

Proficiency tests are different from achievement tests. Proficiency tests are not limited to any single course or curriculum (Brown & Lee, 2015), and they measure the learners' general level of language mastery (Richards et al., 2002). The Test of English as a Foreign Language (TOEFL) and The International English Language Testing System (IELTS) are typical examples of international proficiency tests. These computer-based tests have sections for four skills with standardized multiple-choice items: reading, writing, speaking, and listening (Brown & Abeywickrama, 2021). Proficiency tests are almost always categorized as summative and norm-referenced tests. They play a gatekeeping role in the lives of millions of people aiming to level up for academic or job-related positions.

2.1.2.3. Aptitude Tests

Foreign language aptitude generally refers to the specific talent for learning a foreign or second language (Carroll, 1981; Skehan, 2002). The language aptitude test does not test language skills; instead, it examines the fundamental cognitive abilities that learners of a second language rely on second language learners (*Aptitude Test*, 2020). Language aptitude was studied fairly intensively during the 1950s and 1960s by Carroll (Carroll & Sapon, 1958) and Pimsleur (1966), who each produced test batteries designed to assess such a set of abilities. Dr. Carroll was the lead developer of the Modern Language Aptitude Test (MLAT), while Pimsleur had the Pimsleur Language Aptitude Battery (PLAB). Later on, many more aptitude tests were created with different purposes. Nevertheless, today, these tests are less popular than they once were.

PLAB and MLAT have some similarities and correlations. They correlate in assessing foreign language learning ability with memorization, puzzle-solving, and mimicry. However, these tests may not provide information about how communicative the language will be used and how much success will be achieved by testees in communication (Brown & Abeywickrama, 2021). Clearly, with appropriate learning involvement and motivation, everyone can eventually learn a language by using learning strategies. As a result, standardized aptitude tests are rarely used today.

2.1.2.4. Diagnostic Tests

Diagnostic tests are designed to provide information about L2 learners' strengths and weaknesses so they can be used to diagnose how much a language learner knows at the beginning of a language course (Richards et al., 2002). They provide insight into particular aspects of a language. For instance, a diagnostic grammar test can help instructors determine what should be added to the curriculum to provide better language learning opportunities. Achievement tests are not advised to be used as diagnostic tests because they are for the language features already taught. Diagnostic tests should be prepared specifically for their purpose (Brown & Lee, 2015). They need to elicit information on what students need to work on in the future. This way, the test-makers can determine the aspects of language that students need to develop or that a course should include (Brown & Abeywickrama, 2021).

2.1.2.5. Placement Tests

Placement tests are designed to place testees at an appropriate level in a program or course (Richards et al., 2002). Some achievement and proficiency tests can act as placement tests. They typically include a sampling of course material from the curriculum. The material should be avoided if it is too easy or challenging so the student can be at the appropriate level according to the test result. A practical placement test also functions as a diagnostic test (Brown & Abeywickrama, 2021).

2.2. Speaking Skills in English Language Teaching

Florez (1999) defines speaking as “an integrative process of constructing meaning both its form and meaning depend on the context, the participants, their experiences, the environment and the purpose for speaking” (p.2). While listening and reading are classified as receptive skills, speaking and writing are classified as productive skills. Productive skills require active involvement and mindfulness of learners in the process of using language. Learning to speak in another language is a demanding undertaking (Hinkel, 2006). As pointed out by Brown and Abeywickrama (2021), “Speaking is the production of creative construction, and the speaker has to choose lexicon, structure, and discourse while speaking” (p. 156).

English is the social language of the world as well as the academic language. From a sociolinguistic point of view, to communicate effectively without causing misunderstandings and disconnections and to share the intended message with others, it is necessary to produce English orally in a mindful way (Sudarmo, 2021). Brown and Yule (1983) assert that speaking is the skill that the students will be judged upon most in real-life situations, even if the language teaching has been mainly concerned with the written discourse. For example, a job interview with an undergraduate student or a commercial meeting between representatives of different countries is an essential threshold in which speakers are expected to demonstrate their ability to communicate and speak appropriately to manage what they intend. Thus, having the desired level of English-speaking skills puts individuals at the forefront of the competitive world. However, reading and writing used to be the center of language classes while primarily written discourse was used to get more attention (Carter & McCarthy, 1995), which caused failure in speaking and spontaneous language communication requirements in EFL countries like Türkiye. Apart from that, in the Turkish EFL setting, students are prejudiced towards speaking English traditionally because they lack the opportunity to

practice speaking outside the classroom, and this results in the demotivation of learners when combined with shyness and fear of doing wrong.

To produce language orally, learners need intensive practice and the knowledge of intonation, pronunciation, vocabulary, and grammar structures. Several activities can be used in English lessons To promote speaking, such as role-plays, question-answer games, information gap activities, sing-alongs, brain-storming, discussions, presentations, debates, and story-telling as a group or pair activity (Rao, 2019). Moreover, the examination system should integrate speaking assessment (Sakale, 2012). To make speaking meaningful and sociolinguistically oriented, fillers, pauses, turn-taking, stress, intonation, gestures, organization strategies, and repetitions should be considered while teaching speaking (Zulfugarova, 2018). Authenticity and meaningfulness of the activities in a learner-centered classroom play an essential role in the participation and interests of students of all ages.

According to Brown and Abeywickrama (2021), basic types of speaking performance can be imitative, intensive, responsive, interactive, or extensive (monologue) in the classroom:

- **Imitative:** the ability to simply imitate a part of a sentence
- **Intensive:** the production of short chunks of oral language
- **Responsive:** limited level of interaction and test comprehension during short conversations
- **Interactive:** longer interactions and stretches of oral production with multiple participants
- **Extensive (monologue):** extensive oral production such as story-telling

2.2.1. Assessment of Speaking Skills

Testing has been a neglected area in ELT for decades globally, especially testing productive skills such as writing and speaking (Sarıçoban, 2011). Language testing should include listening, speaking, reading, and writing skills as the samples of communicative competence (McNamara, 2017). Moreover, while doing that, “teachers should test learners writing skills by having them write and their speaking

by having them speak” (Köksal, 2004, p.4). Tests of oral language ability are widely used, and assessing oral language is an area of particular interest in applied linguistics (Chapelle et al., 2013).

2.2.1.1. Types of Oral Assessment

Considering different criteria, oral assessments can be classified into different categories. As a result, different types of classifications were made by Bygate (1987), Luoma (2004), Brown and Abeywickrama (2021), and Nakatsuhara et al. (2022) recently. Figures of different classifications are presented below.

According to Bygate (1987), oral assessment test tasks are divided into two main categories: informational and interactional tasks. Expository talk and evaluative talk are informational tasks (Figure 2.2.). Description, narration, instruction, and comparison are expository talk tasks. Explanation, justification, prediction, and decision are evaluative talk tasks. Interactional tasks include information gaps, role play, interviews, and conversations. They are more authentic and meaningful because learners need to communicate.

A. Informational Tasks

Expository Talk

- Description
- Narration
- Instruction
- Comparison

Evaluative Talk

- Explanation
- Justification
- Prediction
- Decision

B. Interactional Tasks

- Information gap
- Role play
- Interview
- Conversation

Figure 2.2. Types of Tasks in Speaking Assessment According to Bygate

Luoma (2004) classified types of tasks in speaking assessment into three categories, as shown in Figure 2.3. Even if they can be more challenging to score, open-ended tasks give more freedom to students. These tasks are productive and creative with unpredictable outcomes. On the contrary, structured tasks include shorter and predetermined answers. They are easier to score. They test learners' linguistic competence (pronunciation and grammar). Semi-structured tasks are in the middle. They are often more flexible than structured tasks. Learners have to adapt themselves to different social situations and roles, and they have to react accordingly.

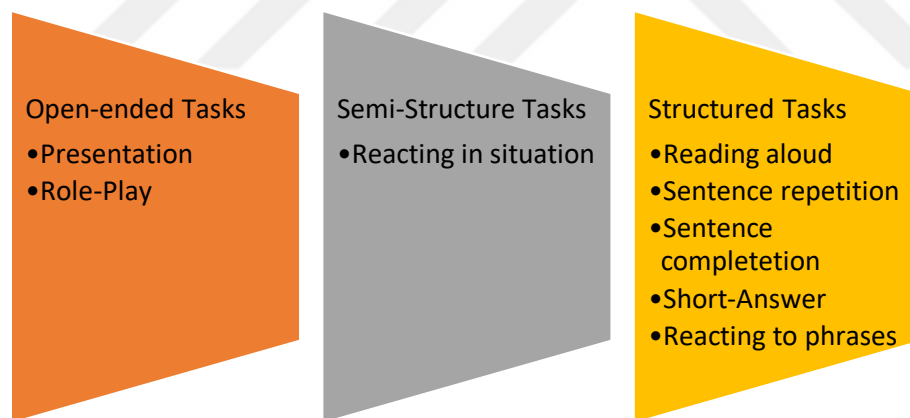


Figure 2.3. Types of Tasks in Speaking Assessment According to Luoma

Brown and Abeywickrama (2021) created six main categories to classify types of tasks in speaking assessment (Figure 2.4.). In imitative tasks, it is not necessary to negotiate or interact. Learners reproduce a word, phrase, or sentence. Pronunciation skills are essential. Teachers can score stress and intonation. Intensive tasks are limited response tasks. As students' production is controlled, scoring these tasks is easy. Interaction with the test administrators is minimal. Responsive tasks require

comprehension and interaction, even if it is limited. Learners comprehend the questions and answer them. They can be described as terse conversations. Interactive tasks require more complex interaction with a pair or group of participants. Learners negotiate and exchange meaning with authentic tasks such as role plays or interviews. Extensive tasks are monologues. Teachers are listeners and do not interfere. Delivering a speech, a presentation, or story-telling are prominent examples. It requires minimal interaction with listeners.

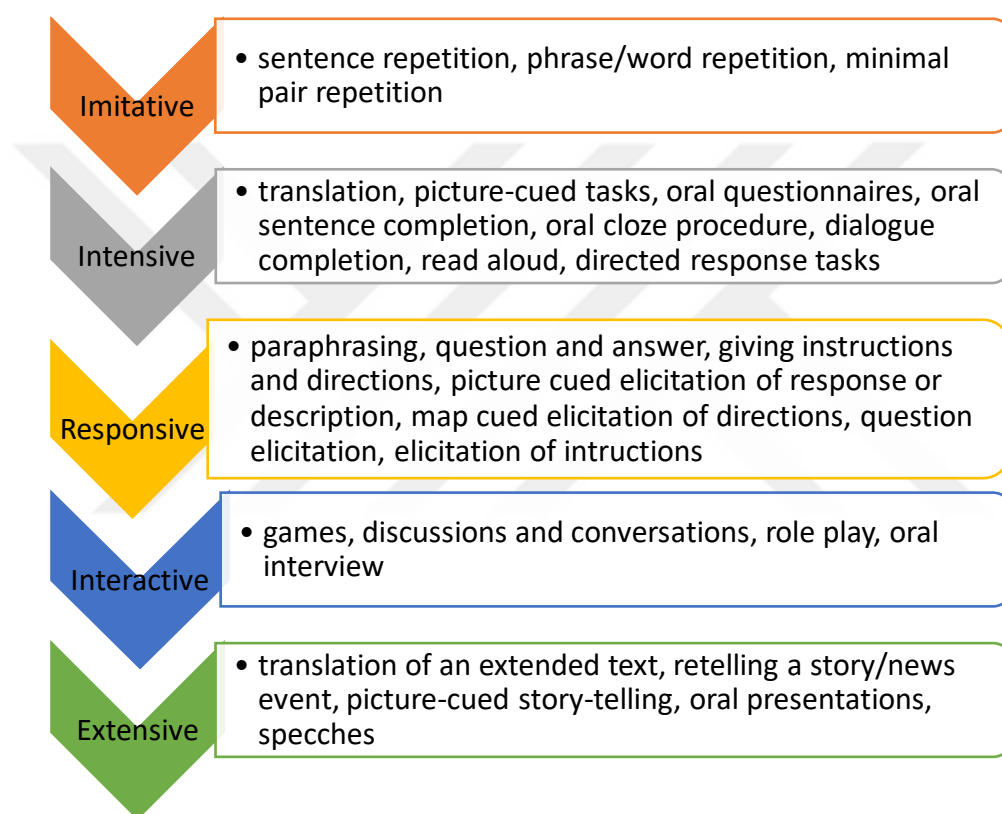


Figure 2.4. Types of Tasks in Speaking Assessment According to Brown and Abeywickrama

Nakatsuhara, Khabbazbashi, and Chihiro (2022) classified the most common speaking tasks into four categories:

- **Elicited Imitation Tasks:** An assessment task in which candidates listen to utterances and repeat them or read aloud written sentences
- **Independent Tasks:** An assessment task in which candidates draw on their ideas or knowledge to answer a question such as question & answer, picture description

- **Integrated Tasks:** An assessment task in which candidates speak about a topic with information from another source, such as computer-assisted language tests like TOEFL IBT.
- **Interactive Tasks:** An assessment task in which candidates are required to be paired or grouped to interact with one another.

2.2.1.2. Discrete Point vs Integrative Tests

According to Chapelle and Plakans (2013), test-makers have come up with combined skills regarding the limitations of the assessment of speaking, reading, writing, and listening individually. Such a four-skill view of language has had some limitations. As a result, test developers had to conceptualize an assessment of integrative skills involving a combination of skills. For instance, a test-taker can read a passage and then write about it or listen to a track and speak about it. This way of assessing language integratively underlines the success of communication and the test-taker's performance.

Assessing speaking skills is a complex process compared to other skills regarding the idiosyncratic characteristics of EFL learners (Ekmekçi, 2016). Four language skills (listening, speaking, reading, writing) and language areas like pronunciation, vocabulary, grammar, and translation can be tested discretely or integratively according to the aim of the language assessment (Harmer, 2022). However, some other researchers support the idea that native speaker norms like pronunciation, vocabulary, grammar, interaction, and fluency upon which the current language testing world (IELTS, TOEFL, and Pearson's Test of English) is built are no longer appropriate, especially for expanding circle countries (Kachru, 1985) like Türkiye (Chopin, 2015; Jenkins, 2020). According to these English as a Lingua Franca (ELF) researchers, communicative effectiveness should be assessed by focusing on performance instead of language norms and form-focused methods. Furthermore, Oller (1980) stated that discrete point tests could not express communicative competence. He proclaimed that language competence could not be tested separately due to its globalness; it requires integrated assessment.

2.2.1.3. Feedback in Oral Assessment

Feedback is one key to successful L2 learning. Quality and quantity of feedback is crucial. Teachers' manner, how they do it, and when they do it affects the quality of feedback. Affective quality feedback is supporting and encouraging, whether verbal or nonverbal. Regarding the quantity, planning when and how often teachers should give feedback is another critical issue. While treating errors, teachers should consider *global* and *local* errors (Brown & Lee, 2015). Hendrickson (1980) suggests that local errors (like slip of the tongue or minor mistakes) not be corrected. They do not interrupt the flow of communication, and the message is transparent to the listener. However, global errors should be treated because the message may be distorted. Even if there is a necessity for corrective feedback, it is not advisable to treat every error (Brown & Lee, 2015) not to discourage learners from speaking.

The preparation and administration process of oral assessment is also significant in administering a reliable and variable test. Some suggestions by Douglas and Fulcher (2014) are as follows:

- Planning and organization should be done carefully.
- Test tasks and assessment criteria should be decided beforehand.
- Test location and time should be arranged and announced to test-takers.
- The testing environment should be clean and comfortable.
- The timetable and the timing of the test should be available to test-takers.
- Teachers should be positive towards students.
- Both oral and written instructions (about the pre-, while, and post-test) should be provided to test-takers so that they are not confused about what to do.
- Materials should be checked before the administration of the test.
- The number of teachers should be enough to administer the test.
- Teachers need to be educated professionally to prepare and administer the test.
- Teachers should collect all the test components when the test is over.
- Scores should be announced after the tests.

- Feedback should be given after the scores are announced.

2.2.1.4. Evaluating and Scoring Speaking Tasks

Evaluating and scoring the speaking tasks demands careful planning as well. Scoring specifications must be cleared before the assessment procedures to avoid reliability problems (Brown & Abeywickrama, 2021). Thus, teachers should prepare and use rating scales (scoring rubrics) to score speaking tasks. They have short descriptions of different levels of language proficiency and serve as a guide for the rating process (Fulcher, 2014). Previously, different types of speaking tasks were presented. Teachers are free to choose and plan their speaking tests by using them. Nevertheless, according to the tasks chosen, the challenge of scoring changes. Brown and Abeywickrama state, "The more open-ended test tasks are, the greater the challenge in scoring as a result of the test taker's freedom of choice" (Brown & Abeywickrama, 2021, p. 157).

Analytics and holistic scales are rating scales (The Council of Europe, 2020). As it sounds, holistic scales give general information, while analytic scales present diagnostic information about the speaking ability of the test-taker. An analytic scale provides part-by-part scoring. "It separates and weights different features of the test taker's performance on a writing or speaking task" (Richards et al., 2002, p. 25). In a speaking task, the commonly analyzed features can be pronunciation, fluency, accuracy, and appropriateness (Richards et al., 2002). Mainly Brown and Lee (2015) reveal the criteria for evaluating speaking tasks as follows:

- Pronunciation
- Fluency
- Vocabulary
- Grammar
- Discourse
- Task objectives

On the contrary, holistic scales work differently. A holistic scale approaches the task as a whole and gives an overall opinion about the test taker's language ability.

They can be practical and easy for teachers; however, they do not provide in-depth information and feedback about the test taker's language ability.

Lastly, considering the "native non-native English teacher dichotomy" in the ELT world, Ekmekçi (2016) revealed no statistically significant difference in the evaluation of speaking skills by native and non-native English language teachers. Regarding the scores the two groups gave, fluency, pronunciation, accuracy, vocabulary, and communication strategies showed no difference, while only pronunciation differed as expected.

2.2.2. Communicative Language Testing and Speaking Skills

Since the introduction of the Communicative Language Teaching approach, there has been a growing awareness of the EFL environment. This awareness is about viewing language not only as a system of forms or linguistic patterns but also as a means of communication with specific functions (Sakale, 2012). 20th-century approaches like the audio-lingual approach to language teaching have underlined oral production and productive skills. Later on, with the communicative approach, the content of language class was rethought and adjusted regarding semantic notions and social functions. With the addition of a sociolinguistic perspective by Hymes, communicative competence has emerged as the most crucial motive after the 1980s (Swain & Canale, 1980). Communicative Language Teaching (CLT) aims to develop Communicative Competence (Hymes, 1972).

Communicative competence can be defined simply as “knowing when and how to say what to whom.” To do this, a learner needs knowledge of functions. According to Hymes, communication is a process affected by psychological and sociological factors in a context that involves the probability of occurrence and appropriateness. Therefore, the aim of language learning and testing should be expanded into the acquisition of communicative competence. According to him, communicative competence comprises four competence areas: linguistic competence, sociolinguistic competence, discourse competence, and strategic competence. Learners with communicative competence can produce sentences and appropriately use them outside of the classroom in a real-life context. After Hymes, Bachman (1990) reconceptualized the Canale’s framework. His model of communicative language ability consists of inclusive components like language competence, strategic competence, and psychophysiological competence (Figure 2.5.).

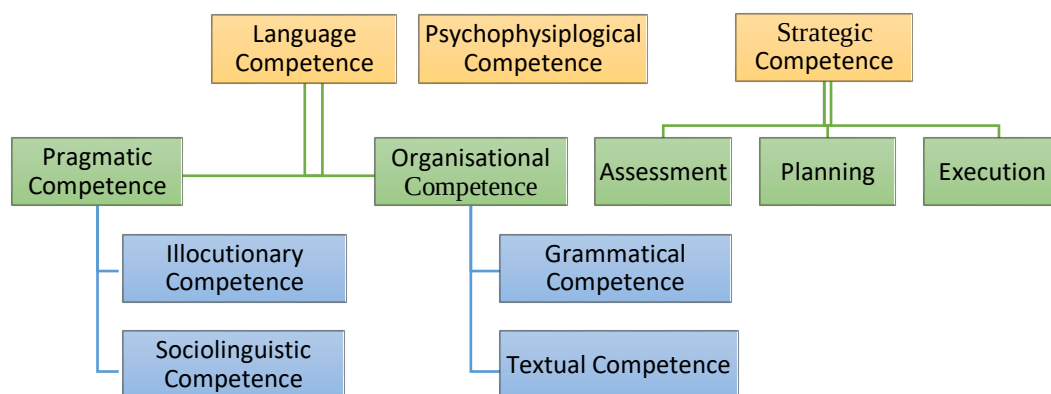


Figure 2.5. Bachman's Communicative Proficiency Model

Teaching cannot be isolated from testing. If it is, learners feel disconnected with a sense of unfairness. For this reason, teachers should avoid teaching-testing mismatches so as not to discourage students from learning languages. Considering the critical position of Communicative Language Teaching in ELT in the last two decades, teachers should provide students with appropriate communicative language tests. In modern language testing, the communicative approach has become dominant (Harding, 2014). Communicative language tests are designed to measure language learners' ability to get involved in communication or to be able to use language in real-life communications. It requires language assessment in a context with authentic tasks and interaction (Davies, 2006). Brown (2004) illustrates that communicative language testing tasks were focused in the language testing field by the mid-1980s, which implies that more research on this area has come up, and it has been recognized more since then (Phan, 2008). With the improvements, the trade-off between the reliability and validity of communicative tests has always been controversial throughout decades because multiple-choice questions cannot reflect real-life situations (validity), and the less validity of a test means, the more reliability it has, unavoidably (Fulcher, 2000). Moreover, according to Harding (2014), there are two main challenges in communicative language testing: utilizing and operationalizing models of communicative ability and theorizing a sufficiently rich construct. However, implementing these tests at an early age may help students become familiar with the types of international tests, such as TOEFL, IBT, or IELTS, that they may encounter in their academic lives at later stages (Phan, 2008).

According to Marrow and Brumfit (1979), a test must meet three criteria to be communicative:

- i. **Communicative tests must involve performance;** test-takers have to produce language in a face-to-face oral interaction and show their ability to deal with the unpredictability of language in an interaction.
- ii. **Communicative tests must be authentic;** a context needs to be created, and the test-taker needs to understand the purpose of the communication while giving appropriate answers. Simplifying inputs or prompts should be avoided, as well.
- iii. **Communicative tests must be scored on real-life outcomes;** the success of the test-taker should be based only on behavioral outcomes. The intended communicative effect has to be the only criterion for success.

Communicative language testing is concerned with how language is used in communication. It is interested in what the test-taker can do with the language instead of what the knowledge of the language test-taker possesses. Additionally, five principles were listed for communicative language test design later on (Marrow et al., 1981). These are:

Principle 1: Know what you are doing.

Focusing on the function, every test should have behavioral objectives. Every test writer should know what they are measuring clearly.

Principle 2: The whole is more than the sum of the parts.

A communicative test must measure the test-taker's ability to deal with discourse because the surrounding context creates the meaning of sentences and decides the appropriateness of utterances in real communication. To provide discourse, test writers must prepare tests with stretches of language, not with isolated sentences or words.

Principle 3: The processes are as important as the forms.

Information gap, choice, and feedback are the components of the real communication process. Creating a test focusing on the communication process with the aforementioned components helps learners engage in unrehearsed communication.

Principle 4: To learn it, do it.

Teachers should set up real-life situations for test-takers to perform linguistic functions and see how they cope with the tasks. This improves the content validity of the tests, and the tests need to be evaluated based on criterion-referenced scales.

Principle 5: Mistakes are not always a mistake.

Understanding the original mistake is a necessity. A grammatical mistake sometimes cannot be a reason for unsuccessful communication. It is possible to have successful communication with minor grammatical mistakes, and teachers should be aware of local ones.

Morrow (1981) explains the basics of CLT methodology with the following words:

A method which aims to develop the ability of students to communicate in a foreign language will aim to replicate as far as possible the process of communication, so that practice of the forms of the target language can take place within a communicative framework (p.62).

Real-life tasks, face validity, authenticity, and performance were used as the keywords of communicative testing (Fulcher, 2000). Teaching speaking, intonation, stress, rhythm, and pronunciation were ignored in the grammar-translation era. However, communicative language tests cover the four language skills and areas based on communicative performance, and communicative performance requires speaking. Marrow (1979) underlines that the communicative test always involves performance (speaking). Lado (1964) highlights that the ability to speak a foreign language is the most prized language skill, even if the testing of speaking is the least developed area. Therefore, he suggests testing speaking through communicative elements such as fluency and accuracy.

2.2.3. Speaking Skills and CEFR

The Common European Framework of Reference for Languages (CEFR) is a norm for teaching, learning, and assessing languages regarding competence in four language skills (reading, writing, speaking, and listening) and a guide to curriculums, materials, and syllabuses. The Council of Europe created it in 2001 based on clear principles. Since then, it has served language education and certification policies in Europe, which respects and promotes plurilingual and pluricultural communities. It also affected language testing as an influential tool in Europe and the world (Mirici & Şengül, 2020). With the help of criteria in the CEFR, learners' level of language competence can be determined in six different levels of language proficiency: A1-A2 as elementary, B1-B2 as intermediate, and C1- C2 as advanced.

Table 2.3. Overall Oral Production

C2	Can produce clear, smoothly flowing, well-structured discourse with an effective logical structure which helps the recipient to notice and remember significant points.
C1	Can give clear, detailed descriptions and presentations on complex subjects, integrating sub-themes, developing particular points, and rounding off with an appropriate conclusion.
B2	Can give clear, systematically developed descriptions and presentations, with appropriate highlighting of significant points, and relevant supporting detail. Can give clear, detailed descriptions and presentations on a wide range of subjects related to their field of interest, expanding and supporting ideas with subsidiary points and relevant examples.
B1	Can reasonably fluently sustain a straightforward description of one of a variety of subjects within their field of interest, presenting it as a linear sequence of points.
A2	Can give a simple description or presentation of people, living or working conditions, daily routines, likes/ dislikes, etc. as a short series of simple phrases and sentences linked into a list.
A1	Can produce simple, mainly isolated phrases about people and places.

Pre-A1	Can produce short phrases about themselves, giving basic personal information (e.g. name, address, family, nationality).
--------	--

Source. Adapted from The Council of Europe (2020).

Oral production and interaction are two critical headlines in the CEFR. The CEFR determines oral production (Table 2.3.) and interaction (Table 2.4.) based on specific criteria according to learners' level of language competencies.

Table 2.4. Overall Oral Interaction

C2	Has a good command of idiomatic expressions and colloquialisms with awareness of connotative levels of meaning. Can convey finer shades of meaning precisely by using, with reasonable accuracy, a wide range of modification devices. Can backtrack and restructure around a difficulty so smoothly that the interlocutor is hardly aware of it.
C1	Can express themselves fluently and spontaneously, almost effortlessly. Has a good command of a broad lexical repertoire allowing gaps to be readily overcome with circumlocutions. There is little obvious searching for expressions or avoidance strategies; only a conceptually difficult subject can hinder a natural, smooth flow of language.
B2	Can use the language fluently, accurately and effectively on a wide range of general, academic, vocational or leisure topics, marking clearly the relationships between ideas. Can communicate spontaneously with good grammatical control without much sign of having to restrict what they want to say, adopting a level of formality appropriate to the circumstances Can interact with a degree of fluency and spontaneity that makes regular interaction, and sustained relationships with users of the target language, quite possible without imposing strain on either party. Can highlight the personal significance of events and experiences, and account for and sustain views clearly by providing relevant explanations and arguments.

B1	Can communicate with some confidence on familiar routine and non-routine matters related to their interests and professional field. Can exchange, check and confirm information, deal with less routine situations and explain why something is a problem. Can express thoughts on more abstract, cultural topics such as films, books, music, etc. Can exploit a wide range of simple language to deal with most situations likely to arise while travelling. Can enter unprepared into conversation on familiar topics, and express personal opinions and exchange information on topics that are familiar, of personal interest or pertinent to everyday life (e.g. family, hobbies, work, travel and current events).
A2	Can interact with reasonable ease in structured situations and short conversations, provided the other person helps if necessary. Can manage simple, routine exchanges without undue effort; can ask and answer questions and exchange ideas and information on familiar topics in predictable everyday situations. Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters to do with work and free time. Can handle very short social exchanges but is rarely able to understand enough to keep conversation going of their own accord.
A1	Can interact in a simple way but communication is totally dependent on repetition at a slower rate, rephrasing and repair. Can ask and answer simple questions, initiate and respond to simple statements in areas of immediate need or on very familiar topics.
Pre-A1	Can ask and answer questions about themselves and daily routines, using short, formulaic expressions and relying on gestures to reinforce the information.

Source. Adapted from The Council of Europe (2020).

The production and interaction activities are also named and classified as in Figure 2.6. While oral production includes monologues, oral interaction activities have discussions and conversations.

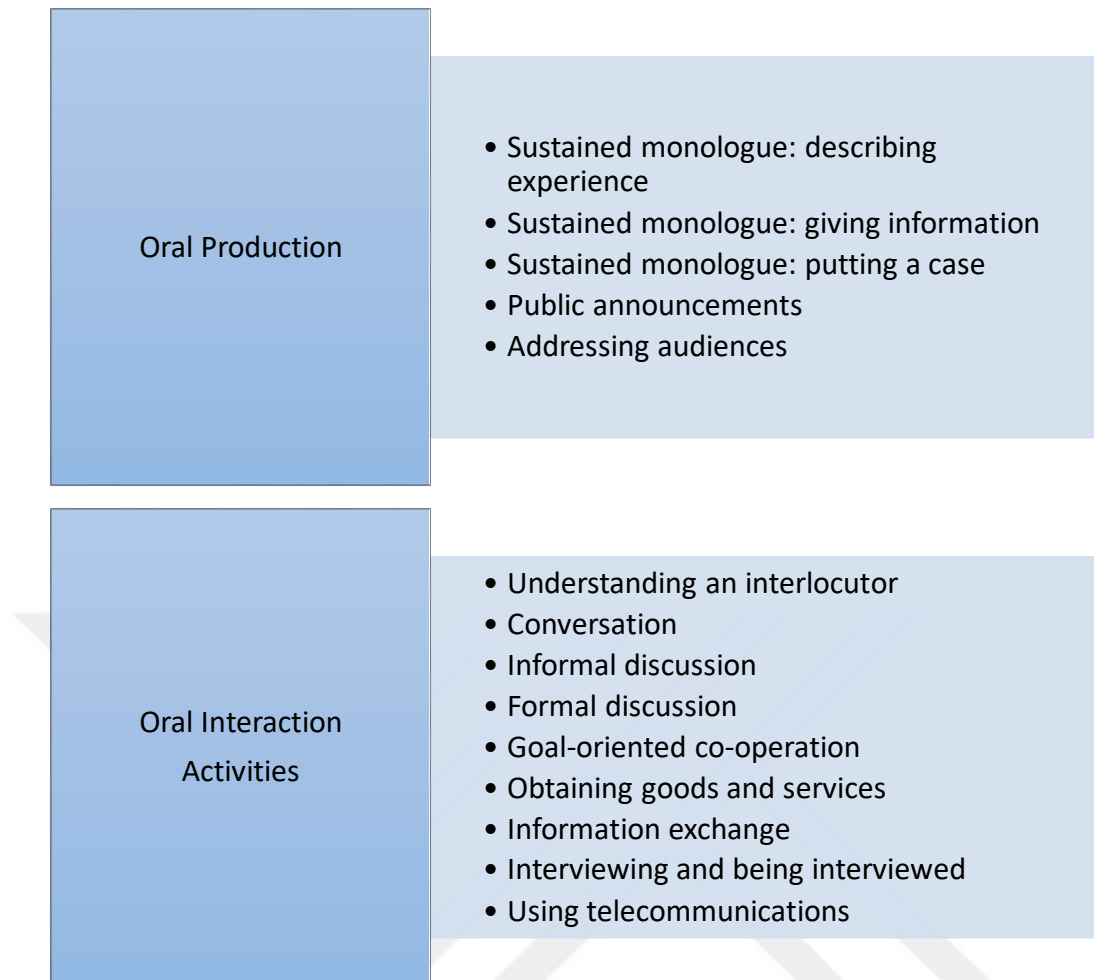


Figure 2.6. Oral Production and Interaction Activities

Source. Adapted from The Council of Europe (2020).

2.3. The New Examination System, Regulations, and Curriculum by Turkish MoNE

The Turkish MoNE integrated the English language course for the first time in the curriculum of primary schools in grades 4 and 5 (Table 2.5.) in 1997 and thus modified the primary school curriculum. They started English lessons for students in the 4th and 5th grades. Following this improvement, another step was taken in 2013. The age of acquaintance with English was declined. With the implementation of the new system, "The 4+4+4 School System", students started having English courses in the 2nd grade for the first time, and the hours of the 5th grade were increased (Table 2.5.). Another regulation in the curriculum and lesson hours came in 2018, which is still in practice.

Table 2.5. Primary Education English Lesson Curricula Weekly Lesson Hours by Year

Year	Grade	Hours per Week
1997	4th, 5th grades	2
	6th,7th,8th grades	4
2013	2nd, 3rd, 4th grades	2
	5th, 6th, 7th, 8th grades	4
2018	2nd, 3rd, 4th grades	2
	5th, 6th grades	3
	7th, 8th grades	4

Table 2.5. shows that lesson hours and curriculums for English courses at secondary schools (5th, 6th, 7th, and 8th grades) have been regulated three times. Currently, 5th and 6th grade students have 3 hours, and 7th and 8th grade students have 4 hours of English course per week. These regulations in the education policy enhanced the integration of communicative language teaching into the curriculum, as well (Kırkgöz, 2007). The English course curriculum has been updated in the last two decades according to CEFR and modern learners ' communicative needs.

English has been a part of high school entrance exams since 2008, along with Turkish, math, science, social sciences, and religion courses in Türkiye. These exams are administered at the end of a school year and contain only multiple-choice questions. This situation pushes students to practice more on multiple-choice tests during the school period to get higher points and open the gates to successful high schools. Naturally, students in the 7th and 8th grades especially expect teachers to help in this direction by paying more attention, effort, and time to multiple-choice tests. As one reality of secondary schools, teachers of English courses prepare their students for these exams with multiple-choice tests. Consequently, quizzes, tests, and practices with multiple-choice questions are common in English lessons.

After starting the 2023-2024 academic year, the Turkish MoNE made some new regulations for the English exams in secondary schools. These changes were profound and concerned many issues, such as preparation, administration, question types, grading, and topics of the exams (Table 2.6.). One main change was the augmentation in the number of exams per term. In the old system, students had two compulsory exams per term. One of them was mid-term, and the other was the final exam. Both

exams were rated out of 100 points and had an equal effect on the student's final grades. With the new regulations, every student has to take six exams per term. Three of them are mid-terms, while three of them are finals. They are all rated out of 100 points. However, to give students a mid-term grade, teachers have to conduct three different sessions with students: a reading-writing exam, a listening exam, and a speaking exam. Each of these exams has different percentages over the last grade given to students. The reading-writing (main) exam has 50%, the listening exam has 25%, and the speaking exam has a 25% effect over the mid-terms and finals. We can understand from this that the Turkish Ministry of National Education's assessment approach to English exams has evolved from a holistic to an analytic perspective.

Another critical point is the preparation of the exams. The new regulations oblige teachers to make grade-wide exams for each level and to prepare these exams jointly with the English department. When it comes to the content of the exams, there are, again, stricter guidelines. The new regulations provide teachers with scenarios to indicate the type and number of questions about the exam topics. Before the new regulations, there used to be no restrictions regarding the exam types; teachers could mix different kinds of questions (fill in the blanks, multiple-choice, odd one out, true-false, open-ended, etc.) in the exam paper. Nevertheless, it has to be only open-ended questions without options now. We can understand from this that the Turkish Ministry of National Education's expectations of students in English exams have evolved from making choices among certain options to a dimension that encourages creative thinking and the expression of one's own ideas and experiences.

At secondary schools, English departments decide on the exam periods at the beginning of the term meetings. Before the regulations, teachers used to administer exams within the exam periods independently, but now, they have to administer the exams on the same day and hour. Besides, the passing grade for the course increased from 45/100 to 50/100. With the new regulations, for each term, teachers have to give each student one more performance grade (three in total).

Table 2.6. The Comparison of the Changes in the Examination System for English Courses in the 2023-2024 Academic Year

The Old System for the Exams	The New Regulations in the 2023-2024 Academic Year
------------------------------	--

Numbers of Exams	Two exams for a term (One mid-term exam, One final exam)	Six exams for a term (One mid-term comprised of listening, speaking, reading-writing main exam, one final exam comprised of listening, speaking, reading-writing main exam)
Preparation	Teachers can prepare their exams individually or cooperate.	The English department at school should prepare a grade-wide exam with the cooperation of all teachers.
Topics	No scenarios were provided.	Two different scenarios for every reading-writing (main) exam of each grade were provided with example questions.
Question Types	No restrictions (multiple-choice questions, true-false questions, matching, odd one out, open-ended questions, fill in the blank, etc.)	Only open-ended questions-answers
Administration	Every teacher can administer exams according to their course hours as long as it is within the date range determined by the department.	The reading-writing (main) exams have to be grade-wide, with the same questions, and at the same time for all students at school.
Grading	The course passing grade is 45 out of 100 points. Two performance grades have to be given for each term. Holistic scoring. The mid-term exam has a 50% effect, and the final exam has a 50% effect.	The course passing grade is 50 out of 100 points. Three performance grades have to be given for each term. Analytic scoring. The mid-term exam has a 50% effect, and the final exam has a 50% effect. However, mid-terms and final exams consist of three sub-exams. The reading-writing (main) exam is 50%, the listening exam is 25%, and the speaking exam has a 25% effect on mid-terms and final exams.
Test Types	Criterion-referenced achievement tests.	Criterion-referenced achievement tests.

2.4. Research on Speaking Skills in Other Countries

Nurhalimah et al. (2020) explored and described English teachers' beliefs and their practices of speaking assessment through a semi-structured interview. The findings indicated that even if participant teachers agreed on their crucial role in assessing students' speaking skills and the benefits of using rubrics, their beliefs and practices in assessing speaking skills differed from each other for multiple reasons, such as teachers' educational background, teaching experience, and school context. One of the main findings revealed that teachers gave more importance to the administrative side of assessment with little consideration of assessment for learning. This clearly explains that a small part of the participants stated that they gave feedback after the assessment. It was underlined that teachers needed some areas for improvement regarding speaking assessment and guidance regarding rubric development and implementing speaking assessment.

Plo et al. (2013) researched the teaching of oral skills in secondary education in a local Spanish context by surveying teachers. Results show that contrary to language policies in Spain, which underline the development of oral skills aligning with CEFR, oral skills need to be sufficiently practiced and assessed. The students are still reluctant to engage in oral activities and must reach the required level in these skills.

Adem and Berkessa (2022) examined EFL teachers' cognition about teaching speaking. The research used sequential mixed methods with a questionnaire and interviews with teachers. The study uncovered that the teachers did not enjoy teaching speaking. Most of the teachers had the idea that speaking was the weakest skill of their learners, and learning speaking was different and more complex than learning other language skills. Even if teachers' self-concept was optimistic, they held pessimistic opinions about student's interest in learning to speak. They thought learner-related factors primarily influenced the teaching of speaking skills. According to them, the relationship between their beliefs and the classroom practices was fair. The teachers stated that group discussion was the most effective classroom activity. Furthermore, it was found that the gender and year of teaching experience of teachers were statistically significant in their beliefs.

Ahmad (2021) investigated the assessment beliefs of internationally educated teachers in assessing speaking skills in Canadian ESL/EAP classrooms through multiple narrative case studies. She found that internationally educated teachers had multi-layered beliefs regarding the purpose and types of assessment, choosing the

assessment tool, and giving feedback. The study discovered numerous contextual factors, such as classroom contexts, institutional requirements, and curriculum expectations, influenced teachers' assessment practices. The study underlines that differentiated professional development opportunities and mentorship programs were needed to integrate better the heterogeneous group of internationally educated teachers in the Canadian education K-12 system.

Austheim (2023) researched Norwegian lower-secondary EFL teachers' perspectives on the relationship between formative assessment and willingness to speak through semi-structured interviews. According to the results, the teachers believed that classroom environment, social relations, and students' beliefs about their competencies affect students' willingness to speak; formative assessment can increase students' willingness to speak while decreasing stress by allowing students to perceive learning as a trial-and-error process.

Gandeel (2016) explored the beliefs and practices of English language teachers in terms of teaching speaking with qualitative multiple-case research. She collected data via semi-structured interviews and classroom observations. She showcased that teaching speaking was course-book based, teacher-centered, and accuracy-oriented (i.e., grammar, vocabulary) with little or no focus on fluency. The results indicated that teachers' beliefs and practices lacked a theoretical basis, and thus, they did not align with contemporary views on teaching speaking.

Brumen and Cagran (2011) used a questionnaire to examine Teachers' perspectives and practices in assessing young foreign language learners in three Eastern European countries (the Czech Republic, Slovenia, and Croatia). It was found that most teachers had yet to receive additional teacher training in assessing pupils' foreign language development. The participant teachers expressed an excellent desire for additional teacher training in assessing young learners. Even though there are significant differences in practice among these three countries, the traditional belief dominates that teachers are the most responsible for assessment in all three countries. Most teachers took a stand to assess young learners. However, they stated that they prefer descriptive assessment rather than numerical methods.

Gopal and Embi (2014) investigated the teacher's beliefs and practices of conducting communicative language testing ELT in quantitative research through a questionnaire with teachers. The results indicate that language teachers hold positive beliefs and practices of communicative language testing and are aware of its principles.

Nevertheless, teachers encounter main constraints during communicative language tests in practice.

Atjonen (2014) researched teachers' views of their assessment practice through a descriptive, mainly qualitative questionnaire. According to the findings, using different assessment methods, an interactive approach, encouraging feedback, and criteria clarification were appreciated as positive factors by teachers. However, improper assessment methods, level of stringency, poorly implemented assessment, and weak grounds for assessment were evaluated as negative factors. According to participants, interpretation of fairness, pupils with special needs, and pupils' heterogeneity made assessment difficult. Besides, versatile assessment methods, curriculum advice, and pupils' good competencies helped the assessment process be easy.

Grada (2014) aimed to search EFL teachers' knowledge of speaking assessment and how their knowledge informed their speaking assessment practices in classrooms in a Libyan secondary school context with qualitative and quantitative methods. As a result of the study, it was clear that EFL novice teachers' current knowledge of speaking assessment was complex. The teachers based their assessment practice on the fact that spoken language is more about linguistic content than communicative effect. They focused on the assessment of content rather than an assessment of the process. The summative assessment was given more importance than the formative assessment. Lastly, their notion of assessment played a more significant role than the context in implementing the speaking assessment.

2.5. Research on Speaking Skills in Türkiye

In recent years, there has been an increase in studies on speaking skills and exams, which have gained significance in Türkiye. This section will discuss several relevant studies conducted in Türkiye.

Özdemir (2018) aimed to investigate the administration process of speaking exams, which became compulsory in high schools in Turkey in 2014, and teachers' attitudes towards speaking exams. In the study conducted in Denizli, 358 students and 22 teachers from Anatolian high schools participated. The findings of the study, in which qualitative and quantitative data were collected, show that teachers and students have a positive attitude towards speaking exams; they state that it is a need and valuable. Teachers informed their students sufficiently about assessment and scoring

beforehand. Teachers evaluated each student individually in the exams. The students expressed that they were satisfied with the duration of the exam. However, the teachers stated that they gave the students more points to show their support for the students in the exam. This situation decreases the reliability of the exam.

In return to previous studies that portrayed speaking English as a national challenge for Turkish speakers of English, Dağtan (2020) investigated the perceptions of Turkish ELT pre-service teachers and lecturers on the learning and teaching English speaking skills in two phases via questionnaires and semi-structured interviews. He focused on problems and suggestions to provide a nationwide portrayal of the problem. The participants were pre-service teachers and lecturers studying/working at seven ELT departments from seven regions of Türkiye. Thirty-four lecturers and 361 pre-service teachers took the questionnaire, and a semi-structured interview was conducted with 38 lecturers and 48 pre-service teachers. Findings showcased that both pre-service ELT teachers and lecturers considered speaking English and instruction on pronunciation to be highly important for personal reasons and for the country to get more international. According to them, they were having pronunciation instruction before the critical age is essential. It was found that, in the first several weeks of ELT instruction, pre-service teachers experience a dreadful adaptation process to speaking English in class, which is the medium of instruction in ELT departments. Also, lecturers were found to be more confident about and less afraid of making mistakes while speaking English than pre-service teachers.

Çalışkan (2013) researched instructors' attitudes and practices in assessing English speaking skills. She used mixed methods research to gather data to understand the relationship between attitudes and practices with the gender, educational background, teaching experience, and English proficiency levels of instructors in three universities in İstanbul. She focused on principles, problems, difficulties, and techniques used by instructors while assessing speaking. To obtain data, 16 instructors had interviews, while 89 had questionnaires. According to the research, even if instructors use generally interactive, authentic techniques, instructors holding MA, BA, and Ph.D. degrees choose different assessment techniques by percentages. The research showed that educational background and experience could be factors in choosing assessment tools by stating that the more experienced the instructors, the more various the assessment techniques.

Güral Akgül (2022) explored the impacts of formative assessment on preparatory students' speaking skills using qualitative methods with several case studies during ten weeks of data collection. The participants were three students in the pre-intermediate level in a preparatory class at a foundation university. Results demonstrated that students profited from formative assessment implementation, gained confidence, lessened anxiety, and became more willing to communicate in English in their listening and speaking lessons. Feedback from peers and instructors helped them improve their English speaking.

Through questionnaires and interviews, Duran (2011) investigated teachers' and students' perceptions of the washback effects of classroom-based speaking tests. According to the results, teachers were not influenced by the speaking tests regarding what they practiced in classes; however, they had positive attitudes towards teaching and testing speaking. The participant teachers believed that speaking tests had a positive effect on their students' speaking ability. Both teachers and students stated that getting ready for speaking tests improves the general speaking skills of students. Furthermore, students were also quite positive towards teaching and testing speaking as a positive washback. The study showed that the students wanted the weight of the speaking exam to be increased, while both students and instructors thought that these speaking tests should remain a component of all the exams.

3. METHODOLOGY

This section has five sub-headings, namely, research design, participants, data collection instruments, data analysis, reliability, and validity, which provide detailed information about the research methodology.

3.1. Research Design

A mixed research design was used in this study. Both qualitative and quantitative data were collected from the participants and analyzed independently, so the data were diversified and enabled to establish the research on a more solid foundation (Yıldırım & Şimşek, 2006). Creswell and Clark (2011) state that “qualitative data provide a detailed understanding of a problem while quantitative data provide a more general understanding of a problem” (p. 8). Considering these benefits, the mixed research design benefited the purpose of the research. Since a mixed research design was chosen, collecting data in two ways was necessary. For this reason, it was decided to collect qualitative data first and then quantitative data. Since no ready-made questionnaire or interview questions were directly related to the subject, the researcher prepared the questions herself.

The data collection process was carried out through a survey for quantitative data and an interview for qualitative data only with secondary school ELT teachers. First, one-on-one interviews were conducted with 11 participants, and then the survey was conducted with 217 participants. The data collection process was carried out within the planned schedule after the ethical permission of the 19 Mayıs University Ethics Committee was taken (Appendix A).

As mentioned earlier, this study seeks answers to these research questions:

RQ1: What are the overall thoughts of EFL teachers towards the recent English speaking exams after the current regulations in the secondary school examination system in Türkiye?

RQ2: What are the feelings of EFL teachers related to the recent English speaking exams?

RQ3: What experiences were encountered by EFL teachers during the implementation of the recent English speaking exams?

RQ4: What are the future thoughts of EFL teachers related to the recent English speaking exams at secondary schools?

3.2. Participants

Eleven EFL teachers (male n=2; female n=9) working at secondary schools participated in the interview part of this study. In order to protect the confidentiality of the participants, the names of the participants were kept confidential, and pseudonyms were given from P1 to P11 according to the order in which they were interviewed to refer them. The demographic information of the teachers who participated in the interview part of this study was demonstrated in Table 3.1. The letters “F” and “M” in the table represent females and males. Considering that English teaching is a female-dominated study group in our country, most participants were expected to be women. All of the teachers in this study were relatively young teachers with undergraduate and graduate degrees and experience ranging from 1 to 15 years. The participants were selected from five cities, mainly Samsun, and all worked in public schools. Since the conditions where teachers work may also affect their experiences with the speaking exams, this information was also collected from the participants. Four participants work in a village, two in the city centers, and five in the towns. Finally, the reason why the participating teachers were selected for this research was that they volunteered to contribute to the research, were comfortable and accessible in communication, and had the potential to enrich the data by adding different perspectives to the research as teachers who had or had not applied the English speaking exam before.

Table 3.1. Demographic Information of the Participants in the Interview Part of This Study

	Gender	Age	City	Experience	Education	Place	Grade
P1	F	33	Samsun	15	Undergraduate	Village	5,6,7,8
P2	F	34	Samsun	10	Undergraduate	City	5,8
P3	F	34	İstanbul	10	Undergraduate	City	6,7,8
P4	F	33	Samsun	11	Undergraduate	Town	5,6
P5	F	34	Çanakkale	8	Undergraduate	Town	7
P6	F	24	Diyarbakır	1	Undergraduate	Village	5,6,7,8
P7	F	33	Çanakkale	11	Undergraduate	Town	5,8
P8	M	40	Samsun	15	Undergraduate	Town	5,6,7,8
P9	M	34	Samsun	11	Graduate	Town	5,6,7,8
P10	F	24	Rize	1	Undergraduate	Village	5,6,7,8

Via the questionnaire presented online, 217 participants who currently work at secondary schools as ELT teachers took part in the survey part of the research. According to the anonymously collected demographic information, 176 respondents were female, and 41 were male. Although there were respondents from different cities in Türkiye, 42.5% of the respondents were from Samsun. 93 teachers working in Samsun participated in the survey. Ankara and İstanbul followed Samsun with 29 participants in total. The majority of these respondents had bachelor's degrees, with 88.7 %. Only one doctoral and one associate degree participated in the survey. Moreover, the number of master's degree graduates was 22 and 10.4 % (Figure 3.1.).

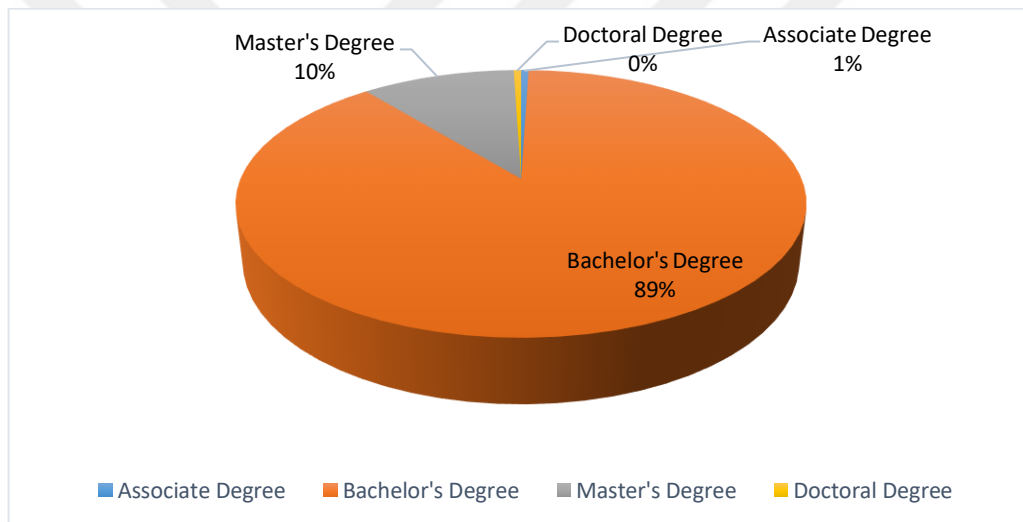


Figure 3.1. Graduate Degrees of the Respondents

In line with the data obtained, when the teaching experiences of the respondents were analyzed, it was seen that teachers with more than six years of teaching experience were the majority of respondents (Figure 3.2.).

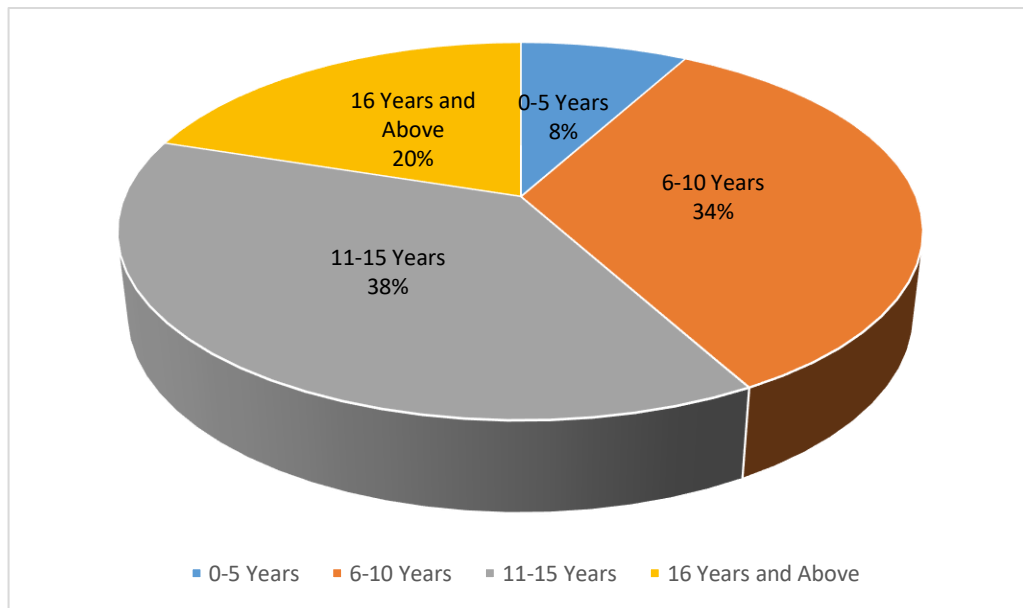


Figure 3.2. Teaching Experiences of the Respondents

According to the data collected via the questionnaire, it is clear that most of the respondents were teaching at different levels in the 2023-2024 academic year at secondary schools. However, due to the high number of courses in the 7th and 8th grades, there is a parallel increase in the number of teachers teaching these grades, as shown in Figure 3.3. One of the points to remember is that in schools with a single teacher, the teacher has to teach at every level, which can influence the answers given.

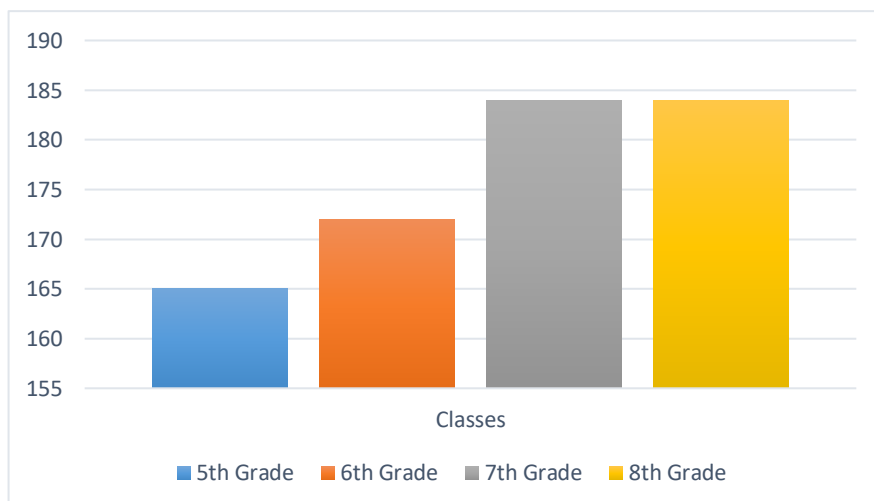


Figure 3.3. The Grades in Which the Respondents Teach

3.3. Data Collection Instruments

Two tools were determined as data collection instruments. One of them was the interview. The researcher found the most suitable time and environment for the participants and conducted the interviews in detail. The other was the questionnaire. Since the questionnaire was requested to be conducted online, it was created using Google Forms and sent to the teachers as a link. Respondents used mobile phones or computers to answer the questions. It took three weeks to conduct the semi-structured interviews, while it took one week for the questionnaire to be answered by the target respondents.

3.3.1. Semi-Structured Interviews

Semi-structured interviews were conducted one-on-one with teachers to add a deeper dimension to this study and increase reliability by triangulating the information. Semi-structured interviews give the researcher more flexibility to ask relevant questions when needed. Before starting the interview, the informed voluntary consent form was read to the participants, and care was taken to ensure that the participants participated in this study with their consent. In these interviews, the confidentiality of the participants was emphasized, and their names and surnames were not requested. All the interviews were audio recorded to be analyzed later on. Five of the interviews were managed online via Zoom meetings, and six were face to face. Teachers explained their experiences, problems, suggestions, and feelings about the subject by giving examples as interview questions were asked. The researcher obtained details by asking extra questions where needed. The first part of the interview included demographic questions. The data about education, years of experience, age, gender, school place, city, and organization type were asked. In the following parts, under the A, B, C, D, and E categories, 29 questions were asked (Appendix B). These questions were about teachers' readiness, thoughts, feelings, observations regarding students, testing and assessment instruments and processes, and future exams. All participants were encouraged to give examples and details. They all answered all the questions and completed the interview. Each interview lasted about 30-40 minutes.

3.3.2. Questionnaire

The survey was conducted to collect information from a larger audience to strengthen the data. The researcher prepared the survey specifically in line with the

research questions. The aim was to keep the survey as understandable and straightforward as possible and to deliver it to as many teachers as possible. Google Forms were used for this, and the survey was shared online as a link and completed within a week. The survey consisted of two main parts. The first part contained five demographic questions (gender, city, educational background, experience, and the current grades the participant teachers teach) and five speaking exam-related multiple-choice questions. The second part contained phrases based on new exam regulations and speaking exams with a Likert scale. The second part contained four subsections. These, in turn, included teachers' thoughts, experiences, students' dimensions, and upcoming exams' dimensions. The first part of the survey contained ten questions, and the second part contained 34 questions (Appendix C). For the questionnaire, pilot research was conducted with three respondents. The pilot study aimed to identify the shortcomings, uncertainties, and points that needed to be corrected before the questionnaire was disseminated. These three teachers helped to improve the questionnaire with their feedback. After disseminating the questionnaire, the targeted participants were reached using the snowball sampling method.

3.4. Data Analysis

For the qualitative data obtained from the interviews, thematic analysis was used (Braun & Clarke, 2006). The researcher first transcribed the interviews, read the transcriptions, and gathered the initial codes. She searched for the themes, defined and named categories, and produced the final report with examples from the raw data. As a second step, another professional's opinion was taken to strengthen the codes. This researcher also went over the coding and reworked the codes. The compatibility of the first codes created by the researcher was checked with a sample section given through the interviews; the researchers exchanged ideas, code edits were made again, and thus, the coding was cross-examined. Thus, the coding was strengthened with a double-check.

The quantitative data were analyzed in detail with the help of Google Forms statistics. According to the statistics, teachers' positive and negative feelings and thoughts were compared in proportion to each other.

3.5 Reliability And Validity

A number of precautions were taken to ensure validity and reliability in the study. The interview and survey questions in this study were prepared by the researcher and strengthened with advice from another professional. In this process, alternative question scenarios that comprehensively addressed the subject were created, and which one produced better answers to the research questions was discussed. Piloting was done before working with actual participants in the qualitative and quantitative aspects of the research. While preparing the survey and interview questions, an effort was made to make the questions comprehensive by associating them with the teachers' thoughts, feelings, experiences, advice, expectations, wishes, and observations related to their students, colleagues, and the curriculum. Additionally, care was taken to ensure that the survey and interview questions were parallel.

A pilot application was carried out with two participants to test the interview questions and three participants to test the survey questions. Following these practices, the survey and interview questions were rearranged and improved. Questions that were not easily understood by the participant or could be considered biased were eliminated. The expressions were neutralized. In addition, the statements were arranged from general to specific so that the participants could follow the research questions in the easiest way.

In the process after the data collection, support was received from a professional regarding thematic analysis of the data obtained through interviews. After the coding was done, this professional examined some of the data and shared her own code, category, and theme ideas. The compatible and incompatible outputs were discussed. As a result of these comparisons, the findings were placed on a more solid basis with double check.

4. FINDINGS

In this part of the study, interview findings and questionnaire findings will be explained separately.

4.1. The Interview Findings

The analysis of the interview data yielded four themes and eight categories. The themes were identified according to the research questions: “teachers’ overall thoughts,” “teachers’ feelings,” “experiences and observations,” and “future thoughts” for the recent English speaking exams (Figure 4.1.). Eight categories and 36 codes were placed under these themes. During this placement, it was considered that the codes did not repeat or overlap. In the following sections, the themes related to each research question will be discussed, categories will be explained, and codes will be presented by giving excerpts specifically.

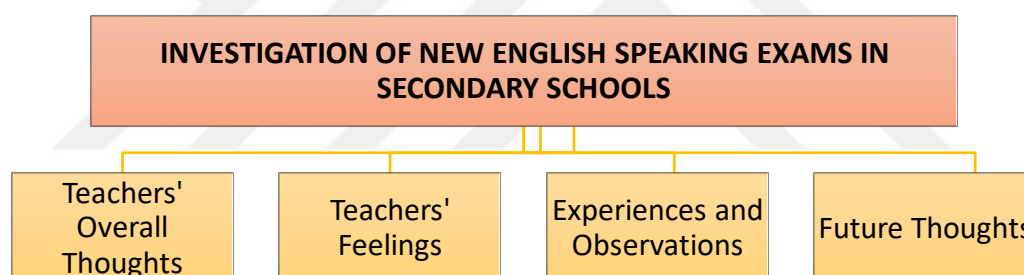


Figure 4.1. Four Themes Created from the Interview Data

There are two categories under each theme (Figure 4.2.). The teachers’ overall thoughts theme (theme 1) has “positive thoughts related to the recent speaking exams” and “negative thoughts related to the recent speaking exams” as categories. The teachers’ feelings theme (theme 2) has “positive feelings related to the recent speaking exams” and “negative feelings related to the recent speaking exams.” The experiences and observations theme (theme 3) has “problems and limitations of the recent speaking exams” and “procedural observations of the new speaking exams.” The last theme, future thoughts (theme 4), has “expectations, beliefs, and requests for speaking exams” and “recommendations to improve the speaking exams.”

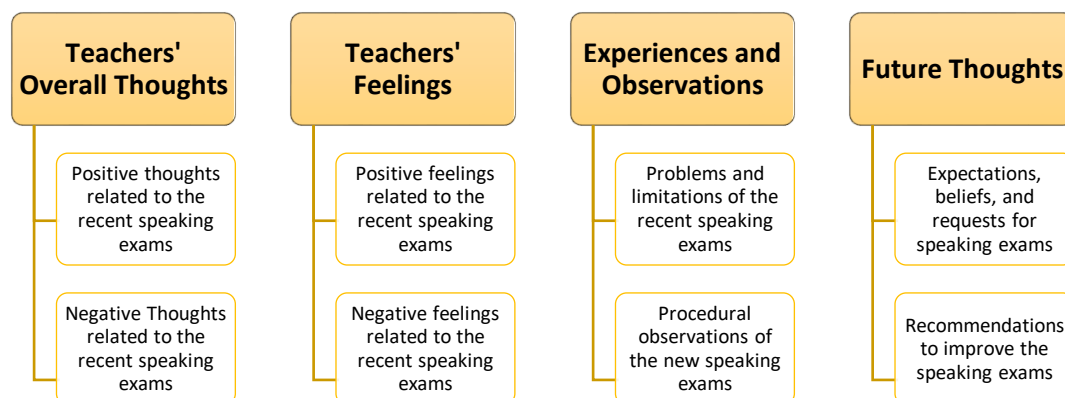


Figure 4.2. The Themes and Categories Created from the Interview Data

As a result of the interviews, 36 codes were created. These codes were grouped by association. The distribution of codes according to categories is shown in Table 4.1.

Table 4.1. Codes, Categories, and Themes

Codes	Categories	Themes
• Necessity of assessing four language skills • More aligned with the curriculum and coursebooks • Motivation for teachers to speak more English • Motivation for teachers to prepare more speaking activities • Increase of students' self-confidence in speaking • Increase of students' awareness of speaking and giving importance to it • Benefits of working together as the department of English • Useful for students' skills learning	Positive thoughts related to the recent speaking exams	Theme 1 Teachers' Overall Thoughts
• Inefficiency of exams • Teachers' lack of readiness • Teachers' lack of guidance • Students' lack of readiness • Unexpected change in the exam policy	Negative thoughts related to the recent speaking exams	
Codes	Categories	Themes
• Motivation for teachers to speak more English • Motivation for teachers to prepare more speaking activities • Increase of students' self-confidence in speaking • Increase of students' awareness of speaking and giving importance to it • A sense of doing the right thing professionally • Benefits of working together as the department of English	Positive feelings related to the recent speaking exams	Theme 2 Teachers' Feelings

• Professional dissatisfaction • Reflection on students' negative feelings toward the teachers • Stress factors related to the 8th grades • Teachers' lack of readiness • Students' lack of readiness • Teachers' feelings of being too busy under time pressure • Teachers' overwhelming emotions	Negative feelings related to the recent speaking exams	
Codes	Categories	Themes
• Teachers' lack of readiness • Students' lack of readiness • Teachers' lack of guidance • Not giving enough space to speaking in the curriculum and textbooks • Inefficiency of exams • Crowded classroom problems • Impracticality of grading and e-school system • Increased workload for teachers • Teachers' feelings of being too busy under time pressure	Problems and limitations of the recent speaking exams	Theme 3 Experiences and Observations
• Differences in practice • Crowded classroom problems • Increased workload for teachers • Inefficiency of exams • Teachers' feelings of being too busy under time pressure • Lack of teacher equipment	Procedural observations of the new speaking exams	
Codes	Categories	Themes
• The high belief that regulations will continue • Suggestions for updates and improvements	Expectations, beliefs, and requests for speaking exams	Theme 4 Future Thoughts
• More hours of English course • Less workload for EFL teachers • Need for extracurricular support • Gradual transition to the new examination system • More curriculum planning with more speaking activities • Regulation in the e-school system and exam scoring system • In-service training and more information • Better classroom environments • A better-planned education calendar	Recommendations to improve the speaking exams	

4.1.1. Theme 1: Teachers' Overall Thoughts

One of the research questions of this study is "What are the overall thoughts of EFL teachers towards the recent English speaking exams after the current regulations in the secondary school examination system in Türkiye?" Two categories were created based on this research question and the data drawn from teachers' interviews: Positive thoughts related to the recent speaking exams and negative thoughts related to the

recent speaking exams. In addition, subcategories have been created to make the categories and codes clearer and more connected. These subcategories were grouped considering that the data obtained concerned teachers, students, and ELT specifically. There are eight positive codes and five negative codes regarding teachers' overall thoughts on the recent speaking exams.

Participant teachers underlined that speaking exams are necessary and useful for skills learning. They think speaking exams benefit students and make the whole learning process more aligned with the curriculum and books because there are speaking objectives and activities in the coursebooks. Speaking exams can motivate students and teachers to speak more English and increase students' self-confidence in speaking activities. Most teachers state they benefit from working as a department and cooperating while preparing and implementing the exams. Furthermore, they state that speaking exams help students increase their self-confidence in speaking. According to them, new regulations about speaking exams give the idea that speaking is more important now, and this directly affects students' understanding of it, which helps them give more importance to it. They suggest that due to the recent speaking exams, students gained awareness about the importance of speaking in English. However, they pointed out problems related to the functionality of the exam. They stated there are many problems that can be related to pre-, while, and post-exam situations. Exam failure to fully fulfill their function is one of the most significant complaints about teachers. Also, the lack of teacher readiness and guidance by MoNE at the beginning of the academic term, the lack of student readiness, and the problems related to the unexpected changes during the academic year are underlined by the participants.

4.1.1.1. Positive Thoughts Related to the Recent Speaking Exams

When the interviews were examined, eight positive thought codes were found regarding teachers' recent speaking exams (Figure 4.3.). Three subcategories were created considering the positive thoughts for teachers, students, and ELT. The participant excerpts that are effective in writing these codes are explained in detail and exemplified below.

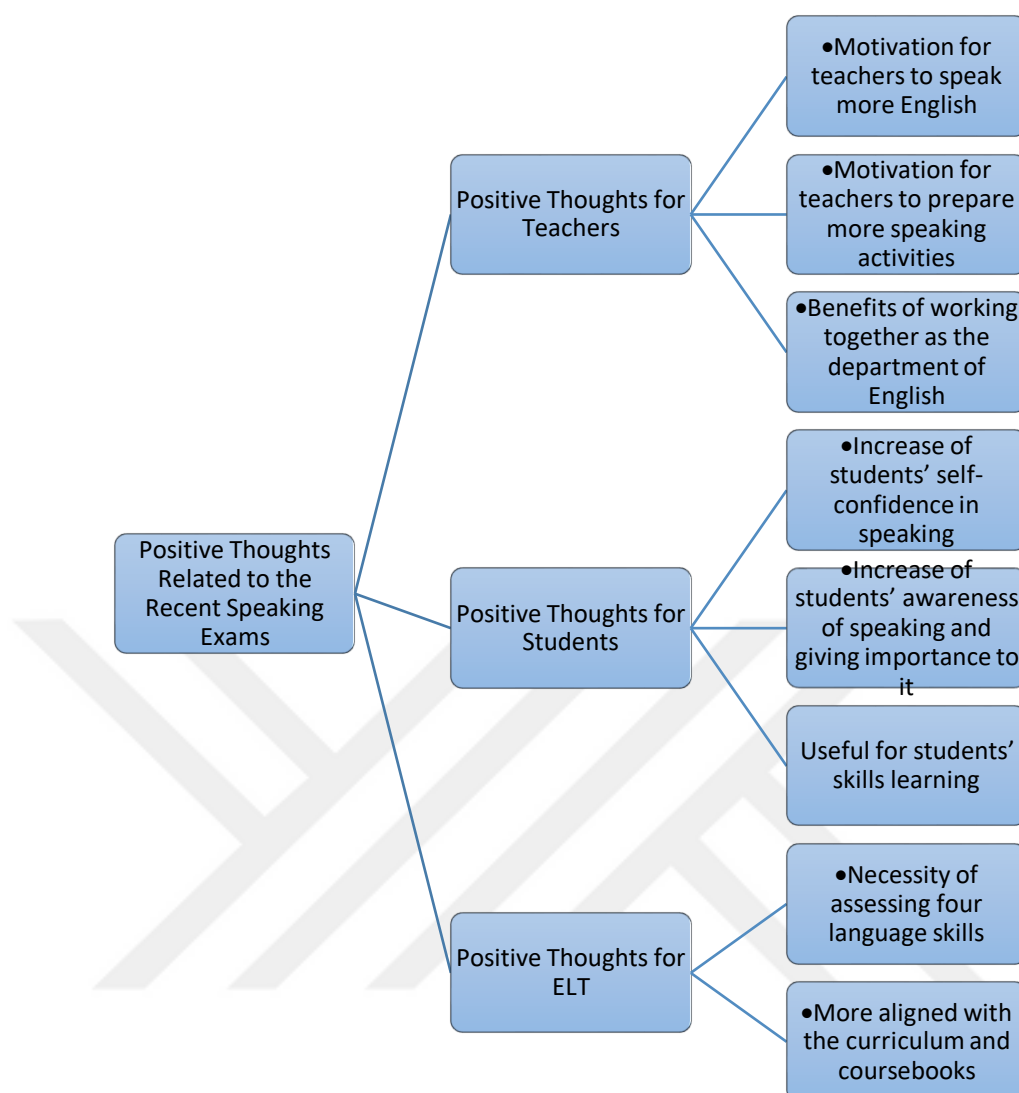


Figure 4.3. The Codes for Positive Thoughts Related to the Recent Speaking Exams

Necessity of Assessing Four Language Skills:

All the participant teachers agreed that speaking exams should be compulsory. Furthermore, they believe four skills are important and should be assessed equally. Some of them support this idea, stating that the curriculum and coursebooks have four skills, and logically, the exams should include them, too. When discussing the necessity of speaking exams, teachers gave some reasons for this. Some teachers considered it a positive and logical step forward in ELT. They said that it is an important step. They also added that speaking exams make sense. According to them, speaking is an essential language skill. The comments supporting these ideas are given below:

Speaking exams should be compulsory because when we learn a language, we have to consider every skill; they all complement each other. (P2)

(Compulsory)

Yes, it should be compulsory because, in the end, if we are teaching this skill, I need to assess it if it is covered under a curriculum. It is necessary. (P3)

(Compulsory, necessary)

I think yes, speaking exams should be compulsory because only going through grammar or listening can help students gain something up to a point. The important thing is to be able to speak a language well if we are learning it. (P10)

(Compulsory, important)

More Aligned with the Curriculum:

According to the participating teachers, the speaking exams introduced by the new regulations strengthen the link between the curriculum, textbooks, and actual practice. For example, while there are speaking objectives in the curriculum, and these activities are practiced in the lessons with the help of books, the lack of criteria in practice creates a contrast. Therefore, the new regulations have made the practices more harmonized. Teachers underlined that the new speaking exams align the English learning process more with the curriculum and coursebooks. The statements supporting this view are exemplified below:

Of course, the achievements in the curriculum should also be reflected in the exams. In this sense, speaking exams sound logical. (P4)

(Reflection, logical)

Having exams compatible with the curriculum is a positive development. If there is an objective, there must be an exam for it. (P11)

(Compatible with the curriculum, necessity)

Motivations for Teachers to Speak More English:

One of the codes is related to the motivation of the new speaking exams for teachers to speak more English. Some teachers stated that they started speaking more

English in class. It is not limited to speaking activities but refers to the language teachers use in the classroom. Thus, teachers are more oriented to speaking in class. On the contrary, the other participating teachers stated that speaking tests did not affect their English speaking orientation in the classroom. Some participants said they were more motivated to prepare speaking activities than last year. According to them, speaking activities have gained more importance with the new regulations. Expressions exemplifying different views are given below:

Speaking was a skill that I ignored in class (mostly due to the lack of class hours). Now I try to give more importance to it. (P5)
(Speak more English, more important)

Speaking exams, so frankly, as I said, they did not motivate me much. In this sense, I thought, yes, there will be a speaking test for a week or two; maybe the rumor is that I should speak more English, but not by much, to be honest, not so much in terms of motivation. (P3)

Motivation for Teachers to Prepare More Speaking Activities:

Teachers expressed their opinions about doing more speaking activities in lessons. The recent speaking exam had some influence on their activity choices. The excerpts expressing this are shared below:

I started to think that I should prepare more speaking activities for students. In fact, it showed me that I, as a teacher, should do more speaking activities rather than the students...I include it more in my activities. However, while I wanted to include all the skills, unfortunately, I cannot complete my weekly plan. (P5)
(More speaking activities)

Increase of Students' Self-Confidence in Speaking:

The increase in students' self-confidence in speaking is created as a code because most of the participant teachers stressed that speaking exams pushed students to speak at least during the exams. As a result of the compulsory speaking time during exams, every student, even the shyest ones, had a chance to speak even if they were reluctant at the beginning. Teachers report that this situation made students realize that

they could speak English and increased their self-confidence. Opinions supporting this are exemplified below:

The children's self-confidence began to increase, and they began to see English as a language. In addition, thanks to speaking exams, students' self-confidence increases, while children who have difficulty expressing themselves in writing can get higher grades. (P1)
(Increasing self-confidence)

It was gratifying to see the students express themselves in the target language and increase their self-confidence. (P8)
(Increasing self-confidence)

Increase of Students' Awareness of Speaking and Giving Importance to It:

According to the teachers' statements, these exams have created an awareness in students about speaking. Students have been made to pay more attention to speaking and take it more seriously. This situation has made the teachers characterize the speaking exams as useful from the teachers' point of view and reflected positively on their attitudes. As a result, the code "increasing students' awareness of speaking and giving importance to it" was created as a positive thought that teachers have. The comments were included below:

Now, when we have speaking activities in the course book, I warn students that they may encounter this section in the speaking exam; at least some of the students pay attention this way, and we practice English in those sections more carefully. (P7)
(Increasing the awareness)

Unfortunately, we were not used to giving students enough opportunity to speak. Maybe this exam can increase this awareness. (P5)
(Increasing the awareness)

Benefits of Working Together as the Department of English:

While discussing the speaking exams, teachers answered the questions about working together, collaborating, cooperating as a team, grade-wide exams, and the exam preparation process with their colleagues. The benefits of working together as the department of English is a code representing positive thoughts of teachers regarding this part of the research. It is evident from their expressions that most of the students benefitted from this working style. Even if the number of exams was multiplied with the new regulations and the workload was increased, they shared that they were happy to divide their responsibilities with colleagues. The crowded departments sometimes shared exam preparation duties and managed the problems. In their opinion, getting an idea and suggestion from another colleague is helpful for a healthy examination system, making the speaking exams more useful. As a result, most teachers feel positive about working together even if some exceptions claim the other way. Both ideas are given below to exemplify different views on this subject:

I find it logical that the exams are shared among the school department. (P7)
(logical)

I believe that it is a positive process that strengthens communication within the English department and uses common sense. (P8)
(positive process)

I see it as problematic. There are five people in our department. Even though I clearly see what my friend in the group did wrong or missing, I cannot say it in order not to hurt him. We say okay, that's fine, and move on, which reduces the quality. If I had prepared an exam on my own, a better-quality exam could have been produced. (P4)

Useful for Students' Skills Learning:

When teachers' overall thoughts are evaluated, usefulness is one of the points. Some participating teachers stated that these exams are useful activities for students to learn language skills. According to them, the more practice in the target language, the more useful the language learning environment. Some participating teachers touched upon points related to communication skills, mentioning that speaking exams develop

language skills and support communication skills. Opinions on this subject are exemplified below:

I find the new regulations really useful. It can give children self-confidence, which is nice in this respect, or they become a little more motivated and more responsive in the lesson. I think it is also useful in this respect. (P10)
(Useful)

Making speaking exams mandatory allows students to improve their communication skills and makes language learning more effective. These exams provide students with the opportunity to practice using language in real life, increase their self-confidence, and contribute to the use of language as a functional tool. Therefore, making speaking tests mandatory can make language learning more useful and effective. (P9)
(Useful)

4.1.1.2. Negative Thoughts Related to the Recent Speaking Exams

Based on qualitative data, five negative thought codes of teachers regarding the recent speaking exams were found (Figure 4.4). Three subcategories related to teachers and ELT were defined. A detailed explanation of these codes and the excerpts taken from the interviews are provided in this section.

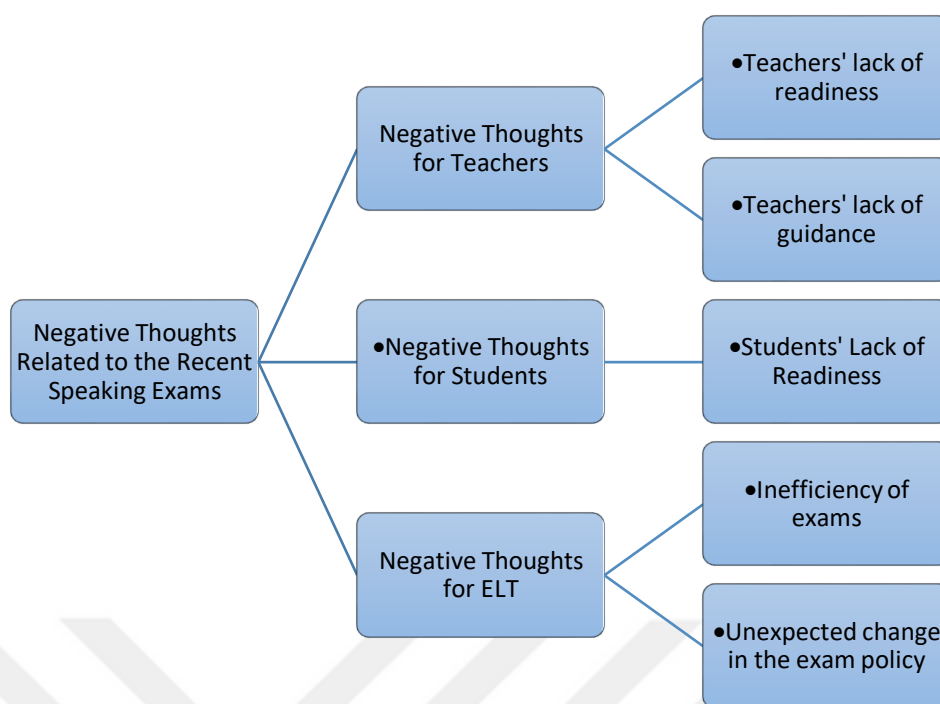


Figure 4.4. The Codes for Negative Thoughts Related to the Recent Speaking Exams

Inefficiency of Exams:

Although teachers make general statements about exams as necessary and useful while answering interview questions, they complain about several issues, particularly regarding the functionality of exams. These include deliberate oversimplification of exams by teachers, ineptly prepared, non-authentic exams, answers with ready-made patterns, rote-memorization of students, unrealistic results with inflated grades that do not reflect reality, falling short of the target, and exams that fail to fulfill their purpose. Teachers reported that they took some actions as a precaution because they were worried about students getting low grades due to new regulations. The increase in the course passing grade from 45 to 50 with the new regulation was effective. Teachers said they did their best because they were worried that this situation might be too challenging for the students whose course passing grades were raised and who would now have to take more exams, and their course success would decrease. These teachers clearly stated that they did not want students to get low grades, that they oversimplified the exams to help them, gave inflated grades, gave students texts so they could memorize them, and gave exam questions in the form of worksheets. This situation seriously affects the ability of an actual exam to serve its purpose, and it is understood from the data collected that those teachers are also aware of this situation.

Participating teachers consider this situation a serious problem and limitation regarding speaking exams. They see the difference between the ideal and the reality, and while sharing their experiences, they also openly expressed their professional dissatisfaction and discomfort with this situation. Expressions related to these problems are provided below:

I work in a crowded school where students come from villages. The success level of students is quite low. We were giving high scores so that children could pass the course. The system was pushing us to do this. But even getting some children to reach 45 was quite difficult. Now we will try to bring it to 50 so that the same children do not fail the classroom. I often witness that teachers worry about this more than students. (P4)
(Inflated grades)

Grades do not reflect students' actual performance. Because the children were given ready-made patterns and the children memorized these patterns. I don't think students assess their performance because they rote memorize it. (P3)
(Rote memorization, ready-made patterns, unrealistic exam results)

Since the transition to the new system was very sudden and in order not to negatively affect the motivation of the students, I felt the need to give speaking topics that were easier than their level. Only in this way was it possible to raise their self-confidence and encourage them to speak. Therefore, the results do not reflect their actual performance. (P11)
(Oversimplification, unrealistic exam results)

Teachers' Lack of Guidance And Unexpected Change in the Exam Policy:

Teachers' lack of guidance and unexpected change in the exam policy were two negative situations teachers frequently mentioned in their interviews. According to the interviews, one of the teachers' general thoughts about the new examination system and speaking exams was that this process was put into practice suddenly, and there was insufficient information and guidance on this subject. Some teachers stated that they had never administered a speaking exam before. They stated that they needed more information and guidance to be efficient before starting the process. They

expressed that they needed time for preparations. They also emphasized that they had negative thoughts about sudden changes in the education system. According to them, sudden changes made without consulting teachers created problems and could not be permanent. Statements supporting this view are shared below:

The mentality that it will get better with time is something that should never be implemented in the education system. Sudden changes in the process are doomed to falter. ...it should be mandatory. Yes, if it is the content we teach in the classroom and we are going to evaluate all the content we teach, exams should be mandatory. But it needs to be guided correctly and the guiding part needs to be done correctly. I think so. (P3)

(Sudden changes, need for guidance)

The most problematic situation about the regulations is that we were informed about this in the middle of the semester, just before the exams, and we were not given a proper guidance. It is not easy to suddenly put into practice something you have not done for years. (P4)

(Unexpected change, lack of guidance, lack of readiness)

I was not ready for this at the beginning of the school year. Maybe if this change had been finalized in the end-of-year seminars of the previous academic year, I could have completed the necessary preparations during the summer vacation. (P11)

(Lack of readiness)

We found ourselves in a process that started with an unplanned, sudden decision. (P8)

(Unexpected change, lack of readiness, lack of guidance)

I find it nice and useful to take a speaking test. But I do not find such a sudden transition right because the students were not prepared for it and neither were the teachers. (P10)

(Unexpected changes, lack of readiness)

If teachers had been included in the process, classes had been observed, deficiencies in the teaching process regarding speaking and listening had been identified, and then the exam had been considered. But we are a society that thinks about the result rather than the process. (P5)

(lack of planning, sudden changes)

Teachers' Lack of Readiness and Students' Lack of Readiness:

It is clear from the statements of the participating teachers that the sudden change of the examination system sometime after the school opened caused the teachers to feel unprepared to face this new situation and to develop negative feelings. Almost all of the teachers stated that they were not ready for this change in the examination system and were very surprised and panicked. Teachers were also asked if the children felt ready. According to the information obtained, no teacher evaluated the students as ready. Therefore, the lack of teacher and student readiness was created as two negative codes. The related statements are exemplified below:

Of course, we were not ready because something like this suddenly came from above; yes, we speak English in the classrooms, yes, to a certain extent, um, this is the language we teach, but the exam for this was to be held, so suddenly, we were informed about a week before the exam week. (P1)

(Lack of teacher readiness)

I was definitely not ready. Since it was a last-minute decision, I did not have the chance to make any preparations. I was not ready either psychologically or professionally. (P4)

(Lack of teacher readiness)

Such a transition had to be made at the end of the year, and the preparation process had to be planned well. It took some time for both the passing score to increase and the children to get used to the new examination system, and the students were never ready. (P9)

(Lack of students' readiness)

4.1.2. Theme 2: Teachers' Feelings

The second research question of this study is, "What are the feelings of EFL teachers related to the recent English speaking exams?". Two categories were created based on this research question, and the data drawn from teachers' interviews related to Positive and negative feelings related to the speaking exams. Considering the situations in which most of the participating teachers expressed their positive feelings, the teachers' motivation to prepare more English activities and to speak more English in the lessons, they observed an improvement in the students' self-confidence in speaking, the feeling of professional accuracy of measuring all four language skills as it supposed to be, the increase in students' awareness of speaking and positive feelings about students starting to give more importance to speaking, and finally positive feelings about working in collaboration with other English teachers in their departments come to the fore.

Teachers' negative feelings were also revealed in the interviews. These are feelings of professional dissatisfaction, the reflection of students' negative feelings about the exam on teachers, stress factors related to the exam groups like the feeling of pressure created by grades such as the 8th grade, teachers and students feeling unprepared for speaking exams, feeling overwhelmed and stressed by the work that needs to be done in a short time under the busy work schedule, and the other challenging feelings that teachers struggle with in their inner worlds.

4.1.2.1. Positive Feelings Related to the Recent Speaking Exams

In the light of teachers' interview data, three subcategories and six positive feeling codes were created about teachers' recent speaking exams (Figure 4.5.). The comments that are effective in creating these codes are explained with example excerpts below.

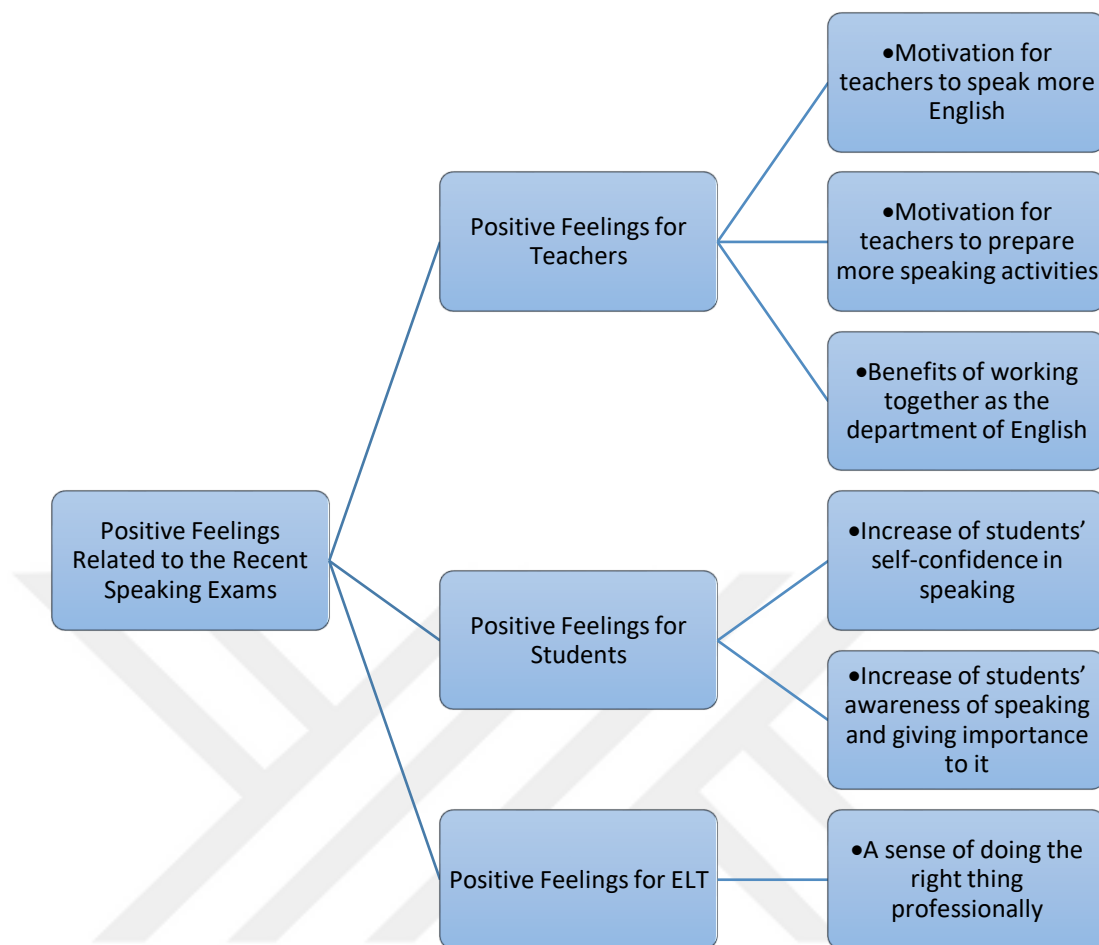


Figure 4.5. The Codes for Positive Feelings Related to the Recent Speaking Exams

The positive feelings are the increasing motivation of teachers and students to speak more English and having more speaking activities, increasing students' self-awareness about their English speaking skills and exams, the English teachers' feeling of doing the right thing professionally, and positive feelings towards the benefits of working together in the department of English. Except for the last code, "A sense of doing the right thing professionally," the first five codes were given in detail and supported by quotations in the previous sections. According to the teachers, doing the right thing professionally with the decision taken created positive emotions in the teachers. The related excerpts are given below:

I feel more at ease that I am doing what I am supposed to do as a teacher.
 Otherwise, something was missing. I can say that it felt better professionally.
 (P5)
 (Professional satisfaction)

I definitely feel like it was a long-overdue decision. I think I finally found the right path. Positive emotions. I feel satisfied. (P1)
 (Professional satisfaction)

4.1.2.2. Negative Feelings Related to the Recent Speaking Exams

As a result of interviews with teachers, two subcategories and seven negative feeling codes were created about teachers' recent speaking exams (Figure 4.6.). The comments that are effective in creating these codes are provided below.

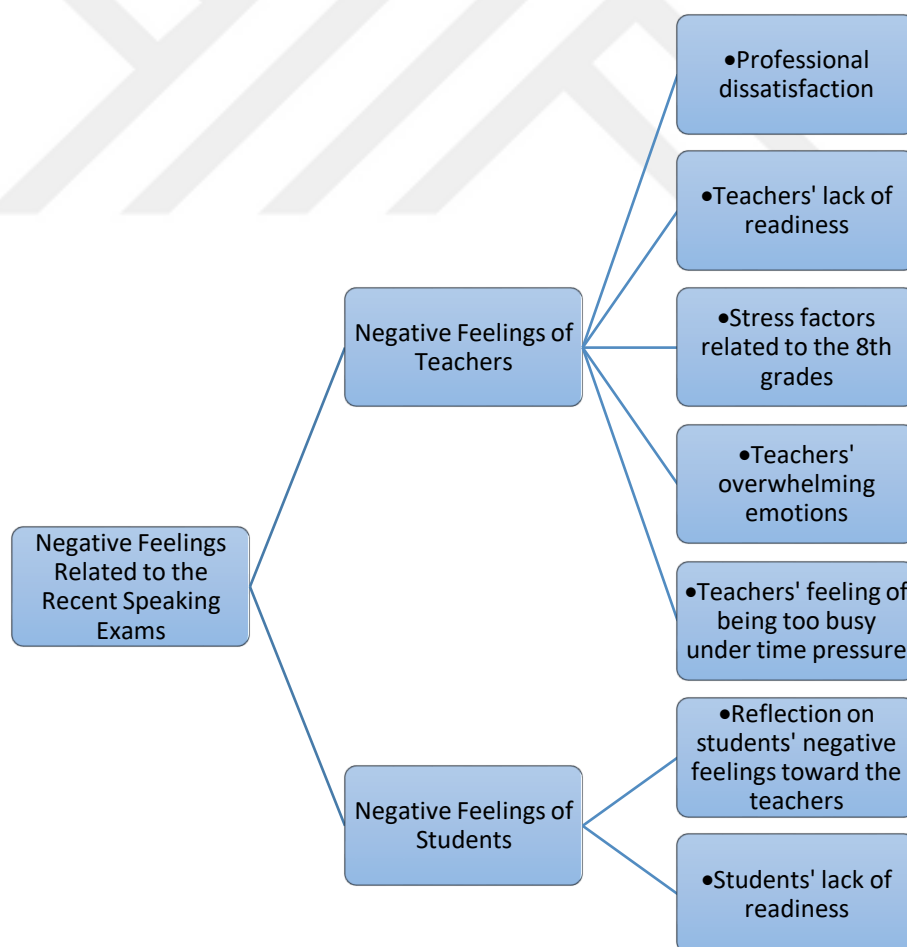


Figure 4.6. The Codes for Negative Feelings Related to the Recent Speaking Exams

Teachers expressed both positive and negative feelings about the newly introduced speaking exams. Not being satisfied professionally, being under the influence of students' negative emotions, the feeling of difficulty in working with students (who take the standardized exams at the end of the 8th grade) called "exam groups", teachers, and students suddenly entering a process for which they do not feel ready, teachers feeling stuck as they have to complete much work in a limited time, and many challenging feelings are among these negative feelings.

Professional Dissatisfaction:

In the interviews with teachers, it is understood that teachers are not satisfied with the speaking exams they deliver. Teachers feel professionally inadequate and dissatisfied due to negative situations such as unsuccessful exam practices, inflated grades, stressful time management, adaptation problems to sudden changes, and student failures. As a result, they often question their professional competence. All of these situations can be categorized as professional dissatisfaction. Opinions on this subject are exemplified below:

Unfortunately, speaking exams did not satisfy me professionally. I do not feel comfortable at all. I have done research on how speaking is taught and how it is evaluated from the sources I follow myself, but I feel very inadequate. Every time, I wish my institution would support me, provide teachers with quality and efficient in-service training, feel more self-confident towards our students, and experience the satisfaction of success in our classrooms. Unfortunately, since this is not the case because our textbooks are not good enough and I cannot feel that I am a professional teacher, I do this internal questioning every exam period, with regret. (P5)

(Professional dissatisfaction)

It did not satisfy me professionally because, as I said, the aim of the speaking exam was to speak more fluently, which was expected from us, but in no way was fluency achieved. On the contrary, the children did not have any background regarding speaking, so we could not go beyond certain patterns. Of course, the fact that I could not go beyond certain patterns made me professionally unsatisfied. (P10)

(Professional dissatisfaction)

The fact that I have never administered the speaking exam since I started my career has made me not fall behind in this sense. I could prepare more original, authentic, and meaningful exams if my knowledge were fresh. Later, I read it myself and remembered it. But our second exams were not much different from the first. Because there are factors here that are beyond the individual effort of the teacher; for example, factors such as the classroom environment, the number of students, and the competencies and equipment of the teachers in the department seriously affect the quality of the exam. Personally, I have zero satisfaction professionally. (P4)

(Professional dissatisfaction)

Reflection on Students' Negative Feelings Toward the Teachers:

Although this study focuses on teachers' emotions, teachers and students cannot be considered independently of each other as beings who act together in the same environment. The positive or negative reactions students give directly affect teachers' feelings. As understood from the interviews, students stated to their teachers that they felt anxious, fearful, low-motivated, bored, panicky, and reluctant due to the increase in exam loads and encountering a new type of exam they had never experienced before. These negative feelings also negatively affected the teachers' emotions. The teachers' comments related to the code "reflection on students' negative feelings towards teachers" are below:

As feedback, students said that of course they would be happier if there was no speaking exam. We can say that it is directly proportional to the students' feelings about how it affected me. At first, I was nervous about the coming of this oral exam, how it would be, how we would prepare, we did not have any preparation for it, how we would develop the conditions. But then somehow we did it. (P3)
(Affected emotions)

Students said that they did not want a speaking exam due to their habit of taking ongoing written exams. This negatively affected my motivation to prepare for the speaking exam. (P6)

(Affected emotion)

I received generally negative reactions from the students because they had prejudices such as "how will it be, how will we do it, we cannot talk about it" and this still motivated me a little negatively because I had the perception of how to break this prejudice and I got a little stressed because there are very few class hours for students to talk. It made me question myself a lot as to how I could break this. (P10)

(Affected emotions)

Stress Factors Related to the 8th Grades:

One of the most critical situations affecting the future of students in secondary schools is the high school entrance exam. Especially 8th-grade students and teachers spend extra hours to get prepared for this standardized exam. Students take support courses and work primarily on multiple-choice exam success to succeed in these exams. It is understood from these interviews that teachers have some reservations about 8th-grade students. Teachers do not want to put more stress on these students. In addition, according to the teachers, it seems appropriate for children to devote more energy and time to courses with a higher exam score coefficient. According to interviews, these teachers' duty to prepare students for the "big exam" seems to take precedence over speaking exams. Teachers feel stuck trying to complete the requirements of the curriculum of 8th-grade students on time and solve many tests for the high school entrance exam preparation. It should be remembered that schools and teachers are evaluated yearly based on how many students they can place in a good high school, which also puts pressure on teachers. Speaking exams are not thought to be necessary for this success. For this reason, some teachers think that speaking exams should not be administered to the 8th-graders, which we call the exam group. The comments related to the code "stress factors related to exam groups" are provided below:

I found the speaking exams challenging for intermediate grades and unnecessary for 8th-graders. It was appropriate to start only with 5th-graders. Or it would have been better if it had started by excluding 8th-graders. Since 8th graders have been focusing on testing for years, it was unnecessary to force them. (P2)

(Unnecessary, challenging, should be excluded)

Expecting the children, who had been preparing for the exams only through tests for eight years, to suddenly answer questions at the lowest comprehension level in their own words created a lot of tension for them. As 8th-graders, we cared about high school entrance exams, and we mostly focused on vocabulary and reading. In class, of course, we do the speaking skills parts in the book, but at home, they study based on reading. Those who have no hope of studying or those who are already away from the course move further away because they do not want to talk at all. 8th-graders might not have been included in the application this year. (P7)

(Should be excluded, unwillingness)

Teachers' Feeling of Being Too Busy Under Time Pressure:

One of the striking points in the interviews was that teachers stated that they were under constant time pressure. This time pressure was related to meeting the exam schedule and completing the syllabus on time. Busy exam week schedule, inability to complete topics on time, and keeping up with the curriculum were some of the major outcomes of the new regulations regarding English exams. Along with the English speaking exams, students must have Arabic (only in İmam Hatip secondary schools) and Turkish. The tripling of the number of exams in two or three language courses directly caused the exam week schedule to become more intense. In this case, according to the teachers' statements, it became difficult even to find a time slot to give the students the exam because grade-wide exams for all courses took away from students' English hours. Regarding children taking 19-20 exams in just two weeks, teachers often stated that they acted under a lot of time pressure with the new exam regulations. All of these reasons cause negative feelings such as stress in teachers. The comments related to the code "teachers' feelings of being too busy under time pressure" are given below:

At İmam Hatip secondary school, children take 19-20 exams in two weeks. Also, in the third week, I have to administer the exams of the children who do not come for the exams on time. When you think about this for the first and second exams in two terms, a significant amount of time of a semester is given to the

exams only. There are ten units in our curriculum, and completing the units in the remaining time was almost impossible especially this year. In the past, when everyone was taking their own exams in their own classes, keeping up with the curriculum was not such a problem. Lesson hours need to be increased, or the units need to be reduced in order to cover every skill. (P4)

(Busy schedule, time pressure, need for improvements)

I wish we had at least two hours of lessons for each skill. Especially in lower classes, three lessons are definitely not enough. Even if we do not administer a speaking test, it is insufficient. Three hours pass before the child adapts and starts to love the language, and therefore we cannot provide efficient learning. (P1)

(Insufficient lesson hours, time pressure)

Teachers' Overwhelming Emotions:

Boredom, exhaustion, fatigue, anger, anxiety, pessimism, stress, hopelessness, helplessness, panic, and resentment are some emotional states teachers describe about the sudden transition to the new examination system. Teachers frequently expressed their negative feelings about this sudden transition and the problems they encountered afterward in interviews. Therefore, these challenging feelings were opened as a code. The participant teachers' comments related to the code "teachers' overwhelming emotions" are given below:

I was angry because it was a decision taken so suddenly and without any basis. (P11)

(Overwhelming emotion)

When I heard the decision, I felt anger, stress, anxiety, and exasperation. Because I am really tired of these last-minute things. I would like everything to be planned, and we would do our jobs more smoothly this way. I am a person who does not like to leave my work to the last minute. It was terrible to find out two weeks before the exam that the entire exam system had changed radically. (P4)

(Overwhelming emotion)

At first, it created despair and pessimism. As a teacher, I felt a panic wondering how we would get out of this situation, what we would do, how we would manage this process. We can call this professional boredom. Suddenly a decision comes, you make preparations accordingly, then you collect the results, then you analyze it, it was a tiring process, but I was relieved because we succeeded in the end. I said it could be done, we did it somehow. But if these were well planned and included in the curriculum, I would feel a little more at peace professionally.
(P3)
(Overwhelming emotion)

4.1.3. Theme 3: Experiences and Observations

The third research question of this study is, "What experiences were encountered by EFL teachers during the implementation of the recent English speaking exams?". Two categories were created based on this research question and the data drawn from teachers' interviews: Problems and limitations of speaking exams and procedural observations of the new speaking exams. When teachers' experiences before, during, and after the speaking exams are analyzed in detail, it is seen that teachers encounter a wide variety of problems and restrictions. These are unexpected changes, lack of teacher and student readiness, teachers' lack of guidance, not giving enough space to speaking in the curriculum and textbooks, inefficiency of exams, crowded classroom problems, impracticality of grading and e-school system, increased workload for teachers, teachers' feelings of being too busy under time pressure.

Another category, "procedural observations of the new speaking exams," focuses on real-life speaking exam situations. What kind of practices the teachers who work in different schools and with students from different backgrounds carry out the exam and what situations they encounter are coded under this heading. When the reality of speaking exams is considered, differences in practice emerge, crowded classroom problems, increased workload for teachers, inefficiency of exams, teachers' feeling too busy under time pressure, and lack of teacher equipment came up.

4.1.3.1. Problems and Limitations of the Recent Speaking Exams

As a result of interviews with teachers, four subcategories and nine problems and limitations codes were created regarding teachers' recent speaking exams (Figure

4.7.). The discourses that are effective in creating these codes are given and explained below with example excerpts.

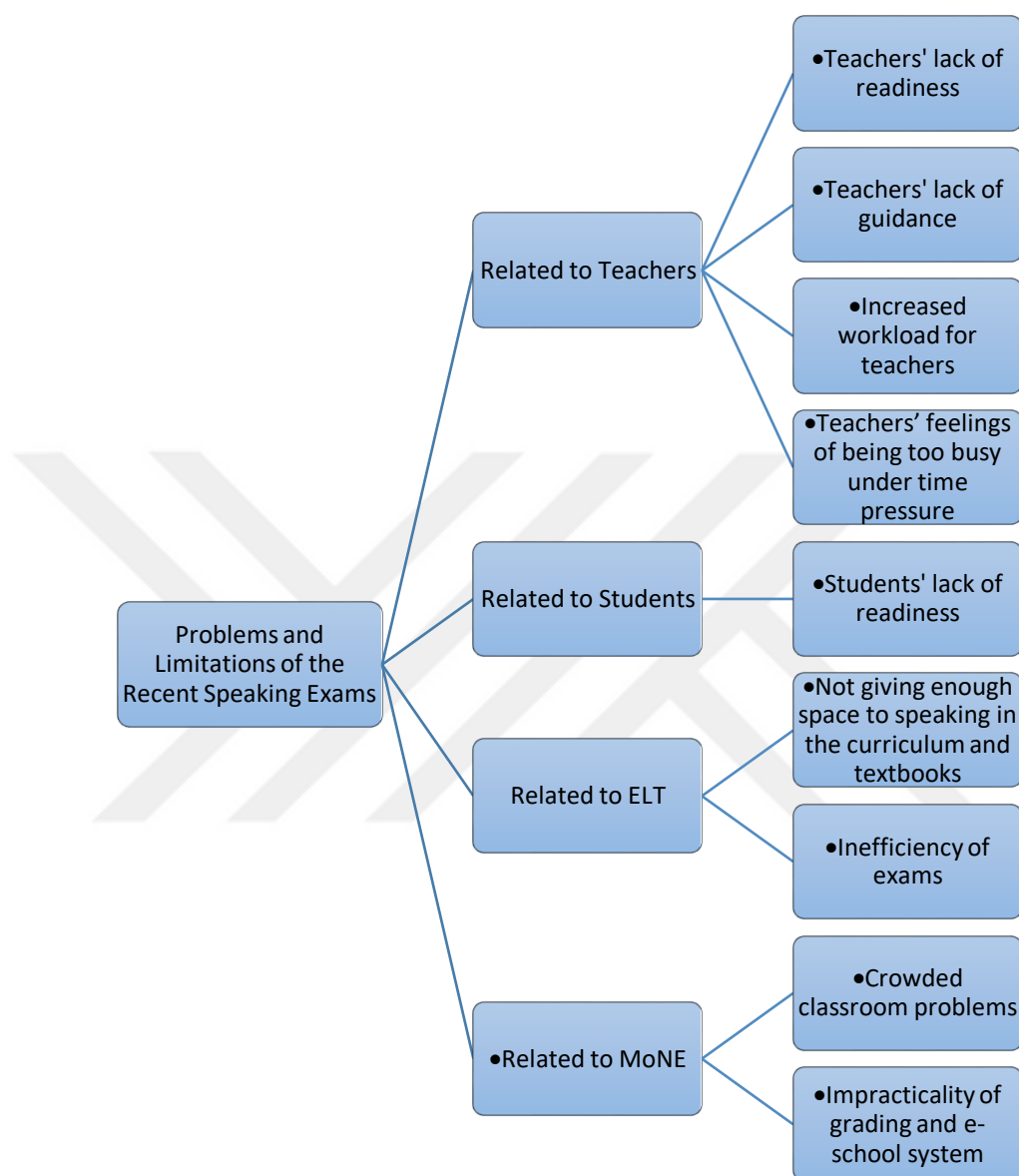


Figure 4.7. The Codes Problems and Limitations of the Recent Speaking Exams

Students' Lack of Readiness:

In this research conducted on teachers, teachers' observations about students' experiences were also asked. Students' negative feelings were grouped under the code "problems and limitations of speaking exams." Stress, anxiety, low self-esteem, adaptation problems, unprepared students, neglect, shyness, panic, fear, and prejudice

were the biggest problems and restrictions highlighted by teachers in the interviews. Teachers' observations about students' problems are exemplified below:

Children who take their grades seriously, and when this examination system was suddenly announced, they panicked. I do not think they are ready to switch to this system. My 8th-graders will take the high school entrance exam anyway, and the written exam was also open-ended for the children preparing for the exam, so we were not allowed to ask for tests, which was announced at the last minute. They panicked about that; this is for Turkish, Arabic, and English, and they all have separate speaking and listening exams, so the children panicked for English and all language classes. (P3)

(Panic, lack of readiness)

Students were very nervous and scared about the exam. Many of them said then that they could not pass this course. Even the best ones got very excited during the practice. (P7)

(Negative feelings, lack of readiness)

They mostly felt fear and anxiety. Even though we wanted them to get used to and make the easiest transition in the face of this rapid change and necessity, we had many students who refused to take the speaking exam. (P11)

(Negative emotions, lack of readiness)

One of the important problems in the application of the speaking test was that extremely shy children did not speak. Some of them also tried to speak by whispering. It may be called low self-confidence. (P4)

(Lack of readiness)

Not Giving Enough Space to Speaking in the Curriculum and Textbooks:

When teachers were asked about their observations and experiences during the process, one of the focal points was that speaking skill needed to be included more in the curriculum and textbooks. Most participants claimed that there were not enough speaking activities, and those that were available were insufficient and did not provide

children with enough speaking practice and fluency. Expressions related to this situation are exemplified below:

We cannot ensure fluency in speaking in classes, our curriculum is not suitable for this. We can make students talk with certain phrases in the curriculum, such as “What is your favorite Sport? My favorite Sports is football”. We cannot go beyond these speech patterns because we neither have enough time nor our curriculum is suitable for this. (P10)

(Curriculum problems)

For speaking courses, courses such as public speaking should be added to the curriculum as in universities, and students should be able to improve themselves in speaking skills such as addressing, fluency, using gestures and facial expressions. But it is not enough right now. (P6)

(Insufficient, curriculum problems)

Although the curriculum is overcrowded, there are not enough speaking activities in it. Time should be spared for practice by getting rid of the curriculum burden. (P8)

(Curriculum problems)

Crowded Classroom Problems:

Since class size is a crucial factor in applied exams and can bring about more than one problem, crowded class problems were examined as a separate code based on the teachers' experiences. According to the information received from teachers, time constraints, noise, discipline problems, and extended exam hours have posed problems in applications conducted in crowded classes. When the statements of the teachers are examined, it is understood that a student is given the right to speak for an average of 5-10 minutes, and completing a speaking exam takes 1-3 class hours in total, depending on the number of students. Teachers state that while they give the student a one-on-one speaking test, the rest of the class distracts them by making unwanted behaviors, such as talking or standing up, and disrupts the discipline in the exam environment. Additionally, exams that cannot be completed in one class hour cause

even greater problems in the tight exam calendar. Comments about this situation are below:

I call the students to my table in the order of the class list one by one. Other students are supposed to listen quietly, but there are always some students who make noise, distract, and stand up. Other children in the class secretly try to help their friends who are taking the exam by mimicking and whispering. This gets harder to control in crowded classrooms while administering the test under time pressure (P4)

(Classroom management problems)

During the speaking exam, the classroom environment is very noisy because there are students who listen to me and their friend during the conversation and at the same time try to produce answers on their own. They ask too many questions to each other, and since I am paying attention to one student during the conversation, I cannot pay attention to the other students, which causes some chaos in the classroom. (P10)

(Classroom management problems)

Problems include students making noise while waiting for other students and the exam taking too much time in line with the class size. Among the students waiting for their turn, some are studying for their grades, some are making their friends talk as if they are taking an exam, and some are trying to misbehave because they think the class is dismissed. (P2)

(Crowded classroom problems)

Impracticality of Grading and E-school System:

One dimension of exam experiences is the electronic school system. Teachers must record each child's grades in this system shortly after taking the exams. When the teachers' experiences were examined in the interviews, it was determined that there were problems with the grading and e-school system. The new exam system has divided the English exam into three stages (listening, speaking, reading+writing), and a percentage value has been dedicated to each of these exams. Accordingly, speaking exams constitute 25% of students' English exam scores. However, interviews show

that teachers were expected to calculate these rates individually for each student's exam. An application to calculate this has not been added to the e-school system. After calculating the average of the student's three exam scores, the teacher must record the final average score in the system. While the teacher wastes much time with these procedures, the students cannot see the grades they received from the individual exams in the system. In addition, the grade obtained after calculating the averages is not always a whole number; small fractions appear, and teachers round these numbers up because they cannot enter the fractions into the system. Example statements regarding these problems are given below:

The inadequacy of e-school is a big problem as exam grades are not uploaded separately into the system, a teacher has to calculate the percentage of all exam grades himself and upload only the average. This situation is both not transparent and confuses children. Since 25% of the exam is taken, the average result is an odd number, such as 43.5 or 33.75. We need to write these scores into the system by rounding them up. The other problem is that students who do not arrive at the exam time and take the exam later create great confusion. We have to calculate the percentage again separately, and the system also shuts down. That's a complete mess. (P4)

(E-school problems)

E-school must be fixed urgently. There is no system in e-school to upload exam grades separately, why does a teacher have to calculate the percentage of all exam grades himself? This is really ridiculous. It is not very difficult to install this software in e-school. But it is not done. In order to enter an exam score of 130 students, I have to take the average of three grades according to different percentages. And some kids question this a lot. 5th and 6th graders have difficulty keeping track of their exam scores individually. When they see their averages, they complain that “I got higher in the speaking exam but my grades in e-school were lower”. They can't understand the concept of average there. An add-on should be made there so that children can see all their grades in the system themselves. Thus, doubts are eliminated. (P1)

(Transparency problem of grades, technique problems)

Increased Workload for Teachers:

With the new regulations, there was a significant increase in the workload of teachers. When the participant teachers were asked about their experiences, they positioned increased workload as a negative factor. When the data is examined, it is seen that teachers are not at all satisfied with the increasing workload. Comments about this are presented below:

With the new assessment and evaluation system, the workload of language teachers has increased greatly. Lesson hours are insufficient to develop basic skills. Current lesson hours are definitely inadequate with the new system. (P8)
(More workload)

The system should be improved by easing the workload of teachers or increasing their salaries in direct proportion to the workload. We do a lot of exam preparation and then exam grading. Exam work has tripled. (P9)
(Paper work, more workload)

Even thinking about the workload is tiring. (P7)
(Mental fatigue, workload)

4.1.3.2. Procedural Observations of the New Speaking Exams

Interviews with teachers led to the creation of three subcategories and six procedural observation codes regarding teachers' recent speaking exams (Figure 4.8.). The comments that are effective in creating these codes are given and explained below with example excerpts.

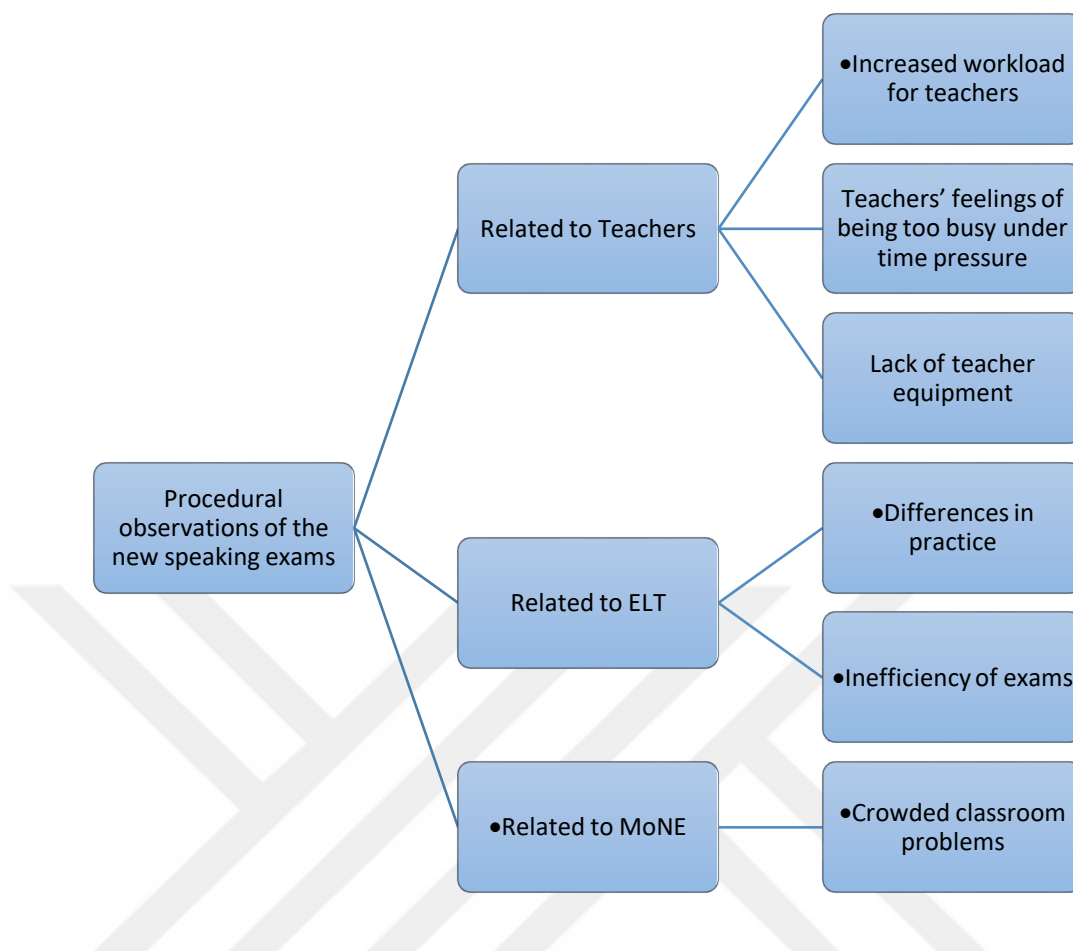


Figure 4.8. The Codes for Procedural Observations of the New Speaking Exams

Differences in Practice:

This study examined what happened before, during, and after the exam practice specifically under this category. The conditions under which and in what way the teachers applied the speaking tests and how they prepared them were also asked, and as a result, differences were detected between the practices. For example, situations have been identified where students are given one-on-one exams, exams are held in groups, exams are evaluated with rubrics or without using rubrics, giving or not giving feedback after the exam, and the exam duration is short or long. In this case, non-standard rubrics, no or limited feedback, different examination styles, variable exam hours for each student, and each classroom can be classified as limitations and problems. Thus, this was named as a differences in practice as a code. Expressions that exemplify this are given below:

I took the students in groups of four, and while creating the groups, I sprinkled one person into each group who was more interested in the lesson and asked them the questions first. Thus, the stress of others was reduced. I asked each person in the group separate but similar questions such as “Do you like football? Why? and Do you like tennis? Why? “I did not ask questions beforehand, but I integrated it by asking similar things during the lesson; that is, I did not ask anything other than what we did in the lesson. I asked three questions and each student answered individually. Except for the welcome, how are you part. In the grading section, I just wrote down their names on a blank sheet of paper so that I would not waste any time and could evaluate them in more detail. (P6)

(No rubric)

I gave the topics in advance, and they memorized them. When that happens, they do not feel like they are taking a speaking test. Since they individually choose the subject, they are best at among the subjects I give and speak about it; they memorize it; that is, they speak by rote memorizing it. If I go out of that context and ask the child what is your name, he/she will just look at me. We applied a scale to grade; there was a rubric. I distributed points to the children in the rubric under the topics I gave. I check pronunciation, fluency accuracy, all of them. I gave a grade accordingly. We also prepared the rubric individually. Because we teach in different classes. We teach no class on grade-wide exam terms with my colleague at school. (P3)

(Oversimplified exam with hand-out questions)

Lack of Teacher Equipment:

It is clear from the data that teachers had experiences where they felt inadequate or that the other teachers they collaborated with were inadequate. In the interviews, wishes, needs, and recommendations such as quality in-service training and the providing more information and assistance by MoNE were also expressed. Comments about it are below:

Teachers may be inadequate in administering exams. It is a problem that it is not known whether teachers are qualified to conduct these exams. I think the authorities are quite unaware of the field... Unfortunately, speaking exams did

not satisfy me professionally. I do not feel comfortable at all. I have done research on how speaking is taught and how it is evaluated from the sources I follow myself, but I feel very inadequate. Every time, I wish my institution would support me, provide teachers with quality and efficient in-service training, feel more self-confident towards our students, and experience the satisfaction of success in our classrooms. Unfortunately, since this is not the case because our textbooks are not good enough and I cannot feel that I am a professional teacher, I do this internal questioning every exam period, with regret. (P5)

(Need for in-service training)

4.1.4. Theme 4: Future Thoughts

The fourth research question of this study is "What are the future thoughts of EFL teachers related to the recent English speaking exams at secondary schools?". Two categories were created based on this research question and the data drawn from teachers' interviews: expectations, beliefs, and requests for speaking exams and recommendations to improve the speaking exams. Teachers' expectations are high that these regulations will continue. Moreover, most teachers anticipate that these exams will be updated and improved over time. Additionally, no teacher wants speaking exams to continue as they are now.

Teachers gave plenty of advice on how speaking tests could be improved. These recommendations were as follows: more weekly hours of English course should be given, less workload for EFL teachers is necessary, need for extracurricular support from families, gradual transition to the new examination systems should be applied, there must be more curriculum planning with more speaking activities, regulation in the e-school system and exam scoring system should be arranged, in-service training and more information, better classroom environments, a better-planned education calendar should be provided.

4.1.4.1. Expectations, Beliefs, And Requests for Speaking Exams

Three codes related to expectations, beliefs, and requests for speaking exams were created through qualitative data (Figure 4.9.). The detailed explanation of these codes and the excerpts taken from the interviews are explained in detail in this section.

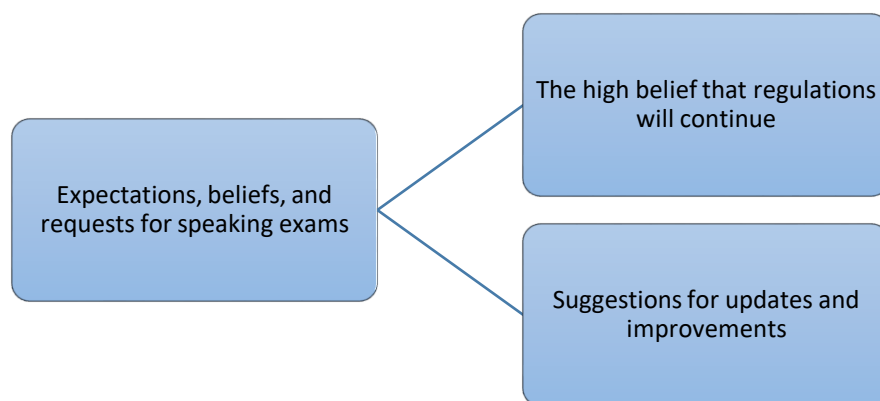


Figure 4.9. The Codes for Expectations, Beliefs, And Requests for Speaking Exams

The High Belief That Regulations Will Continue:

Within the scope of the research, teachers were asked whether these regulations would continue in order to examine their beliefs and predictions about the continuity of this examination system. 8 of the 11 participating teachers clearly stated that they believed that these exam regulations would continue in the future. Although teachers criticized the system's shortcomings, they expressed their belief that the system will continue. Others stated that they did not find the system change permanent, that the system could change with the change of minister, and that the changes occurred quickly. Relevant comments are given below:

I think the exam will continue in the future; I think regulations will be made within the framework of CEFR. (P6)
(Belief in the consistency)

I think this examination system and regulations will continue in the future. (P10)
(Belief in the consistency)

Speaking tests are likely to continue in order to improve language skills through practice and give students experience in real-life language use. Innovations in the field of education, technological developments, and efforts to adapt to student needs will lead to changes in examination systems. (P9)
(Belief in the consistency)

I think this system will not continue in the future. It will change again. Another minister will be appointed, and this system will most likely be abolished. Maybe it will take a year or two. Because a system has never remained the same. I think they will think it is tiring. All language teachers, Arabic, Turkish and English, apply this system. So, I do not think it will continue in the long run. (P3)

Suggestions for Updates and Improvements:

According to the analysis of interviews with teachers, teachers have some expectations and predictions of speaking exams. These include updating, improving, and developing speaking exams in the future. Teachers have both expectations and beliefs on these issues. One of the points teachers had in common during the interviews was that speaking exams should not continue as they are now. When we asked teachers whether they wanted the exams to continue as they are now, the answers we received showed that none of the teachers thought it was appropriate for them to continue in their current state. Relevant comments are given below:

We have already started, it should be continued and improved. (P8)
(Suggestion)

No, the system should not continue as it is; the new system has many shortcomings; it increases the workload of teachers, but there is no positive development or gain in return for teachers and students. Stressful exam weeks are very tiring and challenging for both teachers and students. This exam model should not continue until all these deficiencies are corrected and course hours are increased. It can also be done over time. (P4)
(Suggestion)

Yes, I find it useful to assess listening and speaking in children, so I would like it to continue, but of course, it should be developed a little more. (P10)
(Suggestion)

I do not want it to continue in its current state. (P11)
(Request)

4.1.4.2. Recommendations to Improve the Speaking Exams

As a result of qualitative data analysis, nine codes were created related to developing speaking exams (Figure 4.10.). In this section, these codes are explained and examined with sample quotations.

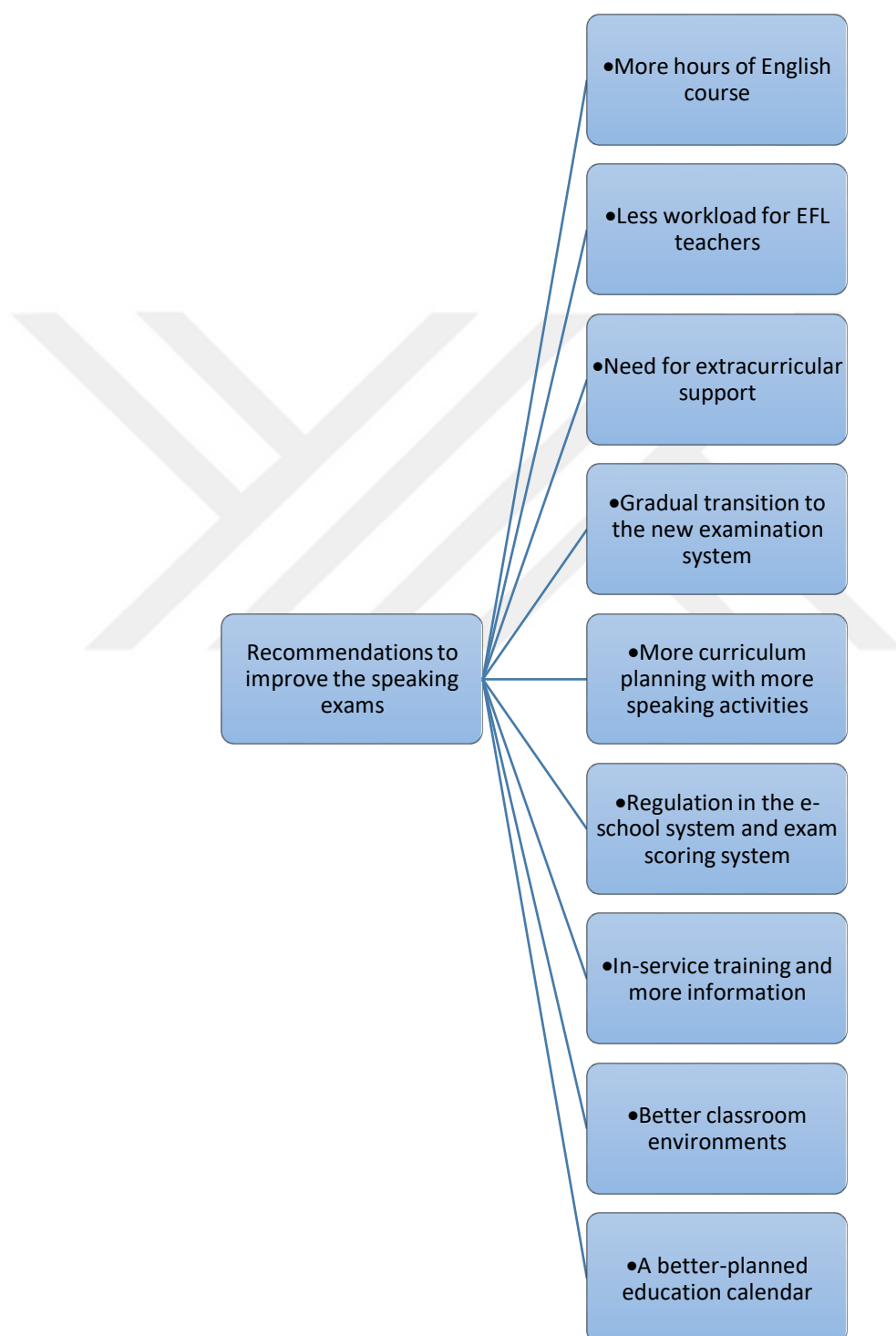


Figure 4.10. The Codes for Recommendations to Improve the Speaking Exams

More Hours of English Course:

Within the scope of the research, suggestions from participating teachers were taken to improve speaking exams. The most prominent of these was the increase in lesson hours, which was mentioned by the majority of participants. Moreover, teachers emphasized that there could be a separate lesson hour for speaking skills to support students' skills more and make the exams more meaningful. Relevant comments are given below:

In order to improve the speaking exams, I think the lesson hours should be increased first because I do not think that teaching certain phrases to students will be of much benefit, that is, I do not think it contributes to the speaking ability. Therefore, if the lesson hours are increased, we will spend more time talking and we will be able to speak more fluently. This way, it will be much better. (P10)

(More lesson hours)

Lesson hours can be increased under the name of “speaking” lessons. In this way, sufficient time is created for long preparations and long exam hours. (P11)

(More lesson hours, speaking lessons)

Less Workload for EFL Teachers:

One of the prominent recommendations to improve speaking exams is that teachers' workload should be reduced. Teachers state that their workload has increased significantly with the new regulations, and they believe this issue should be regulated for the sake of the safety of the exams. Relevant comments are given below:

The system must be improved urgently by alleviating the workload of teachers or increasing their salaries in direct proportion to the workload. At home, I have to constantly grading papers and enter notes. While my workload has increased three times, it is not fair to not see the reward in terms of salary. I think teachers also need financial support to do their jobs properly without teacher burnout. Teachers' discussion of salary and money issues is often neglected. Yes, education and students are important, but it is also very important for a teacher to be compensated for the work he does for the system to function well. (P4)

(Compensations, workload)

Need for Extracurricular Support:

All participating teachers were teachers working in public schools. One of the issues these teachers mentioned was that students needed to receive language learning support outside of school. Children from disadvantaged areas were exempt from family support, private courses, activities, or online or natural environments where they could communicate in the target language. Teachers support the idea that students' speaking skills can be improved with these extracurricular supports and that children will be more enthusiastic about speaking skills and exams. Relevant comments are given below:

In Finland, where I went as part of Erasmus, I saw that students could learn English very easily. When I asked their teachers, I learned that parents expose their students to English through cultural events outside of school. I believe that we will move to a much better level if we encourage parents to provide this opportunity to their students. (P5)

(Opportunity, extracurricular support)

Gradual Transition to the New Examination System:

According to teachers' statements, the sudden introduction of the system after the start of the academic year was one of the underlying problems in the system. Naturally, most teachers advised that this system should be gradual. This gradualness was exemplified by the transition of grade levels to exams over time or the changing of the exam system over the years rather than all at once. Relevant comments are given below:

I think it should have come gradually. Expecting the children, who had been preparing for the exams only through tests for eight years, to suddenly answer the questions at the lowest comprehension level in their own words created a lot of tension for them... 8th graders might not have been included in the application this year. (P7)

(Gradual transition)

I think speaking exams should have been integrated into the system, but they should have been planned with a gradual transition. (P8) (Gradual transition)

The fact that students and teachers were so unprepared and suddenly found themselves in such a system made the process a bit difficult. I wish it could have been progressed more gradually. (P10)

(Gradual transition)

More Curriculum Planning with More Speaking Activities:

Although there are speaking activities and objectives in the curriculum, many participants stated that the curriculum needed to be revised to prepare children for this exam. Participant teachers argue that speaking activities should be increased and the curriculum should be improved. Moreover, recommendations were given, such as increasing the quality of the activities as well as their proportion in the book. Relevant comments are given below:

For speaking courses, courses such as public speaking should be added to the curriculum as in universities, and students should be able to improve themselves in speaking skills such as addressing, fluency, using gestures and facial expressions. (P6)

(More activities in the curriculum)

Speaking and listening are an integral part of language learning. Authorities must have realized this because speaking and listening skills, which were ignored in previous years, are now being included in the process. However, the sad thing is that they see it as right to implement this change directly on the exams. Oh, there is so much to talk about. Speaking and listening activities in textbooks are not good enough. Not enough. (P5)

(Not enough activities)

We cannot go beyond speech patterns because we neither have enough time nor our curriculum is suitable for this...I think the main purpose of the speaking exam is to ensure that students speak effectively, efficiently and fluently, but the level of the students is not sufficient for this because we cannot ensure fluent speaking in the lessons and our curriculum is not suitable for this... And when

we look at the curriculum, we see that speaking is not included much and, as I said, fluency is not achieved... If we are going to make such an exam compulsory, I think we should force the student to speaking skills a little more, this should not be limited to school exams, or we should have a plan where we can allocate time in classes to speak skills more. (P10)

(Curriculum problems)

Regulation in the E-school System and Exam Scoring System:

The parts related to e-school and exam scoring constitute the more technical dimensions of the system. The comments received on these issues underline how important practicality is in teachers' lives and the complaints about workload. According to the interview data analysis, most teachers are dissatisfied with the e-school system. Teachers state that they waste time entering grades and need help with simpler operations. This situation seems to create more problems for teachers working in schools with many students. Additionally, teachers state that students find it confusing that the speaking exam score constitutes 25% of the student's course grade. Problems such as students having difficulty calculating their average grades, students not being able to see their grades separately in the e-school system, teachers doing the math manually to find the average grades, finding small fractional values and uploading them into the system by rounding them were stated by the teachers as the deficiencies of the system that need to be eliminated. Relevant comments are given below:

E-school must be fixed urgently. There is no system in e-school to upload exam grades separately, why does a teacher have to calculate the percentage of all exam grades himself? This is really ridiculous. It is not very difficult to install this software in e-school. But it is not done. In order to upload an exam score of 130 students, I have to take the average of three grades according to different percentages. And some kids question this a lot. 5th and 6th graders have difficulty keeping track of their exam scores individually. When they see their averages, they complain that I got higher in the speaking exam but my grades in e-school were lower. They cannot understand the concept of average there. An add-on should be made there so that children can see all their grades in the system themselves. Thus, doubts are eliminated. (P4)

(E-school problems)

The inadequacy of e-school is a big problem as exam grades are not uploaded separately into the system, a teacher has to calculate the percentage of all exam grades himself and write only the average. This situation is both not transparent and confuses children. Since 25% of the exam is taken, the average result is an odd number, such as 43.5 or 33.75. We need to enter these scores into the system by rounding them up. The other problem is that students who do not arrive at the exam time and take the exam later create great confusion. We have to calculate the percentage again separately, and the e-school system is also closed. That is a complete mess. (P1)

(E-school problems)

In-service Training and More Information:

One of the recommendations given by teachers to improve the speaking exam was that teachers who needed it could receive in-service training. Teachers may naturally feel professionally inadequate over time. It is usual for teachers to need to update their professional knowledge and seek guidance. During the interviews, some teachers expressed that they needed this and that such studies should be planned and presented to teachers. Moreover, teachers stated that more detailed information should be shared with teachers about what is expected from the implemented practice, what is intended, and how the practices should be. Teachers think that exams can be improved with these steps. Relevant comments are given below:

Every time I wish my institution would support me, provide teachers with quality and efficient in-service training, feel more self-confident towards our students, and experience the satisfaction of success in our classrooms. (P5)

A more informative explanation should have been made for friends who have never taken a speaking test before. I found the information incomplete. These deficiencies must be corrected. (P1)

(Need for in-service training)

The fact that I have never given a speaking exam since I started my career has made me not regress in this sense. I could prepare more original, authentic, and meaningful exams if my knowledge were fresh. Later, I did my own readings and remembered... It is the ministry's duty to support teachers in this regard. It would be nice if those who want to benefit from it. For example, videos can be shared via EBA. Anyone who wants can watch and benefit from it. (P4)
(Online in-service education)

Better Classroom Environments:

Most of our participating teachers told us they called the students individually and gave them a one-on-one speaking exam. They claimed that during this process, they struggled with some disciplinary problems and noise in the classroom environment, which made the exam poor quality. Moreover, according to these teachers, conducting applied exams in crowded classes was becoming more difficult. Considering these, according to the teachers' recommendations, improvements to be made in the classroom environment will also increase the exam quality. For example, reducing the number of students in crowded classes, finding solutions to noise and indiscipline, and equipping classrooms with technological equipment can contribute to better speaking exams. Relevant comments are given below:

Students making noise while other students are waiting and the exam taking a lot of time in line with the class size are problems. It will be beneficial for students to talk about any subject with a native English speaker in an environment away from noise. (P2)
(Need for quite classroom)

Other students are supposed to listen quietly, but there are always students who make noise, distract, and stand up. (P4)
(Need for quite classroom)

During the speaking exam, the classroom environment is very noisy because there are students who listen to me and their friend during the conversation and at the same time try to produce answers on their own. They ask too many questions to each other, and since I am paying attention to one student during

the conversation, I cannot pay attention to the other students, which causes some chaos in the classroom. (P10)

(Need for quite classroom)

A Better-Planned Education Calendar:

Teachers also had some recommendations regarding the educational calendar. Teachers suggested that a more detailed and previously clarified timetable and a timely announcement could reduce the problems. According to the participants, with a better-planned calendar, this could be a year in which teachers and students can prepare for the exam more easily, both parties can perform better, and everyone will feel more comfortable psychologically. Comments related to teachers' requests regarding the organization of the academic year are shared below:

You know, you could do a study on this in the summer, or even if you did it during the seminar period, we would talk about it in group meetings and make a plan, but you know, when it happens a week or two before the exam, the speaking exam remains in the air... This work should be planned in more detail and announced in advance... But if it is more planned, next year or so, students will become more familiar with it and it will be more useful. These disruptions and problems are currently caused by lack of system. (P3)

(Better-planned, timing)

We found ourselves in a process that started with an unplanned, sudden decision... More proper planning is needed. (P8)

(Better planning)

I do not think it is right to rush it and make it compulsory before the infrastructure is created. I believe that a more planned, programmed arrangement with clearer lines would be more accurate. (P11)

(Clear lines, better-planning)

I want everything to be planned and we do our jobs more properly. I am a person who does not like to leave my work to the last minute. It was terrible to find out 2 weeks before the exam that the entire exam system had radically changed. (P4)

(Timing, planning)

Such a transition had to be made at the end of the year, and the preparation process had to be well-planned. (P9)

(Planning well)

4.2. The Questionnaire Findings

In this part of the research, the questionnaire results conducted online via Google Forms with 217 secondary school EFL teachers will be analyzed in two parts. Part I relates to the first part of the questionnaire, and Part II relates to the second part.

4.2.1. Part I

While collecting quantitative data, teachers were asked ten questions in the first part of the questionnaire. Five were questions about demographic information, and the other five were about speaking exams in Part I. The questionnaire was conducted with 217 participants (176 female and 41 male teachers). Demographic findings are shown in the methodology section of this study in detail. After the demographic information was collected, the participants were asked whether they had prepared a speaking exam and administered it before the new regulations (Part I, questions 6 and 7). According to the data obtained, 30.8% of the participants (67 teachers) had prepared a speaking exam before, and 69.2% (150 teachers) had not prepared a speaking exam before (Figure 4.11.). 28.8% of the participants (62 teachers) had administered the speaking exam before, but 71.2% (155 teachers) had not administered the speaking exam before the new regulations (Figure 4.12.). In light of these findings, it can be seen that the majority of teachers did not prepare or administer the speaking exam before the new regulations; that is, they lack experience.

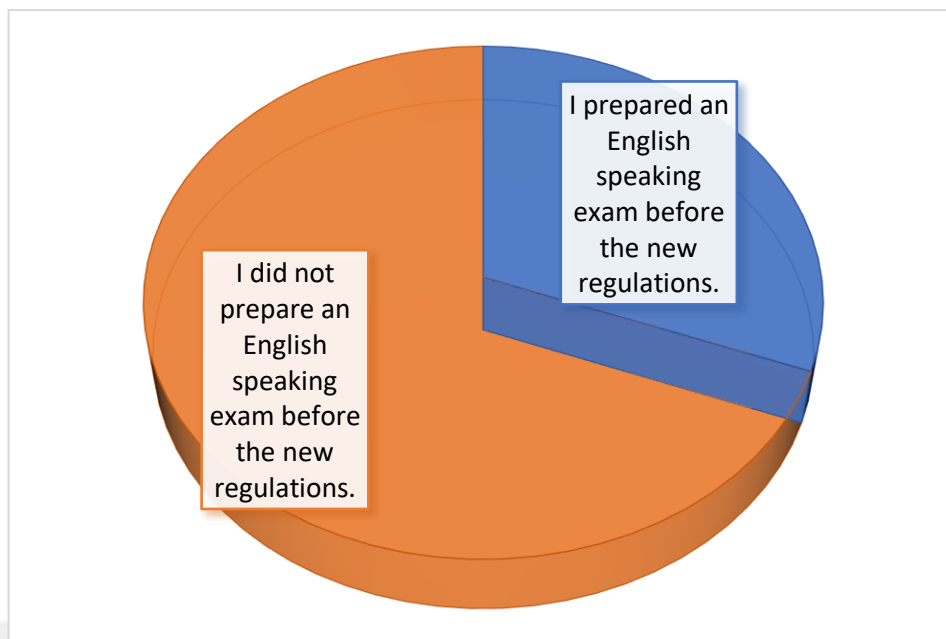


Figure 4.11. The Experience of Teachers Regarding Exam Preparation

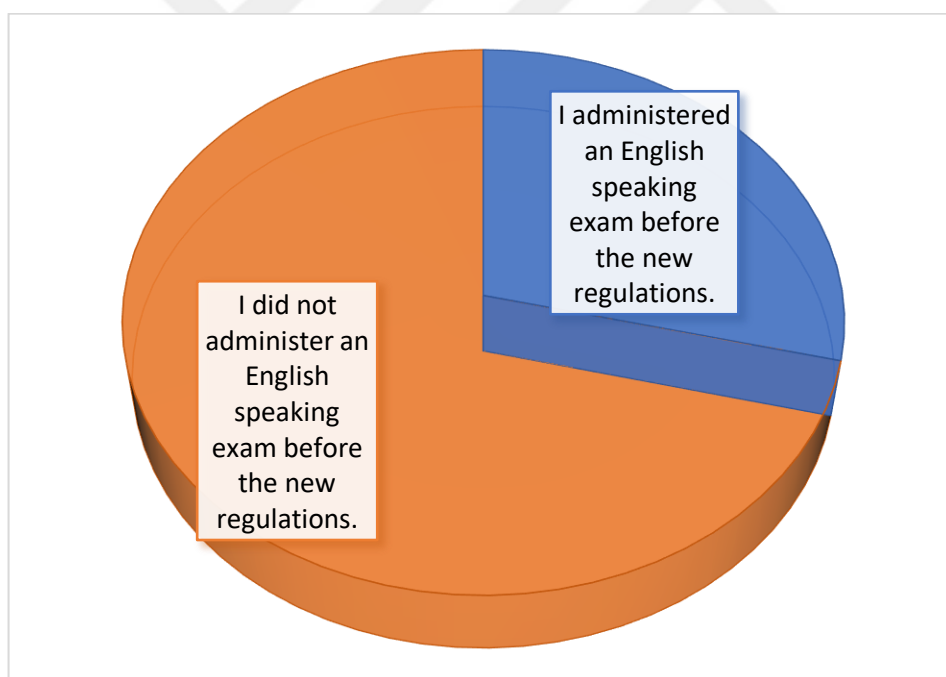


Figure 4.12. The Experience of Teachers Regarding Exam Administration

The participants were asked which techniques they used in their exams when assessing speaking skills in the 2023-2024 academic year (Part I, question 8). They could tick more than one option in the list given. This finding shows which techniques teachers mostly prefer in speaking exams. The data obtained shows that "question and

answer" is the most preferred technique with 94,5%, followed by "giving instructions/explanation/ description with 77%," "presentation with 66,8%," "talking about pictures with 63,6%," and "sentence completion with 55,3%". These are the five top techniques, respectively. The least preferred technique by teachers is the "group discussion with 21,7%". It is followed by "oral interview with 47,9%" and "problem solving with 48,4%". Below is a detailed list of how much teachers prefer each technique (Figure 4.13.).

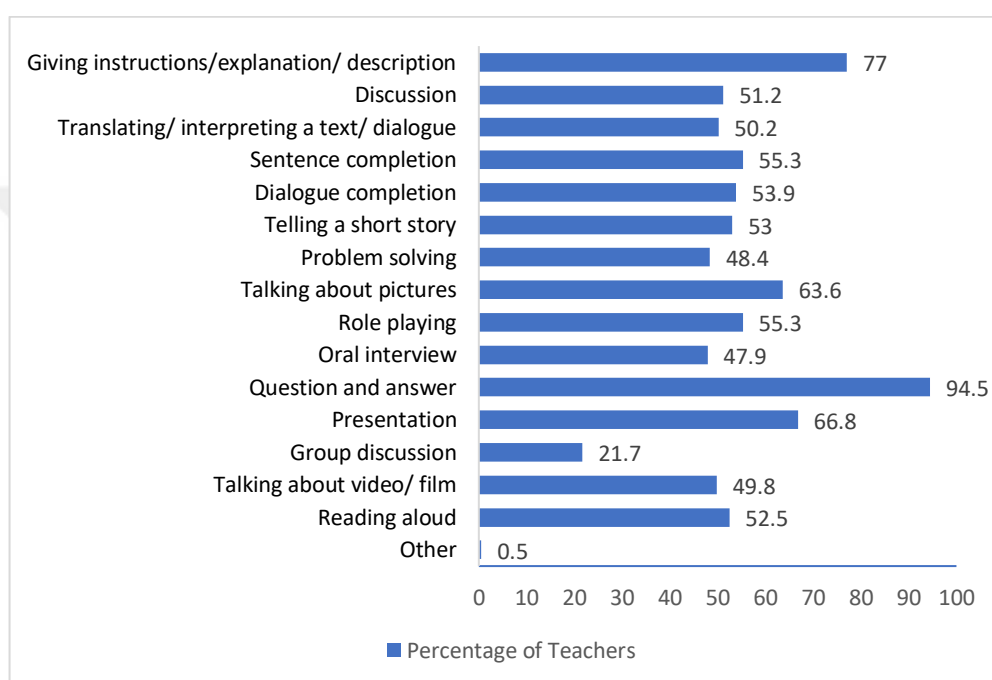


Figure 4.13. The Percentage of Teachers' Choice Regarding Speaking Exam Techniques

In order to show the teachers' practices in the field regarding the speaking exams they administered during the 2023-2024 academic year, teachers were asked whether they used rubrics when administering the speaking exams after the new regulations. According to the answers received, 88.9% of the teachers (193 teachers) used rubrics, while 11.1% (24 teachers) administered the exam without using rubrics (Figure 4.14.).

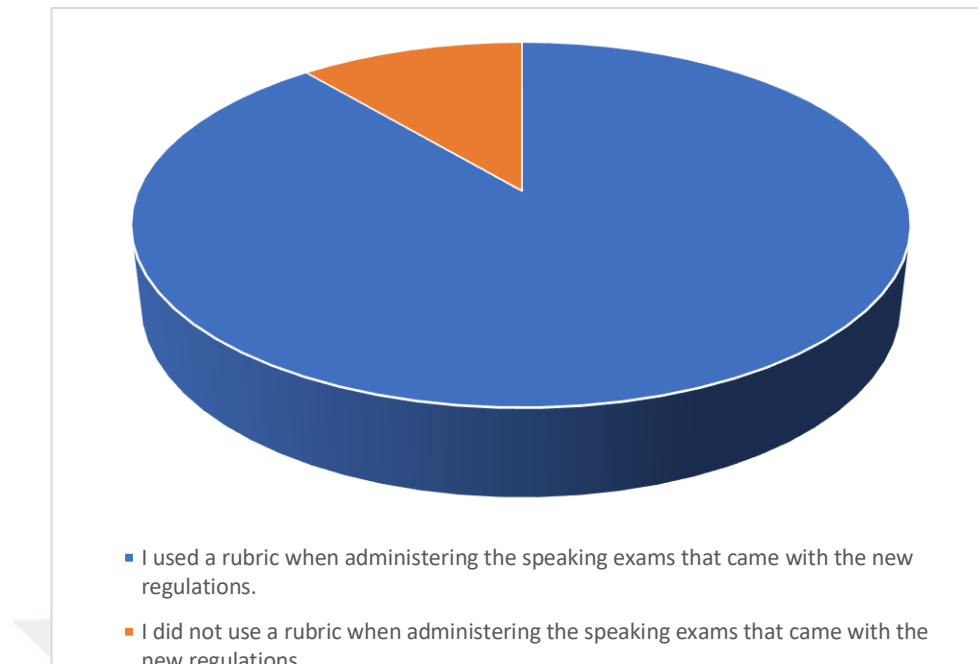


Figure 4.14. The Percentage of Teachers Using a Rubric for the Recent Speaking Exams

Finally, in this section, teachers were asked what criteria they considered when administering speaking exams in the 2023-2024 academic year. Teachers had the right to pick more than one option. When the data is examined, the criterion teachers consider most when applying the speaking exam is vocabulary, with a rate of 90.8%. Pronunciation and fluency were the following criteria, with a rate of 90.3%. Then, grammar, with 88.9%, and content, with 73.7%, are the most considered criteria, respectively. However, body language was the criterion least considered by teachers, with 56.2%, as shown in Figure 4.15.

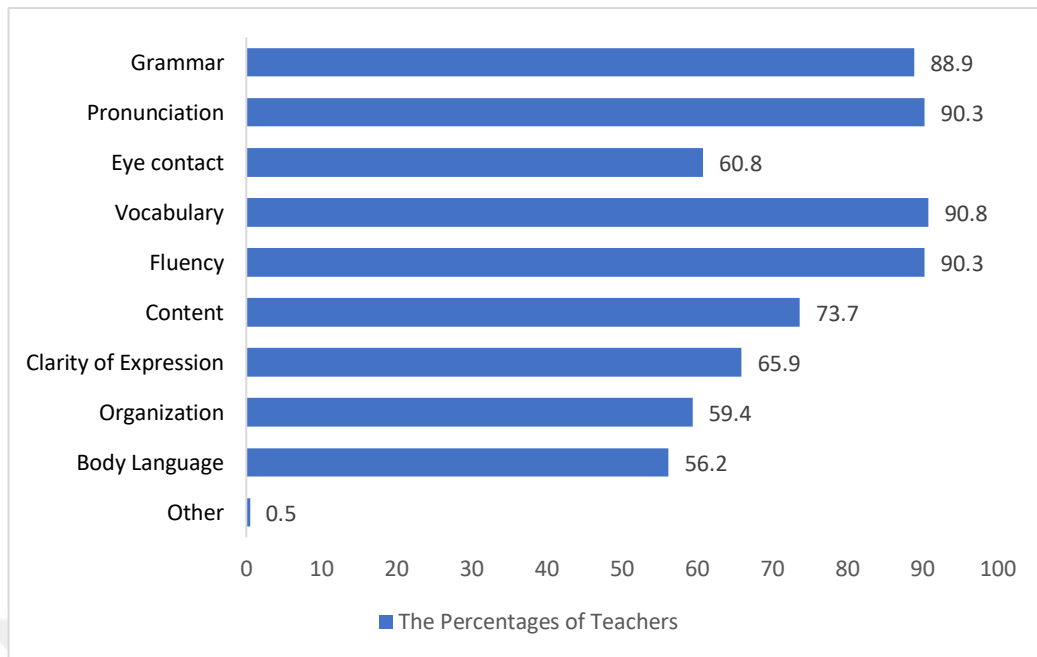


Figure 4.15. The Percentage of Teachers' Criteria for Speaking Exams

4.2.2. Part II

In this part of the analysis, each question will be examined according to the rate of answers given by teachers to these questions in the light of four research questions of this research. While collecting quantitative data, 34 questions were asked of teachers in the second part of the questionnaire. Thirty-three of them were Likert-type questions. Teachers are asked to choose the possibility corresponding to their opinion in the Likert-type questions, and the options are: I disagree, I am undecided, and I Agree. The 34th question was open-ended to gather more detailed data from the participants. There were four main titles in Part II:

- A.** Teachers' Thoughts and Feelings (11 questions)
- B.** Observations Regarding Students (8 questions)
- C.** Experience in Test Preparation, Administration, And Evaluation Processes (11 questions)
- D.** About Future Exams (4 questions)

The following part will examine the questionnaire data with the research questions they relate to.

Table 4.2. The Questionnaire and Research Questions, and the Research Themes They Relate

Research Questions	Themes	Related Questionnaire Questions
RQ1	Teachers' Overall Thoughts	Part II: A1, A4, A5, A8
RQ2	Teachers' Feelings	Part II: A2, A6, A7, A9, A10, A11, B1, B2, B4, B7, C11
RQ3	Experiences And Observations	Part I: 8, 9, 10 Part II: A3, A6, A9, A10, A11, B1, B2, B3, B4, B5, B6, B7, B8, C1, C2, C3, C4, C5, C6, C7, C8, C9, C10, C11
RQ4	Future Thoughts	Part II: D1, D2, D3, D4

Note. RQ stands for “research questions” of the study, Part I and Part II refer to the two main parts of the questionnaire, and expressions such as A1, B7, and D4 refer to the relevant sections and questions in Part II. For example, C11 means question 11 of section C in Part II.

4.2.2.1. RQ1: What are the overall thoughts of EFL teachers towards the recent English speaking exams after the current regulations in the secondary school examination system in Türkiye?

In order to examine teachers' overall thoughts on speaking exams, it is necessary to focus especially on the answers given to questions A1, A2, A5, and A8 in the questionnaire. The survey questions regarding this research question and the proportion of answers given by the participants are provided below:

A1: The new regulations regarding English exams in the 2023-2024 academic year were necessary in secondary schools.

Proportions of Answers: 33,1 % I disagree, 27,2% I am undecided, 39,6% Agree

A4: Increasing the English course passing grade to 50 points from 45 out of 100 was necessary.

Proportions of Answers: 32,7% I disagree, 27,2% I am undecided, 40,1% Agree

A5: I think that English speaking exams are beneficial for students.

Proportions of Answers: 30,4% I disagree, 16,6% I am undecided, 53% Agree

A8: English speaking exams should be mandatory.

Proportions of Answers: 32,7% I disagree, 27,6% I am undecided, 39,7% Agree

When the general opinions of teachers are examined, it is clear that 39.6% of teachers have the opinion that the new regulations regarding English exams in the 2023-2024 academic year are necessary. It should be noted that 33.1% do not think the regulations are necessary, and 27.2% are undecided. However, 53% think that English speaking exams are beneficial for students. According to 40.1% of participants, increasing the English course passing grade to 50 points from 45 out of 100 was necessary. The question of whether the English exam should be compulsory in secondary schools has a similar response distribution to its necessity. While 39.7% support this view, 32.7% disagree, and the rate of undecided voters is 27.6. From a broader view, it is seen that a slightly larger percentage of teachers think that "it should be necessary, useful, compulsory" about English speaking exams and the new regulations at secondary schools. However, the rate of undecided people and those not supporting this idea is still relatively high.

4.2.2.2. RQ2: What are the feelings of EFL teachers related to the recent English speaking exams?

In order to examine teachers' feelings about speaking exams, it is necessary to focus especially on the answers given to questions A2, A6, A7, A9, A10, A11, B1, B2, B4, B7, C11 in the questionnaire. The survey questions regarding this research question and the proportion of answers given by the participants are provided below:

A2: I have a positive feeling about the grade-wide English exams.

Proportions of Answers: 44,3% I disagree, 21,2% I am undecided, 34,6% Agree

A6: As an English teacher, I felt ready to start giving speaking exams at the beginning of this term.

Proportions of Answers: 40,6% I disagree, 20,3% I am undecided, 39,2% Agree

A7: I have positive feelings about the new regulations for speaking exams at secondary schools.

Proportions of Answers: 31,4% I disagree, 28,6% I am undecided, 40,1% Agree

A9: I felt nervous when I first heard that speaking exams would be compulsory in secondary school.

Proportions of Answers: 34,6% I disagree, 17,5% I am undecided, 47,9% Agree

A10: The speaking exams I prepared and administered satisfy me professionally as an EFL teacher.

Proportions of Answers: 33,2% I disagree, 27,2% I am undecided, 39,6% Agree

A11: Administering the speaking tests motivated me to speak more English in the lessons.

Proportions of Answers: 31,3% I disagree, 28,1% I am undecided, 40,6% Agree

B1: Based on my observations, students were ready to switch to this system.

Proportions of Answers: 53% I disagree, 18,4% I am undecided, 28,5% Agree

B2: Students felt worried about the speaking exams.

Proportions of Answers: 27,2% I disagree, 13,8% I am undecided, 59% Agree

B4: I received positive feedback from students regarding the speaking exam.

Proportions of Answers: 45,7% I disagree, 23% I am undecided, 31,3% Agree

B7: Taking the speaking test motivated the students to speak more English in the lessons.

Proportions of Answers: 42,9% I disagree, 23% I am undecided, 34,1% Agree

C11: (Only those with more than one teacher in the English department at school will answer this question) I was pleased to work in collaboration with English department teachers during the exam preparation and administration process.

Proportions of Answers: 30,6% I disagree, 18,4% I am undecided, 51% Agree

When teachers' feelings about the new regulations and recent speaking exams in secondary school are examined, a majority of 40.1% have positive feelings about the new regulations. When more specific questions are asked, the majority of the respondents, 44.3%, do not have positive feelings about the "grade-wide" English exams.

When the speaking exams were explicitly examined, 40.6% of the respondents did not feel ready to start giving speaking exams as an English teacher at the beginning of this term. 47.9% of the participants felt nervous when they first heard that speaking exams would be compulsory in secondary school. The professional dissatisfaction experienced by teachers is also seen in the survey results. It was revealed that a majority of 39.6% were not satisfied with the English speaking exams they prepared and administered. Regarding motivation, it is evident that administering the speaking tests motivated 40.6% of participant teachers to speak more English in the lessons.

When the data were examined in terms of observations about students' feelings, 53% of the participants declared that students were not ready to switch to this system, and 59% stated that students felt worried about the speaking exams. However, positive feelings were also noted. For example, 45.7% of the participating teachers received

positive feedback from students regarding the speaking exam, and 42.9% stated that the speaking test motivated the students to speak more English in the lessons.

Finally, the issue that teachers expressed the most positively about their feelings was in-school collaboration with other EFL teachers. A majority of 51% were pleased to collaborate with English department teachers during the exam preparation and administration process.

4.2.2.3. RQ3: What experiences were encountered by EFL teachers during the implementation of the recent English speaking exams?

In order to examine experiences encountered during the implementation of speaking exams, it is necessary to focus especially on the answers given to questions A3, A6, A9, A10, A11, B1, B2, B3, B4, B5, B6, B7, B8, C1, C2, C3, C4, C5, C6, C7, C8, C9, C10, C11 and also question 8, 9, 10 in Part I in the questionnaire. A6, A9, A10, A11, B1, B2, B4, B7, and C11 will not be repeated as they have been given and explained above. The survey questions regarding this research question and the proportion of answers given by the participants are provided below:

A3: I found the new grading method for English exams, based on percentile values, practical.

Proportions of Answers: 49,3% I disagree, 20,7% I am undecided, 30% Agree

B3: Students experienced adaptation problems to the new speaking exams.

Proportions of Answers: 28,1% I disagree, 14,3% I am undecided, 57,6% Agree

B5: Compulsory English course in secondary school is 3 hours in the 5th and 6th grades and 4 hours in the 7th and 8th grades. Compulsory English course hours in secondary schools were sufficient to prepare students for the speaking exam.

Proportions of Answers: 59% I disagree, 14,3% I am undecided, 26,7% Agree

B6: The grades after the exam reflected the students' in-class speaking performance.

Proportions of Answers: 38,3% I disagree, 27,2% I am undecided, 34,5% Agree

B8: Speaking exams have benefited my students

Proportions of Answers: 31,8% I disagree, 24,4% I am undecided, 43,8% Agree

C1: The courses I have taken during my undergraduate/master's/doctoral studies helped me to prepare for and apply for speaking exams.

Proportions of Answers: 32,2% I disagree, 20,7% I am undecided, 47% Agree

C2: I explained the rules of the speaking exam and how it would happen to the children before the exams.

Proportions of Answers: 26,3% I disagree, 13,8 % I am undecided, 59,9% Agree

C3: I faced main problems during speaking exams.

Proportions of Answers: 33,6% I disagree, 14,7% I am undecided, 51,6% Agree

C4: My classroom environment was optimum during the speaking exams.

Proportions of Answers: 43,3% I disagree, 20,7% I am undecided, 36% Agree

C5: Each student took the exam with fair timing.

Proportions of Answers: 16,6% I disagree, 14,7% I am undecided, 56,7% Agree

C6: One lesson (40 minutes) was enough for me to examine the whole class.

Proportions of Answers: 56,2% I disagree, 14,7% I am undecided, 29% Agree

C7: I gave feedback to the children about their speaking performance after the exams.

Proportions of Answers: 27,2% I disagree, 17,5% I am undecided, 55,3% Agree

C8: (Only those with more than one teacher in the English department at school will answer this question) The new exam system enabled me to co-operate with other English teachers in the school.

Proportions of Answers: 31,1 I disagree, 17,3% I am undecided, 51,6% Agree

C9: (Only those with more than one teacher in the English department at school will answer this question) I used a rubric prepared jointly by all English department teachers at my school to grade my students.

Proportions of Answers: 32,6% I disagree, 15,8% I am undecided, 51,5% Agree

C10: (Only those with more than one teacher in the English department at school will answer this question) The joint preparation and administration of the speaking test by the English department at the school caused problems.

Proportions of Answers: 49% I disagree, 16,8% I am undecided, 34,1% Agree

After the new regulations, when the teachers' experiences and observations about the whole process are considered, it is seen that 49.3% of the participants did not find the new grading method for English exams based on percentile values practical, 59% think that compulsory English course hours in secondary schools were not sufficient to prepare students for the speaking exam. 57.6% observed that students experienced adaptation problems to the new speaking exams. However, the response rates obtained in terms of whether the grades given after the speaking exam reflect the students' actual grades or not are close. 38.3% think it does not reflect students' real performances in class, 34.5% think it does, and 27.2% are undecided. Nevertheless, the majority (43.8%) claim that speaking exams have benefited their students.

47% of the participants think that the courses they have taken during their undergraduate/master's/doctoral studies helped them to prepare for and apply for speaking exams. 47% of the participants explained the rules of the speaking exam and how it would happen to the children before the exams. However, 51.6% faced main problems during speaking exams, and 43.3% stated that their classroom environment was not optimum during the speaking exams. 56.7 of the participants declared that each student took the exam with fair timing, yet 56.2% of the participants stated that one lesson (40 minutes) was not enough for them to examine the whole class. Thus, it is clear that there are exams that last more than one lesson. 55.3% gave feedback to the children about their speaking performance after the exams.

When the cooperation part of the new system, which enables teachers working in the same school to work together more, is examined, it is seen that 51.6% support that the new exam system enabled them to cooperate with other English teachers in the school while 51.5% used a rubric prepared jointly by all English department teachers at their school to grade students. Lastly, 49% state that the English department's joint preparation and administration of the speaking test at the school did not cause problems.

4.2.2.4. RQ4: What are the future thoughts of EFL teachers related to the recent English speaking exams at secondary schools?

In order to examine the thoughts of EFL teachers related to the future of speaking exams, it is necessary to focus especially on the answers given to questions D1, D2, D3, and D4 in the questionnaire. The survey questions regarding this research question and the proportion of answers given by the participants are provided below:

D1: Speaking exams can be improved over time.

Proportions of Answers: 28,1% I disagree, 16,6% I am undecided, 55,3% Agree

D2: I believe that this examination system and regulations will continue in the future.

Proportions of Answers: 32,3% I disagree, 17,5% I am undecided, 50,3% Agree

D3: I would like the new regulations to continue in their current form in the future.

Proportions of Answers: 43,3% I disagree, 22,6% I am undecided, 34,1% Agree

D4: Please write if you have any suggestions for improving the speaking exams.

When opinions about the future of speaking exams are examined, 55.3%, more than half of the participants, think that speaking exams will be improved over time. 50.3% of teachers think this examination system and new regulations will continue as

a prediction. However, looking at the current situation, 43.3% of the respondents do not want the new regulations to continue in their current form. In addition, the teachers we asked for advice on improving the exams by asking open-ended questions stated that the units should be reduced, sample exams should be shared, lesson hours should be increased, speaking activities in the curriculum should be increased, the e-school system should be reorganized, and teachers should be included in the process at every step.



5. DISCUSSION

This chapter will provide a discussion related to the four themes of the study about the recent speaking exams and the new regulations regarding the examination systems in secondary schools.

5.1. Teachers' Overall Thoughts

When the qualitative and quantitative data obtained are brought together and compared, some ideas that the teachers agree on are seen in both parts of the study. For instance, most participating teachers argued that the speaking exam in secondary schools was necessary in both parts of the research. They have the view that all four language skills should be measured and that speaking exams are useful. Dağtan (2020), Duran (2011), and Özdemir (2018) showed that teachers and students had a positive attitude toward speaking exams, as well. Gopal and Embi's (2014) results also showed that language teachers hold positive beliefs and practices of communicative language testing and are aware of its principles. This study found that the teachers' general opinions were positive and that they found the speaking exam important, necessary, and useful, supporting the previously mentioned studies.

When the details of these positive opinions about the recent speaking exams are examined in the current study through interviews, it turns out that teachers can give many reasons for these thoughts. The arguments that support the teachers' views are that the curriculum, textbooks, and exams are more compatible with these exams after the new regulations, and it is important and logical for education planning. Teachers state that teachers and students develop motivation to speak more English. The increase in students' self-confidence and self-awareness in English while giving more importance to English are other supporting thoughts. Finally, the cooperative and collaborative work of English departments after the new regulations is seen as positive by EFL teachers. It is understood from this that the general thoughts of the teachers are consistent in the surveys and interviews.

Although they constitute a smaller part of the survey group, there are also teachers who do not agree with these positive views or state that they are undecided about the importance, necessity, and usefulness of speaking exams. When the negative opinions of these teachers were examined through interviews, several different reasons emerged. The first of these is related to sudden changes. One of the most negative

situations that almost all teachers stated in the interviews was sudden changes in MoNE. Teachers clearly stated that they did not receive enough guidance at the beginning of the term, there were disconnections in the process, and that teachers and students were not ready for this sudden transition. In the following process, they revealed that the exams were far from fulfilling their real functions and the negative situations regarding the functioning of the exams. In this sense, the negative thoughts in the survey data were detailed and supported by the data obtained from the interviews. As Özdemir (2018) stated in his study, inflated grades to encourage students to speak are found to be a problem in the current research. The findings of this study support Özdemir (2018) at this point.

Similarly, Gopal and Embi (2014) indicated that teachers encounter main constraints during the speaking assessment process. In this study, although the general thoughts of the teachers are positive, problems in practice come to the fore, such as badly implemented assessments, as indicated by Atjonen (2014). The current study revealed problems similar to those of Nurhalimah et al. (2020). They underlined teachers' weaknesses regarding speaking assessment and the need for guidance regarding rubric development and implementation of speaking assessment. The same needs were also identified as a result of this research.

5.2. Teachers' Feelings

When teachers' feelings about speaking exams are examined, parallels can be seen between qualitative and quantitative findings. One of these is that most teachers do not feel ready to start speaking exams in 2023-2024. It seems that the majority of both interview and survey participants have difficulty keeping up with the unexpected change. Interview and survey findings support each other in teachers' positive feelings towards speaking exams. According to the data obtained from both channels, most teachers have positive feelings. Nevertheless, one point that both data draw attention to is the extent of stress experienced by teachers. The majority of participating teachers stated that they became stressed when they heard that speaking exams would be mandatory in secondary schools. Thanks to the interviews, these feelings were further detailed, and feelings such as anger, hopelessness, burnout, boredom, and fatigue were also mentioned. Most teachers stated that speaking exams motivated them to speak more English in classes in both the interview and survey. Additionally, teachers clearly showed satisfaction in collaborating with their

colleagues in interviews and surveys. The majority were satisfied with this collaboration, and the qualitative and quantitative data showed parallelism.

Since this situation also affects teachers' emotions, students' feelings were also asked in this study. According to the survey and interviews, it was evident from teachers' responses about students' feelings that students did not feel ready to switch to the new system and felt anxious. Moreover, teachers received negative feedback from their students about these exams and new regulations. Considering these, the data obtained from the two sources support each other.

Lastly, there is also a situation where survey data and survey data do not support each other. Most teachers who participated in the interview stated that the speaking test motivated the students to speak more English in the lessons. In contrast, the opposite was seen in the survey findings. Most survey participants disagree with this idea. Austheim (2023) reported that students' beliefs about their competencies affect their willingness to speak. The current study partially supports this because, according to interview data, students' self-awareness and self-confidence in speaking have increased. In addition, students were more motivated to speak.

Teachers' professional satisfaction is another issue that differs according to the interview and survey findings. While most interview participants stated that they were not professionally satisfied with the speaking exams they prepared and applied, most survey participants stated that they were satisfied. In this case, the findings do not support each other qualitatively and quantitatively.

5.3. Experiences and Observations

Teachers' observations and experiences about speaking exam procedures were also investigated using two different tools. When the interview and survey findings are evaluated separately, some findings support each other. For instance, most teachers did not find the exam grading method introduced with the new system practical. Teachers expressed their complaints and suggestions on this issue in interviews. According to the observation of the majority of teachers, students experienced adaptation problems to speaking exams. The English lesson hours, which are three hours in the 5th and 6th grades and four hours in the 7th and 8th grades, are stated as insufficient by the majority of teachers in both data. According to most teachers, the grades students receive from speaking exams do not reflect their speaking

performance. However, teachers argue that exams benefit students according to both data.

According to the information obtained in both qualitative and quantitative data, it was determined that almost all of the teachers explained to the students how the exam would be before the exam and gave them a fair amount of time. However, according to both data, most teachers agree that one lesson hour is insufficient to complete the exam. Information was obtained from the teachers who participated in the interview: it took two or three lessons. It was learned that more than half of the teachers gave feedback after the exam. Teachers who did not give feedback had explanations such as lack of time in the interviews.

More than half of the teachers stated that the classroom environment was not ideal during the speaking exam. While qualitative and quantitative data support each other, according to interview participants, discipline problems, noise, disorder in crowded classrooms, and extended exam hours can be listed as examples.

When teachers working as single teachers in their schools were excluded, teachers in schools with more than one English teacher worked together and cooperated in the new exam process. Although the teachers stated in the interview and survey that they were happy to work together, the survey findings show that joint preparation and administration of the speaking test by the English department at the school caused problems. The details indicate that most teachers encounter some fundamental problems during the speaking exam, as Gopal and Embi (2014) found, both qualitatively and quantitatively support each other in the current study. Teachers who participated in the interview study detailed these problems and provided examples. When teachers' discourses are examined, these are lack of teacher and student readiness, lack of teacher preparation and guidance (Brumen & Cagran, 2011), not giving enough space to speaking skills in the curriculum and textbooks, inefficiency of exams (Atjonen, 2014; Özdemir, 2018), crowded classroom problems, impracticality of grading and e-school system, increased workload for teachers and lastly teachers' and students' feelings of being too busy under time pressure.

Finally, qualitative and quantitative findings show that most teachers used a rubric jointly prepared by the English department. In the survey, vocabulary, pronunciation, fluency, grammar, and content were found to be the most prominent rubric criteria. Gandeel (2016) showcased that teachers' beliefs and practices lacked a theoretical basis, thus not aligning with the contemporary views on speaking

skills. She underlined that teachers had little or no focus on fluency. Grada (2014) also supported her findings. According to Grada (2014), teachers base their assessment practices considering spoken language mainly on linguistic content. On the contrary, the current study's findings from interviews and surveys show that EFL teachers prioritize fluency as a criterion in rubrics.

In exam administrations, question and answer, giving instructions/explanations/description, presentation, and talking about the pictures are the speaking exam techniques most preferred by teachers. In parallel, teachers using different exam techniques and rubrics were also identified in the interview. Nurhalimah et al. (2020) indicated that teachers' assessment of speaking skills could differ for multiple reasons, such as teachers' educational background, teaching experience, and the school context. Çalışkan (2013) showed that educational background and experience could be factors in choosing assessment tools and techniques for speaking exams. Furthermore, Ahmad (2021) discovered that numerous contextual factors like classroom contexts, institutional requirements, and curriculum expectations influenced teachers' assessment practices. In the current study, there are teachers with different experiences or educational backgrounds who work in villages, towns, or cities, and it is obvious that the teachers who participated in the current study also experienced contextual differences.

5.4. Future Thoughts

When teachers' ideas, expectations, suggestions, and wishes regarding the future of speaking exams, which have started to be implemented in secondary schools, are considered through data obtained through two separate channels, a remarkable similarity can be seen. In the qualitative and quantitative parts of the research, most teachers had similar thoughts about the future of speaking exams. Both findings indicate that the majority of teachers think that speaking tests can be improved over time. Likewise, more than half of the teachers believe that speaking exams and new practices will continue in the future. However, according to both data, most teachers do not want speaking exams and new regulations to continue in their current form.

Recommendations were received from teachers through interviews and surveys on how speaking exams could be improved in the future. When the findings are examined, based on the data collected from two separate channels, the steps most

recommended by teachers are: "more hours of English course, more curriculum planning with more speaking activities, regulation in the e-school system and exam scoring system, in-service training and more information." It appears to be although there are various recommendations in the interviews, four codes, in particular, came to the fore as a result of the surveys conducted via Google Forms. In this sense, data collected in two different ways support each other. Lastly, this study showed teachers' additional training needs to assess speaking, similar to the studies of Brumen and Cagran (2011).



6. CONCLUSION

This chapter includes implications, suggestions, and the conclusion of the study.

6.1. Implications

When the recent speaking exams are examined from teachers' perspectives, it is seen that most teachers are willing, cooperative, and hopeful for the future to perform them correctly, but teachers have various problems currently. When planning such significant changes, it may be good to motivate teachers by including them at every stage of policymakers' decision-making and planning process. Teachers' feeling of being professionally and psychologically prepared can prevent many problems in education before they arise. Additionally, being individuals who are influential in the decision-making mechanism, rather than being informed at the last minute, can affect them positively. It is natural for teachers and students to experience problems in their adaptation to new systems. However, as most teachers state, gradual transitions instead of sudden and sharp transitions can alleviate the challenging effects of these changes and ensure better management of the adaptation process. Determining the educational calendar, decisions, and regulations before the academic year begins may be more welcome than changes made in the middle of the year. Teachers can prepare their materials in this direction on time; students can act with awareness of this situation; thus, education life will run more professionally.

It is vital that teachers are informed in as much detail as possible when changes are made to educational planning. Resources that teachers who need information should be shared on time and be accessible easily. Guidance should be planned, and teachers should benefit from online or on-site in-service training. In this way, teachers may feel more comfortable in completing their shortcomings and can achieve greater professional satisfaction with the exams they prepare and administer. Speaking exams can also function better this way.

While designing materials, speaking activities should be diversified into more meaningful, authentic, and communicative activities in which students will participate more. The rate of speaking activities in textbooks can be increased, and students can be enabled to speak more. Regulations regarding class hours and exam schedules can regulate teachers' concerns about completing subjects and exams on time. Teachers

can work more efficiently when they are not under intense stress. Relieving teachers' workload and offering alternatives to teachers working in different schools under different conditions can relieve teachers because teachers working in crowded classes or schools with heavy course loads, such as Imam Hatip Secondary Schools, work beyond their capacity under time pressure. This arrangement can eliminate problems and increase the quality of the exams.

This research determined that teachers used different exam practices and rubrics that highlighted different criteria. Having more defined frameworks for standardizing speaking exams in secondary schools may yield better results regarding validity and reliability. Additionally, from a more technical point of view, if the factors that slow down the functioning of the e-school system and make teachers' work difficult are regulated, the workflow can become more practical. The exam scoring system may be kept more transparent so that students will not have difficulty understanding it.

Considering all these mentioned above, the findings of this study may be indicative for policymakers to make speaking exams more reliable, valid, standardized, practical, beneficial, and positive from teachers' and students' perspectives. In addition, teachers and other researchers can evaluate their experiences in this field in the light of this study and take initiatives to improve and develop speaking exams.

6.2. Suggestions

This study reflects that many steps can be taken to improve speaking tests from teachers' perspectives. However, the current study is limited to 217 survey and 11 interview participants. All interview participants work in public schools. Expanding the study to different bases may yield better results. Also, this research discussed the speaking exam only from teachers' perspectives. This study may lead to further studies in the coming years. For instance, future research may examine secondary school speaking exams from the students' perspective. This way, the learner dimension can be included more in the subject.

6.3. Conclusion

The current study aims to investigate secondary school EFL teachers' perspectives toward the recent speaking exams introduced with the new regulations in the 2023-2024 academic year by Turkish MoNE while giving implications regarding

how speaking exams could be improvised for the benefit of teachers and students in the following years. Various types of thoughts, experiences, observations, suggestions, and expectations were found through teacher interviews and questionnaire data analysis. The participants' perspectives were grouped into four themes: teachers' overall thoughts, teachers' feelings, experiences and observations encountered, and future thoughts. Mixed research was chosen as the research design, and qualitative and quantitative data were collected. For this, open-ended semi-structured interviews were conducted with eleven secondary school EFL teachers, and a survey was conducted with 217 secondary school EFL teachers. Thematic analysis and statistics were used for data analysis. The analysis has revealed that the teachers' general opinions were positive. The teachers find the speaking exams important, necessary, and useful. According to them, four language skills should be assessed. However, it was also found that although teachers had generally positive feelings about the introduction of speaking exams in secondary schools, they also felt negative feelings such as unpreparedness, stress, burnout, fatigue, and hopelessness due to the unexpected change.

Nevertheless, it was observed that teachers enjoyed and were satisfied with continuing this process in collaboration with their colleagues. With this decision, teachers became more motivated about speaking activities and using English more in the lessons. This study also identified problems such as adaptation problems, ineffective exams, increased workload of teachers, impractical grading and e-school system, and teachers' busy schedules. The part of the study on future speaking exams revealed teachers' expectations, predictions, and aspirations for the future of speaking exams. Accordingly, teachers think that these exams will continue but that they should be developed and improved. In general, teachers believe that the problems with these exams can be solved over time.

All in all, the results show that the EFL teachers favored the speaking exams in secondary schools. In addition, many problems that are expected to be improved over time have been revealed from teachers' perspectives as implications.

REFERENCES

- Adem, H., & Berkessa, M. (2022). EFL teachers' cognitions about teaching speaking skills. *GIST – Education and Learning Research Journal*, 24, 65–94. <https://doi.org/10.26817/16925777.1308>
- Ahmad, S. (2021). *Teachers' Beliefs about Assessment: A Multiple Narrative Case Study of Internationally Educated Teachers' Assessment of Speaking Skills in Canadian ESL/EAP Classrooms* (dissertation).
- Alderson, J. C., Brunfaut, T., & Harding, L. (2017). Bridging Assessment and Learning: A View from Second and Foreign Language Assessment. *Assessment in Education: Principles, Policy & Practice*, 24(3), 379–387. <https://doi.org/10.1080/0969594x.2017.1331201>
- Aptitude Test*. Language Learning and Testing Foundation | Language Learning and Testing Foundation. (2020, April 16). <https://lltf.net/aptitude-tests/>
- Atay, D., & Mede, E. (2017). English language teachers' assessment literacy: The Turkish context. *Dil Dergisi*, 0(168), 43–60. https://doi.org/10.1501/dilder_00000000237
- Atjonen, P. (2014). Teachers' views of their assessment practice. *The Curriculum Journal*, 25(2), 238–259. <https://doi.org/10.1080/09585176.2013.874952>
- Austheim, T.M. (2023). *Norwegian EFL teachers' perspectives on the relationship between formative assessment and willingness to speak* (dissertation).
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.
- Bachman, L. F. (1995). Chapter 7 Validation. In *Fundamental considerations in language testing* (pp. 236–251). Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (2015). *Language testing in practice: Designing and developing useful language tests*. Oxford University Press.
- Berry, V., Sheehan, S., & Munro, S. (2019). What does language assessment literacy mean to teachers? *ELT Journal*, 73(2), 113–123. <https://doi.org/10.1093/elt/ccy055>
- Braun, V., & Clarke, V. (2006). *Using thematic analysis in psychology*. *Qualitative Research in Psychology*, 3(2), 77–101
- Brown, G., & Yule, G. (1983). *Teaching the spoken language*. Cambridge University Press.
- Brown, H. D. (1993). Chapter 10 Language Testing. In *Principles of language learning and teaching* (third edition, pp. 251–274). Prentice Hall Regents.
- Brown, H. D. (2004). *Language assessment principles and classroom practices*. Longman.
- Brown, H. D., & Abeywickrama, P. (2021). *Language assessment: Principles and classroom practices*. Pearson.
- Brown, H. D., & Lee, H. (2015). *Teaching by principles: An interactive approach to language pedagogy*. Pearson Education.
- Brumen, M., & Cagran, B. (2011). Teachers' perspectives and practices in assessing young foreign language learners in three Eastern European countries. *Education 3-13*, 39(5), 541–559. <https://doi.org/10.1080/03004279.2010.488243>
- Bygate, M. (1987). *Speaking*. Oxford University Press.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to Second language teaching and testing. *Applied Linguistics*, 1(1), 1–47. <https://doi.org/10.1093/applin/i.1.1>

- Carroll, J. B. (1981). Twenty-five years of research on FL aptitude. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 83–118). Rowley, MA: Newbury House.
- Carroll, J. B., & Sapon, S. M. (1958). *Modern Language Aptitude Test (MLAT)*. New York: Psychology Corporation.
- Carter, R., & McCarthy, M. (1995). Grammar and spoken language . *Applied Linguistic*, 16, 141–157.
- Chapelle, C., Chapelle, C., & Plakans, L. (2013). In *The Encyclopedia of Applied Linguistics* (pp. 240–244). essay, Wiley-Blackwell.
- Chappuis, S., & Stiggins, R. (2002). Claaroom Assessement for Learning. *Educational Leadership*, 60(1), 40–43.
- Chopin, K. (2015). Chapter 11. Reconceptualizing norms for Language testing: Assessing English language proficiency from within an ELF framework. *Current Perspectives on Pedagogy for English as a Lingua Franca*, pp. 193–204. <https://doi.org/10.1515/9783110335965.193>
- Council of Europe (2020), *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*, Council of Europe Publishing, available at www.coe.int/lang-cefr.
- Creswell, J.W. & Clark, V.L.P. (2011). *Designing and conducting mixed methods research* (second edition). Sage Publications.
- Davies, A. (2006). *Dictionary of language testing*. Cambridge University Press.
- Dağtan, E. (2020). *An investigation into the perceptions of Turkish ELT pre-service teachers and lecturers on the learning and teaching of English speaking skills* (dissertation).
- Douglas, D. (2014). *Understanding language testing* (Second Edition). Routledge.
- Duran, Ö. (2011). *Teachers' and students' perceptions about classroom-based speaking tests and their washback* (dissertation).
- Ekmekci, E. (2016). Comparison of native and non-native English language teachers' evaluation of EFL learners' speaking skills: Conflicting or identical rating behaviour? *English Language Teaching*, 9(5), 98. <https://doi.org/10.5539/elt.v9n5p98>
- Florez , M. C. (1999). *Improving Adult English Language Learners' Speaking Skills*. *ERIC Digest*. <http://www.eric.ed.gov/>. <https://files.eric.ed.gov/fulltext/ED435204.pdf>
- Fulcher, G. (2000). The 'communicative' legacy in language testing. *System*, 28(4), 483–497. [https://doi.org/10.1016/s0346-251x\(00\)00033-6](https://doi.org/10.1016/s0346-251x(00)00033-6)
- Fulcher, G. (2014). *Testing second language speaking* (Second Edition). Routledge.
- Gandeel, A.M. (2016). *English Language Teachers' Beliefs and Practices regarding the Teaching of Speaking* (dissertation).
- Gopal, P., & Embi, M. A. B. (2014). Teachers' Belief and Practices regarding Communicative Language Testing In Malaysian Secondary School. *International Conference on Postgraduate Research 2014*, 1(2), 357–367.
- Grada, T.K. (2014). *An Investigation into Libyan EFL Novice Secondary School Teachers' Current Knowledge and Practice of Speaking Assessment: A Socio-cultural Perspective* (dissertation).
- Güral Akgül, S. (2022). *The impacts of formative assessment implementation on English preparatory class students' speaking skills in EFL context in Turkey: a case study* (dissertation).

- Harding, L. (2014). Communicative language testing: Current issues and future research. *Language Assessment Quarterly*, 11(2), 186–197. <https://doi.org/10.1080/15434303.2014.895829>
- Harlen, W. (2007a). Criteria for evaluating systems for student assessment. *Studies in Educational Evaluation*, 33(1), 15–28. <https://doi.org/10.1016/j.stueduc.2007.01.003>
- Harlen, W. (2007b). *The Quality of Learning: assessment alternatives for primary education* (Ser. Primary Review Research Survey 3/4). University of Cambridge Faculty of Education.
- Harmer, J. (2022). *The practice of English language teaching*. Pearson Education Limited.
- Hendrickson, J. (1980). Error correction in foreign language teaching: Recent theory, research, and practice. In K. Croft (Ed.), *Readings on English as a second language* (2nd ed., pp. 153–173). Cambridge Press.
- Hinkel, E. (2006). Current perspectives on teaching the four skills. *TESOL Quarterly*, 40(1), 109. <https://doi.org/10.2307/40264513>
- Holec, H. (1981). *Autonomy and Foreign Language Learning*. Pergamon Press.
- Hymes, D. (1972). *On Communicative Competence*. In J. Pride, & J. Holmes (Eds.), *Sociolinguistics* (pp. 269–285). Harmondsworth: Penguin Books.
- Jenkins, J. (2020). Where are we with Elf and language testing? An opinion piece. *ELT Journal*, 74(4), 473–479. <https://doi.org/10.1093/elt/ccaa045>
- Kachru, B. (1985). ‘Standards, codification, and sociolinguistic realism: the English language in the outer circle,’ in Quirk, R. and Widdowson, H. G. (eds) *English in the World: Teaching and Learning the Language and Literatures*. Cambridge: Cambridge University Press: pp. 11–30.
- Kadir, J. S., Zaim, M., & Refnaldi, R. (2019). Developing instruments for evaluating validity, practicality, and effectiveness of the authentic assessment for speaking skill at junior high school. *Proceedings of the Sixth of International Conference on English Language and Teaching (ICOELT 2018)*. <https://doi.org/10.2991/icoelt-18.2019.14>
- Kirkgoz, Y. (2007). English language teaching in Turkey: Policy changes and implementations. *RELC Journal*, 38(2), 216–228. <https://doi.org/10.1177/0033688207079696>
- Köksal, D. (2004). Assessing teachers’ testing skills in ELT and enhancing their professional development in testing and assessment through distance learning on the NET. *Turkish Online Journal of Distance Education-TOJDE*, p. 4.
- Lambert, D., & Lines, D. (2013). *Understanding assessment: Purposes, perceptions, practice*. Routledge/Falmer.
- Lado, R. (1964). *Language testing: The construction and use of foreign language tests*. MacGraw-Hill.
- Luoma, S. (2004). *Assessing speaking*. Cambridge University Press.
- Marrow, K. (1979). Communicative Language Testing: Evaluation or Revolution? In C. J. Brumfit (Ed.), *The Communicative Approach to Language Teaching*. essay, Oxford University Press.
- Marrow, K., Marrow, K., & Johnson, K. (1981). Principles of communicative methodology. In *Communication in the Classroom*. essay, Longman.
- McEwen, N. (1995). Educational Accountability in Alberta. *Canadian Journal of Education / Revue Canadienne de l’éducation*, 20(1), 27–44. <https://doi.org/10.2307/1495050>
- McNamara, T. (2017). *Language testing*. Oxford University Press.

- Mirici, İ. H., & Şengül, F. (2020). Assessment in EFL classes based on the CEFR principles. *Bartın University Journal of Faculty of Education*, 9(2), 252–263. <https://doi.org/10.14686/buefad.655985>
- MoNE, Ministry of National Education (2022, September 9). 2021-2022 ÖRGÜN EĞİTİM İSTATİSTİKLERİ AÇIKLANDI. <https://www.meb.gov.tr/2021-2022-orgun-egitim-istatistikleri-aciklandi/haber/27552/tr#:~:text=Bu%20okullar%C4%B1n%2014%20bin%20124,749%20bin%20454%20derslik%20kullan%C4%B1ld%C4%B1>
- Mousavi, S. A. (2009). In *An encyclopedic dictionary of language testing* (4th ed., p. 247). Rahnama Press.
- Nakatsuhara, F., Khabbazzashi, N., & Inoue, C. (2022). *Assessing Speaking*. In *The Routledge Handbook of Language Testing* (Second, pp. 209–223). essay.
- Nurhalimah, N., Husna HR, N., & Hidayat, D. N. (2020). An exploration of English teachers' beliefs and practices of speaking assessment: Bring English teachers into focus. *International Journal of Linguistics, Literature and Culture*, 6(5), 41–48. <https://doi.org/10.21744/ijllc.v6n5.990>
- Oller, J. W. (1980). Language Testing Research (1979–1980). *Annual Review of Applied Linguistics*, 1, 124–150. <https://doi.org/10.1017/s0267190500000544>
- Phan, S. (2008). Communicative Language Testing. *Hawaii Pacific University TESOL Working Paper Series*, 6(1).
- Pimsleur, P. (1966). Aptitude testing. *Language Learning*, 18, 73–77. <https://doi.org/10.1111/j.1467-1770.1968.tb00223.x>
- Plo, R., Hernero, A., & Mur-Duenas, P. (2013). Implementing the teaching/learning of oral skills in secondary education in Aragón: Gauging teachers' attitudes, beliefs and expectations. *International Journal of English Studies*, 14(1), 55–77.
- Rao, P. S. (2019). The importance of speaking skills in English classrooms. *Alford Council of International English & Literature Journal*, 2(2).
- Richards, J. C., & Schmidt, R. W. (2002). In *Longman Dictionary of Language Teaching & Applied Linguistics* (p. 546). Pearson Education Limited.
- Sarıçoban, A. (2011). A study on the English language teachers' preparation of tests. *Hacettepe University Journal of Education*.
- Sarıçoban, G., & Sarıçoban, A. (2012). Atatürk and the history of foreign language education. *The Journal of Language and Linguistic Studies*, 8 (1), 24-49.
- Skehan, P. (2002). Theorizing and updating aptitude. In P. Robinson (ed.), 69–95.
- Stansfield, C. W. (2008). Lecture. *Language Testing*, 25(3), 311–326. <https://doi.org/10.1177/0265532208090155>
- Sudarmo, S. (2021). The importance of speaking in English as a foreign language between skillful and thoughtful competencies: studying sociolinguistics perspectives. *Linguistics and Culture Review*, 5(S1), 113-124. <https://doi.org/10.37028/lingcure.v5nS1.1321>
- Taras, M. (2005). Assessment – summative and formative – some theoretical reflections. *British Journal of Educational Studies*, 53(4), 466–478. <https://doi.org/10.1111/j.1467-8527.2005.00307.x>
- Tomlinson, B. (2005). Testing to learn: A personal view of language testing. *ELT Journal*, 59(1), 39–46. <https://doi.org/10.1093/elt/cci005>
- Tosuncuoglu, I. (2018). Importance of assessment in ELT. *Journal of Education and Training Studies*, 6(9), 163. <https://doi.org/10.11114/jets.v6i9.3443>

- Yıldırım, A., & Şimşek, H. (2006). *Qualitative research methods in the social sciences*. Seçkin Publishing House.
- Zulfugarova, S. (2018). The importance of teaching listening and speaking skills. *World Science*, 7. https://doi.org/10.31435/rsglobal_ws/12062018/5881
- Çalışkan, E. (2013). *Instructors' Attitudes and Practices on Assessing English Speaking Skill* (dissertation).
- Özdemir, C. (2018). *Assessing the speaking skill: an investigation into achievement tests of the 9th grade students in anatolian high schools* (dissertation).



APPENDICES

Appendix A

Social And Human Sciences Research Ethics Committee Decision



ONDOKUZ MAYIS ÜNİVERSİTESİ SOSYAL VE BEŞERİ BİLİMLER ARAŞTIRMALARI ETİK KURUL KARARLARI

KARAR TARİHİ	TOPLANTI SAYISI	KARAR SAYISI
29.12.2023	12	2023-1126

KARAR NO: Üniversitemiz Lisansüstü Eğitim Enstitüsü öğrencisi Merve Nur DİNÇER'in Doç. Dr. Betül BAL GEZEGİN danışmanlığında "Türk Ortaokullarında Konuşma Becerilerinin Değerlendirilmesine İlişkin İngilizce Öğretmenlerinin Bakış Açıları: Yeni Sınav Modeli Sonrası Araştırması (EFL Teachers' Perspectives on Assessing Speaking Skills in Turkish Secondary Schools: A Post-New Exam Model Study)" isimli yüksek lisans tezine ilişkin anket ve mülakat çalışmalarını içeren 152261 sayılı dilekçesi okunarak görüşüldü.

Üniversitemiz Lisansüstü Eğitim Enstitüsü öğrencisi Merve Nur DİNÇER'in Doç. Dr. Betül BAL GEZEGİN danışmanlığında "Türk Ortaokullarında Konuşma Becerilerinin Değerlendirilmesine İlişkin İngilizce Öğretmenlerinin Bakış Açıları: Yeni Sınav Modeli Sonrası Araştırması (EFL Teachers' Perspectives on Assessing Speaking Skills in Turkish Secondary Schools: A Post-New Exam Model Study)" isimli yüksek lisans tezine ilişkin anket ve mülakat çalışmalarının kabulüne oy birliği ile karar verildi.

Appendix B

The Interview Questions

SEMI-STRUCTURED OPEN-ENDED INTERVIEW QUESTIONS FOR EFL TEACHERS

I agree to participate in this interview study with my own consent, and to have my answers recorded as voice recordings and used anonymously in the relevant scientific research.

.../.../20..

.....

Answer the Questions

Years of experience as an EFL teacher.....

The city you work

Education Associate Degree B.A M.A PhD

The grades you currently teach 5 6 7 8

Gender Women Man

Speaking exams became compulsory starting from the 2023-2014 academic year. It was also decided that this exam would be prepared jointly for all classes by the Turkish MoNE. Regarding this;

A. Teachers' Readiness

1. As an English teacher, were you ready to start giving speaking exams at the beginning of the 2023-2024 academic year?

B. Teachers' Thoughts and Feelings

1. What are your general feelings about speaking exams at secondary schools?
2. What do you think about the new regulations for speaking exams at secondary schools?
3. Should speaking exams be mandatory? Why or why not? Explain it, please.
4. How did you feel when you first heard that speaking exams would be compulsory in secondary school?

5. Did the speaking exams you prepared and administered satisfy you professionally as an EFL teacher?
6. According to your experience, did administering the speaking tests motivate you to speak more English in the lessons, in general? Please explain.
7. What do you think about increasing the English course passing grade to 50 points from 45 out of 100?

C. Observations Regarding Students

1. Based on your observations, were students ready to switch to this system?
2. Based on your observation, how did students feel about speaking exams?
3. Did the students experience adaptation problems to the new speaking exams? If so, please describe your observations.
4. What kind of feedback did you receive from students regarding the speaking exam, and how did it affect you?
5. Compulsory English course in secondary school is 3 hours in the 5th and 6th grades and 4 hours in the 7th and 8th grades. Are these lesson hours sufficient to prepare students for the speaking exam? Why or why not?
6. Do the grades after the exam reflect the students' in-class speaking performance? Please explain.
7. According to your experience, did taking the speaking test motivate the students to speak more English in the lessons? Please explain.

D. Testing And Assessment Instruments and Process

1. How do you evaluate the joint preparation and administration of the speaking exam by English departments?
2. What are your ideas about arranging exam dates? How do you set the dates?
3. Did you explain the rules of the speaking exam and how the exam would happen to the children before the exam?
4. What are the main problems you face during a speaking exam?
5. Do you use a scale to grade your students? If yes, is this scale prepared jointly by all English department teachers, or does every teacher at a school prepare it themselves?
6. Can you describe your classroom environment during the speaking exams?
7. How long does the speaking exam take on average for a student?

8. How long does the speaking exam take on average to test the whole class?
9. Did you give feedback to the children about their speaking performance after the exam? If you did, explain how, and if not, explain why not.
10. According to your experiences this academic year, what are the limitations of speaking exams?
11. What are the benefits of speaking exams according to your experience?

E. About Future Exams

1. What would you recommend to improve the speaking exam? In your opinion, how would it be better?
2. Do you think this examination system and regulations will continue in the future?
3. Would you like the new regulations to continue in their current form in the future?

Appendix C

The Questionnaire Questions

QUESTIONNAIRE FOR EFL TEACHERS

Dear participants,

This questionnaire was prepared for the study conducted as part of a master's thesis focusing on English teachers' perspectives on assessing speaking skills after the new exam model and regulations in secondary schools in Türkiye. It aims to examine what teachers have experienced so far in terms of speaking exams in a comprehensive way. In the questionnaire, you are asked to choose the most appropriate option to describe how you, as English teachers, feel before, after, and during the speaking tests and how you feel about your emotional states during this process. Your answers will be recorded by Google Surveys and will be kept confidential and used only for scientific research. The average response time is 10 minutes. Participation in the survey is completely voluntary, and you are free to terminate your participation for any reason.

Sincerely yours,

Merve Nur Dinçer
Ondokuz Mayıs University

PART I: Fill in the information below and answer the questions

- 1-Gender: Woman Man
- 2-The city you work:
- 3-Graduation: Associate Degree B.A M.A Ph.D.
- 4-Years of experience as an EFL teacher:.....
- 5-The grades you currently teach English: 5 6 7 8
- 6-I prepared an English speaking exam before the new regulations.
- ☐ Yes
- ☐ No
- 7-I administered an English speaking exam before the new regulations.
- ☐ Yes
- ☐ No

8-Which techniques did you use in your exams when assessing speaking skills in the 2023-2024 academic year? (You can tick more than one option)

- ☐ Giving instructions/explanation/ description
- ☐ Discussion
- ☐ Translating/ interpreting a text/ dialogue
- ☐ Sentence completion
- ☐ Dialogue completion
- ☐ Telling a short story
- ☐ Problem-solving
- ☐ Talking about pictures
- ☐ Role-playing
- ☐ Oral interview
- ☐ Question and answer
- ☐ Presentation
- ☐ Group discussion
- ☐ Talking about video/ film
- ☐ Reading aloud
- ☐ Other

9- I used a rubric when administering the speaking exams that came with the new regulations.

- ☐ Yes
- ☐ No

10-Which of the following criteria did you consider in your speaking exams in the 2023-2024 academic year? (You can tick more than one option)

- ☐ Grammar
- ☐ Pronunciation
- ☐ Eye contact
- ☐ Vocabulary
- ☐ Fluency
- ☐ Content
- ☐ Clarity of expression
- ☐ Organization

- ☐ Body Language
- ☐ Other

PART II: Choose the possibility corresponding to your opinion

References to "English exams" refer to the whole change in the secondary school English exam system, while references to "speaking exams" refer only to the English speaking exam.

No	Question	I disagree	I am undecided	I agree
A. Teachers' Thoughts and Feelings				
1	The new regulations regarding English exams in the 2023-2024 academic year were necessary in secondary schools.			
2	I have a positive feeling about the grade-wide English exams.			
3	I found the new grading method for English exams, based on percentile values, practical.			
4	Increasing the English course passing grade to 50 points from 45 out of 100 was necessary.			
5	I think that English speaking exams are beneficial for students.			
6	As an English teacher, I felt ready to start giving speaking exams at the beginning of this term.			
7	I have positive feelings about the new regulations for speaking exams at secondary schools.			
8	English speaking exams should be mandatory.			

9	I felt nervous when I first heard that speaking exams would be compulsory in secondary school.			
10	The speaking exams I prepared and administered satisfy me professionally as an EFL teacher.			
11	Administering the speaking tests motivated me to speak more English in the lessons.			
B. Observations Regarding Students				
1	Based on my observations, students were ready to switch to this system.			
2	Students felt worried about the speaking exams.			
3	Students experienced adaptation problems to the new speaking exams.			
4	I received positive feedback from students regarding the speaking exam.			
5	Compulsory English course in secondary school is 3 hours in the 5th and 6th grades and 4 hours in the 7th and 8th grades. Compulsory English course hours in secondary schools were sufficient to prepare students for the speaking exam.			
6	The grades after the exam reflected the students' in-class speaking performance.			
7	Taking the speaking test motivated the students to speak more English in the lessons.			
8	Speaking exams have benefited my students			

C. Experience In Test Preparation, Administration, And Evaluation Processes				
1	The courses I have taken during my undergraduate/master's/doctoral studies helped me to prepare for and apply for speaking exams.			
2	I explained the rules of the speaking exam and how it would happen to the children before the exams.			
3	I faced main problems during speaking exams.			
4	My classroom environment was optimum during the speaking exams.			
5	Each student took the exam with fair timing.			
6	One lesson (40 minutes) was enough for me to examine the whole class.			
7	I gave feedback to the children about their speaking performance after the exams.			
8	(Only those with more than one teacher in the English department at school will answer this question) The new exam system enabled me to co-operate with other English teachers in the school.			
9	(Only those with more than one teacher in the English department at school will answer this question) I used a rubric prepared jointly by all English department teachers at my school to grade my students.			
10	(Only those with more than one teacher in the English department at school will			

	answer this question) The joint preparation and administration of the speaking test by the English department at the school caused problems.			
11	(Only those with more than one teacher in the English department at school will answer this question) I was pleased to work in collaboration with English department teachers during the exam preparation and administration process.			
D. About Future Exams				
1	Speaking exams can be improved over time.			
2	I believe that this examination system and regulations will continue in the future.			
3	I would like the new regulations to continue in their current form in the future.			
4	Please write if you have any suggestions for improving the speaking exams.			

CURRICULUM VITEA

After graduating from Tokat Anatolian Teacher Training High School in 2009, Merve Nur Dinçer graduated from Middle East Technical University, Faculty of Education, Department of English Language Teaching in 2013. In 2022, she enrolled in the Ondokuz Mayıs Foreign Languages Education Department's English Language Teaching Master's program. Following her undergraduate graduation, Merve Nur Dinçer started her career as a Comenius Language Assistant in France and has continued as an English language teacher in many provinces of Türkiye so far. She is fluent in English and French. Her main interests are European youth projects, lifelong learning, photography, and cats.

Contact Information :

ORCID ID : 0009-0008-1322-7921

Publications :

1. Dinçer, M. N., & Bal Gezegin, B. (2024). 1st International Elt, Efl & Ell Conference (Ineeec). In <https://ineec.dpu.edu.tr/>. Kütahya; Oxford. Retrieved from <https://ineec.dpu.edu.tr/index/sayfa/14873/abstract-book>.

Won Awards, Incentives and Scholarships

1. (2024) European Quality Labels
2. (2023) National Quality Labels
3. (2022) Turkish Ministry of National Education Certificate of Achievement
4. (2013) Comenius Language Assistantship Grant
5. (2009) Republic of Türkiye Prime Ministerial Scholarship