



Constructing a Local Rubric for Iranian EFL High School Learners' Speaking Skill: High School Teachers and University Professors' Voices

Seyyed Ayatollah Razmjoo  (Corresponding Author)

Department of Foreign languages and Linguistics, Shiraz University, Iran
arazmjoo@rose.shirazu.ac.ir

Bahman Hassanzadeh

Department of Foreign languages and Linguistics, Shiraz University, Iran
bahasan96@gmail.com

ARTICLE INFO:

Received date:

2024.01.23

Accepted date:

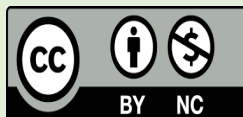
2024.03.25

Print ISSN: 2251-7995

Online ISSN: 2676-6876

Keywords:

rubric construction, rubric, collaborative rubric, speaking skill, instructors' voices, teachers' voices



Abstract

Reviewing the literature, there is an evident lack of appropriate research on the potential difficulties of assessing speaking skill and the urgent need of more rigorous assessment of speaking skill and its intricacies. Therefore, this study aimed to develop a collaborative local rubric for assessing Iranian high school students' speaking skill constructed through the school-university members' collaboration. The participants of this study were 3 high school English teachers and 3 university instructors. They had at least 5 years of experience in English teaching and familiarity with rubrics and how to use and develop them. To gather data, a semi-structured interview was conducted with the participants. After coding and analyzing the data, it was revealed that the participants mentioned some similar criteria such as the importance of pronunciation, grammar, vocabulary, fluency and the use of a 0-4 assessment scale. However, some contradictions arose between the participants about the scores and the strictness/flexibility of those levels. The currently-developed speaking rubric, which is a simulation of the IELTS and TOEFL speaking rubrics, intended to help high school teachers undertake a more thorough assessment resulting in students' enhanced learning of speaking. This study proposed practical implications for teachers and instructors in particular test constructors and examiners.

DOI: 10.22034/elt.2024.61421.2637

Citation: Razmjoo, A. & Hassanzadeh, B; (2024). Constructing a Local Rubric for Iranian EFL High School Learners' Speaking Skill: High School Teachers and University Professors' Voices. *Journal of English Language Teaching and Learning*, 16(34), 343-366. DOI: 10.22034/elt.2024.61421.2637

Introduction

In recent years, the scope of assessment in higher education has expanded to be more than just a simple assessment tool (Pui et al., 2021). Nowadays, the emphasis is more on applying evaluation to enhance student learning and to evaluate underlying thinking processes (Pui et al., 2021). Individuals, whether intentionally or unintentionally, generate rubrics — standards for decision-making and appraisal — mentally on a daily basis (Quinlan, 2012). These mental rubrics can be helpful in making informed decisions aligned with our existing knowledge and pre-determined goals (Ulker, 2017).

Rubrics are sets of multidimensional scoring guidelines and instructions that can ensure consistency in the assessment of students' work (Dawson, 2017). Rubrics are deemed important for showing students what qualities they and their work should have. Moreover, rubrics are commonly used to help students gain accurate perceptions of the learning goal and success criteria (Jonsson, 2014). That is the possible reason behind the usefulness of rubrics in teaching, correlating teaching and assessment, and facilitating students' learning (Black & Wiliam, 1998). English Professor, Elbow (2012) explained the fact that when teachers offer their students a rubric, they are actually telling their students that they deserve to know more about their teacher's important writing aspects and his value as their reader. It is important to pinpoint here that performance-based assessment encompass both writing and speaking skills. Thus, developing rubrics for them requires more or less similar processes (Wigglesworth & Frost, 2017).

Rubrics are frequently a fundamental component of a robust, impartial, and efficient evaluative system in the realm of education (Ulker, 2017). This is mainly due to their usefulness for both formative, learning-oriented, and summative, learning-assessment-oriented, purposes. The primary function of rubrics is to aid students in their learning journey by directing their efforts and providing a more precise comprehension of their learning progress (Chowdhury, 2019). They serve as markers of the achievement of learning objectives. In other words, rubrics ensure that learning outcomes are in correspondence with the evaluation process (Ulker, 2017). A rubric is not only an evaluative tool for giving high-stakes tests, but also an instructional tool that helps students select suitable learning techniques, assist teachers in developing successful instruction tactics, and improve the reliability and validity of the assessment. Additionally, rubrics help teachers evaluate talents through an interpretative process that is only limited by the rubric wording (Nkhoma & Thomas, 2020).

It is crucial to have an objective evaluation — to assess students fairly. However, subjective evaluation enables the assessment of students' analytical abilities, creativity, etc. In such instances, rubrics play a pivotal role in the evaluation process within the educational sector. There are various reasons why rubrics are highly recommended by scholars. First, they provide students with detailed feedbacks during and after the assessment. Second, they guide students in their learning journey, which is crucial to the overall quality of an educational institution. It should be mentioned that the design and use of rubrics may be demanding, but they offer considerable benefits to educational stakeholders (Ulker, 2017).

Widely considered as a key feature of learners' language proficiency whose assessment is of utmost concern of language assessors; speaking is a vital linguistic parameter in determining proficiency and an important factor in efficacious communication. Moreover, Tahir (2015) asserted that a learner is deemed accomplished in acquiring a foreign language if they possess the capacity to speak the language. Speaking is the most difficult ability to evaluate impartially and dependably since it encompasses a blend of diverse factors that may be unrelated or only slightly connected to each other (Abedini & Chalak, 2017). To avoid or, at least, lessen the subjectivity of the evaluation, all of the aforementioned factors may need to be assessed individually. One of the most challenging components of language acquisition is mastering spoken language which has caused abundant struggles for language learners. According to Leong and Ahmadi (2017), foreign language learners frequently face the difficulty of adequately expressing themselves in the target language. People are always employing this beneficial ability throughout their communications, life's pleasures, and every other wonderful day-to-day moment. Moreover, effective communication leads to enriched understanding. Therefore, to have effective communication and avoid communicative complications a great deal of competence is required (Akhter et al., 2020). Despite the immense importance of speaking in second language learning, it has been widely disregarded in schools and other educational centers for a variety of reasons like focus on grammar (Clifford, 1987).

The more important English becomes for worldwide communication, the more remarkable speaking skill is needed (Kargar Behbahani et al., 2024; Nazara, 2011). According to Akhter et al. (2020), there would be no communication if there were no language since the purpose of language is to communicate ideas, emotions, and thoughts. Language is the splendor of communicative existence and speech is the necessary piece to complete that puzzle. Although speech is mostly neglected in testing due to its difficulties for students, EFL students work hard to develop such an oral productive ability. There is no civilization without language and there is no human language without speaking. Language without speaking is like a theater without characters (Akhter, et al., 2020).

As stated, speaking is one of the most essential language skills. People are always dealing with it in everyday situations like doing daily routines and addressing their needs. Children learn to speak prior to reading and writing because it is a necessity in communication with others. In simple terms, the survival element for individuals in any society is knowing how to speak. Since speaking has such huge impacts on people's lives, rigorous examination of students' speaking ability is essential for language teachers. To determine their students' pedagogical requirements, teachers should examine students' speaking abilities. Hence, such a proper examination is very critical for students' successful learning.

What has been evident in the previous research is that the challenges of assessing the speaking ability of second language learners have not been deeply investigated (Akhter, et al., 2020; Nazara, 2011; Ulker, 2017). In other words, assessing speaking skill requires more detailed evaluations. Therefore, designing an appropriate comprehensive rubric encompassing proper criteria for assessing L2 speaking is necessary.

This research intends to construct a comprehensive national rubric for assessing Iranian high school students' speaking through the cooperative work of high school teachers and university instructors. In line with the main objective, this study tries to answer the following questions:

1. What components should be involved in assessing high school students' speaking skill based on the university instructors' viewpoint?
2. What components should be involved in assessing high school students' speaking skill based on the high school teachers' viewpoint?
3. How are high school teachers' perceptions different from those of university instructors?
4. How could the components of assessing high school students' speaking skill be used to develop a rubric?

Due to the fact that speaking is one of the most important language skills, adequate evaluation of students' speaking ability is essential for language teachers (Leong & Ahmadi, 2017). This evaluation allows teachers to understand students' needs more properly so that they are able to supply students with more adequate instructions.

Iranian high-school learners are mostly exposed to a grammar-based and structurally approach in the language classes they attend (Fazilatfar & Kargar Behbahani, 2018; Kargar Behbahani & Khademi, 2022; Rahimi, 2005). Therefore, it could be inferred that some Iranian researchers have cast doubt on the suitability of IELTS and TOEFL speaking rubrics for Iranian high school students. So, developing a rubric which is suitable for students who are educating in the context of Iranian high schools is very important. This rubric attempt to consider every detail of students and even teachers' needs to help them teach and learn better. In other words, since speaking is one of the most important language skills, fully assessing students' speaking ability is essential for foreign language teachers. This assessment allows teachers to better understand their students' needs so they can provide more appropriate instruction.

The potential implications of this study extend beyond the confines of academic research, offering tangible benefits to both educators and learners within the Iranian EFL high school context. By constructing a collaborative rubric tailored specifically to the speaking skills of high school students, this study has the potential to revolutionize language assessment practices in the region. Firstly, the development of a localized rubric acknowledges the unique linguistic and educational landscape of Iranian high schools, ensuring that assessment criteria are culturally relevant, pedagogically sound, and contextually appropriate. This not only enhances the validity and reliability of speaking assessments but also promotes a more inclusive and equitable learning environment for all students. Additionally, the collaborative nature of this endeavor fosters a sense of ownership and investment among educators, empowering them to actively engage in the assessment process and contribute their expertise towards the development of effective instructional strategies. Ultimately, by bridging the gap between theory and practice, this study holds the potential to enhance both teaching and learning outcomes, paving the way for more meaningful and impactful language education experiences for Iranian EFL high school learners.

1. Literature Review

In this section, the variables of this study, rubrics and speaking skill, are defined. Moreover, the related literature regarding the study variables is thoroughly explored and reviewed.

1.1. Rubric

The term 'rubric' has its roots in Latin meaning 'red'. According to the Merriam-Webster Dictionary (2022), a rubric can refer to an "authoritative rule" and also to a "listing of specific standards used to evaluate academic papers, projects, or tests". Technically speaking, a rubric is a way of analyzing and grading students' work, which measures achievement standards for all of the components. In another definition, rubrics are viewed as a valuable instructional and evaluative tool which can assist students in both preparation and evaluation of their own presentations (Phan & Phuong, 2017).

Rubrics can be divided into two types. First, holistic rubrics collect a variety of assessment criteria and group them under grade headings or proficiency levels. Such rubrics are basically single-criterion (one-dimensional) through which students' overall performance is evaluated. The second type is analytic (two-dimensional) rubrics which analyze participants' performance based on several criteria. They separate distinct evaluation criteria and deal with each of them thoroughly (Facione & Facione, 1994). Several experts (Berger, 2011; Brookhart, 2013; Orlich et al., 2010; Walvoord, 2010) make a distinction between two main elements of rubrics: criteria and scoring scales (standards). Criteria define "what is being evaluated" while standards specify "the level of accomplishment and task engagement required to achieve that level" (Orlich, et al., 2010, 34).

Rubric-based assessment is not effective in educational settings where performance is measured by answering questions such as multiple-choice or matching-item tests which have only one correct (right) answer. However, in some rare cases, test performance can also be evaluated using rubrics. For instance, the adequacy, relevance, or completeness of the answer to a question can be assessed using a rubric (Ulker, 2017).

1.2. Empirical studies on rubric

The existing literature underscores the importance of rubrics as an evaluation tool in education. Becker (2016) found that rubrics improve judgment consistency, scoring consistency, and judgment validity. Hima (2017) showed that involving students in developing rubrics leads to higher performance and broader understanding of assessment criteria, a finding echoed by Helmanda and Nisa (2019). Bulut (2022) validated analytical rubrics for assessing speaking and writing skills among elementary students, confirming their reliability and content validity. Similarly, Alias and Osman (2015) developed a reliable oral skills rubric, validated through expert consensus.

Pui et al. (2021) demonstrated that rubrics enhance students' critical thinking and comprehension of assessment criteria. Mahmoudi and Bugra (2020) emphasized the need for rubric instruction to improve writing skills, noting that personalized feedback aids self-assessment. Ghaffar (2020) found that co-constructing rubrics with students increases class interaction and self-evaluation.

However, some studies reported challenges. Hay (1995) highlighted difficulties in consensus on rubric interpretation, while Andrade (2001) noted gender-based differences in rubric use. Additionally, some students struggled to accurately predict their scores using rubrics (Andrade, 2001; McCormick et al., 2007). These mixed results suggest that while rubrics are generally effective, their impact can vary.

1.3. Speaking Skill

Speaking, as described by Burns and Joyce (1997), is a collaborative process of generating, receiving, and analyzing information to create meaning, heavily influenced by contextual factors like participants, their status, physical setting, and conversation purpose. Despite the focus on developing speaking skills from early education, many EFL learners face challenges in L2 production, impacting their job opportunities and interactions with native speakers (Akhter et al., 2020; Farrokhi et al., 2017). Proficient speakers gain academic and personal benefits through active participation in discussions, which fosters their thinking and knowledge.

Definitions of speaking vary among scholars. Webster's New World Dictionary (2022) defines it as orally saying words to communicate, while Chaney (1998) views it as producing and transmitting meaning through vocal symbols in different contexts. Brown (1994) and Burns & Joyce (1997) describe speaking as an interactive process of meaning construction involving information production, receiving, and processing. Flucher and Reiter (2003) focuses on conveying ideas through spoken language, and Darmuki et al. (2018) emphasize ensuring the interlocutor understands the transmitted thoughts, ideas, and emotions. Irawati (2014) sees speaking as verbally producing words and sentences for interaction, and Khorashadyzadeh (2014) highlights the need for both linguistic and sociolinguistic competence, including understanding the context of interaction (Zuhriyah, 2017). Thornbury (2005) notes the unique aspects of spoken language, stressing the need for speakers to adapt language in real-time based on relationships and the dynamic nature of conversations.

1.3.1. Acquiring Speaking Skill

Mazouzi (2013) emphasizes that effective language learning activities must balance accuracy and fluency, both critical for successful communication. Classroom practice helps students grasp the functioning of the language system, with a focus on correctness and comprehensiveness in their linguistic output (Mazouzi, 2013). Fluency is crucial for engaging in smooth conversations, preventing communication breakdowns due to listener disinterest (Hughes, 2003). Thornbury (2005) adds that correct language use involves the complexity, organization, and vocabulary choice in utterances, while pronunciation, including stress, intonation, and pitch, is foundational for clear communication (Thornbury, 2005; Leong & Ahmadi, 2017).

To address speaking challenges, Tuan and Mai (2015) suggest focusing on performance circumstances, listening skills, and feedback. Performance circumstances, such as time restrictions and task preparation, impact speaking ability (Nation & Newton, 2009). Emotional factors like motivation, self-esteem, and anxiety also play a significant role (Krashen, 1982; Oxford, 1990). Effective listening is key to speaking skills, as comprehension is necessary for interaction (Doff, 1998; Shumin, 2002). Constructive feedback is crucial, as excessive criticism can demotivate learners (Harmer, 1991; Baker & Westrup, 2003). Topical knowledge stored in long-term memory enhances speaking performance by enabling contextual language application

(Bachman & Palmer, 1996). Attitudes toward language learning, influenced by personality and social context, also affect speaking performance (Zeinivand, 2015; Dehbozorgi, 2012; Zakian, 2021). Incorporating storytelling in speaking classes can improve fluency, grammar, pronunciation, vocabulary, and content knowledge (Mahdavivand et al., 2024).

1.3.2. Speaking Skill Difficulties

Students encounter several challenges when developing speaking skills, such as hesitation, insufficient knowledge, limited engagement, and reliance on their first language (L1) (Tuan & Mai, 2015). Inhibition, a major issue, arises from fear of making mistakes and being judged, leading to insecurity and embarrassment (Leong & Ahmadi, 2017). This inhibition is fueled by factors like negative evaluation, lack of confidence, anxiety, and an unsupportive classroom environment (Abedini & Chalak, 2017). Anxiety, often tied to fear of criticism, significantly impacts speaking performance (Tok, 2009). Al-Lawati (1995) noted that Omani EFL learners struggle with vocabulary, grammar, pronunciation, and discourse due to weak foundational L2 skills. Park and Lee (2005) found that higher anxiety levels correlate with lower oral performance scores, while Horwitz et al. (1986) identified communication anxiety, test anxiety, and grade anxiety as key sources of language anxiety, with fear of negative evaluation being particularly common in language classes.

Students also struggle with temporary forgetfulness, especially when discussing uninteresting or unfamiliar topics (Rivers, 1968; Baker & Westrup, 2003). This issue is worsened in large classes where limited speaking time and dominant students reduce opportunities for others (Leong & Ahmadi, 2017). Additionally, students often use L1 out of convenience, usually due to a lack of understanding of the topic and insufficient emphasis on L2 communication (Tuan & Mai, 2015; Harmer, 1991). Dewi (2016) identified factors like limited practice time, lack of vocabulary, unengaging teaching methods, and difficulty relating material to real-world situations as contributors to low speaking proficiency. Lopez (2011) stressed the importance of teaching speaking strategies to improve L2 communication, especially for beginners, who need strategies to address issues like hesitation (Melendez et al., 2014). Encouraging students to read texts can also help alleviate these speaking challenges.

1.3.3. Speaking Skill Assessment

The hardest skill to accurately and objectively evaluate is speaking. It has a variety of uncorrelated or slightly correlated variables, each of which needs independent evaluation to prevent or at least lessen assessment subjectivity. "A speaker may produce all the appropriate sounds but not make any sense or have significant difficulties with phonology and grammar and yet be able to get the point across" (Kitao & Kitao, 1996, p. 23). It is one of the most difficult language abilities to analyze, given that learning, teaching, and evaluation all take place individually (Bachman & Palmer, 1996; Luoma, 2004). This problem may be one of the reasons why speech evaluation has received little attention in higher educations throughout the world; instead, the emphasis has been mostly on grammar and vocabulary (Egan, 1999). Berger (2011) advises beginning speaking assessment by reviewing preexisting rubrics when creating a new one. Today, ready-made rubrics can be found very quickly and easily as a byproduct of modern technology. Examining other people's rubrics can help you develop one that perfectly suits your students' needs, your curriculum, and learning task. Assessors and students should find answers

to the following inquiries to create a successful and effective assessment rubric: What is the purpose of assessment? What are the objectives of the assessment? What is the most suitable task for the assessment? How is the assessment procedure carried out? What are the most suitable criteria for scoring learners' performance? What are the most suitable scales for grading the assessment? (Ulker, 2017). According to the caliber of the students' performance, a qualitative scale of assessment provides valuable information on the accomplishment level labeled for each student. This evaluation method can build up a vivid image of learners' accomplishments and reveal the caliber of knowledge. For instance, the label "satisfactory" indicates that a learner passed the test but that further development is required (Ulker, 2017).

To develop the most suitable set of criteria, it is crucial to understand the framework for which the set was created. Knight (1992) offered a set of potential criteria in assessing speaking, from which teachers could select the ones that fit their contexts well (i.e., the needs of their students, the curriculum, etc.). According to Knight (1992), the list comprises eight different criteria: grammar, vocabulary, fluency, pronunciation, conversational skill, sociolinguistic skill, nonverbal (eye contact and body language), content, and pragmatic competence. The speaking evaluation criteria in the most well-known tests in the world, such as Cambridge ESOL, IELTS, and TOEFL, are grounded on a list of pre-set factors: "vocabulary range, accuracy, speaking delivery, pronunciation, ability to respond immediately and adequately, and use of various language instruments" (Ulker, 2017, p. 13).

The technique which is used to evaluate oral interactive skills is dependent on the evaluation's goals. According to Luoma (2004), there are two ways to evaluate speaking abilities. In the first approach, called the observational approach, students' actions are discreetly surveyed and evaluated. However, in the structured approach, students are required to carry out one or more particular oral interactive task(s). The performance on the task is then assessed. The task can be given one-on-one or in a group. In the current study, we used a structured process to interview each learner one-on-one when the course was completed. Structured and observational approaches can make use of a diverse of rating systems.

Holistic and analytical speaking assessment scales can be distinguished in general. The holistic scale can provide a general picture of the learner's performance. A primary trait score can measure a learner's ability to accomplish a particular communication goal like persuading a listener into taking a particular viewpoint. On the other hand, analytical scales can be used to account for the learner's performance on differing communication-related factors like language, organization, delivery, and content. Rating systems might depict different competency levels along a scale or may show the existence of a feature (Luoma, 2004).

A reliable evaluation of speaking proficiency leads to greater performance in speaking, whereas an inadequate assessment of speaking proficiency leads to poor performance in speaking (Bachman & Palmer, 1996). According to a recent study of educational measures in today's teaching and learning methodology (Masters, 2015), 'assessment of learning' has been replaced by 'assessment for learning,' which is known as 'formative assessment' (Khodabakhshzadeh et al., 2018). According to Romer (2017), a speaking assessment scale survey revealed that spoken English lexicogrammar's primary components are not properly represented in spoken evaluation practice. Non-holistic scoring methods aim to explicitly identify the overlap between vocabulary

and grammar and prevent assigning separate ratings to closely- related abilities (Romer, 2017). In his study, Romer (2017) sought to represent features of actual speech by, for example, recognizing the value of discourse markers, coherence ties, and collocations. However, the sharp divide in evaluation criteria between ‘vocabulary’ and ‘grammar’ appears to remain untouched (Romer, 2017). This move from evaluation of learning to assessment of learning paved the way for the development of rubrics with more precise descriptions of next level learning (Masters, 2013). In reality, the evaluation of speaking proficiency should be carried out to improve teaching and learning rather than to rate just task performance (Alberola Colomar, 2014).

Reviewing the literature shows that there have been some previous studies on the issues related to co-constructing a rubric. Also, regarding the importance of speaking evaluation, many studies have been conducted on assessing speaking and its probable difficulties. However, the issue of co-construction of rubrics is a brand-new concept and most of the existing research on this area is about writing skill in higher proficiency levels (Ghaffar, 2020). There is also the issue of the context. The previous studies were not conducted in the context of Iran and did not consider Iranian high school learners. These students have no experience with English especially when it comes to speaking. So, the speaking rubrics of IELTS and TOEFL cannot be completely suitable for these students. Therefore, due to the importance of speaking and its evaluation and the insufficient number of studies on this area, there is an urgent need to qualitatively investigate the collaborative construction of speaking rubrics while considering lower-level learners.

2. Method

In this part, the research methodology used in this study is explained. In addition, information regarding the participants, data collection procedure, and data analysis are presented.

2.1. Design and Participants

To answer the research questions, this study made use of a type of qualitative paradigm, which was grounded theory. This research method was selected because it properly suits the scope of the research. Moreover, such a method is more compatible with the limitation of the low number of participants.

In this study, we used purposive sampling. This kind of sampling was chosen because the participants needed to be experts and have experience in rubrics and developing them. Also, they had to be familiar with high schools and students (Ary et al., 2019).

The participants of this study were three high school teachers and three university instructors. The English teachers were in constant contact with high school students in schools. They all had at least 5 years of experience in English teaching. They also had some sorts of familiarity with rubrics and how to develop or at least use them. University instructors were the language experts who could provide a great help in developing the current rubric. These experts have knowledge and experience of rubrics and their rating. It is worth mentioning that the university instructors also had some high school teaching experience.

The age range of the participants was between 27 to 50. Participants included both men and women. The process of selecting the participants was limited by some factors like availability and having the time to participate in the research. The participants were from Shiraz and Tehran.

2.2. Instrument

The main instrument of this study was a semi-structured interview. The reason behind using this kind of interview was that although some questions were designed and asked during the course of the interview, it allowed the participants to delve deeper into matters and go through all the needed information. Also, this kind of interview is very helpful when we do not have the necessary time, resources, or access to hold the second or more interview. Due to the difficulty of the face-to-face interview for some of the participants, the interview was carried out on Adobe Connect, an online social platform. English was the language of both the interview questions and interview procedure. To check the credibility and make sure that the interview questions were what the research needed and they measured what the research wanted to measure, the researchers double-checked the developed questions with previous literature and their colleagues. To check the trustworthiness of the interview, another rater was asked to go through the interviews and report the results and there was more than 90 percent of inter-coder association based on the calculated Cronbach's alpha.

2.3. Data Collection Procedures

Based on Ary et al. (2019), a series of steps needs to be taken to conduct the interview. The first step was to identify the interviewees and the time and the place of the interview. One of the researchers was the interviewer and the interview was held online on Adobe Connect software. The interview involved 6 main questions and each interview lasted about 30 minutes. Analyzing rubrics made by others may help to develop the one that mostly suits the needs of the learner, learning task and program (Ulker, 2017). Following were the main interview questions based on Ulker (2017) and Ary et al. (2019):

1. What is the purpose of the assessment of Iranian EFL high school learners' speaking skill?
2. What are the most suitable scoring criteria for assessing Iranian EFL high school learners' speaking skill?
3. What are the most suitable grading scales for assessing Iranian EFL high school learners' speaking skill?
4. According to the integrated speaking rubric of TOEFL iBT, the factors involved in assessing speaking are language use, delivery, general description, and topic development. How can we modify these factors in order to suit the assessment of Iranian EFL high school learners' speaking skill?
5. According to speaking rubrics of the IELTS, coherence and fluency, lexical resources, pronunciation, grammatical range, and accuracy are the factors influencing speaking assessment. Are these factors suitable for assessing Iranian EFL high school learners' speaking skill?
6. According to the integrated speaking rubric of TOEFL iBT, the rating scale goes from 0 to 4. On the other hand, according to the IELTS speaking rubric, rating scale goes from 0 to 9. Are these scales appropriate for assessing Iranian EFL high school learners' speaking skill? If not, how should such scales be changed to suit the context?

The second step was introductory in nature which allowed the participants to get to know the interviewer and the purpose of the study. Such a step involves asking some warm-up questions

to reduce emotional tensions and create a relaxed atmosphere. At this stage, the participants were informed of their rights as research participants and their consent forms were obtained regarding their willingness to record their voices for later analysis.

The third step was the conduction of the interview. The follow-up questions were made based on the needed essential details regarding the research objectives. Because it was a semi-structured interview, the interviewees could talk freely and answer any questions that emerged during the interview.

The final step was to finish the interview by asking some concluding questions like 'is there anything else you want to add?' In the end, the interviewees were assured that the information would be kept confidential. Moreover, their phone numbers were taken in case of possible future questions.

2.4. Data Analysis

First, the audio-recordings of the interviews were transcribed based on Marcuschi's (2007) and Bucholtz's (2000) transcribing standards. Transcriptions were done word-for-word to avoid potential biases. The process of transcribing can be based on pragmatic logic in order to aid the reader's comprehension. For instance, if the interviewees talk very quietly, it is unlikely for an audio-recorder to capture all the words in a comprehensible way. As a result, the researcher will have to educate the reader about the parts of the transcript, which can be demonstrated by statements like 'incomprehensible' or 'not clear' (Marcuschi, 2007). According to Bucholtz (2000), transcription should not be viewed as a fixed and precise procedure in which there is only a specific path to follow and a specific result to be obtained (Bucholtz, 2000). As a result of the fact that transcribing is a process that is both objective and subjective, it needs to be stated that reflexivity is an intrinsic part of the transcription process (Bucholtz, 2000).

The primary stage of analysis in this qualitative research was familiarization and organization. At this point, the researchers familiarized themselves with the data by reading, re-reading, and reviewing the notes and transcriptions. After familiarization, the researchers started organizing the data through compiling a list of participants and information about them, resources, and questions.

The next stage of analysis was data coding and reducing. Coding is a framework utilized for categorizing particular behaviors as they happen and reducing is the classification of codes into categories and categories into themes. According to Ary et al. (2019), the first step in coding was called "open coding". In this step, the researchers read and reread the gathered data then organized them based on recurring and significant shared units of meaning such as words, phrases, sentences, or even phenomena. Each of these coding units were understandable. In the beginning, the researchers created as many codes as possible and, later on, they reduced them. The next step was axial coding which is the practice of intra- and inter-connecting categories. Items labeling or coding was done in order to find similarities and differences in the data, leading to the development of higher-level categories. These categories were also modified due to the continuing analysis and the emergence of a 'big picture'. During these processes, the researchers used analytic strategies and thinking techniques provided by Corbin and Strauss (2008), Creswell (2007) and Creswell and Poth (2018). The last step was taken after establishing the categories. At this stage, the researchers considered whether any of the categories were linked or not to

create quite abstract high-level categories (called themes). Such major categories or themes were created by linking lower-level categories.

The last stage of analysis was interpreting and representing. Interpreting involves displaying participants' verbal or non-verbal data and abstracting significant understandings from them. It is an inductive process in which broad generalizations are made based on the commonalities between the categories. In this study, empathic interpretation was used. According to Ricoeur (1996), empathic interpretations are motivated by a desire to get as close to the meaning of a text as possible by trying to understand it inside out.

To check the trustworthiness and credibility of the coding, the researchers asked another expert to code the transcriptions.

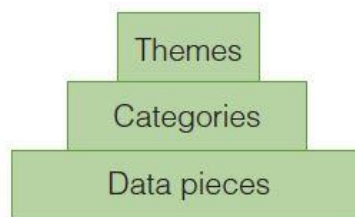


Figure 1. Data Pyramid (Ary et al, 2019, P. 462)

3. Results and Discussion

In this section, first, the research results are explained in detail. Moreover, research findings are compared to those of the previous studies on this topic. The results of the analyzed data were divided into four parts to better provide answers to the research questions. All the approving and disapproving opinions of the participants, who were university instructors and high school teachers, are presented below.

3.1. University instructors' viewpoints on what needs to be included in assessing high school students' speaking skill

This part answers the first question of the study: What are the main components involved in assessing high school students' speaking skill based on the university instructors' viewpoint?

The first point that the participants mentioned was about the overall purpose of assessing the speaking skill of Iranian EFL high school students. All the university instructors pointed out **the importance of the assessing speaking in Iranian high schools** and believed that there was not enough emphasis on speaking in Iran's high schools. According to what one of the participants mentioned:

I think since in an Iranian school, speaking is not usually worked on and this skill is not treated as something important at schools I guess it's very important for teachers to assess students in this skill as well as vocabulary grammar and reading (participant 2).

Another participant said, "We can't deny the importance of speaking in learning English. This cannot be achieved appropriately without a proper assessment" (participant 4). Also, participant 2 added some points about how this assessment can prepare students:

I guess the purpose for the teachers and the system is actually to get to the point that how suitable the books and the educational system was for the students to be prepared for communicating in English, not just reading some passages and answering some questions so I guess the purpose of assessing the students is just to make them ready to be able to communicate with others and as a part of class activity it can be paid attention to too (participant 2).

The next main section of the results is about the rubric criteria. All university instructors believed that **pronunciation, vocabulary, grammar and fluency should be included in the assessment**. Participant 2 said “I guess the criteria should be grammar, vocabulary, pronunciation and how they are able to use all these three together or having a fluent speaking.” In a similar vein, participant 4 mentioned “I believe the factors should be vocabulary, grammar and pronunciation; and after those fluency is a factor that should be considered.” In agreement with what said by others, participant 6 stated “How they can develop the topic with proper vocabulary with the good use of grammar is very important. Pronunciation is very critical to have a fluent speaking.”

The other important aspect of a rubric is its **grading scales**. One of the university instructors believed that the scale should go from 0 to 10 to have a more appropriate and more comprehensive evaluation:

I think 0 to 10 would be ideal. Of course, 0 would be for someone who won't participate and 10 would be for the perfect performance. This way we can examine with more cautious and have a better understanding of the performances (participant 6).

Other instructors insisted on a scale from 0 to 4 / 5. They believed this scale is more suitable for the proficiency level under study. One of them said “I think for this level of students we do not need very detailed evaluation and assessment so a scale from 0 to 4 would be enough” (participant 4). In addition, participant 2 who thought a scale from 0 to 4 / 5 is proper for speaking assessment mentioned the use of half-marks to show the details when it is necessary:

I think it should be 1 to 5. I myself when I'm assessing my students, I don't assess them based on the general knowledge of someone who is studying English, I assess them based on what they have learnt so far. So, what is important to me is that I know that they know all these grammatical points or vocabulary related to some specific topics and I know what they have learnt so far. So, I focus on what they have known and for example have a level on my mind for those who know all those things and are able to use them properly. So, the language use for me is so important, how they can use vocabulary and grammar that they have learnt during this term or the school year and this is 5 in my mind. If the students can reach that level so they use all the vocabulary they have learnt, their speaking is coherent and the use of the grammatical points they have learnt so far and with the correct pronunciation. If someone reach this level the score would be 5 and I guess since they do not know much 0 to 5 will be enough. In my grading scale I have this half of the marks too to be more precise (participant 2).

The next part of the results is about IELTS and TOEFL speaking rubrics and how they can be modified to create a rubric that can serve the goal of this study. One of the most important differences shown between IELTS and TOEFL speaking rubrics is the importance of pronunciation as a factor in assessing speaking. All university instructors stated that pronunciation must be a main element of the rubric, which can be vividly witnessed in the IELTS speaking rubric. For example, participant 2 mentioned “I’m so happy that they paid attention to pronunciation, so I think pronunciation should be considered as a separate part and very important part.” In the same line, participant 4 also stated “I think pronunciation is a part that we cannot miss. It is a very important aspect of speaking and it can show us the level of student’s speaking.”

Most of the university instructors believed that titles used in IELTS speaking rubric, in most cases, were not as vague as titles in TOEFL. However, in some rare cases, the instructors preferred titles used in TOEFL speaking rubric due to their plainness. Also, one of the instructors mentioned that teachers can combine grammar and vocabulary under the title of language use, which is borrowed from TOEFL speaking rubric:

It is obvious that we need to evaluate grammar, vocabulary, pronunciation and fluency of the speaking but to make a comparison between IELTS and TOEFL; I prefer topic development from TOEFL because it tells you exactly what the rubric wants, then I will use fluency and coherence with grammatical range (participant 6).

Similarly, participant 2 said that:

Grammatical range, lexical resource, and accuracy and coherence and fluency are the parts I like but as a comparison between IELTS and TOEFL I like topic development in TOEFL because it conveys the meaning better... I guess lexical resource and grammar can both be put in language use. But for high school students, I guess fluency is not that much of the case because they haven’t been taught regarding speaking skill. But regarding Coherence, Pronunciation, lexical resources and Grammatical range together they are important, and again as I said for TOEFL, I guess not exactly the same format but the details are so important because they draw teacher’s attention to some specific part of language learning which should be paid attention to enough but I think they are still strict.

3.2. High school teachers’ viewpoints on what needs to be included in assessing high school students’ speaking skill

This part answers the second question of the research: What are the main components involved in assessing high school students’ speaking skill based on the high school teachers’ viewpoint?

The first noticeable point mentioned by the teachers was **the difficulty of level placement existing in IELTS and TOEFL speaking rubrics**. All the teachers believed that rubrics should be more flexible and less strict since EFL learners who have recently started their learning cannot be easily assessed using IELTS and TOEFL speaking rubrics. According to the teachers, students should not be afraid of being evaluated; instead, they should feel motivated when exposed to

evaluative environments. For example, participant 5 said “the levels need some balancing. They should not be irritating for students.” Similarly, participant 3 stated:

I'm not really sure about the scores, they are a little bit difficult for a student who has not a lot of experience regarding English ... I think both of these exams are great and serve their purposes but they can be less strict for children who may not have much experience with English and how to speak it. So, I think the levels should be fairer, I mean for example to get the best score, we should not expect the students to be perfect in every way.

Participant 1 also stated that:

You know when we want to assess them, we shouldn't really be so hard and cruel towards them because as you know some teachers expect them to be very careful about grammar for instance. But we know that they are just starting and we all have some weak points. When we want to teach a kid to run, we cannot expect them to run in the first week, little by little. So, when we want to assess these amateurs, we should not expect complex structures regarding grammatical points and vocabulary; just the fact that they know the vocabulary should be enough. Some teacher might expect pronunciation.

There were some themes in line with what the university instructors mentioned. For example, one of the teachers mentioned that the developed rubric could be used to evaluate what has been taught during the class to see how much the teaching was successful:

We can look at it as such a criterion to see whether after an education year they have learnt what was taught, so this criterion would be helpful. Sometimes it's also good to challenge them. Some of them has potential that we are not aware. For example, they are playing video games all the time and when it came to studying, they did not push themselves to their limits that they have. So, if we make it a little bit challenging for them it would be helpful because we let them know what they are capable of. If I want to put it in a nutshell, to be kind but at the same time fair and strict (participant 1).

The teachers also believe that **grading scales should go from 0 to 4 or 5 instead 0 to 9**. For instance, participant 5 mentioned that “I think the scales should be from 0 to 4 like TOEFL because it is more suitable for high school student and decrease the pressure.”

The teachers also talked about proper criteria for the rubric. They consider **vocabulary, grammar, and fluency as the main factors to be evaluated**. Except for one of them, they also consider pronunciation as the main element of the rubric:

Some students try to convey their points to tell audience but the Pronunciation might not be very correct in the beginning. It might be understandable but not as it should be. So, in the early stages, we should not really consider Pronunciation a must but rather a plus ... When we look at IELTS we see that its strength point is its titles, because they are very understandable and not vague and we see that also in terms of being friendly, IELTS is more friendly

because it can close its eyes to some little mistakes that will occur in almost all of the students (participant 1).

Participant 3 also said similar points:

First of all, I think Pronunciation is very important and it should be a main part of the assessment; so, it is the first thing I would adopt from IELTS ... As I mentioned before Pronunciation should be a main part of the assessment. Then we should have Vocabulary and Grammar. After that we should pay attention to Fluency and Coherence.

3.3. Differences in High school teachers' viewpoint and instructors' viewpoints

This part has answered the third question of the study: How are the high school teachers' perceptions different from those of university instructors?

High school teachers' perceptions were not too distinct from those of university instructors. Most of the points mentioned by university instructors and teachers were the same. However, as the old proverb goes "tastes differ," some highly intricate differences can always be seen when more than one individual's viewpoint is involved.

One of the high school teachers believed that pronunciation is not important for the evaluation of the students at this level and background but all the instructors consider pronunciation a very important and independent factor in the speaking rubric.

An important difference that can be seen is the mention of the difficulty of the levels of the rubrics of ILETS and TOEFL by teachers. They believe that these levels should be more flexible and not too strict.

As it was mentioned, there were a lot of similarities in the viewpoints of high school teachers and university instructors. They all believe **the grading scale should be between 0 to 4 and they also consider grammar, vocabulary and fluency to be factors that should be included in the speaking rubric.**

3.4. Developing a rubric

This part answers the fourth question of the study: How could the identified components be used to develop a rubric?

After analyzing the data and comparing the results, the first draft of the rubric was developed. An expert was asked to do the same thing and provide a rubric according to the results. After developing the rubric, it was checked by an expert who has experience in rating and developing rubrics and also working with them. There were also some consultations through talks and interview about the final rubric. Some changes were made after analyzing and comparing the data that was extracted from the interviews with the experts of rubrics and ratings.

When the rubric was ready, it was the time to test this rubric. Rubric was given to an examiner to test the use of this rubric.

Based on what has been achieved through the analysis of interview transcriptions, the following collaborative rubric was developed. It is important to mention that while a few participants referenced scales ranging from 0 to 5 or 0 to 4.5, the rubric provided is based on a

scale of 0 to 4. This because we standardized the scores by converting all participant-referenced scales to the 0 to 4 rubric scale used in the study. This involved proportional scaling to ensure that the ratings remained consistent and comparable. For instance, a score of 4.5 on a 0 to 5 scale was converted to 3.6 on a 0 to 4 scale by applying a simple ratio conversion. This approach maintained the integrity of the original assessments while ensuring uniformity across all data points for accurate analysis and comparison.

Score	Pronunciation	Vocabulary Range	Grammar Use	Fluency & Coherence
4	<ul style="list-style-type: none"> -Using an immense variety of phonological features, conveying clear meaning -Able to maintain proper rhythm. Flexibility in using stress and intonation for longer utterances, although there are occasional stutters -Can be easily understood 	<ul style="list-style-type: none"> -unlimited resource flexibly and easily used to discuss every topic and convey clear meaning. -Skillful use of less common figurative items in spite of infrequent inaccuracies in collocation and word choice. -Effective use of paraphrasing as required. 	<ul style="list-style-type: none"> The answer shows Efficient grammar use. It indicates a quite high level of automaticity with acceptable control of primary and complicated structures. Some small errors are noticed but do not complicate meaning. 	<ul style="list-style-type: none"> The flow is Mostly well-paced (expressing fluidly). Clear speech, may involve minor struggling with pronunciation or intonation patterns, minor lapses, which do not influence overall comprehensibility.
3	<ul style="list-style-type: none"> -Using a variety of phonological features with variable control. -Generally Proper Chunking, but a rapid speech rate or a lack of precise stress timing affect rhythm. -Can generally be understood 	<ul style="list-style-type: none"> - Sufficient resource used to discuss topics for quite a long time. -Improper vocabulary use but with clear meaning. -Generally being able to produce successful paraphrase. 	<ul style="list-style-type: none"> The answer shows quite automatic and successful grammar use. Response might show some imprecise or erroneous use of grammatical structures or somehow use of limited range of structures. 	<ul style="list-style-type: none"> Generally clear Speech, with some fluid expressions, though minor struggling with pace, pronunciation, and intonation are evident which at times need listener's effort.
2	<ul style="list-style-type: none"> -Using some satisfactory phonological features, limited range -Producing some correct chunking, but frequent lapses are evident in overall rhythm 	<ul style="list-style-type: none"> -Sufficient resource for known topics but conveying only primary meaning on unknown topics. -Repeated use of inappropriate words with Errors in word choice. 	<ul style="list-style-type: none"> The answer shows the limited control of and use range of grammar. Mostly, only primary sentence structures are used accurately 	<ul style="list-style-type: none"> Basically, intelligible speech, though vague articulation, strange intonation, and uneven rhythm or pace require listener's effort; complications in meaning in places.

	-Understanding with lots of efforts and incomprehensibility of some parts of speech.	-Rare efforts to paraphrase.	and are spoken with fluidity.	
1	-Uses few acceptable phonological features -Overall problems with delivery -Often unintelligible	-Very restricted resource. isolated lexical items or memorized sentences are evident in utterances. -Very little successful communication without using mime or gesture.	Very restricted range and control of grammar severely impede expressing opinions and associations among them. Some inferior responses might depend greatly on formulaic expressions or practiced chunks.	Constant struggling with stress, pronunciation, and intonation need hard listener's effort; delivery-pace is uneven, telegraphic, and broken; there are also repeated pauses and hesitations.
0	Does Not Attend /non-attendance			

3.5. Discussion

The present study investigated the development of an appropriate rubric for assessing the speaking skill of EFL learners in high schools of Iran. The research findings suggested that the rubrics should be flexible and not very strict because they may negatively impact high school students and their performance. To have such a rubric, participants, especially high school teachers, believed that many studies have been conducted on developing rubrics and their uses to improve learning. However, developing and then using a speaking rubric in Iranian high schools was a matter awaiting discussions and finally possible solutions. This study's findings align with the results of some studies on the importance of speaking evaluation and its important role of this in learning process (Leong & Ahmadi, 2017; Kitao, 1996). The results of the study are also consistent with the findings of the Ulker's (2017) study which investigated the development and use of speaking evaluation rubrics. Moreover, this study supported the points about the importance of proper assessment of speaking skill mentioned by Bachman & Palmer (1996) and Masters (2015).

Also, the results of this study declined the other empirical studies on the use of rubrics reported minor contradictory changes in students' performance. For example, Andrade (2001), revealed contradictory results regarding students' use of rubric in the issue of gender (Andrade, 2001). Hay (1995), however, indirectly described a difficulty in reaching a consensus regarding the identification and interpretation of the marking criteria, just as stated in other studies (Hay, 1995).

The study reaffirms the pivotal role of rubrics as effective tools for language assessment, particularly in speaking skills. The distinction between holistic and analytic rubrics (Facione & Facione, 1994) is well-documented, with the latter providing a more detailed evaluation. The current study's emphasis on developing analytic rubrics for Iranian high school students aligns with the findings of Becker (2016) and Bulut (2022), who highlighted the benefits of rubrics in

enhancing judgment consistency, scoring reliability, and validity in performance assessments. The robustness of analytic rubrics in educational settings, as validated by these studies, underscores the importance of using detailed assessment criteria to capture the complexities of speaking skills.

The process of co-constructing rubrics emerged as a significant theme in this study, demonstrating its potential to enhance student understanding and engagement with assessment criteria. This finding supports the work of Hima (2017) and Ghaffar (2020), who found that involving students in the rubric development process leads to higher mean scores and increased class interaction, respectively. The collaborative approach not only clarifies assessment expectations but also fosters a more student-centered learning environment, promoting self-assessment and deeper engagement. These outcomes are consistent with the theoretical framework of formative assessment, which emphasizes using assessment as a tool for learning (Masters, 2013). The present study reinforces the value of integrating formative assessment practices, such as rubric co-construction, to promote deeper learning and self-reflection among students.

The study's focus on developing rubrics tailored to the specific context of Iranian high school learners addresses a critical gap in the literature. Speaking, as described by Burns and Joyce (1997), is a complex, interactive process that involves generating, receiving, and analyzing information. This complexity makes the assessment of speaking particularly challenging. The findings align with existing research that underscores the importance of balancing accuracy and fluency in language learning (Mazouzi, 2013; Hughes, 2003). The study contributes to the body of knowledge by providing detailed assessment criteria that capture the multifaceted nature of speaking skills, including grammar, vocabulary, pronunciation, fluency, and conversational skills (Thornbury, 2005; Luoma, 2004). These criteria are crucial for evaluating the comprehensive performance of learners in speaking tasks.

The study also sheds light on the various challenges students face in developing speaking skills, such as hesitation, anxiety, and limited engagement (Tuan & Mai, 2015; Leong & Ahmadi, 2017). The findings emphasize the need for supportive and constructive feedback during speaking activities, which can help mitigate anxiety and promote a more positive learning experience (Harmer, 1991; Baker & Westrup, 2003). This approach aligns with the literature on the impact of emotional factors, such as motivation and self-esteem, on speaking performance (Krashen, 1982; Oxford, 1990). The study's focus on rubric co-construction helps create a supportive learning environment that encourages student participation and reduces the fear of negative evaluation.

Conclusion and Implications

In this part the concluding remarks of the study are stated. Practical implications resulting from the current research are also mentioned.

According to results of this study viewpoints of the two groups of the participants, high school teachers and university instructors, were not very different. Both groups considered grammar, vocabulary and fluency as the main factors of the rubric. All the participants believed that the grading score should be between 0 to 4.

There were also some differences in the viewpoints of high school teachers and university instructors. All the instructors believed that pronunciation should be a main factor in the rubric but one of the teachers did not consider the pronunciation as a factor that should be examined. Also, all the teachers mentioned the difficulty of the levels of IELTS and TOEFL speaking rubrics. They believed that these levels are too strict and they should be more flexible.

After analyzing the data and comparing the answers a rubric was developed by the researchers. An expert was asked to develop another rubric. A final rubric was developed from these two rubrics. The final rubric was tested by an examiner to evaluate the use of this rubric and how it can help students and teachers.

The culmination of this research, in the form of a finalized rubric, stands as a valuable tool for high school educators, offering a systematic approach to assessing students' speaking skill. This instrument represents more than a mere assessment framework; rather, it serves as a conduit for facilitating a deeper understanding of students' learning needs and guiding educators in tailoring their instructional practices to meet those needs effectively. By leveraging the developed speaking rubric, educators can gain insights into optimal pedagogical strategies that foster meaningful language acquisition experiences for their students.

The significance of this research extends beyond the mere creation of a rubric; it lies in its potential to enhance the quality of speaking assessments in Iranian high schools, aligning with the unique educational context of EFL instruction. Through the utilization of this refined assessment tool, educators can evaluate students' speaking abilities with greater relevance and precision, ensuring that assessment practices resonate with the specific linguistic and cultural nuances inherent to the Iranian educational landscape.

Furthermore, the adoption of this speaking rubric empowers educators to gain a more nuanced understanding of their students' proficiency levels, enabling them to tailor interventions and support mechanisms that address individual learning trajectories effectively. By fostering a deeper comprehension of students' linguistic competencies, educators can orchestrate targeted interventions that nurture language development and promote meaningful progress.

The benefits of implementing such a rubric extend beyond the realm of educators, permeating into the student experience as well. Through the structured feedback provided by educators using the rubric, students gain valuable insights into their strengths and areas for improvement, thereby facilitating self-awareness and self-directed learning. By harnessing the feedback loop facilitated by the rubric, students can embark on a journey of continuous improvement, iteratively refining their speaking skills in alignment with their individual learning goals.

In essence, the development and deployment of this speaking rubric herald a paradigm shift in educational assessment practices, underscoring the symbiotic relationship between assessment and pedagogy. Through its judicious application, this rubric has the potential to catalyze transformative learning experiences, fostering linguistic growth and proficiency among students while empowering educators to navigate the complexities of language instruction with precision and efficacy.

This study has potential limitations. The results that are discussed in this study were based on the analyses of the ideas mentioned by the participants during the interviews. These analyses are

subject to biases. Also, due to time constraints, the selection and sampling of the participants were limited.

Since the participants of this study were limited, the further studies can be done with a wider range of participants. This study tried to determine the criteria for speaking skill assessment in the context of high schools of Iran; further studies can be done in other contexts and other levels of proficiency.

References

- Abedini, F., & Chalak, A. (2017). Investigating the inhibitive factors in the speaking of Iranian EFL learners. *Journal of Applied Linguistics and Language Research*, 4(6), 82-97.
- Akhter, S., Haidov, R., Rana, A. M., & Qureshi, A. H. (2020). Exploring the significance of speaking skill for EFL learners. *PalArch's Journal of Archaeology of Egypt/Egyptology*, 17(9), 6019-6030.
- Alberola Colomar, M. P. (2014). A classroom-based assessment method to test speaking skills in English for Specific Purposes. *Language Learning in Higher Education*, 4(1), 9-26.
- Alias, A., & Osman, K. (2015). Assessing oral communication skills in science: A rubric development. *Asia Pacific Journal of Educators and Education*, 30(1), 1-18.
- Al-Lawatiyah, M. T. (1995). *A diagnostic study of the difficulties encountered by Omani secondary school students in their oral production of English* [Doctoral thesis, Sultan Qaboos University].
- Andrade, H. (2001). The effects of instructional rubrics on learning to write. *Current Issues in Education*, 4(4), 1-28.
- Ary, D., Jacobs, L.C, Sorensen, C.K., & Walker, D. (2019). *Introduction to research in education*. (10th ed.). Wadsworth/Cengage Learning.
- Bachman, L., & Palmer, A. (1996). *Language testing in practice*. Oxford University Press.
- Baker, J., & Westrup, H. (2003). *Essential speaking skills: A handbook for English language teachers*. Bloomsbury Publishing.
- Becker, A. (2016). Student-generated scoring rubrics: Examining their formative value for improving ESL students' writing performance. *Assessing Writing*, 29, 15-24.
- Berger, J. (2011). Evaluating the effectiveness of instruction using principles of adult learning. In V.C.X. Wang (Eds.), *Assessing and evaluating adult learning in career and technical education* (pp. 173-190). Information Science Reference.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, 5(1), 7-74.
- Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. ASCD.
- Brown, H. D. (1994). *Teaching by principles: An interactive approach to language pedagogy*. Prentice Hall Regents.
- Bucholtz, M. (2000). The politics of transcription. *Journal of Pragmatics*, 32(10), 1439-1465.
- Bulut, B. (2022). Analytical rubric development study for assessment of the speaking and writing skills at primary school level. *International Journal of Progressive Education*, 18(6), 180-193.
- Burns, A., & Joyce, H. (1997). *Focus on speaking*. National Center for English Language Teaching and Research.
- Cantera, M. A., Arevalo, M. J., García-Marina, V., & Alves-Castro, M. (2021). A rubric to assess and improve technical writing in undergraduate engineering courses. *Education Sciences*, 11(4), 146.
- Chaney, A. (1998). *Teaching oral communication in grades K-8*. Allyn and Bacon.
- Chowdhury, F. (2019). Application of rubrics in the classroom: A vital tool for improvement in assessment, feedback and learning. *International Education Studies*, 12(1), 61-68.
- Clifford, R. T. (1987). Language teaching in the federal government: A personal perspective. *The Annals of the American Academy of Political and Social Science*, 490(1), 137-146.

- Creswell, J. W. (2007, May). *Exploring the dialectic tensions in the discourse in mixed methods: What is mixed methods research* [Paper presentation]. Congress of Qualitative Inquiry Conference 2007, Urbana-Champaign, United states.
- Creswell, J. W., & Poth (2018). *Qualitative inquiry and research design: Choosing among five traditions* (4th ed.). Sage Publications.
- Darmuki, A., Andayani, A., Nurkamto, J., & Saddhono, K. (2018). The development and evaluation of speaking learning model by cooperative approach. *International Journal of Instruction*, 11(2), 115-128.
- Dawson, P. (2017). Assessment rubrics: towards clearer and more replicable design, research and practice. *Assessment & Evaluation in Higher Education*, 42(3), 347-360.
- Dehbozorgi, E. (2012). Effects of attitude towards language learning and risk-taking on EFL student's proficiency. *International Journal of English Linguistics*, 2(2), 41-48.
- Dewi, H. (2016). Project based learning techniques to improve speaking skills. *English Education Journal*, 7(3), 341-359.
- Doff, A. (1998). *Teach English: A training course for Teacher*. Cambridge University Press.
- Egan, K. B. (1999). Speaking: A critical skill and a challenge. *CALICO Journal*, 277-293.
- Elbow, P. (2012). Good enough evaluation: When is it feasible and when is evaluation not worth having. *Writing assessment in the 21st century: Essays in honor of Edward M. White*, 303-325.
- Facione, P. & Facione, N. (1994). *The holistic critical thinking rubric*. California Academic Press.
- Farrokhi, F., Zohrabi, M. & Chehr Azad, M. (2017). The effect of the corrective feedback on Iranian EFL learners' speaking accuracy and breakdown fluency. *Language Horizons*, 1(2), 107-129.
- Fazilatfar, A. M., & Kargar Behbahani, H. (2018). The effect of read-aloud method on Iranian EFL learners' reading comprehension. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*, 4(2), 81-100. <https://doi.org/10.22049/jalda.2018.26255.1070>
- Fulcher, G., & Reiter, R. M. (2003). Task difficulty in speaking tests. *Language testing*, 20(3), 321-344.
- Ghaffar, M. A. (2020). Co-constructing writing rubrics to enhance students' autonomy and collaboration. *Journal of English Language Teaching and Applied Linguistics*, 2(5), 43-47.
- Ghaffar, M. A., Khairallah, M., & Salloum, S. (2020). Co-constructed rubrics and assessment for learning: The impact on middle school students' attitudes and writing skills. *Assessing Writing*, 45, 100-468.
- Harmer, J. (1991). *The practice of English language teaching*, 3rd edition, Longman.
- Hay, I. (1995). Communicating geographies: Development and application of a communication instruction manual in the geography discipline of an Australian university. *Journal of Geography in Higher Education*, 19(2), 159-175.
- Helmanda, C. M., & Nisa, R. (2019). The analysis of students' oral performances by using speaking scoring rubric. *Jurnal Dedikasi Pendidikan*, 3(2), 76-86.
- Hima, A. N., & Saputro, T. H. (2017). *Developing Scoring Rubric: Do the Students Need It? ELT-Echo*, 2 (2), 101-113.
- Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern language journal*, 70(2), 125-132.
- Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge University Press.
- Indramawan, A., & Akhyak, D. (2013). Improving the students' English-speaking competence through storytelling (Study in Pangeran Diponegoro islamic college (STAI) of Nganjuk, East Java, Indonesia). *International Journal of Language and Literature*, 1(2), 18-24.
- Irawati, I. (2014). Improving students' speaking ability through communicative language games. *Magistra*, 87, 25-36.
- Jonsson, A. (2014). Rubrics as a way of providing transparency in assessment. *Assessment & Evaluation in Higher Education*, 39(7), 840-852.

- Kargar Behbahani, H., & Khademi, A. (2022). The concurrent contribution of input flooding, visual input enhancement, and consciousness-raising tasks to noticing and intake of present perfect tense. *MEXTESOL Journal*, 46(4), n4.
- Kargar Behbahani, H., Namaziandost, E., & YarAhmadi, M. (2024). The effect of 4/3/2 technique on Iranian EFL learners' speaking fluency: The moderating role of working memory. *Journal of English Language Pedagogy and Practice*, 1(33). <https://doi.org/10.71586/jal.2024.05191120204>
- Khodabakhshzadeh, H., Kafi, Z., & Hosseinnia, M. (2018). Investigating EFL teachers' conceptions and literacy of formative assessment: Constructing and validating an inventory. *International Journal of Instruction*, 11(1), 139-152.
- Khorashadyzadeh, A. (2014). Why to use short stories in speaking classes. *International Journal of Foreign Language Teaching in the Islamic World*, 2(1), 9-15.
- Kitao, S. K., & Kitao, K. (1996). Testing speaking (Report No.TM025215). *ERIC Document Reproduction Service No. ED398261*. <http://eric.ed.gov/?q=Testing+Speaking+Kita>
- Knight, B. (1992). Assessing speaking skills: a workshop for teacher development. *ELT Journal*, 46(3), 294-302.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Leong, L. M., & Ahmadi, S. M. (2017). An analysis of factors influencing learners' English-speaking skill. *International Journal of Research in English Education*, 2(1), 34-41.
- Li, C. (2018). *Constructing and applying rubrics in college-level EFL writing assessment in China* [Doctoral thesis, Syracuse University]. ProQuest Thessis Publishing
- López, M. M. (2011). Speaking strategies used by BA ELT students in public universities in Mexico. *MEXTESOL Journal*, 35(1), 1-22.
- Luoma, S. (2004). *Assessing speaking*. Cambridge University Press.
- Mahdaviand, S., Zohrabi, M. & Zarei, P. (2024). Exploring the facilitate-listen-engage (FLE) model's role on EFL learners' communication willingness. *Iranian Journal of English for Academic Purposes*, 13(1), 48-62.
- Mahmoudi, F. & Buğra, C. (2020). The effects of using rubrics and face to face feedback in teaching writing skill in higher education. *International Online Journal of Education and Teaching*, 7(1). 150-158.
- Mahmoudi, L. A. (2007). *Análise da conversação*, 6th editon, Ática.
- Marcuschi, L. A. (2007). *Análise da conversação* [Conversation analysis], 6th ed, São Paulo, Brazil: Ática.
- Masters, G. (2015) Rethinking formative and summative assessment. *Teacher Magazine*, 20 May 2015. <https://www.teachermagazine.com.au/geoffmasters/article/rethinking-formative-and-summativeassessment>
- Mazouzi, M. S. (2013). *Analysis of some factors affecting learners' oral performance. A Case Study: 3rd Year Pupils of Mena'a's Middle Schools* [Master's thesis, Mohamed Khider University].
- McCormick, M. J., Dooley, K. E., Lindner, J. R., & Cummins, R. L. (2007). Perceived growth versus actual growth in executive leadership competencies: An application of the stair-step behaviorally anchored evaluation approach. *Journal of Agricultural Education*, 48(2), 23-35.
- Menéndez-Varela, J. L., & Gregori-Giralt, E. (2018). Rubrics for developing students' professional judgement: A study of sustainable assessment in arts education. *Studies in Educational Evaluation*, 58, 70-79.
- Nation, I. S. P., & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. Routledge.
- Nazara, S. (2011). Students' perception on EFL speaking skill development. *JET (Journal of English Teaching)*, 1(1), 28-43.
- Nkhoma, C., Nkhoma, M., Thomas, S., & Le, N. Q. (2020). The role of rubrics in learning and implementation of authentic assessment: A Literature review. In M. Jones (Ed.), *Proceedings of*

- InSITE 2020: Informing Science and Information Technology Education Conference* (Vol. 4606, pp. 237-276). Informing Science Institute.
- Orlich, D. C., Harder, R. J., Callahan, R. C., Trevisan, M. S., & Brown, A. H. (2010). *Teaching strategies: A guide to effective instruction* (9th ed.). Cengage Learning.
- Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Newbury House Publishers.
- Park, H., & Lee, A. R. (2005, August). L2 learners' anxiety, self-confidence and oral performance. In *Proceedings of the 10th Conference of Pan-Pacific Association of Applied Linguistics* (pp. 107-208).
- Phan, T. X., & Phuong, H. Y. (2017). Using analytic rubric for speaking self-assessment: EFL students' perceptions and challenges. *IOSR Journal of Research & Method in Education*, 7(3), 34-39.
- Pui, P., Yuen, B., & Goh, H. (2021). Using a criterion-referenced rubric to enhance student learning: A case study in a critical thinking and writing module. *Higher Education Research & Development*, 40(5), 1056-1069.
- Quinlan, A. M. (2012). *A complete guide to rubrics: Assessment made easy for teachers of K-college* (2nd ed.). R&L Education.
- Rahimi, M. (2005). An investigation into the factors affecting the use of language learning strategies by Persian EFL learners. [Unpublished doctoral dissertation]. Shiraz University, Shiraz, Iran.
- Ricoeur, P. (1996). On interpretation. In R. Kearney & M. Rainwater (Eds.), *The continental philosophy reader* (pp. 138-56). Psychology Press.
- Rivers, W. (1968). *Teaching foreign language skills*. University of Chicago Press.
- Römer, U. (2017). Language assessment and the inseparability of lexis and grammar: Focus on the construct of speaking. *Language Testing*, 34(4), 477-492.
- Shumin, K. (2002). Factors to consider: Developing adult EFL students' speaking abilities. *Methodology in language teaching: An anthology of current practice*, 12(35), 204-211.
- Strauss, A., & Corbin, J. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Sage Publications.
- Tahir, S. Z. A. (2015). Improving students' speaking skill through Yahoo messenger at University of Iqra Buru. *International Journal of Language and Linguistics*, 3(3), 174-181.
- Thornbury, S. (2005). *How to teach speaking*. Longman.
- Tok, H. (2009). EFL Learners' communication obstacles. *Electronic Journal of Social Sciences*, 8(29), 84-100.
- Tuan, N. H., & Mai, T. N. (2015). Factors affecting students' speaking performance at LE Thanh Hien High School. *Asian Journal of Educational Research*, 3(2), 8-23.
- Ulker, V. (2017). The design and use of speaking assessment rubrics. *Journal of Education and Practice*, 8(32), 135-141.
- Walvoord, B. E. (2010). *Assessment clear and simple: A practical guide for institutions, departments, and general education*. John Wiley & Sons.
- Wigglesworth, G. & Frost, K. (2017). Task and performance-based assessment In E. Shohamy., I. G. Or., & S. May (Ed.), *Language Testing and Assessment* (3rd ed). Springer.
- Zakian, M. (2021). The effects of recast on the speaking accuracy and willingness to communicate (WTC) of Iranian EFL learners across gender. *International Journal of English Language and Translation Studies*, 9(1), 75-80.
- Zeinivand, T., Azizifar, A., & Gowhary, H. (2015). The relationship between attitude and speaking proficiency of Iranian EFL learners: The case of Darrehshehr city. *Procedia-Social and Behavioral Sciences*, 199, 240-247.
- Zuhriyah, M. (2017). Storytelling to improve students' speaking skill. *English Education: Jurnal Tadris Bahasa Inggris*, 10(1), 119-134.