Instructors' Views on the Assessment and Evaluation of the Speaking Skill in Turkish as a Foreign Language (TFL) Classes

Ahmet Akkaya [i]
Adıyaman University, Adıyaman, Turkey

Mehmet Yalçın Yılmaz [ii]
İstanbul University, İstanbul, Turkey

Gülnur Aydın [iii]
Adnan Menderes University, Aydın, Turkey

Abstract

Speaking is viewed as a skill more often used in communication and interaction in particular. However, the speaking skill is measured in various ways. The aim of this study is, therefore, to determine the views of Turkish as a Foreign Language (TFL) instructors (lecturers and teachers) on measuring and evaluating speaking skills. To this end, the research data were obtained with the help of 51 TFL instructors who volunteered to fill out a semi-structured interview form titled "Instructors' Views on the Assessment and Evaluation of the Speaking Skill in Turkish as a Foreign Language" created as a Google Docs form. The obtained data were then analyzed through content analysis technique, which revealed that the instructors perform their speaking skill assessments on basis of the criteria including grammar rules, pronunciation/ articulation, fluency, vocabulary knowledge, starting/sustaining dialogues, meaning construction, consistency, self-expression, language learning level, level of achieving outcomes, using phrases, style, length of speech, individual differences, exemplification, chronological narration, avoiding repetition, stress-intonation, body language, rate of speech, and effective use of time. It can be concluded that assessing and evaluating the Turkish speaking skill with a rubric to be developed based on these criteria will help minimize any potential rater subjectivities involved in such measurement and evaluation.

Key words: Assessment and evaluation, lecturers, speaking skill, teachers, Turkish as a foreign language (TFL), views.

---

[i] **Ahmet Akkaya,** Assoc. Prof. Dr., Adıyaman University, Faculty of Education, Department of Turkish Education, Adiyaman/Turkey.

Correspondence: ahmet23akkaya@gmail.com

[ii] **Mehmet Yalçın Yılmaz,** Assist. Prof. Dr., İstanbul University, Research Institute of Turkology, Department of Turkish Language and Literature. İstanbul/Turkey

[iii] **Gülnur Aydın,** Assist. Prof. Dr., Aydın Adnan Menderes University, Faculty of Education, Department of Turkish Education, Aydın/Turkey.

Introduction

Each of the language skills is important in itself for foreign language learning. However, beyond just mastering the language rules, since the speaking skill is directly observable during interaction/communication with other people, it must be paid utmost attention, as pointed out by Alperen (1991) that "language is mostly learned by speaking" (p.60). Recognition of the vowels and consonants of a language, vocabulary knowledge level, grammar rules, meaning construction, stress-intonation, and even body language are all directly related to the competence in speaking skills. Asserting that the primary objective of language teaching is ensuring communication, İşisağ and Demirel (2010, p. 193), underscore the speaking skill as the most important skill among all four language skills.

In addition, speaking is a physical and mental process in which a commonly shared and agreed-upon set of signs and sounds turn into meanings and messages in the minds of the interlocutors using them (Erdem, 2013, p. 181). In other words, "speaking is the way of conveying, explaining and expressing emotions, thoughts, and wishes through visual and auditory elements" (Taşer, 2000, p. 27). Whether for native Turkish speakers or for foreigners learning Turkish, "the main purpose of teaching speaking is to enable someone to express his/her feelings and ideas before a person or community with or without preparation" (Temizyürek, 2007, p. 121). However, it is also stated that students find speaking as the most difficult skill to master where they have the slowest progress (Köksal and Dağ-Pestil, 2014, p. 315). It is obvious that TFL learners have serious problems in improving their speaking skill, and some steps need to be taken to remedy the inadequacies they experience.

The collective action required by the globalization has brought certain standards in language teaching, and when measuring and evaluating foreign language competencies, the criteria in "Common European Framework of Reference for Languages" (CEFR), as an international standard focusing on communicative approach, should be taken into account (Demirel, 2004; Köse, 2007; Özdemir, 2011; Yağmur-Şahin and Aydın, 2014). This can also lay the groundwork for national partnerships as well as internationalization. Indeed, an analysis of the TFL activities in Turkey reveals that the current assessment instruments and practices vary quite extensively. Increasing the efficiency of the instructional process and the quality of learning requires taking time to implement systematic assessment.

In education, *assessment* is defined as "observing a quality and expressing the result of the observation with numbers or other symbols" (Turgut, 1986, p. 3); and *evaluation* is described as "judging a student's success by comparing measurements against a criterion or criteria" (Tan, 2005, p. 180). Kavcar, Oğuzkan and Sever (2004, p. 103) emphasize that assessment and evaluation aim not only to give grades, but also to identify and remedy students' shortcomings. Thus, language teaching assessment should create meaningful feedback to determine the level of instructional attainment, that is, to find out whether the targeted learning outcomes have been achieved. Written tests, short-answer tests, pairing tests, multiple-choice tests and oral exams can be used for this purpose. According to Demir (2015, p. 323), oral exams are the ideal measurement tools that can be used to assess general language or speaking subskills. Such exams, which can be carried out in the form of independent speaking, and dialogues or interviews with the examiner, must be planned well when measuring the TFL speaking skill. Göçer (2014) also states that cognitive, emotional and kinetic outcomes must be taken into account in the evaluation of the speaking skill.

The CEFR classifies the speaking skill into two types: participation in a conversation and performing uninterrupted speech. Correspondingly, Köse (2008, p. 40) categorizes the oral exams as dialogues and verbal expressions. The CEFR (Telc, 2013, p. 33) also focuses on variety/domain, accuracy, fluency, interaction and consistency to assess oral performance. However, our analysis of the related literature has revealed that some studies (Çerçi, 2015; Hamzadayı and Dölek, 2017) focus on assessing Turkish native speakers' speaking skill. Although some research on TFL learners' speaking skill (İşisağ and Demirel, 2010; Sallabaş, 2012a; Sallabaş, 2012b; Göçer, 2015a; Kurt, 2017;

Kurudayıoğlu and Güngör, 2017) exists, the number of studies regarding the assessment of this skill is very limited, two of which focus on scale development (Aksu-Kurtoğlu and Eken, 2011; Arıcı, Sallabaş and Başaran, 2017), and one (Göçer, 2015b) makes theoretical suggestions only. By determining a more rigorous criteria set to assess TFL speaking, the current study aims to make a contribution by filling this extant gap in the literature.

Purpose of the research

This study aims to find out the views of TFL instructors (lecturers and teachers), about the assessment and evaluation of the speaking skill. As part of this general objective, answers to the following research questions were sought:

1. Do the instructors perform assessment and evaluation of the speaking skill?

2. How much time do instructors spend to assess and evaluate the speaking skill?

3. What are the criteria instructors use to assess and evaluate the speaking skill?

4. According to instructors, what are the things to be aware of when assessing and evaluating the speaking skill?

5. What are the problems encountered by instructors in assessing and evaluating the speaking skill?

Method

Research design

In line with the nature of the research purpose, this study was designed as a *descriptive qualitative study*. "Qualitative research is an approach that aims to create theories and to investigate and understand social phenomena in the environment they are part of. ... This requires the researcher to be flexible, to reshape the research process according to collected information, and to follow an approach that is based on the research design and an inductive approach in analyzing the collected data" (Yıldırım, 1999, p. 10). Qualitative descriptive studies are quite helpful in elaborating on the multivariate assessments of the sample within its real-life context. In the current study, the activities and practices for "evaluating and assessing learners' speaking skills in Turkish as a foreign language" are described under sub-headings classified according to the questions in the forms filled by the instructors (lecturers and teachers).

Participants

This study was carried out with the participation of 51 TFL instructors from various organizations between August 4 and 10, 2018, who completed a Google Docs form via web-based virtual office. In selecting the research participants, *convenience sampling* as a purposive sampling method was used. In convenience sampling, the researcher identifies the sample to be studied by considering variables such as cost, time and easy accessibility (Cohen, Manion & Morrison, 2007, p. 109). The characteristics of the volunteering instructors contributing to the research are as follows:

Table 1. Characteristics of the participants

		f	%
Gender	Female	26	50.98
	Male	25	49.02
	Total	51	100
Age group	22-25	2	3.92
	26-30	24	47.06
	31-35	13	25.49
	36-40	7	13.73
	41-45	4	7.84
	46+	1	1.96
	Total	51	100
Work Organization	Adıyaman University TÖMER	5	9.80
	Aydın Adnan Menderes University TÖMER	2	3.92
	İstanbul Aydın University TÖMER	1	1.96
	Çanakkale Onsekiz Mart University TÖMER	1	1.96
	Harran University TÖMER	1	1,96
	İstanbul University Language Center	20	39.22
	Kızılay Education Camp	2	3.92
	Ministry of National Education (MEB)	1	1.96
	Nevşehir Hacı Bektaş Veli University TÖMER	1	1.96
	RET International	9	17.65
	UNICEF	1	1.96
	Yunus Emre Institute	1	1.96
	Presidency for Turks Abroad and Related Communities	2	3.92
	Yuva Association	4	7.84
	Total	51	100
Experience in teaching TFL	1-5 years	34	66.66
	6-10 years	15	29.41
	11-15 years	2	3.92
	Total	51	100

When Table 1 is examined, it can be seen that 26 of the participants are female and 25 of them are male. Two of these are aged 22-25, 24 are in the 26-30 age group, 13 are in the 31-35 age group, 7 are in the 36-40 age group, 4 are in the 41-45 age group, and one is 46+ years old. Of the participants from 14 different workplaces, 34 reported that they have been teaching Turkish as a foreign language between 1 and 5 years, 15 of them 6-10 years, and two reported 11-15 years of teaching experience. Patton (2014, p. 244) states that there is no rule for determining the sample size in qualitative research, and that the sample size depends on what the researcher aims to know, the purpose of the research, what is on the agenda, what is useful, what is credible and what can be done with the time and resources available.

Our study required that the participants included in the survey had to have taken part in the TFL exams before they could make accurate assessments. The information about the exams that the participants reported to have played a role in are as follows:

Table 2. The Turkish tests in which the participants played a role

Test type	f
Placement test	32
Level completion (achievement) test	48
Turkish proficiency test	37
End-of-unit tests	31
Quizzes	36

Looking at Table 2, 32 participants can be seen to have involved in the preparation and administration of a placement test, 48 in a level completion test, 37 in a Turkish proficiency test, 31 in an end-of-unit test, and 36 in quizzes.

Data Collection

In the present study, a semi-structured interview form prepared on the Google Docs web-based virtual office was used as the data collection tool. In studies carried out in the field of social sciences, especially when considering limitations on time and material resources, the ability to collect data from an online application offers significant advantages for researchers.

During the creation of the interview form, first, the relevant literature was reviewed, and then aligning it with the theoretical framework, a draft form with two subsections was created. The form was then revised in the light of the feedback given by three experts who are specialized in TFL. Finally, a pilot study was conducted with six volunteer instructors to finalize the form, increasing its content validity.

The first section of the form titled "Instructors' Views on the Assessment and Evaluation of the Speaking Skill in Turkish as a Foreign Language" includes demographical information about the participants (*gender, age, workplace organization, TFL teaching experience and involvement in the preparation and administration of TFL tests*), and the second section covers two yes/no questions regarding the views and comments on the assessment and evaluation of TFL speaking skill (*whether or not the participant was involved in speaking skill assessment and evaluation, and if the response was negative, why such an assessment and evaluation was not performed*), and four open-ended questions (*the time allotted for the assessment and evaluation of the speaking skill, by what criteria the participants scored student performance, what needs to be paid attention during the assessment, the problems encountered in the assessment and evaluation*). The form was then sent to the TFL instructors between August 4 and 10, 2018, and thus the research data were obtained.

Data analysis

The data obtained from the research were analyzed by content analysis. "Content analysis is defined as a systematic, replicable technique in which certain words of a text are summarized with smaller content categories using codes based on certain rules" (Büyüköztürk, Kılıç-Çakmak, Akgün, Karadeniz & Demirel, 2012, p. 240). Such content analysis includes the steps of (i) coding the data, (ii) identifying emerging themes, (iii) organizing the codes and the themes, (iv) describing and interpreting the results (Yıldırım and Şimşek, 2006). Accordingly, the interview forms were carefully analyzed, and various meaningful statements were identified and coded. This process was independently performed for each form by each of the three researchers. In this way, by identifying and evaluating the items on which the researchers of the study agreed and disagreed, inter-rater reliability was ensured. The formula (Reliability = agreement / [agreement + disagreement] X 100) proposed by Miles & Huberman (1994) was used to calculate the coding reliability of the researchers and inter-rater reliability was found to be 87%. Since a rate of 70% and higher inter-rater reliability is generally accepted as adequate, the coding relability in this study can be viewed as quite high. In the next step, categories/themes were identified and classified according to the similarities or differences among the previously identified codes. Each category was tabulated with the frequency values, and the Tables were further supported by quotations taken directly from the views reported by the research participants.

Results

Firstly, the instructors were asked whether or not they had participated in the assessment and evaluation of the speaking skill as part of TFL test preparation and/or administration.

Table 3. Instructors' experience in speaking skill assessment and evaluation

Assessment and evaluation	f
Yes	50
No	1
Total	51

Table 3 shows that only one of the 51 instructors participating in the study (Participant 15) stated that she did not perform assessment and evaluation of the speaking skill. In the justification (for why not) part of this question, the participant reported that this was because such assessment was "*a challenging process*". The findings of the study were thus based on the data from the remaining 50 participants who reported having performed speaking skill assessment and evaluation.

Instructors reported the amount of time they spend on measuring and evaluating the speaking skill in the following way:

Table 4. The time spent by the instructors for speaking skill assessment and evaluation

Duration/process	Participants	f
1-5 min.	P4, P8, P23, P27, P28, P29, P32, P39, P40, P51	10
6-10 min.	P2, P3, P12, P17, P30, P33, P34, P35, P36, P38, P41, P42, P43, P47, P48, P49	16
11-20 min.	P1, P5, P6, P7, P9, P10, P11, P13, P14, P18, P22, P24, P44, P45, P50	15
30 min.	P21	1
60 min.	P26, P31, P37, P46	4
Process-oriented	P19, P20, P25	3
Total		49

Table 4 demonstrates that only one (Participant 16) among the 50 instructors failed to indicate the amount of time allowed for assessing the speaking skill, and the remaining 49 participants gave answers as shown in the Table. Thus, of these instructors, 10 spend 1-5 minutes, 16 of them 6-10 minutes, 15 spend 11-20 minutes, 1 spends 30 minutes, 4 spend 60 minutes, and 3 have a process-oriented approach. Some answers given by the instructors are as follows:

"I spend 10 minutes per student" (P3)

"A minimum of 15 minutes for each student..." (P5)

"Two hours a day..." (P19)

"5 minutes for each student..." (P23)

"Half an hour." (P21)

"1 hour." (P46)

When the data obtained through the open-ended questions in the data collection form of the study are examined, it is obvious that since most of the instructors reported multiple views on a certain question item, an instructor is shown as a participant for several items on a Table. Thus, the criteria reported to be applied by the instructors for speaking skill assessment and evaluation are as follows:

Table 5. The criteria used by instructors for speaking skill assessment and evaluation

Criteria	Participants	*f*
Grammar rules	P2, P4, P7, P9, P10, P12, P13, P16, P19, P20, P22, P24, P28, P29,	32
Pronunciation / Articulation	P2, P4, P5, P6, P8, P10, P12, P13, P16, P20, P22, P24, P28, P29, P32,	27
Fluency	P2, P4, P6, P7, P12, P13, P28, P32, P33, P34, P35, P37, P38, P39,	24
Vocabulary knowledge	P2, P4, P7, P8, P10, P11, P13, P22, P28, P33, P38, P39, P40, P41,	22
Meaning construction	P2, P5, P6, P11, P12, P13, P22, P23, P26, P28, P37, P39, P40, P42,	15
Starting/sustaining dialogues	P1, P2, P4, P10, P14, P23, P33, P34, P38, P41, P43, P44, P47, P49.	14
Consistency	P4, P28, P34, P35, P39, P41, P43, P44, P47, P48, P49, P50	12
Self-expression	P1, P5, P6, P8, P9, P14, P23, P27, P29, P31, P37, P46	12
Language proficiency level	P21, P24, P44, P48, P50, P51	6
Level of achieving outcomes	P3, P18	2
Using phrases	P24, P25	2
Style	P25, P49	2
Length of speech	P42	1
Individual differences	P3	1
Exemplification	P24	1
Chronological narration	P26	1
Avoiding repetition	P26	1
Stress-intonation	P10	1
Body language	P10	1
Rate of speech	P46	1
Effective use of time	P49	1
Total		179

When the instructors' criteria for speaking assessment are analyzed, 47 instructors were found to have provided valid responses, and three instructors (P17, P30, P36) were found to have given invalid answers. Ragarding frequency values, the instructors mostly emphasized the "grammar rules" (f: 32), "pronunciation/ articulation" (f: 27), "fluency" (f: 24), "vocabulary knowledge" (f: 22), "starting/sustaining dialogues" (f: 14), "meaning construction" (f: 15), "consistency" (f: 12), "self-expression" (f: 12) and "language proficiency" (f: 6) items. Regarding the high-frequency items in particular, some answers given by the instructors are as follows:

"I score according to some criteria such as correct use of the grammar, elocution, fluency, sustaining a dialogue, vocabulary knowledge, and listening comprehension." (P2)

"Listening comprehension, ability to express self, pronunciation." (P6)

"Fluency, consistency, a high level of vocabulary knowledge, using the labguage accurately, fluent speaking, correct pronunciation, giving logical answers to the questions asked." (P28)

"The words used, meaning and context framework, fluency, pronunciation, consistency, language knowledge." (P39)

"Grammar knowledge, vocabulary knowledge, fluency, ability to communicate, consistency in pronunciation…" (P41)

"First of all, I take into account which level (language structures in the proficiency level) the student is at." (P21)

"Self-expression, grammar, fluent speaking" (P46)

In addition to searching for the criteria used by the TFL instructors in the speaking exams, the answer to the question "What are the things to be aware of when assessing and evaluating the speaking skill?" was also sought, and determining the ideal assessment and evaluation according to the instructors was another aim of the study. The views given by the instructors on this item are as follows:

Table 6. Instructor views regarding the things to be aware of when assessing and evaluating the speaking skill

Views	Participants	*f*
Grammar rules	P2, P5, P9, P10, P13, P29, P32, P34, P35, P37, P40, P41,	19
Proficiency level	P3, P4, P19, P21, P22, P24, P28, P32, P34, P35, P38, P39,	18
Fluency	P5, P6, P10, P12, P13, P28, P33, P35, P37, P39, P41, P42,	17
Vocabulary knowledge	P2, P8, P10, P39, P40, P41, P42, P43, P45, P48, P49, P50,	13
Meaning construction	P2, P5, P6, P11, P12, P13, P17, P34, P35, P37, P39, P40,	13
Pronunciation	P5, P6, P12, P13, P14, P16, P28, P29, P33, P39, P40, P42	12
Self-expression	P5, P6, P9, P11, P29, P33	6
Student's test anxiety	P2, P21, P23, P24, P27, P38	6
Starting/sustaining dialogues	P4, P21, P41, P43, P49	5
The structural, contextual, and gradual	P1, P21, P22, P24, P28, P38	5
Individual differences	P1, P3, P19, P38	4
Giving enough time	P1, P7, P28, P49	4
Behaviors of the examiner	P1, P28, P31	3
Consistency	P28, P49, P50	3
Scoring according to a specific rubric	P24, P28	2
Student's approach to language learning	P25, P48	2
Rate of speech	P14, P48	2
Visual reading	P30, P36	2
Length of speech	P27, P42	2
Avoiding repetition	P26	1
Prepared-unprepared speech	P2	1
Intelligibility	P35	1
Level of achieving outcomes	P3	1
Body language	P10	1
Characteristics of the test environment	P23	1
Style	P49	1
Stress and intonation	P14	1
Free speech	P29	1
Total		147

46 instructors provided their views on the question of what needs to be paid attention during the assessment and evaluation of the speaking skill. The remaining 3 instructors (P15, P20, P44) did not answer this question, and 1 instructor (P18) gave an invalid answer. In terms of frequency values, the instructors highlighted the "grammar rules" (f: 19), "language proficiency" (f: 18), "fluency" (f: 17), "vocabulary knowledge" (f: 13), "meaning construction" (f: 13), "pronunciation" (f: 12), "self-expression" (f: 6), "student's test anxiety" (f: 6), "starting/sustaining dialogues" (f: 5), "formal, contextual, gradual character of questions" (f: 5), "individual differences" (f: 4) and "giving enough time" (f: 4) items in their responses. Especially concerning the high-frequency items, some responses provided by the instructors are as follows:

"Proper sentence construction, use of words, and body language are important. Also, it should be a bit smoother without making intermittent speaking by making sounds like hmmm, erm..." (P10)

"Questions need to be based on dialogues. For this, speaking subjects need to be taken from topics that students are interested in. However, students should feel comfortable and not feel under pressure. Most importantly, whatever level of speaking skill is to be measured while preparing the questions, language structures outside of that level should not be included. " (P21)

"*Comprehension and self-expression, making sentences that are proper for the structure of the Turkish language, using the affixes of Turkish as an agglutinitive language correctly, pronunciation, speaking fluently and accurately.*" (P5)

"*Can he/she pronounce the words correctly? Is his/her answer relevant? Is there continuity? Does he/she use vocabulary and grammar that are proper to his/her proficiency level? These should be noted.*" (P40)

"*Whether there are communication problems and (the scope of) vocabulary knowledge should be paid attention to.*" (P43)

"*Negative situations arising from the personality of the assessed student, test anxiety, incompatibility of the native language phonetics with Turkish, or from his/her own speech impairments (stutter, lisp etc.) must be noted during the assessment. These factors should not be negatively reflected in the assessment. Care should be taken to ensure that students are asked level-appropriate questions that progress from easier to harder to help alleviate their test anxiety.*" (P38)

Instructors' views regarding the problems they experienced during the speaking skill assessment and evaluation are as follows:

Table 7. Problems experienced by instructors during the speaking skill assessment and evaluation

	Views	Participants	f
Learner-related problems	Test anxiety	P4, P7, P8, P18, P21, P23, P24, P27, P28, P29, P31, P34, P35, P37, P39, P41, P48, P49, P50, P51	20
	Pronunciation inadequacy	P1, P5, P6, P9, P19, P39, P40, P42, P43, P48	10
	Vocabulary knowledge inadequacy	P5, P6, P11, P16, P35, P39	6
	Memorized speaking	P24, P25, P35, P39, P40, P41	6
	Giving short answers	P23, P24, P27, P28, P39	5
	Grammar inadequacy	P5, P7, P35, P46	4
	Failure to build meaning	P5, P6, P14, P26	4
	Failure to express self	P5, P6, P14	3
	Failure to adjust voice tone	P14, P23, P48	3
	Individual differences	P5, P43, P49	3
	Failure to speak fluently	P5, P39	2
	Remaining silent/irresponsive	P23, P28	2
	Failure to participate in dialogues	P34, P48	2
	Using informal language	P40	1
	Fatigue resulting from speaking being the final skill being measured	P4	1
	Speaking Turkish out of school	P29	1
	Failure to speak in a logical order	P26	1
	Failure to adjust the rate of speech	P14	1
	Failure to use body language	P14	1
	Total		76
Instructor- and learning and environment-related probems	Lack of any standard/objective measurement instrument	P2, P3, P45	3
	Inadequacy in instructors' knowledge and behavior	P3, P12	2
	Inadequacy of speaking courses and activities	P2, P37	2
	Overcrowded classrooms	P29, P38	2
	Inadequacies in question/subject quality	P34, P49	2
	Lack of standardization in test practices	P3	1
	Teaching students at different language levels together	P29	1
	Inadequacy of time devoted to assessment	P32	1

and evaluation		
Undeservedly transferring students to one higher level despite their failure	P49	1
Total		15

Regarding the problems they encountered during the speaking skill assessment and evaluation, six instructors (P13, P15, P17, P20, P22, P44) did not respond, one instructor (P33) gave an invalid answer, and three instructors (P10, P30, P36) reported that they did not encounter any problems. The remaining 40 instructors' views were analyzed in two parts: "learner-related problems" and "instructor- and learning environment-related problems." When considered in terms of frequency values, in the "learner-related problems" part, instructors emphasized the "test anxiety" (f: 20), "pronunciation inadequacy" (f: 10), "vocabulary knowledge inadequacy" (f: 6), "memorized speaking" (f: 6), "giving short answers" (f: 5), "grammar inadequacy" (f: 4) and "failure to build meaning" (f:4) items. As for the "instructor- and learning environment-related problems" section, the instructors underlined the "lack of standard/ objective measurement instruments" (f: 3), "inadequacy in instructors' knowledge and behavior" (f: 2), "inadequacy of speaking courses and activities" (f: 2), "overcrowded classrooms" (f: 2) and "inadequacies in question/subject quality" (f: 2) questionnaire items. Some answers from the instructors, especially about the high frequency items, can be exemplified as follows:

"*When students make mistakes, they lose their motivation, they do not use the tenses and the topics taught properly. " (P7)*

"*Students feel pressured and accordingly students do not speak comfortably." (P21)*

"*Fluency and pronunciation problems, as well as not wanting to say complicated sentences, not being able to speak according to the context, preferring simple words and phrases, being hesitant to talk due to psychological reasons, students' reluctance to speak." (P39)*

"*The basic problem is that the speaking skills of the students are weak according to the level they are in, that they get overly nervous when they are speaking with a different teacher and they do not talk. They answer every question with a short answer like "yes" or "no." They memorize answers related to the most common topics for the speaking exam (money or health, harms and benefits of internet usage, etc.)" (P24)*

"*Unavailability of adequate measuring instruments, not having standard exams, lecturers' not being interested or concerned about measuring the speaking skill, not giving adequate assessment training to the lecturers." (P3)*

"*Overcrowded classrooms, having mixed-level students together in the same classroom, students' reluctance to talk in the presence of their classmates, students' preference for keeping silent due to their fear of making mistakes, students' not speaking Turkish much outside the school and their lack of exposure to Turkish." (P29)*

"*Students' anxiety, the questions asked to the students not being interesting to them, students' failure in their own native languages." (P49)*

Discussion and Conclusion

The increasing internationalization, globalization along with the growing awareness of different cultures has made communication skills much more important than ever before, and communication is primarily performed through putting speaking skills into use. In addition to performing regular course activities by following particular syllabi, effective teaching of such speaking skills depends on determining the level of achievement of the targeted outcomes as well as eliminating the identified deficiencies, which requires accurate assessment and evaluation.

Of the 51 instructors, 50 reported performing assessment and evaluation by particularly focusing on the speaking skill, and one instructor reported not testing speaking because it is "a challenging process". Based on their research on the TFL curriculum, Kan, Sülüşoğlu and Demirel (2013) report that lecturers and students agree that skill-specific tests are required to measure each language skill separately, indicating a consensus and supporting the instructor views in our own study.

When the time that the instructors devoted to measuring-evaluating the speaking skill was examined, it was found that 16 instructors spent 6-10 minutes, 15 instructors 11-20 minutes, 10 instructors 1-5 minutes, and 8 instructors spent more than 20 minutes on the speaking exam. The answers to the longer-term or process-oriented assessment were limited. Çerçi (2015) found that Turkish teachers generally agreed that the time allowed for measuring the speaking skill was insufficient. Göçer (2015b) points out the necessity of using the process approach in evaluating the speaking skill. However, there is a similarity between the fact that there is a 20-minute limit for speaking skills in examinations where English proficiency is measured such as TOEFL (https://www.ets.org/tr/toefl/ibt/about/content) and the common view expressed by the instructors that the time given for speaking skill assessment in the Turkish proficiency tests should not be more than 20 minutes.

The criteria used by the instructors in performing speaking skill assessment and evaluation include "grammar rules", "pronunciation/ articulation", "fluency", "vocabulary knowledge", "starting/sustaining dialogues", "meaning construction", "consistency", "self-expression", "language proficiency", "level of achieving outcomes", "using phrases, "style", "length of speech", "individual differences", "exemplification", "chronological narration", "avoiding repetition", "stress-intonation", "body language", "rate of speech", and "effective use of time." Arhan (2007) found that nearly all of the Turkish teachers were doing "assessment and evaluation" of speaking subskills with non-scientific criteria. Hamzadayı and Dölek (2017) found that Turkish teachers paid more attention to the speaking outcomes that required low-level cognitive processing such as gesture-mimicry, voice adjustment, pronunciation, and correct and proper use of words. However, the CEFR (2013) focuses on variety/domain, accuracy, fluency, interaction and consistency to assess the oral performance. It is clear that there is an agreement between these instructor views and the CEFR.

Regarding what they should be careful about when assessing the speaking skill, the instructors mostly focused on "grammar rules", "language proficiency", "fluency", "vocabulary knowledge", "meaning construction", "pronunciation", "self-expression", "student's test anxiety", "starting/sustaining dialogues", "formal, contextual, gradual character of questions", "individual differences" and "giving enough time." It is important to standardize these views, which are quite diverse in terms of the criteria they use in measuring and evaluating speaking skills as well as the points to be considered, so that a consensus can be reached. To do so, the Common European Framework of Reference for Languages (CEFR) can serve as starting point. Köse (2007) emphasized that the CEFR enriches the curriculum and assessment and evaluation methods and makes them more effective in terms of ensuring student achievement and achieving the specified goals. İşisağ and Demirel (2010) concluded that using the CEFR is effective in achieving a positive attitude toward speaking and maintaining success. Göçer (2015a) further stressed the importance of preparing an experiential learning environment for Turkish learners to allow them to learn by doing and thus helping them speak successfully by effectively using the speaking skill, and underscored the benefits to be gained from communicative and task-based language learning practices in achieving targeted speaking skill outcomes.

Problems experienced by the instructors in assessment and evaluation related to speaking skills identified some "problems with learners" and "problems with the instructor and learning environment". As for the problems related to the learners, the instructors especially emphasized the "student test anxiety", followed by "pronunciation inadequacy", "vocabulary knowledge inadequacy", "memorized speaking", "short answer" and "grammar inadequacy". As regards the problems with the instructor and learning environment, "lack of any standard/ objective measurement instruments", "inadequacy in instructors' knowledge and behavior", "inadequacy of speaking courses and activities",

"high number of students" and "inadequacies in question/subject quality" were highlighted. Suggesting that scoring is an important challenge in oral tests, Demir (2015) points out that scoring and scorer bias is highly likely in testing speaking, and emphasizes the need to use detailed rubric or well-described criteria to minimize the risk of such bias. Köksal and Dağ-Pestil (2014) emphasize the importance of language as true to life in speaking classes and the connection of subjects to the students and their lives.

Suggestions

Based on the findings of the current study, the following suggestions can be made:

- What is missing in this study focusing on TFL instructors' views on the assessment and evaluation of the speaking skill, is not including a research question on what type of assessment tools the instructors use. Other researchers can design a future study that fills this gap in our study.

- In the assessment of TFL speaking subskills, detailed criteria can be defined by field experts based on the language level descriptors of the CEFR, which then can be used at the national level.

- Instructors' shortcomings in speaking skill assessment and evaluation can be solved by various pre-service and in-service training courses.

- A rubric can be developed by revising and arranging the items in the findings identified by the current study.

References

Aksu-Kurtoğlu, Ö. & Eken, N. T. (2011). Development of an analytical rubric to determine the ability of Turkish conversational speech at A2 level. Leyla Uzun, Ümit Bozkurt (Ed.). In *Theoretical and applied research on Turkish language teaching* (pp. 533-544). Essen: Die Blue Eule.

Alperen, N. (1991). Turkish eloquence, reading, and writing teaching guide. İstanbul: MEB.

Arhan, S. (2007). *Speaking education at the second level of elementary schools according to teachers' opinions* (Unpublished master's thesis). Gazi University Institute of Educational Sciences, Ankara.

Arıcı, A. F., Sallabaş, M. E. & Başaran M. (2017). B2 level speaking skills of Turkish as a foreign language learners: A rubric development study. *International Online Journal of Educational Sciences, 9*(2), 516-526.

Büyüköztürk, Ş., Çakmak, E., Akgün, Ö., Karadeniz, Ş. & Demirel F. (2012). *Scientific research methods* (11th Edition). Ankara: Pegem Akademi.

Cohen, L., Manion, L. & Morrison, K. (2007). *Research methods in education* (6th Edition). London: Routledge.

Çerçi, A. (2015). Opinions of Turkish teachers on the assessment and evaluation of students' speaking skills. *International Social and Educational Sciences Journal, 2*(4), 186-204.

Demir, E. (2015). Exam preparation in Turkish as a foreign language teaching. Arif Sarıçoban (Ed.) In *Teaching Turkish as a foreign language methodology* (pp. 288-331). Ankara: Pegem Akademi.

Demirel, Ö. (2004). *Foreign language teaching, language passport, language biography, language file*. Ankara: Pegem Akademi.

Erdem, İ. (2013). Speaking education. Abdurrahman Güzel, Halit Karatay (Ed.), In *Turkish teaching manual* (pp. 181-220). Ankara: Pegem Akademi.

Göçer, A. (2014). *Assessment and evaluation in Turkish education.* Ankara: Pegem Akademi.

Göçer, A. (2015a). Equipping learners with the ability to speak Turkish as a foreign language. *Trakya University Social Sciences Journal, 17*(2), 21-36.

Göçer, A. (2015b). Determining the development status of speaking skills in teaching Turkish as a foreign language. *Erzincan University Social Sciences Institute Journal,* ÖS-II, 47-56.

Hamzadayı, E. & Dölek, O. (2017). Approaches of Turkish language teachers to evaluating speaking skills. *Journal of Language Education and Research, 3*(3), 135-151.

İşisağ, K. U. & Demirel, Ö. (2010). Use of the Common European Framework of Reference for languages in the development of speaking skills. *Education and Science, 35*(156), 190-204.

Kan, M. O., Sülüşoğlu, B. & Demirel, Ö. (2013). Evaluation of the views of students and lecturers about Turkish curriculum as a foreign language. *Mustafa Kemal University Social Sciences Institute Journal, 10*(24), 29-42.

Kavcar, C., Oğuzkan, F. & Sever, S. (2004). Teaching Turkish (7th Edition). Ankara: Engin.

Köksal, D. & Dağ-Pestil, A. (2014). Teaching Turkish as a foreign language speaking. Abdullah Şahin (Ed.), In *Teaching Turkish as a foreign language* (pp. 294-316). Ankara: Pegem Akademi.

Köse, D. (2007). The effect of teaching Turkish in accordance with the CEFR on success and attitude. *Manas University Social Sciences Journal,* (18), 107-118.

Köse, D. (2008). Preparation and administration of exams in Turkish language teaching. *Language Journal,* (139), 36-47.

Kurt, B. (2017). Turkic students' speaking in Turkey Turkish analysis according to the "conversational skills assessment scale": The Gagauz sample. *Journal of History Culture and Art Research, 6*(2), 358-384.

Kurudayıoğlu, M. & Güngör, H. (2017). An analysis of the speech self-efficacy of Turkish language learners by various variables. *International Journal of Turkish Literature Culture Education, 6*(2), 1105-1121.

Miles, B. M. & Huberman, A. M. (1994). Qualitative data analysis. London: Sage.

Özdemir, O. (2011). A study on the application of CEFR for Languages at B1 level in the teaching of Turkish as a foreign language. *Journal of Turkish Language Education and Teaching Research,* (1), 234-245.

Patton, M. Q. (2014). *Qualitative research and evaluation methods* (Translation Editors: Mesut Bütün, Selçuk Beşir Demir). Ankara: Pegem Akademi.

Sallabaş, M. E. (2012a). Evaluation of the speech self-efficacy of those who learn Turkish as a foreign language. *Turkish Journal of Social Research, 16*(2), 269-290.

Sallabaş, M. E. (2012b). Evaluation of the speaking anxiety of those who learn Turkish as a foreign language. *Turkish Studies International Periodical for the Languages, Literature and History of Turkish or Turkic, 7*(3), 2199-2218.

Tan, Ş. (2005). *Planning and evaluation of teaching*. Ankara: Pegem Akademi.

Taşer, S. (2000). *Speaking training.* İstanbul: Papirüs.

Telc Language Tests (2013). *Common European Framework of Reference for Languages* (Second Edition). Frankfurt: telc GmbH. (retrieved from https://www.telc.net/fileadmin/user_upload/Publikationen/Diller_iain_Avrupa_Ortak_oneril er_AEeraevesi.pdf on 11.06.2018).

Temizyürek, F. (2007). Developing speaking skills at the second level of primary education. *Ankara University Journal of Educational Sciences Faculty, 40*(2), 113-131.

TOEFL IBT Examination Content (retrieved from https://www.ets.org/tr/toefl/ibt/about/content on 28.06.2018).

Turgut, F. (1986). *Assessment and evaluation methods in education.* Ankara: Saydam.

Yağmur-Şahin, E. & Aydın, G**.** (2014). Common European Framework of Reference for Languages. Abdullah Şahin (Ed.), In *Teaching Turkish as a forign language* (pp. 455-472). Ankara: Pegem Akademi.

Yıldırım, A. (1999). Basic characteristics of qualitative research methods and their importance in educational research. *Education and Science, 23*(112), 7-17.

Yıldırım, A. & Şimşek, H. (2006). *Qualitative research methods in social sciences.* Ankara: Seçkin.